American Psychological Association

Journal of Experimental Psychology: General

ISSN: 0096-3445

2020, Vol. 149, No. 6, 1097–1115
http://dx.doi.org/10.1037/xge0000705

Autistic Adults Anticipate and Integrate Meaning Based on the Speaker's Voice: Evidence From Eye-Tracking and Event-Related Potentials

Mahsa Barzy, Jo Black, David Williams, and Heather J. Ferguson
University of Kent, Canterbury

Typically developing (TD) individuals rapidly integrate information about a speaker and their intended meaning while processing sentences online. We examined whether the same processes are activated in autistic adults and tested their timecourse in 2 preregistered experiments. Experiment 1 employed the visual world paradigm. Participants listened to sentences where the speaker's voice and message were either consistent or inconsistent (e.g., "When we go shopping, I usually look for my favorite wine," spoken by an adult or a child), and concurrently viewed visual scenes including consistent and inconsistent objects (e.g., wine and sweets). All participants were slower to select the mentioned object in the inconsistent condition. Importantly, eye movements showed a visual bias toward the voice-consistent object, well before hearing the disambiguating word, showing that autistic adults rapidly use the speaker's voice to anticipate the intended meaning. However, this target bias emerged earlier in the TD group compared to the autism group (2240 ms vs. 1800 ms before disambiguation). Experiment 2 recorded ERPs to explore speaker-meaning integration processes. Participants listened to sentences as described above, and ERPs were time-locked to the onset of the target word. A control condition included a semantic anomaly. Results revealed an enhanced N400 for inconsistent speaker-meaning sentences that was comparable to that elicited by anomalous sentences, in both groups. Overall, contrary to research that has characterized autism in terms of a local processing bias and pragmatic dysfunction, autistic people were unimpaired at integrating multiple modalities of linguistic information and were comparably sensitive to speaker-meaning inconsistency effects.

Keywords: spoken language comprehension, pragmatics, visual world paradigm, event related brain potentials, autism

The process of inferring meaning from language is strongly influenced by the wider context, including verbal frame, tone of voice, gestures, and body language, and therefore falls within the pragmatics domain of language processing (Martin & McDonald, 2003). Pragmatic language use has recently been conceptualized within an extended account of situated language processing, known as the "social Coordinated Interplay Account" (sCIA; Münster & Knoeferle, 2018). This account proposes that characteristics of both the comprehender and speaker, including their mood, education level, and social stereotypes, are taken into account online when interpreting language (Rodríguez, Burigo, & Knoeferle, 2016; Van Berkum, De Goede, Van Alphen, Mulder, & Kerstholt, 2013; Van Berkum, Van den Brink, Tesink, Kos, & Hagoort, 2008). Hence, the social context is integrated with linguistic input in real-time when we process language. A much-debated question remains *when* these characteristics (context dependent) and the sentence's message (i.e., meaning of individual words, context free) are integrated to extract meaning, and which cognitive and social mechanisms underpin these processes.

Early research in this area postulated that individuals first extract the sentence's message using syntax and semantics, and only refer to pragmatics to integrate the speaker's identity at a later stage of processing (Cutler & Clifton, 1999; Lattner & Friederici, 2003; Osterhout, Bersick, & McLaughlin, 1997). For example, Lattner and Friederici (2003) recorded event related brain potentials (ERPs) while participants listened to sentences in which the gender of the speaker either matched or mismatched the meaning of the sentence in its usual/prototypical context (e.g., "I like to wear lipstick" spoken by a female or male). They observed a posterior P600 effect when the speaker gender and sentence meaning mismatched. This posterior P600 effect has been interpreted as a marker for the detection of pragmatic violations (i.e., reintegrating information in the presence of an inconsistency between pragmatics and meaning inferences; Osterhout et al., 1997; Spotorno, Cheylus, Van Der Henst, & Noveck, 2013), and is distinct from the more widespread centrally distributed P600 component that is typically elicited by syntactic violations (Gouvea, Phillips, Kazanina, & Poeppel, 2010). Indeed, Lattner et al. associated this late

This article was published Online First November 11, 2019.
Mahsa Barzy, Jo Black, David Williams, and Heather J. Ferguson, School of Psychology, Keynes College, University of Kent, Canterbury.
All analysis procedures were pre-registered, and the full experimental materials, datasets and analysis scripts are available on the Open Science Framework web pages (see https://osf.io/7hna3/).
This work was carried out with the support of a grant to Heather J. Ferguson and David Williams from the Leverhulme Trust (Ref: RPG-2014–298). The authors would also like to thank Jos Van Berkum for sharing his materials from Van Berkum et al. (2008), which inspired the development of items for our experiments.
Correspondence concerning this article should be addressed to Heather J. Ferguson, School of Psychology, Keynes College, University of Kent, Canterbury, Kent CT2 7NP, England, United Kingdom. E-mail: h.ferguson@kent.ac.uk

posterior positivity (in the absence of any earlier effects in the N400 range) with participants using pragmatics at a later stage to integrate the speaker-related information (i.e., after processing the sentence's message), and thus concluded that it supports the two-step account. However, these conclusions are somewhat limited by design features of the task, including an absence of filler sentences with syntactic violations, which could have provided a baseline measure of a syntactic P600 to contrast with the pragmatic P600 effect reported here. In addition, the gender stereotype violations were always sentence-final, meaning that the speaker-meaning effects were likely to be influenced by more global "wrap up effects" (i.e., an increase in processing time at sentence end due to semantic integration processes; Schacht, Sommer, Shmuilovich, Martíenz, & Martín-Loeches, 2014; Stowe, Kaan, Sabourin, & Taylor, 2018).

An alternative view has been proposed, which suggests that the linguistic input and context are processed in a single step ("one step model"), as a joint action (Clark, 1996; Perry, 1997). Clark proposes that the nonverbal cues provided by the linguistic context (e.g., gestures, body language etc.) are processed in parallel with the linguistic input. This one step account is supported by empirical evidence from Van Berkum et al. (2008), who recorded ERPs while participants listened to sentences in which speaker and meaning were either consistent or inconsistent. In Van Berkum et al.'s study, speaker voices were manipulated in three ways: (a) age: child versus adult (e.g., "I cannot sleep without my <u>teddy</u> in my arms"), (b) social class: lower versus higher class accent (e.g., "I have a large <u>tattoo</u> on my back"), and (c) gender: male versus female (e.g., "On weekends I usually go <u>fishing</u> by the river"). Note that critical words (underlined in the above examples) were always presented midsentence, which allowed sufficient time for participants to infer the speaker's characteristics and avoided wrap-up effects. Van Berkum and colleagues examined effects on the N400 ERP component; a centroparietal negative-going deflection that is sensitive to stimulus predictability and semantic integration processes (Kutas & Hillyard, 1980; Nieuwland et al., 2018). Results revealed a larger N400 effect for inconsistent compared to consistent sentences, with effects emerging as early as 200 ms after the onset of the critical word, thus showing that speaker-related information is integrated at an early stage. These findings therefore support the one-step model of language processing, by demonstrating that interpretation of the sentence meaning is influenced concurrently by inferences about the speaker characteristics and the explicit message (i.e., "who is saying what").

The rapid influence of social pragmatic information on meaning was subsequently replicated by Van den Brink and colleagues (2012), using ERPs. Importantly, Van den Brink et al. revealed that social information processing was enhanced among people who self-reported high levels of empathy, using the Empathizing Questionnaire (Baron-Cohen & Wheelwright, 2004). In contrast, people who self-reported low levels of empathy were consistently impaired in using information about social stereotypes during sentence comprehension. This pattern is consistent with previous research showing that high empathizers are better at predicting other people's actions and responding to them appropriately (Saxe & Baron-Cohen, 2006). Moreover, it suggests that pragmatic processing can be influenced by individual preferences for bottom-up (i.e., language first) or top-down (i.e., rapid integration of voice-based information) language processing.

All of the issues discussed so far are relevant for our understanding of autism. Autism spectrum is a developmental disorder, diagnosed on the basis of behavioral difficulties in social communication, and restricted and repetitive behaviors/interests (American Psychiatric Association, 2013; Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997; Shah & Frith, 1993). Some researchers have proposed that the ability to empathize with others is impaired among autistic people[1] (Baron-Cohen & Wheelwright, 2004), however this finding has been challenged more recently by evidence that the ability to deploy empathizing abilities depends on the context. Thus, autistic people do not lack empathy but they may experience a specific difficulty empathizing with TD individuals (and vice versa), since the two groups have different world experiences (Milton, Heasman, & Sheppard, 2018; Nicolaidis, Milton, Sasson, Sheppard, & Yergeau, 2018).

Importantly, communication difficulties in autism are separable from basic language impairments; semantic language comprehension and syntactic preferences seem to be relatively spared among high functioning autistic individuals (e.g., Allen, Haywood, Rajendran, & Branigan, 2011; Hopkins, Yuill, & Keller, 2016; Howard, Liversedge, & Benson, 2017a; Tager-Flusberg & Joseph, 2003). However, some studies have shown that even when structural language skills are intact autistic people show deficits in processing linguistic information *in context* (i.e., successfully extracting the intended meaning), including difficulty using the sentence context to distinguish homographs (e.g., pronouncing *tear* in, "In her eye/dress there was a big tear," Frith & Snowling, 1983) or process nonliteral utterances (e.g., "He drew a gun," where the verb could mean drawing or pulling out, Jolliffe & Baron-Cohen, 1999; see also Connolly, 2001; Deliens, Papastamou, Ruytenbeek, Geelhand, & Kissine, 2018; Vulchanova, Saldaña, Chahboun, & Vulchanov, 2015). The validity and generalizability of these context impairments, however, have been questioned in recent years (e.g., Brock & Bzishvili, 2013; Brock & Caruana, 2014; Hahn, Snedeker, & Rabagliati, 2015). Moreover, eye-tracking research has revealed that autistic adults are delayed relative to age and IQ-matched TD peers in detecting passage level anomalies in text (i.e., where global coherence is required; Au-Yeung, Kaakinen, Liversedge, & Benson, 2018), and in detecting implausible words in a sentence (Howard, Liversedge, & Benson, 2017b). These findings suggest that subtle differences may exist in the speed with which context is accessed and influences language processing in autism (cf. Black, Barzy, Williams, & Ferguson, 2019; Black, Williams, & Ferguson, 2018; Ferguson, Black, & Williams, 2019).

Traditionally, these pragmatic deficits have been linked to general difficulties integrating information in context (known as "weak central coherence," WCC; Booth & Happé, 2010; Frith, 1989; Martin & McDonald, 2003), given that autistic people tend to show a local, rather than global, processing bias (Frith, 1989; Frith & Happé, 1994; Happé & Frith, 2006). In turn, atypical attention distribution in autism (i.e., allocating attention to details and ignoring the context) has been attributed to impaired metalearning abilities (known as the "predictive coding theory of autism" or the "Bayesian brain"; Van Boxtel & Lu, 2013; Van de Cruys et

[1] We acknowledge recent debates about the terminology used to describe autism, and in this article adopt the identity-first language preferred by autistic adults who took part in the study by Kenny et al. (2016).

al., 2014), which disrupts the ability to distinguish between important and less important prediction errors. These weaker priors mean that autistic individuals struggle to contextualize sensory input and make predictions based on experience, which is likely to affect many aspects of cognition, including language, memory, emotions, and motor skills (Pellicano & Burr, 2012). These weaker expectations of how people behave therefore mean that autistic people find it harder to process social information during communication and are likely to show delays generating appropriate responses. Despite these converging accounts, there is little agreement on how a detail-focused cognitive style and weaker predictive processing style might influence the quality of social interactions. This raises the question of whether the mechanisms involved in *integrating* social pragmatic information and language meaning are disrupted among autistic individuals who experience impaired use of context and atypical social inferencing. This is an important topic to investigate, because as well as further informing theoretical models of pragmatic language comprehension and shedding light on the nature of these social impairments, it has the potential to help practitioners develop specific interventions or learning shortcuts to improve the quality of social interactions in autism.

In this article, we present two fully preregistered experiments that used eye-tracking (Experiment 1) and ERP (Experiment 2) methods to investigate whether and how real-time pragmatic processing of spoken language is affected when global coherence and social abilities are compromised. Specifically, we tested whether autistic adults differ significantly from matched neurotypical controls in the timecourse with which they anticipate meaning based on a speaker's characteristics (i.e., their age, gender or social status), and whether they manifest equivalent disruptions during language integration when speaker and meaning information are inconsistent.

Experiment 1 examined the timecourse with which listeners *predict* meaning based on characteristics inferred from the speaker's voice. We used the classic visual world paradigm to address this question by recording participants' eye movements around a visual scene that contained images depicting objects/events that were consistent or inconsistent with the speaker's voice (Cooper, 1974; Tanenhaus Spivey-Knowlton, Eberhard, & Sedivy, 1995). The visual world paradigm has been used extensively in psycholinguistic research to show that participants incorporate cues from syntax, semantics and world knowledge to constrain the available set of objects, and move their eyes to an appropriate visual object *before* it has been mentioned in the audio (e.g., Altmann & Kamide, 2007, 2009; Kamide, Lindsay, Scheepers, & Kukona, 2016). For example, it has been shown that participants are more likely to look at an empty glass of wine compared to a full glass of beer when hearing the sentence "the man has drunk all of. . . ," and vice versa for "the man will drink all of . . ." (Altmann & Kamide, 2007). This paradigm therefore provides a valuable implicit measure of expectation in real-time, though it has never before been used to examine the timecourse with which listeners infer meaning from a speaker's voice characteristics. In the current study, we tested whether participants' predictive eye movements toward visual objects (e.g., a shaver vs. a car) were modulated by inferences from the speaker's voice (e.g., whether an adult vs. child said, "On my last birthday, I got an expensive electric . . ."). This paradigm enabled us to examine for the first time whether and how autistic adults implicitly integrate pragmatic cues to predict meaning, and how these processes compare to those engaged by age, IQ, and gender matched TD adults. Participants' explicit ability to infer meaning from a speaker's voice was measured using the "Reading the Mind in the Voice" task (Golan, Baron-Cohen, Hill, & Rutherford, 2007; Rutherford, Baron-Cohen, & Wheelwright, 2002), and their local/global processing bias was measured using a sentence completion task (Booth & Happé, 2010).

Experiment 2 sought to explore the timecourse with which listeners *integrate* semantic and pragmatic cues, and respond to inconsistencies in speaker and meaning. To this end, we replicated Van Berkum et al.'s (2008) study, using ERPs to compare the brain's electrophysiological responses to words that were consistent or inconsistent with characteristics inferred from the speaker's voice (e.g., "I cannot sleep without my teddy in my arms," spoken by a child or an adult), among adult participants with and without autism. In addition, we extended the paradigm to include a semantic anomaly condition using the same content (e.g., "I cannot sleep without my pizza in my arms"), that provided a baseline measure of anomaly detection N400 responses in each participant group. The addition of this semantic anomaly condition serves to overcome the possible limitation of Van Berkum et al.'s study, which tested the N400 effect to semantic anomalies in a completely different set of sentences.

First, if the linguistic input and context are processed in a single parallel step, we expected TD participants in Experiment 1 to initiate anticipatory eye movements toward the image that was consistent with the speaker's voice long before the disambiguating target word was uttered (e.g., shaver/car). In Experiment 2, we predicted an enhanced N400 effect for inconsistent sentences relative to consistent ones, which would be comparable in timecourse to the N400 elicited by semantically anomalous sentences. In contrast, a two-step account would predict that effects of pragmatic fit would be delayed, as lexical-semantic fit would be prioritized in the early stages in processing.

Second, we considered how these processes may be influenced among autistic people, and compared predictions for accounts that characterize autistic people as having a general deficit in contextual integration (e.g., Behrmann, Thomas, & Humphreys, 2006; Happé & Frith, 2006; Koldewyn, Jiang, Weigelt, & Kanwisher, 2013), with the predictions of accounts that imply global integration ability is not universally impaired in autism (e.g., Mottron, Burack, Iarocci, Belleville, & Enns, 2003; Plaisted, Dobler, Bell, & Davis, 2006; Van der Hallen, Evers, Brewaeys, Van den Noortgate, & Wagemans, 2015; following the results of Black et al., 2018, 2019; Ferguson et al., 2019). Based on the former, we predicted that in Experiment 1 autistic individuals would be slower than TD individuals to direct anticipatory gaze to the speaker-relevant image and would experience greater interference from the semantic competitor (i.e., a weaker target bias). In Experiment 2, we predicted that the autism group would show a delayed, reduced, or absent N400 response when integrating inconsistent speaker-meaning information. Alternatively, if pragmatic processing is largely spared in autism (as it appears to be for semantic processing), then no between-groups differences in the anticipation or integration of social pragmatic meaning should emerge.

Experiment 1

All methodological procedures were preregistered on the Open Science Framework (OSF) web pages (see https://osf.io/7hna3/).

Method

Participants. Participants, including those with and without autism were recruited using the Autism Research at Kent (ARK) database. Participants on the database were initially recruited from a community sample in the areas of Kent, Essex and London in the U.K., using a variety of recruitment strategies (e.g., newspaper adverts, contacting local groups, autism support groups and word-of-mouth). We deliberately avoided using university students to minimize differences in socioeconomic status between the groups. A total of 50 adult participants were initially recruited, but two were excluded from both experiments: one due to technical errors during EEG recording, and one due to excessive noise during EEG recording (i.e., >25% data loss). Hence, both Experiments 1 and 2 included 24 autistic adults and 24 TD adults, which is in accordance with our preregistered sample size. These sample sizes were chosen a priori based on the sample size used in Van Berkum et al.'s study (2008; $N = 24$), and to be comparable or even exceed the sample sizes used in previous research that has examined eye movements in autistic and TD adults (e.g., Au-Yeung, Kaakinen, & Benson, 2014, 2018; Black et al., 2018, 2019; Brock, Norbury, Einav, & Nation, 2008; Ferguson et al., 2019; Howard et al., 2017a, 2017b). Post hoc calculations of power were conducted given the current sample size using the simr package in R (Green & MacLeod, 2016), and returned an estimated power of 100% with the significance level of $\alpha = .05$ on 80% of occasions (as suggested by Cohen, 1988) for Experiment 1.

Groups were matched on age, verbal IQ[2] and gender (as measured by the Wechsler Abbreviated Scale of Intelligence; WASI; Wechsler, 1999; see Table 1 for demographic information), were native English speakers, and did not have a diagnosis of dyslexia or reading comprehension impairment. Participants in the TD group did not report any current psychiatric diagnoses. All participants completed the Autism-spectrum Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) to measure self-reported autistic traits.

In accordance with *DSM–IV* or 5, all autistic participants had a formal diagnosis of Autistic Disorder, Asperger's Syndrome, or Pervasive Developmental Disorder Not-Otherwise Specified (American Psychiatric Association, 2013). To assess the current autistic characteristics, all the autistic participants were also assessed on module 4 of the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) by a trained, research-reliable researcher, and videos were double coded to ensure reliability of scoring (see Table 1; interrater reliability was found to be excellent with intraclass correlation of .89). Eleven individuals in the autistic group scored higher than 7 on the ADOS (i.e., the cut off score).

Materials.

Eye-tracking task. Twenty-four experimental sentences were created based on those used in Van Berkum et al. (2008). Each item described a person's preferences, or activities. The final word in each sentence was manipulated across two conditions so that the lexical content either matched a specific speaker's stereotypical characteristics or not (speaker-consistent vs. speaker-inconsistent). For example, the sentence "When we go shopping, I usually look for my favorite sweets" is consistent with social stereotypes for a child, but the sentence "When we go shopping, I usually look for my favorite wine" is inconsistent with expectations for a child. Each experimental sentence was recorded by two contrasting speakers, resulting in four versions of each item, with social stereotypes manipulated in three ways- 1) Age: child versus adult (see above example), 2) Class: higher versus lower class accent (e.g., "I never smoke inside, because my wife doesn't like the smell of cigars/rollies), 3) Gender: female *vs* male (e.g. "Before starting my new job, I need to buy a new skirt/tie"). Twenty-four filler sentences were also included (e.g., "It was Valentine's Day so I bought her a bunch of red roses"), which didn't include any inconsistent content.

Ten different speakers were recruited to record the sentences. One female and one male adult speaker read eight items in the 'gender' category (four sentences per item). Two children (one female and one male, aged 6 and 8, respectively) and two adults (one female and one male) read eight items in the 'age' category. Finally, four professional actors (2 females and 2 males) were recruited from local drama groups to read eight items in typically high or low socioeconomic British accents for the 'class' category. Audios were recorded in a soundproof room using a digital voice recorder. One female and one male adult speaker read the filler sentences (12 sentences each). All speakers were native speakers of English.

To verify that listeners inferred the intended social stereotypes from speaker's voices, we conducted a posttest, in which 22 TD participants (10 males, 12 females) listened to each item then used a 5-point sliding scale to rate "how normal or strange do you think it is to have the speaker say this particular thing" (1 = *completely normal,* 5 = *very strange*). Overall, inconsistent speaker-meaning combinations were rated as significantly more strange than consistent speaker-meaning combinations ($M = 2.39$ vs. 1.76, $t(77) = 10.22$, $p < .001$). In addition, we tested the effect of consistency separately for each speaker type (i.e., age, gender and class). This analysis confirmed that inconsistent speaker-meaning combinations were rated as significantly more strange than consistent speaker-meaning combinations in all three speaker categories: Age, $t(77) = 9.09$, $p < .001$, Class, $t(77) = 2.62$, $p = .011$, and Gender, $t(77) = 8.23$, $p < .001$.

Each of the 24 experimental sentences was paired with an image that depicted four different objects (see open materials on OSF, https://osf.io/7hna3/). Two objects in each image were semantically relevant to the sentence (e.g., edible objects for the supermarket example). One of these was consistent with social stereotypes about the speaker (subsequently referred to as the target picture, e.g., a picture of 'sweets' when the sentence was read by a child), and the other was inconsistent with social stereotypes about the speaker (subsequently referred to as the competitor picture, e.g., a picture of 'wine' when the sentence was read by a child). The remaining two pictures depicted distractor objects that were irrelevant to the sentence content (e.g., a house, a lake). Filler

[2] Note that the autistic group scored significantly higher on PIQ. Therefore, in addition to the full-sample analyses, we ran analyses among subsamples of autistic and TD participants that were matched for PIQ (by excluding one participant from each group with the highest and lowest PIQ scores). Crucially, none of the statistical results from the experimental task changed substantively with this smaller matched sample (i.e. no p value changed from significant to non-significant or vice versa).

Table 1
Demographic Information (Means and Std. Errors) of Participants in Each Group

Variable	Autistic (n = 24)	Typically developing (n = 24)	t-value	p-value	Cohen's d
Sex (m:f)	18:6	18:6	—	—	—
Age (years)	32.58 (2.23)	31.75 (2.21)	.27	.792	.08
Verbal IQ	105.46 (2.51)	101.46 (1.80)	1.29	.202	.37
Performance IQ	112.75 (3.84)	102.29 (2.36)	2.32	.025*	.67
Total AQ	30.92 (1.75)	18.05 (1.64)	5.35	<.001***	1.58
ADOS2 Module4	7.79 (.99)	—	—	—	—

Note. ADOS = autism diagnostic observation schedule; AQ = autism-spectrum quotient.
* $p < .05$. *** $p < .001$.

items were also paired with images that included four pictures, but only one picture matched the lexical content of the sentence (e.g., red roses in the example above). Each individual picture measured 400x400 pixels, with the complete image comprising four pictures on a white background measuring 960x720 pixels, with the position of target, competitor and distractor pictures counterbalanced across items.

Revised Reading the Mind in the Voice task (RMIV). Participants' explicit recognition of meaning from voices was assessed using the RMIV task. In this task, developed by Golan et al. (2007), participants listened to 25 different excerpts of speech and had to judge how each person was feeling (only based on their voice) from a choice of four options (e.g., "angry, derogatory, resentful, or nostalgic"). There was no time limit for participants to respond, although they were encouraged to respond as quickly as they could. Participant's accuracy was recorded.

Linguistic central coherence task. Participants' local processing bias during language processing was measured using a sentence completion task. In this task, participants were asked to complete 14 sentences that required global sentence completions. For example, the sentence fragment, "in the sea there are fish and. . . ." could be completed with a locally biased word "chips," or with a globally biased word like "sharks" or "crabs." Participants' responses and their RTs were recorded.

Procedure. The Psychology Research Ethics Committee at the University of Kent granted approval to conduct this study. For the eye-tracking task, participants' dominant eye was tracked with an EyeLink 1000 Plus eye-tracker and participants listened to the sentences through headphones. Head movement was minimized with the use of a fixed chin rest. Images were presented on a VDU approximately 70 cm in front of the participants' eyes. Calibration was performed using a 9-point procedure. Before each trial, a central drift correction was conducted to verify the calibration accuracy. Participants were asked to listen to each sentence and look at the images, and used the mouse to click on the picture that was mentioned in the sentence as quickly as possible. Images appeared on screen 1000 ms before the onset of related audio and stayed onscreen until the participant clicked the mouse to move on. Participants' picture selection accuracy, RTs (time-locked to the onset of the target picture), and eye-movements across the whole trial were recorded. The next trial began following a 500 ms blank screen. The first two items were filler trials to ensure participants understood the task. Following presentation of these, the 24 experimental items were randomly interleaved with 22 filler items, with a break offered half way through. Participants saw each item once, in one of the four conditions. Item order and condition was randomized across four lists, and the presentation of each list was randomized among participants. Each participant completed the eye-tracking and RMIV tasks on the same day as the EEG task reported in Experiment 2. The whole testing session took about 2 hr including EEG setup and breaks.

Results

All analysis procedures were preregistered, and the full experimental materials, data sets and analysis scripts are available on the Open Science Framework web pages (see https://osf.io/7hna3/).

RMIV task. Accuracy scores were analyzed using a generalized linear mixed model, using the lme4package in RStudio software Version 1.1.453 (Bates, Mächler, Bolker, & Walker, 2015; R Core Team, 2016). Group (autistic vs. TD) was included in the model as a fixed effect and was contrast coded: (–.5 vs. .5). We applied the maximal random effects structure, by including participants and items as random effects, and Group as a random slope on items (as suggested by Barr, Levy, Scheepers, & Tily, 2013). The analysis revealed that autistic participants were significantly less accurate at explicitly recognizing speakers' emotions based on their voice compared to TD participants (M = 65% vs. M = 70%; Est = .41, SE = .20, z = 2.07, p = .038).

Linguistic central coherence task. Similar to Booth and Happé (2010), a 3-point scoring system was used to analyze responses. Two points were given if participants provided a global sentence completion word/phrase within 10 s, and 1 point was assigned if they took longer than 10 s or provided no response. If they used a local sentence completion word/phrase, then 0 points were assigned. Response scores were analyzed using a linear mixed model. Group (autistic vs. TD) was included in the model as a fixed effect and was contrast coded: (–.5 vs. .5). The maximal random effects structure included participants and items as random effects, and Group as a random slope on items. The analysis revealed no difference between groups in terms of global/local sentence completion bias (autistic vs. TD; M = 1.75 vs. M = 1.73; $Est.$ = −.02, SE = .10, t = −.21, p = .832).

Eye-tracking task.

Accuracy. Accuracy of selecting the mentioned picture was analyzed using a generalized linear mixed model, with Group (autistic vs. TD) and Condition (consistent vs. inconsistent) as contrast coded fixed effects (−.5 vs. .5). The maximal random effects structure that fit the data included participants and items as random effects, with Condition as a random slope on items and participants. Participants were highly accurate at choosing the mentioned picture (autistic vs. TD, $M = 97\%$ vs. 98%), and this did not differ between groups (*Est.* = 1.57, $z = 0.82$, $p = .412$) or conditions (*Est.* = −7.56, $z = -1.29$, $p = .196$).

Reaction times. Only trials on which participants accurately clicked on the mentioned object were included in the analysis. In addition, response times that fell more than 2.5 standard deviations from the individual's mean reaction time (RT) were excluded from analysis. These steps removed 4.25% of the original data. Statistics were performed using a linear mixed model, including the same fixed effects structure as the accuracy analysis, and the maximal random effects structure to fit the data (Group and Condition as random slopes on items, and Condition as a random slope on participants). Mean response times per condition are shown in Figure 1.

Results showed that participants were faster to select the mentioned object when the speaker characteristics were consistent with the mentioned object than when the speaker characteristics were inconsistent (M = 1572 ms vs. 1729 ms; *Est* = 158.81, *SE* = 45.81, $t = 3.47$, $p = .002$). Reaction times did not differ by Group (*Est* = −137.02, $t = .80$, $p = .427$), nor did Group modulate the effect of Consistency (*Est* = −10.77, $t = .14$, $p = .888$).

Eye movement data processing. Eye movements were time-locked to the onset of the sentence-final disambiguating word (e.g., 'sweets' or 'wine'), and were analyzed in two separate time periods: anticipatory period (eye movements in the 3000 ms *before* disambiguating word onset, reflecting listeners' expectations about forthcoming language input) and integration period (eye movements in the 1000 ms *after* disambiguating word onset, reflecting the ease with which incoming language is integrated with expectations). Four areas of interest (AOIs) were defined around the pictures of objects in each visual scene: target (the object that matched both the semantic context of the sentence and the speaker's voice), competitor (matched the semantic context but not the speaker's voice), and two distractors (did not match either semantic context or the speaker's voice).

Eye movements during the anticipation period were analyzed across consistency conditions, since listeners had not yet heard the consistent/inconsistent critical word, so expectations should be solely driven by inferences from the speaker's voice. Thus, anticipatory analyses tested whether participants in each group differed in their likelihood of fixating the speaker-relevant target picture or speaker-irrelevant competitor picture, and whether these preferences emerged over a different time course for each group. To fulfil this aim, fixations during the 3000 ms anticipatory period were broken down into 20 ms time bins, and the spatial coordinates were mapped onto AOIs as a function of time. Visual preferences to target or competitor pictures were represented by a binary term in each 20 ms time bin, where "1" indicated a fixation on the target/competitor and "0" indicated no fixation. The resulting data was analyzed separately for target and competitor biases using generalized mixed models and growth curve analysis (Mirman, Dixon, & Magnuson, 2008), using the "lme4"and "eyetrackingR" packages in RStudio. We note that our preregistration proposed to analyze the probability of fixating the target and competitor images as a function of time using permutation and cluster analysis and did not specify the use of growth curve analysis. We chose to use growth curve analysis to examine anticipatory effects of linguistic context (permutation and cluster analyses were used to examine integration, as detailed below) following more recent statistical norms in the field. Fitting models to the data to test different shapes of visual bias over time allows us to capture effects of group as the sentence unfolded, while also testing for variance between and within individuals. In this study, third-degree orthogonal polynomials, incorporating intercept, linear, quadratic and cubic components, were used to model the time-course of anticipatory bias over the 3000 ms period (see Mirman et al., 2008). Thus, final models included a contrast coded fixed effect for Group (–.5 vs. .5) alongside the time polynomials, and random effects of participants and items. The final model also included Group as a random slope within items. Resulting statistical effects are shown in Table 2.

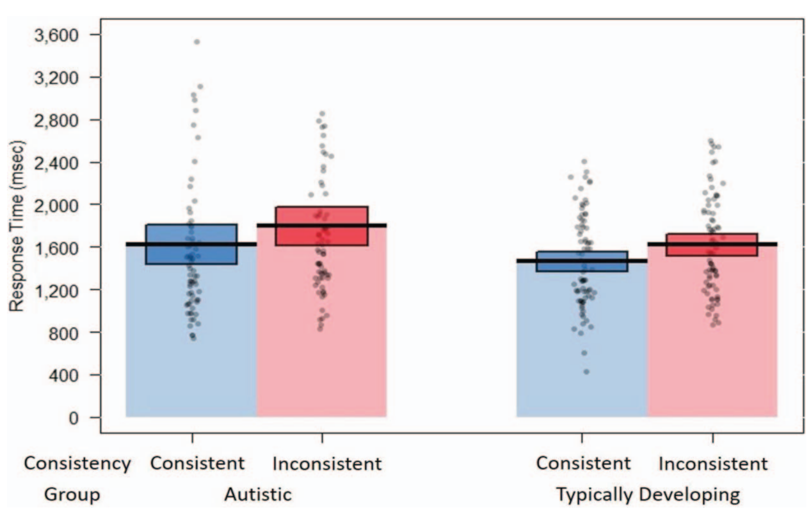

Figure 1. Target selection response times for each condition and group, Experiment 1, showing raw data points, a horizontal line reflecting the condition mean, and a rectangle representing the Bayesian highest density interval. See the online article for the color version of this figure.

Table 2
Statistical Results From the Growth Curve Analysis Examining Anticipatory Fixations Toward the Target and Competitor Objects in Experiment 1

	Target			Competitor		
Variable	*Est.*	*SE*	*z-value*	*Est.*	*SE*	*z-value*
Group	.06	.08	.77	−.04	.09	−.39
ot1	2.11	.07	28.45***	1.51	.11	13.49***
ot2	.49	.07	6.54***	−.44	.11	−3.96***
ot3	.47	.07	6.30***	−.03	.11	−.24
Group × ot1	.2	.15	1.35	−.96	.15	−6.27***
Group × ot2	.06	.15	.4	.18	.15	1.2
Group × ot3	−.64	.15	−4.28***	−.26	.15	−1.73

Note. Ot1, ot2, and ot3 refer to linear, quadratic, and cubic models of time, respectively.
*** $p < .001$.

Follow-up analyses explored *whether* and *when* anticipatory biases to the target or competitor picture exceeded chance level (i.e., .25) for each group. Thus, we ran cluster-based permutation analysis by participants (Maris & Oostenveld, 2007) to compare the proportion of target or competitor fixations during the anticipatory period to chance, using the "eyetrackingR" package in RStudio. First, we computed a 1-sample test statistic for each of the 20 ms timebins, comparing each sample to chance (.25). Next, we clustered together adjacent timebins for which the test statistic was significant at the .05 level, and calculated a cluster-level test statistic as the sum of the test statistics for the individual timebins within a particular cluster. Finally, a simulation with 2000 randomly permuted samples was run to determine the likelihood of obtaining a significant cluster by chance. Permutation analyses included random effects for participants.

Eye movements during the integration period examined *when* participants in each group identified the consistent/inconsistent word, and how quickly they were able to switch their attention away from the target image to the competitor image in the inconsistent condition. To this end, fixations during the 1000 ms integration period were broken down into 20 ms time bins, and the spatial coordinates were mapped onto AOIs as a function of time. Visual preferences to target or competitor pictures were represented by a binary term in each 20 ms time bin, where '1′ indicated a fixation on the target/competitor and '0′ indicated no fixation. The resulting data was analyzed separately for each group, and for target and competitor biases, using a similar cluster-based permutation analysis approach to that described for the anticipation period. Crucially, here we used paired-samples *t* tests to compare the proportion of target or competitor fixations in each 20 ms sample between consistent and inconsistent conditions. This allowed us to identify when a significant difference in visual biases emerged between consistent and inconsistent conditions in each group. Permutation analyses included random effects for participants. Statistical effects for the permutation analyses, for both anticipatory and integration periods, are shown in Table 3.

Figure 2 plots the proportion of fixations to the target and competitor pictures in each group for every 20 ms time bin from 3000 ms before disambiguating word onset. Figure 3 plots the proportion of fixations to the target and competitor pictures in each consistency condition and group for every 20 ms time bin from disambiguating word onset until 1000 ms.

Anticipatory fixations toward target. As is clear in Figure 2, preference to fixate the target object increased over the 3000 ms anticipatory period, reflected in significant effects on the linear, quadratic and cubic fit curves. More importantly, Group significantly interacted with the cubic fit, revealing that while both groups clearly exhibited an increasing target preference prior to disambiguation, participants in the TD group exhibited shallower curvature- a slower rate of target bias increase-compared to the autistic group. Permutation tests confirmed that TD participants first showed a significant bias to fixate the target from 2240 ms before disambiguation (*SumT* = 115.40, $p < .001$), but this subsequently plateaued between 1500 ms and 1000 ms, then rapidly increased from 960 ms onward (*SumT* = 182.83, $p < .001$). In contrast, autistic participants showed a sustained and increasing bias to fixate the target from 1800 ms before disambiguation onward (*SumT* = 392.02, $p = .001$).

Anticipatory fixations toward competitor. A significant effect on the linear fit curve revealed that overall preference to fixate the competitor object increased over the 3000 ms anticipatory period. Importantly, this linear fit interacted significantly with Group, showing that the autistic group exhibited a steeper rise in looks to the competitor compared to the TD group. Permutation tests revealed that while TD participants never fixated the competitor above chance level during the 3000 ms anticipation period (all $ps > .05$), autistic participants showed a significant bias to fixate the competitor between 1000 ms and 300 ms before disambiguation (*SumT* = 22.31, $p = .001$).

Integration fixations toward target. The timecourse plots in Figure 3 reveals that looks to the target continued to rise when the mentioned object was consistent with the speaker characteristics but showed a steep decrease when the mentioned object was inconsistent with the speaker characteristics. Permutation analysis showed that in the autistic group, a significant difference in fixations toward the target emerged between consistent and inconsistent conditions from 300 ms after the disambiguating word (*SumT* = 90.68, $p < .001$). In contrast, the TD group showed this same effect from 400 ms after the disambiguating word (*SumT* = 102.81, $p < .001$).

Integration fixations toward competitor. The timecourse plots in Figure 3 reveals that looks to the competitor rose steeply when the mentioned object was inconsistent with the speaker characteristics but decreased when the mentioned object was consistent with the speaker characteristics. Permutation analysis revealed that a significant difference in fixations toward the competitor emerged between consistent and inconsistent conditions from 400 ms after the disambiguating word in both the autistic (*SumT* = −93.07, $p < .001$) and TD group (*SumT* = −105.70, $p < .001$).

Summary

The results of Experiment 1 revealed that participants in both groups accurately used the speaker's voice to anticipate the speaker's intended message. Participants were slower to select the mentioned object when it was inconsistent with the speaker's voice than when it was consistent with the speaker's voice. The influence of speaker expectations was also evident in the eye movement data as participants in both groups showed a strong and increasing

Table 3
Statistical Results From the Permutation t-Test Analyses Comparing Anticipatory and Integratory Biases Toward the Target and Competitor Objects to Chance in Experiment 1

	Target				Competitor		
Cluster No.	Start Time	End Time	SumT	Cluster No.	Start Time	End Time	SumT
				Anticipation			
Autistic							
1	−3000	−2940	7.13	1	−2200	−1900	8
2	−2300	−2160	17.46	2	−1500	−1200	8.02
3	−2100	−2080	2.03	3	−1000	−300	22.31**
4	−2060	−2000	6.9	4	−200	−100	2.15
5	−1960	−1940	2.32				
6	−1840	−1820	2.11				
7	−1800	0	392.02***				
Typically developing							
1	−2680	−2660	2.14	1	−1700	−1500	4.11
2	−2240	−1520	115.4***	2	−1100	−700	9.61
3	−1500	−1400	10.98	3	−600	−500	2.43
4	−1360	−1320	4.49	4	−200	0	5.38
5	−1020	−1000	2.28				
6	−960	0	182.83***				
				Integration			
Autistic							
1	300	1000	90.68***	1	400	1000	−93.07***
Typically developing							
1	400	1000	102.81***	1	400	1000	−105.7***

** $p < .01$. *** $p < .001$.

preference to fixate the object that was consistent with speaker's voice (i.e., the target) long before hearing the disambiguating word. Importantly, the nature and timing of these visual biases showed subtle differences between groups. Specifically, the target bias emerged earlier among participants in the TD group (TD: 2240 ms vs. autistic: 1800 ms prior to disambiguation), but showed shallower curvature, as the bias stalled before a final rapid increase from 960 ms before the disambiguation point. In contrast, participants in the autistic group showed a consistent steep increase in the visual bias toward the speaker-consistent object from 1800 ms before the disambiguation point. Interestingly, only the autistic group showed an above-chance bias to fixate the competitor during this anticipatory period. As expected, following the disambiguating word, participants in both groups made increasing fixations toward the mentioned object, regardless of whether it was consistent or inconsistent with the speaker expectations. As in the anticipatory period, some subtle differences emerged between groups; the autistic group were faster to switch *away from* the target in the inconsistent condition compared to the TD group (300 ms vs. 400 ms respectively). Both groups were equally fast to switch *to* the competitor in this inconsistent condition.

Taken together, these findings provide strong evidence that participants used the voice to infer characteristics of the speaker, and rapidly anticipated their intended meaning. This finding provides further evidence for the one step model of language processing by showing that the relevant knowledge and social context are processed hand-in-hand with semantics to facilitate language processing (Clark, 1996; Perry, 1997). The fact that these online voice-based inferences of meaning were generated by autistic adults is important because this is in contrast with several prominent theories of autism, including the WCC theory, suggesting that autism is associated with a tendency to process the information locally first and only later switching to global processing and using the context, including the social context (Booth & Happé, 2010; Frith & Happé, 1994). Nevertheless, the subtle differences in timing and strength of effects revealed by eye-tracking suggest that the autistic group had weaker speaker-meaning expectations, perhaps due to greater interference from the competitor or having weaker social stereotypes. In addition, TD participants' eye movements showed a clear cubic pattern of looks to the target over time (i.e., an increasing bias toward the target, followed by a plateau, then a final increase until disambiguation), though they never fixated the competitor object above chance. It is possible that this temporary reduction of the target bias reflects greater exploration of the visual scene and the irrelevant distractor objects among participants in the TD group compared to the autistic group (Heaton & Freeth, 2016).

In Experiment 2, we sought to further examine how people integrate these social contrasts using event-related potentials (ERPs). ERPs were recorded while participants listened to sentences that were either consistent or inconsistent with the speaker's characteristics. We predicted that there will be a larger N400 effect while participants hear the sentences in the inconsistent condition compared to the consistent one. In other words, they will show greater difficulty to integrate the sentence when there is a social contrast. We predicted that if autistic individuals have problems to integrate the information from context, then they would show less sensitivity while hearing these social contrasts (i.e., an absent or a reduced N400 effect in this group). We also included semantic

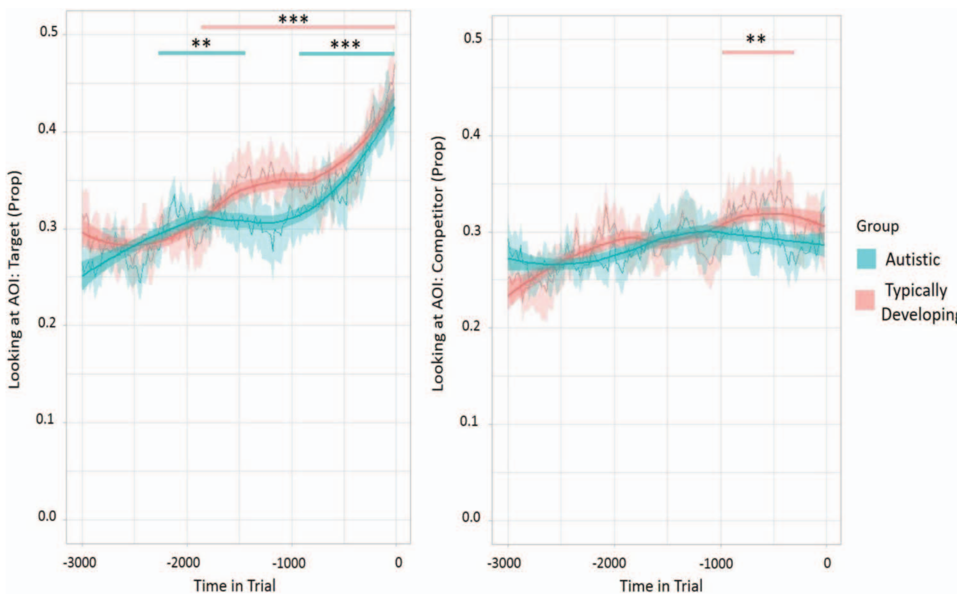

Figure 2. Timecourse of anticipatory fixations toward the target (left panel) and competitor (right panel) pictures for each group, in Experiment 1, showing the best fit curves for the data and 95% confidence interval shadow. Horizontal lines show clusters of time where the fixations toward the target exceeded chance (** $p <$ .01. *** $p < .001$). AOI = areas of interest. See the online article for the color version of this figure.

anomalous sentences as a baseline measure of the anomaly detection N400 effect.

Experiment 2

Method

This experiment was conducted concurrently with Experiment 1, hence the participants were identical to those described in Experiment 1 ($N = 48$). This sample size was defined a priori to match that used (for each group) in Van Berkum et al. (2008), and it is comparable to or exceeds the sample sizes of previous studies that have used EEG to study language in autism (e.g., Coderre, Chernenok, Gordon, & Ledoux, 2017; Korpilahti et al., 2007; Lartseva, Dijkstra, Kan, & Buitelaar, 2014; Pijnacker, Geurts, Van Lambalgen, Buitelaar, & Hagoort, 2010). Nevertheless, post hoc power calculations showed an estimated power of approximately 38% to detect a significant 4-way interaction. We would have needed more than 135 participants (i.e., ~68 autistic individuals, as well as ~68 age- and IQ-matched controls) to reach 80% power, which would not be feasible using these complex methods and given the difficulties associated with recruiting and testing autistic people.

All methodological procedures were preregistered on the Open Science Framework (OSF) web pages (see https://osf.io/7hna3/).

Materials. The experimental and filler sentences used in this study were based on those used in Van Berkum et al.'s (2008) study. One hundred sixty speaker-consistent and speaker-inconsistent experimental sentences were translated from Dutch to English, and adapted to ensure they matched English sociocultural stereotypes, names and places. Each sentence included a single, sentence medial, critical word that was either consistent or inconsistent with the speaker (critical words are underlined in the following examples). There were 40 sentences in the age category: 20 adult versus 20 child type sentences (e.g., "I drink a glass of <u>wine</u> every night before I go to sleep," "I cannot sleep without my *teddy* in my arms"), 40 sentences in the class category: 20 stereotypical high class versus 20 stereotypical lower class type sentences ("Every month, we go to the <u>*opera*</u> for a night out," "I have a large *tattoo* on my back") and 80 sentences in the gender category: 40 stereotypical female versus 40 stereotypical male type sentences (e.g., "I bought a very comfortable <u>bra</u> from an expensive shop," "Every week I trim my *beard* with a small pair of scissors"). A third semantic anomaly condition was created by replacing the critical word in each sentence with a semantically anomalous word (e.g., "I cannot sleep without my *pizza* in my arms"), matched in length and syllables to the consistent/inconsistent conditions. This condition provides a within-subjects baseline measure of the anomaly detection N400 effect (note that this differs from Van Berkum's study that tested semantic anomaly sentences in a separate experiment). In addition, 60 filler sentences were created to balance the number of sentences presented with anomalous/inconsistent content (as in Van Berkum et al., 2008). Thus, 30 sentences described 'true' events (e.g., "The dog usually sleeps in his basket in the living room") and 30 sentences described 'semantically correct' information (e.g., "The Sahara is a place that is very dry and hot").

Sentences were recorded by 14 different speakers. Sentences in the age category were read by four speakers: 2 adult speakers (one female and one male) and two child speakers (one female age 6 and one male age 8). Four adult speakers, 2 females and 2 males, read the sentences in the gender category, and four professional actors (2 males and 2 females) were recruited to imitate the stereotypical higher versus lower class British accents for the

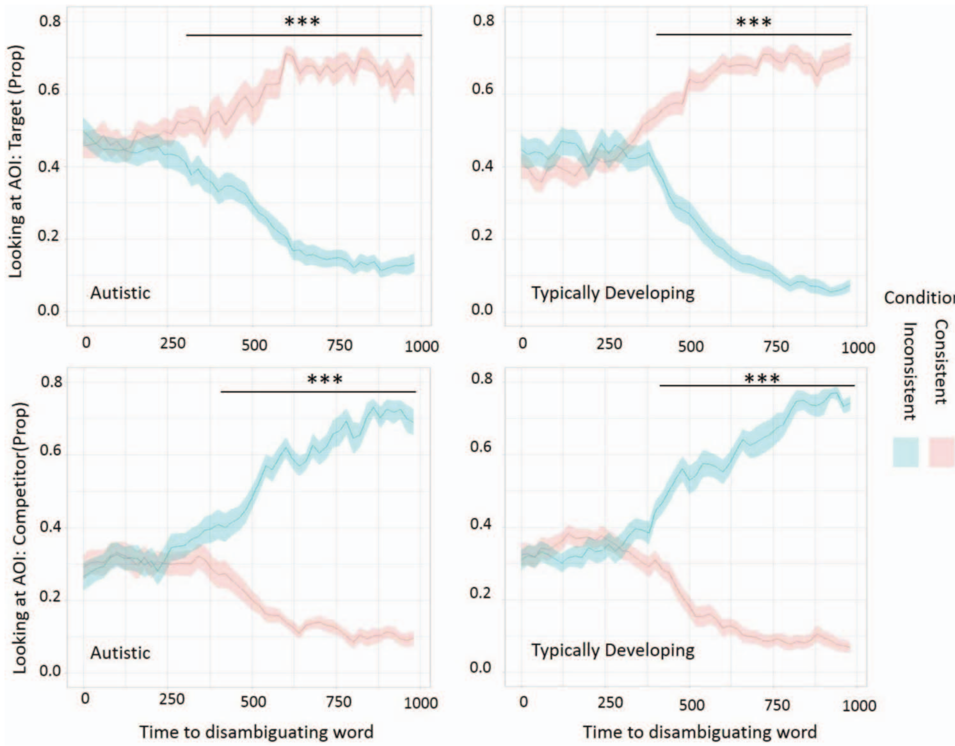

Figure 3. Timecourse of integration fixations toward the target (top panels) and competitor (bottom panels) for each consistency condition and group, Experiment 1. The horizontal lines above them show the points at which the fixations toward the areas of interest (AOI) in different condition first became significant. *** $p < .001$. See the online article for the color version of this figure.

sentences in the class category (one male and one female to each class category). Audios were recorded in a soundproof room using a digital voice recorder. Two further adult speakers (one female and one male) read the filler sentences. All speakers were native speakers of English.

To ensure the validity of our items and speakers, we conducted a posttest, in which 12 TD males and 12 TD females were asked to rate the plausibility of each experimental audio on a 5-point scale: "how normal or strange you think it is to have the speaker say this particular thing" (1 = *completely normal,* 5 = *very strange*). A 1-way ANOVA testing the effect of consistency (consistent vs. inconsistent vs. semantic anomaly) revealed a significant effect of consistency, $F(2, 142) = 340.11$, $p < .001$, $\eta^2 = .83$, with participants rating the semantic anomalous ($M = 3.80$ vs. 1.47, $t = 22.09$, $p < .001$) and inconsistent ($M = 2.23$ vs. 1.47, $t = 11.85$, $p < .001$) audios as less plausible, compared to the consistent ones. Semantic anomalous audios were also rated as less plausible compared to the inconsistent ones ($M = 3.80$ vs. 2.23, $t = 15.98$, $p < .001$). To verify that this consistency effect held for all three speaker types, we conducted separate 1-way ANOVAs for each speaker type (i.e., age, gender and class). This revealed a significant effect of consistency for all three speaker types: (age: $F(2, 142) = 226.58$, $p < .001$, $\eta^2 = .76$; gender: $F(2, 142) = 338.38$, $p < .001$, $\eta^2 = .83$; class: $F(2, 142) = 278.93$, $p < .001$, $\eta^2 = .80$), reflecting the same pattern of lower plausibility ratings for semantic anomalous and inconsistent audios compared to consistent audios.

Three presentation lists were created, with each list containing one hundred and 60 experimental items, 53 or 54 in each of the three conditions. The one hundred and 60 experimental items in each list were interspersed randomly among 60 unrelated filler sentences to create a single random order and each subject only saw each target sentence once, in one of the three conditions. Participants were randomly assigned to read each list.

Procedure. Participants were informed about the EEG procedure and experimental task. After electrode application they were seated in a booth where they listened to the spoken sentences through speakers, while a fixation cross was presented on a computer screen (presented using E-Prime software). There were two practice trials to familiarize participants with the procedure, after which the experimenter answered any questions. Each trial began with the presentation of a single centrally located red fixation cross for 500 ms to signal the start of a new trial. After this time, a white fixation cross appeared for 500 ms. The target sentence was then presented auditorily, with the white fixation cross remaining onscreen throughout. A 1000 ms blank-screen interval followed each item. There was no secondary task. Trials appeared in five blocks of 44 sentences, each lasting ~6 min. Each block was separated by a break, the duration of which was determined by the participant. The EEG task, including setup, took approximately 60 min to complete.

EEG recording and data analysis. A Brain Vision Quickamp amplifier system was used with an ActiCap cap for continuous recording of electroencephalographic (EEG) activity from 30 active electrodes over midline electrodes Fz, Cz, Pz, and Oz, over the left hemisphere from electrodes Fp1, F3, F7, FC1, FC5, C3, T7, CP1, CP5, TP9, P3, P7, O1, and from the homologue electrodes over the right hemisphere. EEG and EOG recordings were sampled at 500 Hz, and electrode impedance was kept below 10 kΩ. Off-line, all EEG channels were recalculated to an average mastoid reference.

Prior to segmentation, EEG and EOG activity was band-pass filtered (.05–70 Hz, 12 dB/oct), and EEG activity containing blinks was corrected using a semiautomatic ocular ICA correction approach (Brain Vision Analyzer 2). The continuous EEG record was then segmented into epochs of 2000 ms, starting 500 ms prior to the onset of the target word (e.g., 'teddy' in the sentence "I cannot sleep without my teddy in my arms"). Thus, the poststimulus epoch lasted for a total duration of 1500 ms. Semiautomatic artifact detection software (Brain Vision Analyzer 2) was run, to identify and discard trials with nonocular artifacts (drifts, channel blockings, EEG activity exceeding $\pm$ 50 μV). This procedure resulted in an average of 43 trials retained for analysis, per condition.

Procedures for the analysis of EEG data replicated those used in Van Berkum's study. First, the signal at each electrode site was averaged separately for each experimental condition, time-locked to the onset of the target word, and aligned to a 200 ms pretarget baseline. Mean ERP amplitude was determined in five time widows, replicating those used in Van Berkum et al. (2008) and in line with our preregistered analysis plans: 100-200 ms, 200-300 ms, 300-500 ms, 500-700 ms, and 200-700 ms. ERP amplitudes over lateral electrodes sites were analyzed using four regions of interest (ROIs). Lateral electrodes were divided along a left-right dimension, and an anterior-posterior dimension. The two ROIs over the left hemisphere were: left-anterior (Fp1, F7, F3, FC5, FC1), and left-posterior (CP5, CP1, P7, P3, O1); two homologue ROIs were defined for the right hemisphere. ERP amplitudes over midline electrodes (Fz, FCz, Cz, Pz, Oz) were analyzed in a single AOI, calculated by averaging data over the five electrodes, and analyzed separately from data recorded over lateral electrode sites.

Results

All analysis procedures were preregistered, and the full experimental materials, data sets and analysis scripts are available on the Open Science Framework web pages (see https://osf.io/7hna3/).

Linear mixed models and lmer in the lme4 package in RStudio software were used to analyze the ERP data (Bates et al., 2015; Version 1.1.453, R Core Team, 2016). We note that our preregistration planned to use ANOVAs to analyze the ERP data, replicating Van Berkum et al. (2008), however in line with analyses for Experiment 1 and more recent statistical norms in the field, we adapted this plan to use linear mixed models since this allowed us to include random effects for both participants and items, and a maximal random effects structure. Thus, over lateral electrodes, each model included fixed effects of Group, AntPos, Hemisphere and Condition, and random effects for items and participants. Over the midline electrodes, each model included fixed effects of Group and Condition, and random effects for items and participants. Fixed effects with two levels (i.e., Group, Hemisphere, AntPos) were contrast coded (–.5 vs. .5). To accommodate the three levels of Condition, we used deviation coded contrast schemes to compare each of the experimental conditions to the consistent reference level: Consistent versus Inconsistent (Consistent (–.33), Inconsistent (.66), Anomalous (–.33)) and Consistent versus Anomalous (Consistent (–.33), Inconsistent (–.33), Anomalous (.66)).

The maximal random effects structure over lateral electrodes included crossed random slopes for Group, AntPos, Hemisphere and Condition within items, and crossed random slopes for AntPos, Hemisphere and Condition within participants. Over midline electrodes, the maximal random effects structure included crossed random slopes for Group and Condition within items, and a random slope for Condition within participants. Some of the random slopes were removed later due to the nonconvergence of the model (as suggested by Barr et al., 2013). The final models used to analyze the data across the different time windows are presented in the supplementary material. Note that due to space constraints, only significant or marginal ($p \leq .06$) effects are presented in the text. Full statistical effects for each time window are summarized in Table 4, and grand average waveforms for each condition/group are shown in Figure 4.

100–200 ms. Analyses revealed a significant effect of AntPos, with a more negative waveform over anterior electrode sites ($M = -.24$ μV) compared to posterior electrode sites ($M = -.05$ μV). More importantly, the speaker-inconsistency effect was significant over the midline electrodes, revealing a more negative wave in the inconsistent condition ($M = -.45$ μV) compared to the consistent condition ($M = -.09$ μV). The semantic anomaly effect was marginally significant over the midline electrodes, showing a more negative wave in the semantic anomalous condition ($M = -.32$ μV) compared to the consistent condition ($M = -.09$ μV). The effect of Group was not significant, and Group did not interact with any other variables.

200–300 ms. A significant effect of AntPos once again showed a more negative waveform over anterior electrode sites ($M = -.43$ μV) than posterior electrode sites ($M = -.13$ μV). The speaker-inconsistency effect was significant over lateral electrodes and marginal over the midline, revealing a more negative wave in the inconsistent condition ($M_{lateral} = -.40$ μV, $M_{central} = -.59$ μV) compared to the consistent condition ($M_{lateral} = -.16$ μV, $M_{central} = -.28$ μV) Over the midline electrodes there was a significant effect of Group, reflecting a more negative wave in the TD group ($M = -.62$ μV) compared to the autistic group ($M = -.27$ μV). There was also a significant interaction between Group and AntPos. Post hoc comparisons revealed a more negative wave over anterior ($M = -.45$ μV) compared to posterior electrodes ($M = .02$ μV) in the autistic group (*Est.* = 0.45, $SE = 0.17$, $t = 2.60$, $p = .015$), but no difference in the TD group, $t = 1.16$, $p = .258$. None of the remaining effects or interactions involving Group or semantic anomaly were significant.

300–500 ms. The effects of speaker-inconsistency and semantic anomaly were marginally significant over the lateral electrodes, and the semantic anomaly effect was significant over midline electrodes. As expected, the N400 was more negative for the inconsistent ($M = -.51$ μV) and semantic anomaly ($M = -.52$ μV) conditions, compared to the consistent condition ($M = -.29$ μV). The effect of Group was marginal over the midline electrodes, showing a larger overall N400 in the TD group ($M = -.84$

Table 4
Statistical Results From the Analysis of N400 Effects Over Lateral and Midline Electrodes in Experiment 2

	100–200 ms			200–300 ms			300–500 ms			500–700 ms			200–700 ms		
Variable	*Est.*	*SE*	*t-value*	*Est.*	*SE*	*t-value*	*Est.*	*SE*	*t-value*	*Est.*	*SE*	*t-value*	*Est.*	*SE*	*t-value*
Lateral															
AntPos	.19	.08	2.43*	.3	.06	4.93***	.11	.07	1.56	.24	.08	2.91**	.22	.07	3.41***
Hemisphere	−.07	.05	−1.24	−.24	.06	−.39	.11	.07	.16	−.02	.08	−.3	−.01	.07	−.13
Group	−.05	.12	−.38	−.16	.12	−1.33	−.13	.15	−.87	−.22	.16	−1.33	−.17	.14	−1.26
Consistent vs. Anomaly	−.12	.11	−1.13	−.13	.08	−1.72	−.25	.13	−1.95$^{\cdot}$	−.44	.1	−4.38***	−.31	.15	−2.02*
Consistent vs. Inconsistent	−.23	.12	−1.82	−.23	.08	−3.02**	−.22	.12	−1.86$^{\cdot}$	−.2	.1	−1.96$^{\cdot}$	−.22	.14	−1.51
Ant-Pos × Hemisphere	.04	.11	.4	.02	.12	.15	−.03	.14	−.22	−.11	.16	−.73	.01	.13	.08
Ant-Pos × Group	−.01	.15	−.8	−.34	.12	−2.74**	−.26	.14	−1.91$^{\cdot}$	−.35	.16	−2.12*	−.26	.13	−1.97*
Hemisphere × Group	.12	.11	1.09	.03	.12	.21	.07	.14	.54	.15	.16	.9	.1	.13	.73
Ant-Pos × Consistent vs. Anomaly	−.11	.13	−.81	−.09	.15	−.57	.02	.17	.1	−.1	.2	−.5	.04	.16	.26
Ant-Pos × Consistent vs. Inconsistent	−.01	.13	−.12	−.001	.15	−.57	.02	.17	.14	−.18	.2	−.92	−.08	.16	−.48
Hemisphere × Consistent vs. Anomaly	0	.13	0	0	.15	−.01	.03	.17	.18	.06	.2	.31	.05	.16	.3
Hemisphere × Consistent vs. Inconsistent	.08	.13	.6	.09	.15	.61	.2	.17	1.15	.02	.2	.11	.1	.16	.62
Group × Consistent vs. Anomaly	.05	.17	.33	−.16	.15	−1.05	−.17	.26	−.67	−.26	.2	−1.27	−.23	.23	−.99
Group × Consistent vs. Inconsistent	.19	.21	.88	0	.15	−.03	−.26	.23	−1.12	.01	.2	.07	−.12	.22	−.55
Ant-Pos × Hemisphere × Group	.2	.21	.95	.18	.25	.73	.12	.28	.43	.1	.33	.3	−.12	.26	.96
Ant-Pos × Hemisphere × Consistent vs. Anomaly	−.16	.26	−.61	−.16	.3	−.53	−.2	.34	−.59	−.13	.4	−.32	.2	.33	.19
Ant-Pos × Hemisphere × Consistent vs. Inconsistent	−.04	.26	−.16	−.12	.3	−.41	−.19	.34	−.57	−.32	.4	−.79	.06	.33	−.78
Ant-Pos × Group × Consistent vs. Anomaly	.06	.26	.23	−.12	.3	−.4	−.2	.34	−.58	.12	.4	.29	−.26	.33	.4
Ant-Pos × Group × Consistent vs. Inconsistent	.03	.26	−.13	.05	.3	.17	−.07	.34	−.21	.23	.4	.56	.13	.33	.14
Hemisphere × Group × Consistent vs. Anomaly	.21	.26	.79	.23	.3	.77	.36	.34	1.05	.63	.4	1.57	.05	.33	1.44
Hemisphere × Group × Consistent vs. Inconsistent	.32	.26	1.22	.45	.3	1.48	.41	.34	1.2	.53	.4	1.31	.46	.33	1.4
Ant-Pos × Hemisphere × Group × Consistent vs. Anomaly	.05	.52	.1	.19	.61	.31	.13	.68	.18	−.05	.81	−.06	.52	.65	.79
Ant-Pos × Hemisphere × Group × Consistent vs. Inconsistent	.14	.52	.26	.17	.61	.28	−.02	.68	−.03	−.07	.81	−.09	−.05	.65	−.08
Midline															
Group	−.19	.12	−1.57	−.35	.14	−2.43*	−.36	.18	−1.98$^{\cdot}$	−.51	.19	−2.62*	−.42	.16	−2.55*
Consistent vs. Anomaly	−.24	.13	−1.82$^{\cdot}$	−.19	.15	−1.25	−.34	.16	−2.12*	−.57	.19	−2.91**	−.40	.16	−2.55*
Consistent vs. Inconsistent	−.37	.15	−2.48*	−.31	.17	−1.82$^{\cdot}$	−.16	.16	−1.02	−.23	.19	−1.26	−.22	.16	−1.41
Group × Consistent vs. Anomaly	.21	.27	.78	−.27	.33	−.83	−.22	.32	−.68	−.10	.37	−.26	−.18	.31	−.57
Group × Consistent vs. Inconsistent	.41	.30	1.34	.08	.36	.22	−.31	.32	−.97	.19	.37	.5	−.04	.31	−.12

$^{\cdot}$ $p < .1$. * $p < .05$. ** $p < .01$. *** $p < .001$.

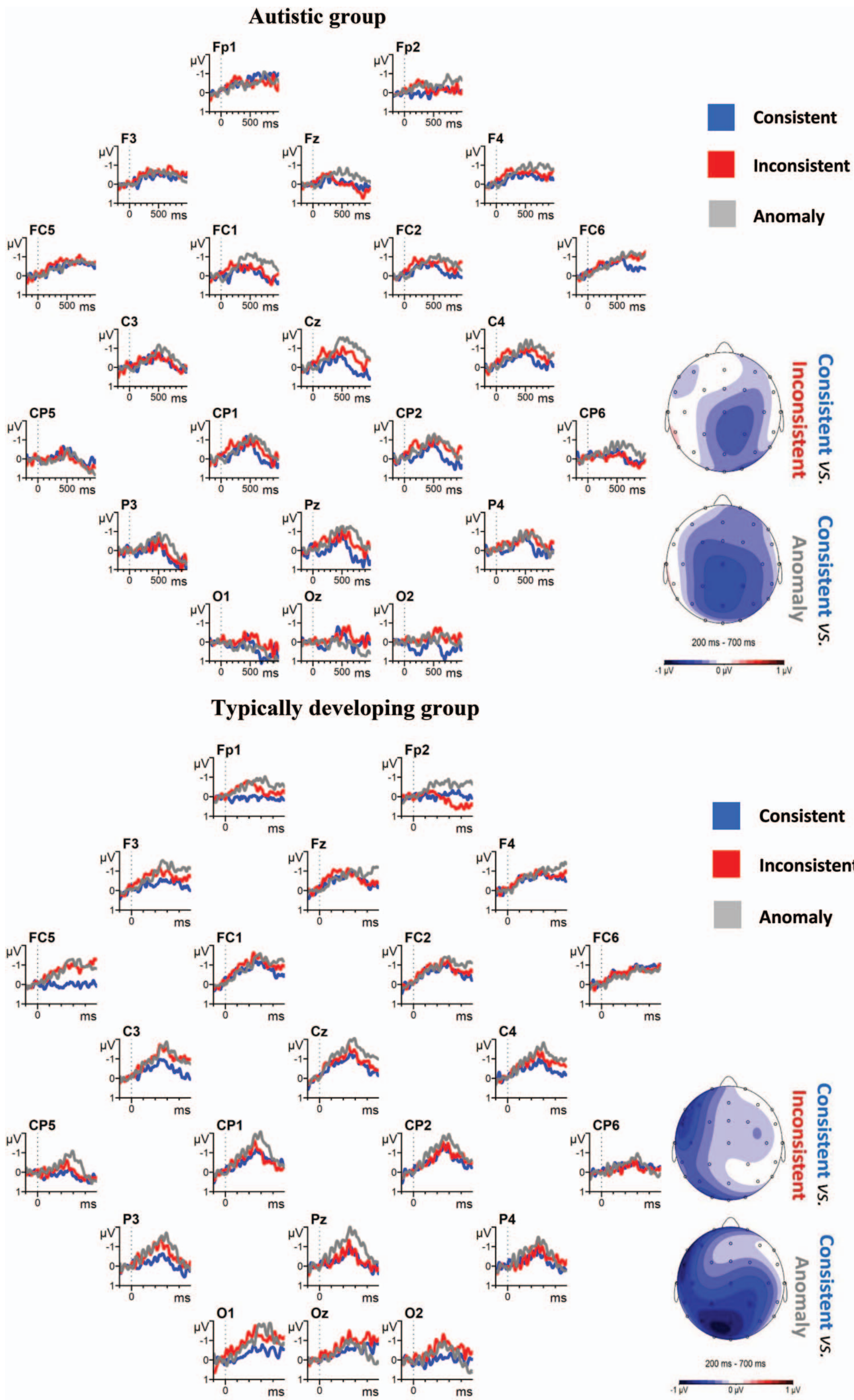

Figure 4. Grand-average ERPs elicited by critical words in the autistic group (top panel) and TD group (bottom panel) for Consistent, Inconsistent and Anomalous speaker conditions. Note that negativity is plotted upward. See the online article for the color version of this figure.

μV) than the autistic group ($M = -.49$ μV). None of the remaining effects or interactions involving Group reached significance.

500–700 ms. The effect of speaker-inconsistency was significant over the lateral electrodes, with a larger N400 in the inconsistent ($M = -.50$ μV) compared to consistent condition ($M = -.30$ μV). The effect of semantic anomaly was significant over both lateral and midline electrodes, reflecting a larger N400 in the semantic anomaly condition ($M_{lateral} = -.73$ μV, $M_{central} = -1.00$ μV) compared to the consistent condition ($M_{lateral} = -.30$ μV, $M_{central} = -.44$ μV). Once again, the effect of Group was significant over the midline electrodes, with larger N400 effects in the TD group ($M = -.95$ μV) than in the autistic group ($M = -.45$ μV). Post hoc comparisons revealed a more negative wave over anterior ($M = -.62$ μV) compared to posterior electrodes ($M = -.21$ μV) in the autistic group (*Est.* = 0.40, SE = 0.19, $t = 2.10$, $p = .048$), but no difference in the TD group, t = 0.43, $p = .669$. None of the remaining effects or interactions involving Group reached significance.

200–700 ms. Analyses over lateral electrodes revealed a significant effect of AntPos, with a more negative N400 over anterior electrode sites ($M = -.55$ μV) compared to posterior electrode sites ($M = -.33$ μV). The semantic anomaly effect was significant over both lateral and midline sites, reflecting a larger N400 in the semantic anomaly condition ($M_{lateral} = -.56$ μV, $M_{central} = -.83$ μV) compared to the consistent condition ($M_{lateral} = -.27$ μV, $M_{central} = -.43$ μV). Over the midline, the N400 was significantly more negative in the TD group ($M = -.84$ μV) compared to the autistic group ($M = -.43$). There was also a significant interaction between Group and AntPos. Post hoc comparisons revealed a more negative wave over anterior ($M = -.54$ μV) compared to posterior electrodes ($M = -.18$ μV) in the autistic group (*Est.* = 0.34, SE = 0.16, $t = 2.10$, $p = .046$), but no difference in the TD group, t = 0.89, $p = .383$. None of the remaining effects or interactions involving Group reached significance.

Summary

First, the results of Experiment 2 replicated Van Berkum et al.'s findings (2008), showing that individuals integrated their world knowledge (voice-based inferences in our study) and the semantics of the sentence to detect an inconsistency between the speaker's voice and meaning as early as 200-300 ms after hearing the critical word. Second, the speaker inconsistency effect emerged within a comparable timeframe to the semantic anomaly effect, perhaps even earlier. Thus, the results provide further evidence for the notion that language processing goes beyond processing the linguistic input, and that pragmatic processing can be activated immediately (Berkum, Hagoort, & Brown, 1999; Just & Carpenter, 1980; Zwaan, 2003). Importantly, autistic individuals took the speaker's voice into account as quickly as TD individuals, showing that they were as fast to integrate pragmatics and semantics. This pattern contrasts with theories that suggest autistic individuals have difficulties in using context while processing language (Tager-Flusberg, Paul, & Lord, 2005).

General Discussion

In two preregistered experiments we investigated the timecourse with which autistic and TD adults understand a speaker's meaning based on characteristics inferred from the speaker's voice. Experiment 1 used the visual world paradigm to capture the timecourse of *anticipated* meaning while participants listened to spoken sentences in which the speaker's voice and message were either consistent or inconsistent (e.g., "When we go shopping, I usually look for my favorite *wine*," spoken by an adult or a child). Experiment 2 recorded ERPs to examine *integration* of meaning while participants listened to spoken sentences that were either consistent or inconsistent in terms of voice and message, or semantically anomalous (e.g., "I cannot sleep without my *pizza* in my arms"). These experiments allowed us to test the general question of whether inferences about pragmatic meaning are activated online during language comprehension (i.e., linguistic input and context are processed in a single incremental step), or whether these pragmatic inferences are delayed to a second step of language processing (i.e., individuals first extract the sentence's message using syntax and semantics, and only integrate the speaker's identity at a later stage of processing). Moreover, by comparing real-time pragmatic processing of spoken language among autistic and TD people we investigated whether and how these processes are affected when global coherence and social abilities are compromised. Thus, we tested whether autistic adults would show disrupted use of context to infer meaning (i.e., replicating Happé, 1997; Jolliffe & Baron-Cohen, 1999), or whether aspects of contextual language comprehension and perspective-taking are intact among autistic people (as seen in Au-Yeung et al., 2014, 2018; Black et al., 2018, 2019; Ferguson et al., 2019; Williams & Happé, 2010).

Results provided converging evidence that listeners rapidly and accurately anticipate a speaker's intended meaning based on inferences from their voice. In Experiment 1, participants were faster to select the mentioned object when it was consistent with the speaker's voice than when it was inconsistent. More importantly, eye movement data revealed a strong and increasing preference to fixate the object that was consistent with the speaker's voice (i.e., the target) long before this object was disambiguated in the auditory input (~2000 ms before). These incremental expectations were further evidenced in Experiment 2, as the N400 revealed that participants detected an inconsistency between the speaker's voice and meaning as early as 200 ms after hearing the critical word. This speaker inconsistency effect emerged within a comparable timeframe to the semantic anomaly effect. This suggests that listeners used the inferred speaker context to constrain their expectations about forthcoming language, in a similar way that semantics and linguistic discourse context constrain expected meaning (see Van Berkum, 2009).

These findings therefore provide novel insights into the timecourse of social language understanding. In line with hypotheses from the one-step model of language processing, our data support the proposal that social context (voice of speaker here) and the linguistic input are taken into account concurrently when we process language (Clark, 1996; Perry, 1997). This early and incremental anticipation was particularly evident in Experiment 1, where eye movements provided a novel measure of predictive processing and showed that voice-related processes are activated even before hearing the socially relevant contrasts (e.g., wine/sweets). Here, participants inferred characteristics of the speaker based on their voice (i.e., their age, gender or social class), and directed their eye movements to objects in the visual scene that were consistent with this prediction, and relevant to the content of their unfolding utterance. Importantly, Experiment 1 showed that pragmatic inferences about the speaker modified constraints based on lexical-semantic input. In other words, while both the sweets and wine fit the semantic constraints of objects that one can buy at the supermarket, world knowledge provided cues for participants to distinguish the most relevant option for the particular speaker (e.g., *adults are more likely to buy wine than sweets*). This suggests that pragmatic inferences about a speaker (based on their voice) have a strong and early influence on predictive language processing, and that this is comparable to the effects seen when world knowledge constraints have been explicitly defined in the language input (e.g., "The girl will ride the carousel/motorbike"; Kamide, Altmann, & Haywood, 2003).

Further evidence of these rapid pragmatic inferences was seen in the ERP data in Experiment 2, which replicated and extended the results from Van Berkum et al. (2008)'s study. Here, the N400 was amplified for inconsistent speaker-meaning sentences relative to consistent speaker-meaning sentences. Indeed, speaker inconsistency effects emerged as early as 200 ms after critical word onset. This suggests that listeners already had strong predictions about the unfolding language, and the sorts of objects the speaker was most likely to mention, based on world knowledge constraints activated by the speaker's voice. This pattern provides further evidence that these social stereotypes can overrule lexical-semantic processing, since both critical words are semantically appropriate to the sentence context, and one only becomes incongruent when meaning is interpreted based on inferred knowledge about the specific speaker. Moreover, our design allowed direct comparison of this pragmatic N400 effect with a semantic anomaly condition and revealed that the brain's response to pragmatically infelicitous language is indistinguishable from that elicited by semantic fit. This shows that language comprehension is a dynamic process whereby people can rapidly access and integrate information based on the explicit and inferred context (including words, sentence, discourse, and world knowledge), then flexibly shift between these different constraints as appropriate.

Importantly, similar patterns of anticipation and integration based on speaker-meaning fit were found among autistic and TD people, despite the autistic group showing a significant impairment in explicitly recognizing the emotions of speakers from their voice (in the RMIV task). This finding provides evidence that autistic adults do not experience a general deficit in inferring social characteristics of speakers, or integrating information in context (as seen in Black et al., 2018, 2019; Ferguson et al., 2019; Koldewyn et al., 2013; Mottron et al., 2003; Plaisted et al., 2006; Van der Hallen et al., 2015). In Experiment 1, autistic participants successfully inferred the pragmatic context from the speaker's voice and directed their visual attention to anticipate mention of the speaker-relevant target object nearly 2000 ms before disambiguation. In Experiment 2, autistic individuals inferred the spoken utterance's pragmatic meaning as quickly as TD individuals, evidenced by deflections on the N400 to inconsistent speaker-meaning sentences within 200 ms of hearing the critical word. These patterns provide a clear indication that autistic people are aware of social stereotypes and can infer and apply these in real-time to constrain language comprehension. This is in line with previous research showing intact social knowledge in autism when judging attributions, such as race, age, social status and so forth from faces or bodies (Frith, 2007; Saldaña & Frith, 2007; White, Hill, Winston, & Frith, 2006).

This unimpaired anticipation of speaker meaning is unexpected based on accounts that characterize autism in terms of a reduced drive for global coherence (WCC account; Frith, 1989; Frith & Happé, 1994; Happé & Frith, 2006), disordered processing of complex information (Minshew & Goldstein, 1998; Minshew, Goldstein, & Siegel, 1997; Minshew, Williams, & McFadden, 2008), or atypical pragmatic integration (Happé, 1997; Jolliffe & Baron-Cohen, 1999; Nuske & Bavin, 2011). These accounts predict that autistic people would show impairments in using the context (the speaker's voice here) to predict language. For example, the complex information processing theory suggests that autistic individuals struggle with integrating the information from multiple sources or components, so these individuals would struggle completing complex tasks that involve combining information from different components (Minshew et al., 1997). Yet data from both experiments here showed clear effects of speaker inferences among both groups of participants. Thus, the current results are consistent with recent research that has used implicit methods to show that autistic adults have an intact ability to integrate information online during language comprehension (Au-Yeung et al., 2018; Black et al., 2018, 2019; Ferguson et al., 2019; Howard et al., 2017b), and extend this by showing that global coherence of information in autism can go beyond 'what is said' to assess 'who is saying what.' The intact contextual integration seen in Experiments 1 and 2 is also consistent with the results of the linguistic central coherence task, which did not find any evidence of a local processing bias among our autistic participants (cf. Booth & Happé, 2010). Importantly, the fact that autistic individuals were impaired at explicitly inferring emotions from a speaker's voice in the explicit RMIV task, suggests that although autistic individuals are unimpaired at integrating social stereotypes online, they struggle with extracting more complex information offline, such as emotions, supporting the previous literature (Jones et al., 2011; Philip et al., 2010). Hence, future studies, should examine how these individuals process complex information, including emotions or mental states online, while extracting the meaning from language.

Nevertheless, Experiment 1 revealed some subtle differences in the timecourse and strength of voice-based pragmatic inferences among TD and autistic participants, which might suggest that autistic people activated weaker speaker-meaning expectations or were less bound to these social stereotypes. First, the eye-tracking data showed that participants in the TD group biased their visual attention to the target object earlier than the autistic group (2240 ms vs. 1800 ms before disambiguation), though this anticipatory bias in the TD group subsequently declined prior to a rapid increase (960 ms before disambiguation), whereas the autistic group showed a consistent increase in target bias from 1800 ms before the disambiguation point. Second, only the autistic group in Experiment 1 showed significant interference from the competitor object (i.e., the object that was semantically, but not pragmatically relevant to the context) during the anticipation period. Finally, analysis of the period after disambiguation (i.e., integration) showed faster switches *away from* the target in the inconsistent condition among the autistic group compared to the TD group (300 ms vs. 400 ms respectively).

Taken together, these findings could suggest that autistic individuals are more likely to adopt a bottom-up (i.e., semantics first) approach to pragmatic language processing (Van den Brink et al., 2012), which means that they are less able to ignore pragmatically irrelevant information. This explanation is in line with the predictive coding theory of autism (Van Boxtel & Lu, 2013), which suggests that autistic people attribute greater weight to bottom-up errors due to metalearning impairments, and consequently contextualize sensory signals in a less automatic way, especially when facing complicated unexpected input. Support for this predictive coding theory of autism is particularly evident in the anticipatory data from Experiment 1, where autistic adults were successfully able to predict the speaker's meaning based on their voice but were slower to do so and exhibited weaker biases to the speaker-relevant target. In an experimental setting, these subtle differences in timing and strength of predictions are not sufficient to disrupt comprehension, however it is likely that in real-world settings, where conversation is more fast-paced and involves greater distracting sensory input, these weaker top-down predictions can have a cumulative impact on social communication. Alternatively, the different patterns might reflect a more flexible use of social stereotypes among autistic individuals compared to their TD peers. Previous research has established that autistic individuals are able to recognize and use social stereotypes (including age and social status) despite profound difficulties in mental state reasoning (Hirschfeld, Bartmess, White, & Frith, 2007; White et al., 2006). Our data might then demonstrate that autistic people are less constrained in automatically assigning meaning according to these usual/prototypical contexts (see Zalla et al., 2014). These subtle differences between groups, despite intact overt understanding of social stereotypes in autistic individuals, are analogous to recent neuroimaging research that has shown distinct patterns of brain activation during speaker-meaning integration, among autistic people and their TD peers (Groen et al., 2010; Tesink et al., 2009). Based on these findings, researchers have proposed that autistic people recruit atypical brain areas to integrate social information, and may rely on compensatory mechanisms to integrate social contrasts.

Finally, we note some limitations with the current experiments. First, it is possible that we simply did not have sufficient power to detect the 3- and 4-way interactions that were tested in Experiment 2. Our sample size was chosen a priori to achieve comparable participant numbers in each group to the total sample size used in Van Berkum et al. (2008; $N = 24$), and to match or exceed the sample sizes used in previous studies in these areas, however post hoc power analyses suggested that at least 68 participants would be needed in each group to reach the desired 80% power. Nevertheless, concerns about power are alleviated somewhat by our use state-of-the-art statistical methods which meant that analyses were run on individual data points rather than data aggregated across participants (thus improving power; Baayen, Davidson, & Bates, 2008), which also allowed us to control for by-participant and by-item variation in a single analysis. Moreover, given that results from Experiment 2 replicated the patterns seen in Van Berkum et al. (2008), and that group did not modulate speaker consistency effects in *any* of the five preregistered analysis time windows (in either the midline or lateral analyses), we can feel relatively confident that the reported findings are reliable. Nevertheless, as a field, research on autism should continue to aim for larger sample sizes, ideally recruiting participants with a diverse representation on the autism spectrum to ensure generalizability of results. Another point to consider is that this study did not test whether there were any differences between groups in terms of attitudes toward social stereotypes. For example, previous studies have shown that gender dysphoria is more prevalent among autistic than TD individuals, which could influence their attitudes toward gender stereotypes (Van Der Miesen, Hurley, & De Vries, 2016). Thus, future research should consider whether norms and expectations differ between autistic and TD individuals. Furthermore, since our autistic participants were impaired at explicitly recognizing others' emotions from their voices, future research should investigate whether subclinical emotional conditions, such as Alexithymia (prevalent among autistic people), correlates with the ability to understand external emotions in autism.

In conclusion, the two experiments reported here employed complementary measures to assess online processing of spoken language among autistic and TD adults. Together they provide strong evidence that language is processed in a single step, by showing that speaker-related information (i.e., social context) is processed in parallel with the linguistic input, and can override salient lexical-semantic input to influence listeners' expectations of an unfolding utterance in real-time. Moreover, this ability to anticipate and integrate language meaning based on social inferences about the speaker was unimpaired among autistic people. This shows that autistic people are aware of social stereotypes, and can infer and apply these automatically to constrain language comprehension. Nevertheless, we observed subtle differences in the timecourse with which these processes are activated among autistic individuals, which might indicate a preference for bottom-up (i.e., language first) processing, or more flexible use of social stereotypes in this group. Further research is needed to determine how these social contrasts are applied in real life, where language is less structured and social cues may be less salient.

References

Allen, M. L., Haywood, S., Rajendran, G., & Branigan, H. (2011). Evidence for syntactic alignment in children with autism. *Developmental Science, 14,* 540–548. http://dx.doi.org/10.1111/j.1467-7687.2010.01001.x

Altmann, G. T., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language, 57,* 502–518. http://dx.doi.org/10.1016/j.jml.2006.12.004

Altmann, G. T., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: Eye movements and mental representation. *Cognition, 111,* 55–71. http://dx.doi.org/10.1016/j.cognition.2008.12.005

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM–5)*. Washington, DC: Author.

Au-Yeung, S. K., Kaakinen, J. K., & Benson, V. (2014). Cognitive perspective-taking during scene perception in autism spectrum disorder: Evidence from eye movements. *Autism Research, 7,* 84–93. http://dx.doi.org/10.1002/aur.1352

Au-Yeung, S. K., Kaakinen, J. K., Liversedge, S. P., & Benson, V. (2018). Would Adults with Autism be less likely to bury the Survivors? An Eye Movement Study of Anomalous Text Reading. *The Quarterly Journal of Experimental Psychology, 71,* 280–290. http://dx.doi.org/10.1080/17470218.2017.1322621

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412. http://dx.doi.org/10.1016/j.jml.2007.12.005

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. *Child Psychology & Psychiatry & Allied Disciplines, 38,* 813–822. http://dx.doi.org/10.1111/j.1469-7610.1997.tb01599.x

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34,* 163–175. http://dx.doi.org/10.1023/B:JADD.0000022607.19833.00

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31,* 5–17. http://dx.doi.org/10.1023/A:1005653411471

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255–278. http://dx.doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*. Advance online publication. http://dx.doi.org/10.18637/jss.v067.i01

Behrmann, M., Thomas, C., & Humphreys, K. (2006). Seeing it differently: Visual processing in autism. *Trends in Cognitive Sciences, 10,* 258–264.

Berkum, J. J. V., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11,* 657–671.

Black, J., Barzy, M., Williams, D., & Ferguson, H. (2019). Intact counterfactual emotion processing in autism spectrum disorder: Evidence from eye-tracking. *Autism Research, 12,* 422–444. http://dx.doi.org/10.1002/aur.2056

Black, J., Williams, D., & Ferguson, H. J. (2018). Imagining counterfactual worlds in autism spectrum disorder. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44,* 1444.

Booth, R., & Happé, F. (2010). "Hunting with a knife and . . . fork": Examining central coherence in autism, attention deficit/hyperactivity disorder, and typical development with a linguistic task. *Journal of Experimental Child Psychology, 107,* 377–393.

Brock, J., & Bzishvili, S. (2013). Deconstructing Frith and Snowling's homograph-reading task: Implications for autism spectrum disorders. *The Quarterly Journal of Experimental Psychology, 66,* 1764–1773. http://dx.doi.org/10.1080/17470218.2013.766221

Brock, J., & Caruana, N. (2014). Reading for sound and reading for meaning in autism. In J. Arciuli & J. Brock (Eds.), *Communication in autism* (pp. 125–145). Amsterdam: the Netherlands: John Benjamins Publishing.

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,* 896–904. http://dx.doi.org/10.1016/j.cognition.2008.06.007

Clark, H. H. (1996). *Using language*. New York, NY: Cambridge University Press.

Coderre, E. L., Chernenok, M., Gordon, B., & Ledoux, K. (2017). Linguistic and non-linguistic semantic processing in individuals with autism spectrum disorders: An ERP study. *Journal of autism and developmental disorders, 47,* 795–812.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York, NY: Routledge Academic.

Connolly, J. H. (2001). Context in the study of human languages and computer programming languages: A comparison. In V. Akman, P. Bouquet, R. Thomason, & R. Young (Eds.), *Modeling and Using Context* (pp. 116–128). Berlin, Germany: Springer.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6,* 84–107. http://dx.doi.org/10.1016/0010-0285(74)90005-X

Cutler, A., & Clifton, C., Jr. (1999). Comprehending spoken language: A blueprint of the listener. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 123–166). New York, NY: Oxford University Press.

Deliens, G., Papastamou, F., Ruytenbeek, N., Geelhand, P., & Kissine, M. (2018). Selective pragmatic impairment in autism spectrum disorder: Indirect requests versus irony. *Journal of Autism and Developmental Disorders, 48,* 2938–2952. http://dx.doi.org/10.1007/s10803-018-3561-6

Ferguson, H. J., Black, J., & Williams, D. (2019). Distinguishing reality from fantasy in adults with autism spectrum disorder: Evidence from eye movements and reading. *Journal of Memory and Language, 106,* 95–107. http://dx.doi.org/10.1016/j.jml.2019.03.001

Frith, C. D. (2007). The social brain? *Philosophical Transactions of the Royal Society B: Biological Sciences, 362,* 671–678.

Frith, U. (1989). *Autism. explaining the enigma*. Oxford, England: Blackwell.

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind." *Cognition, 50*(1–3), 115–132. http://dx.doi.org/10.1016/0010-0277(94)90024-8

Frith, U., & Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *British Journal of Developmental Psychology, 1,* 329–342. http://dx.doi.org/10.1111/j.2044-835X.1983.tb00906.x

Golan, O., Baron-Cohen, S., Hill, J. J., & Rutherford, M. D. (2007). The 'Reading the Mind in the Voice' test-revised: A study of complex emotion recognition in adults with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders, 37,* 1096–1106. http://dx.doi.org/10.1007/s10803-006-0252-5

Gouvea, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes, 25,* 149–188. http://dx.doi.org/10.1080/01690960902965951

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution, 7,* 493–498. http://dx.doi.org/10.1111/2041-210X.12504

Groen, W. B., Tesink, C., Petersson, K. M., van Berkum, J., van der Gaag, R. J., Hagoort, P., & Buitelaar, J. K. (2010). Semantic, factual, and social language comprehension in adolescents with autism: An FMRI study. *Cerebral Cortex, 20,* 1937–1945. http://dx.doi.org/10.1093/cercor/bhp264

Hahn, N., Snedeker, J., & Rabagliati, H. (2015). Rapid linguistic ambiguity resolution in young children with autism spectrum disorder: Eye tracking evidence for the limits of weak central coherence. *Autism Research, 8,* 717–726. http://dx.doi.org/10.1002/aur.1487

Happé, F. G. (1997). Central coherence and theory of mind in autism: Reading homographs in context. *British Journal of Developmental Psychology, 15,* 1–12. http://dx.doi.org/10.1111/j.2044-835X.1997.tb00721.x

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

Heaton, T. J., & Freeth, M. (2016). Reduced visual exploration when viewing photographic scenes in individuals with autism spectrum disorder. *Journal of Abnormal Psychology, 125,* 399–411. http://dx.doi.org/10.1037/abn0000145

Hirschfeld, L., Bartmess, E., White, S., & Frith, U. (2007). Can autistic children predict behavior by social stereotypes? *Current Biology, 17*(12), R451–R452. http://dx.doi.org/10.1016/j.cub.2007.04.051

Hopkins, Z., Yuill, N., & Keller, B. (2016). Children with autism align syntax in natural conversation. *Applied Psycholinguistics, 37,* 347–370. http://dx.doi.org/10.1017/S0142716414000599

Howard, P. L., Liversedge, S. P., & Benson, V. (2017a). Benchmark eye movement effects during natural reading in autism spectrum disorder. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43,* 109–127. http://dx.doi.org/10.1037/xlm0000289

Howard, P. L., Liversedge, S. P., & Benson, V. (2017b). Investigating the use of world knowledge during on-line comprehension in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 47,* 2039–2053. http://dx.doi.org/10.1007/s10803-017-3129-x

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,* 149–185. http://dx.doi.org/10.1016/S0010-0277(99)00022-0

Jones, C. R., Pickles, A., Falcaro, M., Marsden, A. J., Happé, F., Scott, S. K., . . . Charman, T. (2011). A multimodal approach to emotion recognition ability in autism spectrum disorders. *Journal of Child Psychology and Psychiatry, 52,* 275–285. http://dx.doi.org/10.1111/j.1469-7610.2010.02328.x

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329–354. http://dx.doi.org/10.1037/0033-295X.87.4.329

Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49,* 133–156. http://dx.doi.org/10.1016/S0749-596X(03)00023-8

Kamide, Y., Lindsay, S., Scheepers, C., & Kukona, A. (2016). Event processing in the visual world: Projected motion paths during spoken sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42,* 804–812. http://dx.doi.org/10.1037/xlm0000199

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the U. K. autism community. *Autism: The International Journal of Research and Practise, 20,* 442–462. http://dx.doi.org/10.1177/1362361315588200

Koldewyn, K., Jiang, Y. V., Weigelt, S., & Kanwisher, N. (2013). Global/local processing in autism: Not a disability, but a disinclination. *Journal of Autism and Developmental Disorders, 43,* 2329–2340. http://dx.doi.org/10.1007/s10803-013-1777-z

Korpilahti, P., Jansson-Verkasalo, E., Mattila, M. L., Kuusikko, S., Suominen, K., Rytky, S., . . . Moilanen, I. (2007). Processing of affective speech prosody is impaired in Asperger syndrome. *Journal of Autism and Developmental Disorders, 37,* 1539–1549. http://dx.doi.org/10.1007/s10803-006-0271-2

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207,* 203–205. http://dx.doi.org/10.1126/science.7350657

Lartseva, A., Dijkstra, T., Kan, C. C., & Buitelaar, J. K. (2014). Processing of emotion words by patients with autism spectrum disorders: Evidence from reaction times and EEG. *Journal of Autism and Developmental Disorders, 44,* 2882–2894. http://dx.doi.org/10.1007/s10803-014-2149-z

Lattner, S., & Friederici, A. D. (2003). Talker's voice and gender stereotype in human auditory sentence processing—Evidence from event-related brain potentials. *Neuroscience Letters, 339,* 191–194. http://dx.doi.org/10.1016/S0304-3940(03)00027-2

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., . . . Rutter, M. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30,* 205–223. http://dx.doi.org/10.1023/A:1005592401947

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164,* 177–190. http://dx.doi.org/10.1016/j.jneumeth.2007.03.024

Martin, I., & McDonald, S. (2003). Weak coherence, no theory of mind, or executive dysfunction? Solving the puzzle of pragmatic language disorders. *Brain and Language, 85,* 451–466. http://dx.doi.org/10.1016/S0093-934X(03)00070-1

Milton, D. E. M., Heasman, B., & Sheppard, E. (2018). Double empathy. In F. R. Volkmar (Ed.), *Encyclopedia of autism spectrum disorders,* 1–8. New York, NY: Springer.

Minshew, N. J., & Goldstein, G. (1998). Autism as a disorder of complex information processing. *Mental Retardation and Developmental Disabilities Research Reviews, 4,* 129–136. http://dx.doi.org/10.1002/(SICI)1098-2779(1998)4:2<129::AID-MRDD10>3.0.CO;2-X

Minshew, N. J., Goldstein, G., & Siegel, D. J. (1997). Neuropsychologic functioning in autism: Profile of a complex information processing disorder. *Journal of the International Neuropsychological Society, 3,* 303–316.

Minshew, N. J., Williams, D. L., & McFadden, K. (2008). Information processing, neural connectivity, and neuronal organization. In A. W. Zimmerman (Ed.), *Autism* (pp. 381–405). Totowa, NJ: Humana Press. http://dx.doi.org/10.1007/978-1-60327-489-0_18

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,* 475–494. http://dx.doi.org/10.1016/j.jml.2007.11.006

Mottron, L., Burack, J. A., Iarocci, G., Belleville, S., & Enns, J. T. (2003). Locally oriented perception with intact global processing among adolescents with high-functioning autism: Evidence from multiple paradigms. *Journal of Child Psychology and Psychiatry, 44,* 904–913. http://dx.doi.org/10.1111/1469-7610.00174

Münster, K., & Knoeferle, P. (2018). Extending situated language comprehension (accounts) with speaker and comprehender characteristics: Toward socially situated interpretation. *Frontiers in Psychology, 8,* 2267. http://dx.doi.org/10.3389/fpsyg.2017.02267

Nicolaidis, C., Milton, D., Sasson, N. J., Sheppard, E., & Yergeau, M. (2018). An expert discussion on autism and empathy. *Autism in Adulthood, 1,* 4–11. http://dx.doi.org/10.1089/aut.2018.29000.cjn

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., . . . Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *eLife, 7,* e33468. http://dx.doi.org/10.7554/eLife.33468

Nuske, H. J., & Bavin, E. L. (2011). Narrative comprehension in 4–7-year-old children with autism: Testing the Weak Central Coherence account. *International Journal of Language & Communication Disorders, 46,* 108–119.

Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory & Cognition, 25,* 273–285. http://dx.doi.org/10.3758/BF03211283

Pellicano, E., & Burr, D. (2012). When the world becomes "too real": A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences, 16,* 504–510. http://dx.doi.org/10.1016/j.tics.2012.08.009

Perry, J. (1997). Indexicals and demonstratives. In Robert Hale and Crispin Wright (Eds.), *A companion to the philosophy of language* (pp. 586–612). Oxford, England: Blackwell.

Philip, R. C. M., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., . . . Hall, J. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychological Medicine, 40,* 1919–1929. http://dx.doi.org/10.1017/S0033291709992364

Pijnacker, J., Geurts, B., van Lambalgen, M., Buitelaar, J., & Hagoort, P. (2010). Exceptions and anomalies: An ERP study on context sensitivity in autism. *Neuropsychologia, 48,* 2940–2951. http://dx.doi.org/10.1016/j.neuropsychologia.2010.06.003

Plaisted, K., Dobler, V., Bell, S., & Davis, G. (2006). The microgenesis of global perception in autism. *Journal of Autism and Developmental Disorders, 36,* 107–116. http://dx.doi.org/10.1007/s10803-005-0047-0

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rodríguez, A., Burigo, M., & Knoeferle, P. (2016). Visual constraints modulate stereotypical predictability of agents during situated language comprehension. *Proceedings of the 38th Annual Cognitive Science Society Meeting, Philadelphia, PA*.

Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the mind in the voice: A study with normal adults and adults with Asperger syndrome and high functioning autism. *Journal of Autism and Developmental Disorders, 32,* 189–194.

Saldaña, D., & Frith, U. (2007). Do readers with autism make bridging inferences from world knowledge? *Journal of Experimental Child Psychology, 96,* 310–319. http://dx.doi.org/10.1016/j.jecp.2006.11.002

Saxe, R., & Baron-Cohen, S. (2006). The neuroscience of theory of mind Social Neuroscience, *1,* 1–9.

Schacht, A., Sommer, W., Shmuilovich, O., Martíenz, P. C., & Martín-Loeches, M. (2014). Differential task effects on N400 and P600 elicited by semantic and syntactic violations. *PLoS ONE, 9*(3), e91226. http://dx.doi.org/10.1371/journal.pone.0091226

Shah, A., & Frith, U. (1993). Why do autistic individuals show superior performance on the block design task? *Child Psychology & Psychiatry & Allied Disciplines, 34,* 1351–1364. http://dx.doi.org/10.1111/j.1469-7610.1993.tb02095.x

Spotorno, N., Cheylus, A., Van Der Henst, J. B., & Noveck, I. A. (2013). What's behind a P600? Integration operations during irony processing. *PLoS ONE, 8*(6), e66839. http://dx.doi.org/10.1371/journal.pone.0066839

Stowe, L. A., Kaan, E., Sabourin, L., & Taylor, R. C. (2018). The sentence wrap-up dogma. *Cognition, 176,* 232–247. http://dx.doi.org/10.1016/j.cognition.2018.03.011

Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 358,* 303–314. http://dx.doi.org/10.1098/rstb.2002.1198

Tager-Flusberg, H., Paul, R., & Lord, C. (2005). In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders: Vol. 1. Language and communication in autism* (pp. 335–364). Hoboken, NJ: Wiley.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268,* 1632–1634. http://dx.doi.org/10.1126/science.7777863

Tesink, C. M., Buitelaar, J. K., Petersson, K. M., van der Gaag, R. J., Kan, C. C., Tendolkar, I., & Hagoort, P. (2009). Neural correlates of pragmatic language comprehension in autism spectrum disorders. *Brain: A Journal of Neurology, 132,* 1941–1952. http://dx.doi.org/10.1093/brain/awp103

Van Berkum, J. J. (2009). The neuropragmatics of "simple" utterance comprehension: An ERP review. In U. Sauerland & K. Yatsushiro (Eds.), *Semantics and pragmatics: From experiment to theory* (pp. 276–316). Basingstoke, England: Palgrave Macmillan.

Van Berkum, J. J., De Goede, D., Van Alphen, P. M., Mulder, E. R., & Kerstholt, J. H. (2013). How robust is the language architecture? The case of mood. *Frontiers in Psychology, 4,* 505. http://dx.doi.org/10.3389/fpsyg.2013.00505

Van Berkum, J. J., Van den Brink, D., Tesink, C. M., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience, 20,* 580–591.

Van Boxtel, J. J., & Lu, H. (2013). A predictive coding perspective on autism spectrum disorders. *Frontiers in Psychology, 4,* 19.

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review, 121,* 649–675. http://dx.doi.org/10.1037/a0037665

van den Brink, D., Van Berkum, J. J., Bastiaansen, M. C., Tesink, C. M., Kos, M., Buitelaar, J. K., & Hagoort, P. (2012). Empathy matters: ERP evidence for inter-individual differences in social language processing. *Social Cognitive and Affective Neuroscience, 7,* 173–183. http://dx.doi.org/10.1093/scan/nsq094

Van der Hallen, R., Evers, K., Brewaeys, K., Van den Noortgate, W., & Wagemans, J. (2015). Global processing takes time: A meta-analysis on local-global visual processing in ASD. *Psychological Bulletin, 141,* 549–573. http://dx.doi.org/10.1037/bul0000004

Van Der Miesen, A. I., Hurley, H., & De Vries, A. L. (2016). Gender dysphoria and autism spectrum disorder: A narrative review. *International Review of Psychiatry, 28,* 70–80. http://dx.doi.org/10.3109/09540261.2015.1111199

Vulchanova, M., Saldaña, D., Chahboun, S., & Vulchanov, V. (2015). Figurative language processing in atypical populations: The ASD perspective. *Frontiers in Human Neuroscience, 9,* 24. http://dx.doi.org/10.3389/fnhum.2015.00024

Wechsler, D. (1999). *Manual for the Wechsler abbreviated intelligence scale (WASI)*. San Antonio, TX: The Psychological Corporation.

White, S., Hill, E., Winston, J., & Frith, U. (2006). An islet of social ability in Asperger Syndrome: Judging social attributes from faces. *Brain and Cognition, 61,* 69–77. http://dx.doi.org/10.1016/j.bandc.2005.12.007

Williams, D., & Happé, F. (2010). Representing intentions in self and other: Studies of autism and typical development. *Developmental Science, 13,* 307–319. http://dx.doi.org/10.1111/j.1467-7687.2009.00885.x

Zalla, T., Amsellem, F., Chaste, P., Ervas, F., Leboyer, M., & Champagne-Lavau, M. (2014). Individuals with autism spectrum disorders do not use social stereotypes in irony comprehension. *PLoS ONE, 9*(4), e95568. http://dx.doi.org/10.1371/journal.pone.0095568

Zwaan, R. A. (2003). The immersed experiencer: Toward an embodied theory of language comprehension. *Psychology of Learning and Motivation, 44,* 35–62. http://dx.doi.org/10.1016/S0079-7421(03)44002-4

Received May 28, 2019
Revision received September 17, 2019
Accepted September 19, 2019 ■