CrossMark

Eye-Tracking Measurements of Language Processing: Developmental Differences in Children at High Risk for ASD

Meia Chita-Tegmark[1] · Sudha Arunachalam[2] · Charles A. Nelson[3,4,5] ·
Helen Tager-Flusberg[1]

Abstract To explore how being at high risk for autism spectrum disorder (ASD), based on having an older sibling diagnosed with ASD, affects word comprehension and language processing speed, 18-, 24- and 36-month-old children, at high and low risk for ASD were tested in a cross- sectional study, on an eye gaze measure of receptive language that measured how accurately and rapidly the children looked at named target images. There were no significant differences between the high risk ASD group and the low risk control group of 18- and 24-month-olds. However, 36-month-olds in the high risk for ASD group performed significantly worse on the accuracy measure, but not on the speed measure. We propose that the language processing efficiency of the high risk group is not compromised, but other vocabulary acquisition factors might have lead to the high risk 36-month-olds to comprehend significantly fewer nouns on our measure.

Keywords Autism · Development · Word comprehension · Processing speed · Eye tracking

✉ Meia Chita-Tegmark
meia@bu.edu

[1] Department of Psychological and Brain Sciences, Boston University, 64 Cummington Street, Boston, MA 02215, USA

[2] Department of Speech, Language, and Hearing Sciences, Boston University, Boston, MA, USA

[3] Division of Medical Sciences, Harvard Medical School, Harvard University, Cambridge, MA, USA

[4] Laboratories of Cognitive Neuroscience, Division of Developmental Medicine, Children's Hospital Boston, Boston, MA, USA

[5] Harvard Graduate School of Education, Harvard University, Cambridge, MA, USA

Introduction

Autism spectrum disorder (ASD) often appears in conjunction with language deficits that include poor receptive language skills (Barbaro and Dissanayake 2012; Luyster et al. 2008, 2011; Landa and Garrett-Mayer 2006; Yirmiya et al. 2007; Zwaigenbaum et al. 2005). These deficits in receptive language are, however, not limited to individuals with ASD. Children who have a sibling with ASD but do not receive a diagnosis of ASD themselves have been found to score worse than typically developing children on language measures (Messinger et al. 2013; Mitchell et al. 2006; Toth et al. 2007). This shows that language deficits can also extend to non-syndromic frist degree relatives. (Note that approximately 13 % (Sandin et al. 2014) to 20 % (Ozonoff et al. 2011; Elsabbagh and Johnson 2010) of children that have a sibling with ASD will develop ASD themselves, placing them at high risk for ASD and associated risk for developing language deficits.)

There are two categories of explanations for these results: either children at high risk for developing ASD are unable to show proficiency on these standardized tasks (for example because of the high response demand or because of the required interaction with the experimenter, which they might find difficult), or indeed they have language deficits. One candidate factor for language deficits is poor processing abilities. Fernald and colleagues have found language processing speed to be a good correlate and predictor of language abilities in other at-risk populations like children from low SES or late talkers (Fernald and Marchman 2012; Fernald et al. 2013). Very little is known about how children at high risk for ASD process language in real time compared to low risk controls: do they process language slower? Does this cause them to miss out on language learning opportunities?

The use of eye-gaze measures could offer some important clues about the underlying causes of the possible deficits in receptive language that standardized tests have flagged. First, by lowering the response demand and circumventing the communication demand, eye-gaze measures can rule out the possibility that lower scores on receptive language are simply by-products of task demands. Second, by measuring language processing in real time this type of measure can determine whether language deficits are caused by poor processing efficiency itself (possibly a genetic liability) or are instead related to other factors such as social and communication problems.

Receptive Language Abilities and High Risk for ASD

Studies have shown that at 24 and 36 months, children at high risk for ASD with no ASD outcome score significantly lower than low risk controls on the receptive language scale of the Mullen Scales of Early Learning (MSEL; Mullen 1995), an individually administered direct assessment measure. Mitchell et al. (2006) and similarly, Toth et al. (2007) reported differences between high-risk and low-risk children on the MSEL receptive language scale at 24 months and between 18 and 27 months respectively, with high-risk no ASD outcome children obtaining significantly lower scores. Using the same measure, Messinger et al. (2013) found significant group effects for the Mullen Verbal Developmental Quotients (DQ) at 36 months. However, even though as a group they had lower DQ scores, when further classifying the high-risk children into subclasses based on their ASD symptoms and developmental functioning, they found that only 21 % of them were characterized by low developmental functioning. This suggests that only a subset of the high-risk children had language problems. Hudry et al. (2014) similarly found a subclass of high-risk ASD children with no ASD outcome that showed an atypical language profile at 24 months, namely lack of receptive advantage over concurrent expressive language.

Similar results have been found in studies of children at high-risk for ASD (based on having an older sibling with ASD) whose outcome was not ascertained. Bedford et al. (2013) found significantly lower scores for verbal ability on the MSEL (using its T-scores for both receptive and expressive language) for high-risk children with unknown outcome when compared to low-risk controls at 24 months. Other offline measures such as the Reynell Developmental Language Scales (Reynell and Grubber 1990) and the Clinical Evaluation of Language Fundamentals—Preschool (Wiig et al. 1992) yielded similar results. Yirmiya et al. (2007) reported a significant difference between the number of high-risk children with unknown outcome and the number of low-risk children who scored one and two standard deviations below the mean on the Reynell Developmental Language Scales at 24 months. Similarly, at 36 months, more high-risk children with unknown outcome than low-risk children scored one or two standard deviations below the mean on the receptive scales of the Clinical Evaluation of Language Fundamentals—Preschool measure.

Using the MacArthur-Bates Communicative Development Inventory (MCDI; Fenson et al. 2007), a parent report measure of vocabulary and language knowledge, Ference and Curtin (2013) reported that low-risk infants understood more words than high-risk infants with unknown outcome at 12 months. However Mitchell et al. (2006), found no significant difference between high-risk children with no ASD outcome and low-risk children at either 12 or 18 months on selected variables from this measure. Comparing high-risk children with no ASD outcome to high-risk children with positive outcome, Mitchell et al. (2006) found significant differences in language abilities between the two groups both on the MSEL and the MCDI, with the positive outcome group scoring significantly lower than the no ASD outcome group.

Online Versus Offline Measures of Receptive Language

The studies above show that high-risk children obtain poorer receptive language scores. However, the measures used in these studies are exclusively offline measures, which evaluate comprehension by assessing complex behaviors that children make in response to language input after this input ends, and not while they are listening to it and trying to make sense of it. As a consequence, these measures might be missing some of the subtle real-time properties of comprehension. Very little is known about the real-time behavior related to language processing of both children with ASD and children at high risk for ASD. While standardized tests such as the MSEL reflect a "look-back" approach from the point of view of the tester, eye gaze measures of receptive language reflect a real-time one.

Measures of real-time language processing, such as eye gaze to the visual referent of a spoken word used as an index of comprehension and processing speed, have several advantages over offline measures of receptive language and may offer more information about the mechanisms underlying receptive language deficits.

First, online eye gaze measures of receptive language have a minimal response demand: to show comprehension, the child simply needs to look at the image that is being named. In contrast, in offline measures, children's comprehension is judged based on complex non-verbal

responses to words—such as pointing or executing a command. The complexity of the behavior required by these measures might prevent very young children from showing the true extent of their knowledge. This could lead to underestimation of their receptive language abilities. But behaviorally demanding tests of comprehension can also lead to overestimation of language abilities, since the context might influence the measured behavior, and reactions to the context may be mistaken for responses to the word itself (Houston-Price et al. 2007). Online eye gaze measures of receptive language may be more sensitive to children's word knowledge because of their low task demands and constrained context.

Second, online eye gaze measures of receptive language, unlike most standardized receptive language tests, require no interaction with an experimenter or parent, and virtually no social or pragmatic skills. This is an important feature considering that children at high risk for ASD often have social pragmatic deficits that might mask their receptive language abilities in interactive assessment tasks (Messinger et al. 2013).

Third, online eye gaze measures of receptive language are more sensitive to subtle language processing differences, like differences in speed of processing (Fernald et al. 2006) which has been found to be atypical in other populations at risk for language difficulties (Fernald and Marchman 2012; Fernald et al. 2013). It is therefore important to investigate language behaviors as they happen on shorter time scales for children at high risk for ASD as well. Online measures of language allow this by tracking gaze behavior indicating comprehension while the language processing is happening.

Finally, another difference worth mentioning between online eye gaze measures and (at least some) offline ones is that online eye gaze measures do not, in our paradigm require the child to have a rich lexical representation. Since for these measures the child chooses between only two images to look at when one of them is named, the child does not have to know exactly what the word means, but just that the target is a better fit than the distractor. For example, if the target word is "coffee" and the distractor is "quiche", the child requires only limited knowledge about the meaning of "coffee" in order to look at the correct image, such as that it refers to something liquid that comes in a cup. These eye gaze measures also allow children to use mutual exclusivity: a child who knows the name for one of the pictures may simply prefer the other picture on hearing an unknown word. Two year-olds at high risk for ASD have been shown to be able to use mutual exclusivity (Bedford et al. 2013).

Online eye gaze measures therefore allow us to capture even early stages in children's acquisition of any given word, but at the same time constrain the scope of what we can learn about children's word knowledge—children can have the same score on these measures without having the same level and depth of word knowledge.

Language Risk Factors and Online Gaze Measures

Online gaze measures of receptive language have proven useful in studying language processing in both typically developing children and children with language risk factors including children from households with a lower socioeconomic status (SES) and late talkers (Fernald and Marchman 2012; Fernald et al. 2006, 2008, 2013; Marchman and Fernald 2008). Fernald and colleagues have used this type of measure to test comprehension accuracy and speed of language processing. In this paradigm, which they call "looking-while-listening," children look at a pair of images and hear a noun label corresponding to one of them embedded in an instruction or question, e.g., "Where's the cookie?" If they understand the word, they are expected to fixate the target image in response.

This method of measuring language processing yields assessments of language that are developmentally sensitive and that have captured early atypicalities in processing speed for some populations at risk for receptive language deficits. Fernald and colleagues have shown that the speed and accuracy of speech processing increased between 15 and 25 months for typically developing children, and the magnitude of increases in speed and accuracy were correlated with gains in vocabulary both at 25 months and at age 8 years (Fernald et al. 2006; Marchman and Fernald 2008). Increased speed of processing at 18 months for late talkers was associated with accelerated vocabulary growth over the following year (Fernald and Marchman 2012). In other studies differences in processing efficiency between children from higher- and lower-SES families that eventually developed into a 6-month gap between the two groups were detected at 18 months using the looking-while-listening paradigm (Fernald et al. 2013).

Venker et al. (2013) have recently extended the use of the looking-while-listening paradigm to children with ASD and have shown that it can offer similar insights about patterns of language processing in this population as it does for typically developing children. They tested children with ASD between the ages of 3 and 6 on two measures: accuracy (based on the proportion of time children spent looking at the target image as opposed to the distractor) and speed of processing (the latency for the children to shift their gaze from the distractor to the target after hearing the word) and found that the two measures were correlated and that online accuracy was related to children's vocabulary comprehension on the MCDI three years earlier, as well as strongly correlated with offline language comprehension as measured by the Preschool Language Scale 4th edition

(Zimmerman et al. 2002) Auditory Comprehension raw score.

To the best of our knowledge, no one has yet used this paradigm to investigate language processing by comparing children at high versus low risk for ASD with no ASD outcome. Here, we report on data collected cross-sectionally using an online eye gaze measure of receptive language, more specifically a version of the looking-while-listening paradigm, adapted for an automatic eye tracker, to compare the performance of children at high risk for ASD with no ASD outcome and low-risk controls at 18, 24 and 36 months—a period of rapid growth in vocabulary during which eye-tracking measures provide a valid measure of language comprehension in typically developing children (Fernald et al. 2006). Using this adapted method with children at high risk for ASD with no ASD outcome might allow us to better map the ability continuum between children with ASD, studied by Venker et al. (2013), and the typically developing children studied by Fernald and colleagues.

The primary aim of this study is to detect potential differences between children at high risk for ASD with no ASD outcome and low-risk controls in vocabulary knowledge and speed of lexical processing. To probe these abilities, we chose nouns typically acquired early and late in childhood and tested three different age groups of children (18, 24 and 36 months) using two measures—an accuracy measure and a reaction time measure. Our accuracy measure provides an overall index of whether children understand the words and was operationalized in terms of the proportion of time children spent looking at the target image after word onset. The reaction time measure indicated the speed of language processing and represented the latency with which the children shifted their gaze to the named target picture after the word onset.

Based on the findings of Fernald et al. (2006) and Venker et al. (2013), we expected all children to be more accurate in directing gaze to the target picture when the target noun was a word typically acquired earlier in childhood and for accuracy to increase with age. Also, we expected words typically acquired earlier to be processed more quickly by both groups and for processing efficiency to increase with age.

In terms of group differences, four patterns of results with different implications are possible. First, there may be no differences between the two groups on either measure, in which case we would infer that offline measures may underestimate the abilities of children at high risk for ASD for some or all of the reasons mentioned above: demanding task, reliance on social and pragmatic skills, testing of rich lexical representations, or that eye gaze measures may mask the differences in richness of lexical representations between the two groups. Another possibility is that the high-risk group performs worse on both measures. In this case, slow processing speed and poor comprehension might be related, and the slow processing speed could be causing high-risk children to miss opportunities to learn new words. This relation between the two measures seems highly plausible given recent findings that link processing speed and vocabulary growth both in typically developing children (e.g., Fernald et al. 2008; Marchman and Fernald 2008) and children at risk for language difficulties such as late talkers (Fernald and Marchman 2012) or children from lower SES families (Fernald et al. 2013).

However, additional skills beyond fast processing of speech input are required for vocabulary acquisition, so it would not be surprising to see a dissociation between our measures. A third possibility is therefore that the high-risk children perform worse on the accuracy measure but equally well on the reaction time measure. This would suggest that their language processing speed is not inherently compromised, but that processing speed alone is not sufficient to acquire a large vocabulary and that high risk children might lack skills beyond processing speech input that are required for vocabulary acquisition such as social and communication skills. Finally another way to see a dissociation between the two measures is if high-risk children perform worse on the reaction time measure but not the accuracy measure. This would suggest that word acquisition relies on more than fast processing of the language input and that these children must be using alternate skills for word learning to compensate for slow processing. Therefore, the interplay between the accuracy and processing speed measures as well as how the two groups perform compared to each other, can provide evidence for or against some of the possible sources of language deficits in children at high risk for autism. Applying this measure at three different ages can further offer clues about the developmental trajectory of abilities such as processing speed and vocabulary knowledge for the two groups. Fernald et al. (2006) have shown that both accuracy and the speed of language processing are increasing between 18 and 25 months in typically developing children. This might not be the case for children at high risk for ASD whose developmental trajectory might show delays, regress or compensation, with patterns of results at one age not holding at another.

Methods

Participants

Two groups of children (divided into three age subgroups) participated in this study and contributed usable data (see Table 1): children at high risk for ASD (18-month-olds

Table 1 Participant demographics, by age and group

| | 18 months | | | | |
	High risk ASD	Low risk control			
n[a]	24	49			
Age in days: mean (SD)	561 (10)	561.47 (10.59)			
Gender (M:F)	19:15	26:23			
MSEL scores: mean (SD)			t value	df	p value
Verbal DQ	102.68 (18.93)	109.08 (15.23)	−1.52	68	0.13
Non-verbal DQ	104.07 (10.56)	107.15 (1052)	−1.14	67	0.26
	24 months				
	High risk ASD	Low risk control			
n[a]	20	36			
Age in days: mean (SD)	742 (13.76)	743.02 (9.96)			
Gender (M:F)	12:8	18:18			
MSEL scores: mean (SD)			t value	df	p value
Verbal DQ	109.84 (11.60)	115.6 (12.11)	−1.69	52	0.09
Non-verbal DQ	105.22 (11.99)	108.9 (15.31)	−0.9	52	0.37
	36 months				
	High risk ASD	Low risk control			
n[a]	21	17			
Age in days: mean (SD)	1131.9 (43.67)	1121.35 (26.95)			
Gender (M:F)	12:9	9:8			
MSEL scores: mean (SD)			t value	df	p value
Verbal DQ	*103.08 (12.78)*	*115.42 (7.88)*	*−3.36*	*33*	*<0.002*
Non-verbal DQ	101.27 (14.34)	109.58 (12.24)	−1.82	33	0.08

Eight children (4 high risk ASD, 4 low risk control) did not contribute MSEL data

Italics represent statistically significant differences

[a] The number of participants who contributed usable eye-tracking data

$n = 24$, 24 month-olds $n = 20$, 36-month-olds $n = 21$, for a total of $n = 52$[1]) and low-risk controls (18-month-olds $n = 49$, 24 month-olds $n = 36$, 36 month-olds $n = 17$ for a total of $n = 67$[2]). The participants were enrolled in a larger IRB approved longitudinal study conducted jointly by Boston University and Boston Children's Hospital. The larger study aimed at identifying markers, predictors and developmental trajectories of ASD. All children were

screened for exclusionary criteria (primary languages other than English, prematurity, extended stay in the neonatal intensive care unit, maternal drug or alcohol use during pregnancy or family history of genetic disorders associated with ASD). Additional data points ($n = 36$, 19 high risk ASD, 17 low risk control) were excluded from analyses due to incomplete sessions because of fussiness or technical difficulties.

The children were classified as either high risk ASD or low risk control based on their siblings' profile. The high-risk children were selected based on having one or more older siblings with a diagnosis of ASD, Asperger Syndrome, or Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS). The diagnostic information

---

[1] One child contributed data at all three ages and 11 children contributed data at two ages. The total number of children reflects this overlap between the age sub-groups.

[2] Five children contributed data at all three ages and 26 children contributed data at two ages. The total number of children reflects this overlap between the age sub-groups.
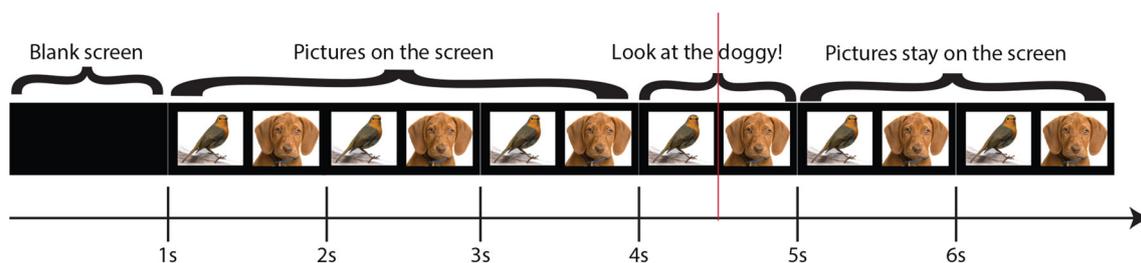
Fig. 1 Order and duration of stimuli presentation. A blank screen is shown for one second, followed by a pair of pictures that the child explores freely for three seconds, after which a recorded voice instructs the child to "Look at the doggy (target noun)!".The pictures remain on the screen for two additional seconds

was obtained through parent report and was confirmed using the Social Communication Questionnaire (SCQ; Rutter et al. 2003) for older siblings that were at least 48 months old, or through the Autism Diagnostic Observation Schedule—Revised (ADOS; Lord et al. 2000) for younger children. The low risk control children had a typically-developing older sibling and no known first-degree relatives with ASD or other neurodevelopmental disorder based on a screening interview. Since this study was concerned with children at high risk for ASD who do not later have an ASD diagnosis, additional data from children who met criteria for ASD on the ADOS and clinical evaluation at either 24 or 36 months (15 high risk ASD, 2 low risk control) were not included in this study.

Procedure

We used a similar paradigm to that of Fernald et al. (2006) to measure accuracy and speed of language processing at the three different ages (18, 24 and 36 months) for the two groups (high risk ASD and low risk control). In each session, children saw 20 image pairs consisting of one target and one distractor image, digitized photographs of objects presented side-by-side on a plain background, and heard ten recordings of spoken words, target nouns associated with the target picture (see "Appendix" for a list of the nouns associated with the pictures).

The target nouns were selected such that they could be categorized into two groups: words expected to be acquired early (*bottle*, *doggy*, *baby*, *car*, *shoe*), and words expected to be acquired late (*weasel*, *coffee*, *nail*, *kiwi*, *bow*) according to the lexical development norms for young children (Dale and Fenson 1996), which were derived from a norming study using the MCDI. The nouns expected to be acquired early were words on the MCDI understood by 90 % of children according to the comprehension norms at 18 months, while words expected to be acquired late were words on the MCDI understood by less than 42 % of children at 18 months, or words that were not even part of the MCDI (Dale and Fenson 1996). Having these two

groups of words in our measure allowed us to test the effects of experience on accuracy and speed of processing.

The target nouns appeared twice and were paired each time with a different distractor picture depicting the referent of a distractor noun. The typical age of acquisition of the distractor nouns varied according to the lexical development norms based on the MCDI. The pairs of pictures were the same for every child; they were presented in random order and the appearance of the target image on the left or right side of the screen was counterbalanced.

The presentation of the visual and auditory stimuli followed a pre-established script (see Fig. 1) for all 20 trials. A blank screen was presented for 1000 ms followed by a pair of pictures that the child viewed freely for 3000 ms. After the 3000 ms of free exploration, each child heard pre-recorded spoken instructions to look at one of the two images shown on the computer screen. The instructions were given by saying "Look at the…," followed by the noun matching the target image, after which the response was considered accurate if the child looked at the target image. The pre-recorded audio instructions were equal in length (1000 ms) and the duration of the spoken target nouns was the same for all trials (500 ms).

We used an automated infrared eye-tracker, the Tobii T60 (Tobii Technology, Stockholm, Sweden) to acquire the gaze data. This made it possible to work with higher time resolution than the one given by conventional recording cameras: one recording roughly every 16–20 ms. The visual and auditory stimuli were presented using the E-prime Extensions for Tobii (Psychology Sotware Tools, Sharpsburg, PA) software and the data analysis was performed using Mathematica 9 (Wolfram, Champaign, IL, USA).

The children were seated 60 cm from the eye-tracker on the parent's lap and the visual stimuli presented were approximately 9–12 cm in width (12 degrees of visual angle) and 4–6 cm apart. The calibration procedure we used was, a 5 point infant calibration within Tobii Studio 3.03 (Tobii Technology, Stockholm, Sweden) prior to the start of the experiment. The eye-tracker allows for large

eye-movements as well as head movement without losing data. For blinks and head movements that go outside of the screen area, gaze coordinates were not collected and an output code was generated. These portions of the gaze data did not enter the analyses.[3]

Measures

We defined two measures: accuracy to quantify word comprehension and reaction time to quantify the speed of language processing when children understood the words.

Accuracy as the Proportion of Looking Time at the Target Picture 300–1800 ms After Word Onset

This measure represents the total amount of time the children spent looking at the target picture as a percentage of total time looking at either picture during the 300–1800 ms time interval after word onset. This measure represents an average of each of the curves in Fig. 2, over a time period of 1500 ms that begins 300 ms after word onset. This measure was used by Fernald et al. (2006) and the rationale for it is that if children understood the word, they would spend a longer time looking at the image representing it (the target image) than at the other image. Fernald et al. (2006) have shown that indeed children spend a longer time looking at the target picture when the target noun is understood and Venker et al. (2013) have replicated these findings for children with autism. Therefore, for this study the expectation was that children initially (at the onset of the target noun) would perform at chance, with about 50 % of them looking at the target image and 50 % looking at the distractor. If children comprehended the word, after some processing time, estimated to be around 300 ms (Fernald et al. 2001; Swingley et al. 1999) the number of children looking at the target as opposed to the distractor would increase, and they would spend more time looking at the target than the distractor image during the 300–1800 ms time interval after the onset of the word.

Reaction Time as the Mean Latency to Shift to Target After Word Onset

This measure represents the mean amount of time it took children to switch to the target image in the time-window of 300–1800 ms after the noun onset. This measure was calculated for all the trials on which the children were not looking at the target image at the word onset and on which

the children showed understanding of the word by switching gaze successfully to the target in the 300–1800 ms time interval following word onset. The rationale for this measure is that the faster the children processed the language input (and were able to map the spoken word to the picture referent), the faster they would switch their gaze to the target image.

Data Analysis and Results

We adopted a conservative approach to selecting and analyzing the data, implementing several exclusion criteria. As mentioned above, we excluded sessions for which the data were incomplete, meaning that the child did not have 20 intact trials due to the child being fussy or due to technical difficulties. The first four trials of every session were also eliminated because of a clear learning effect: the correct response (looks to the target image following the noun onset) kept increasing over the first trials, and only stabilized after the first four, so these trials were treated as practice trials. Also, trials were eliminated if data was recorded for less than 5300 ms after the onset of the picture presentation since our measure depended on the gaze patterns during this time window. This left an average of 14.9 trials per child. Since the randomization process was independent of whether a child was in the high risk for autism or low risk control groups, this leaves the results unbiased. As expected, the fraction of early developing words presented to both groups is close to 50:50.6 % for low risk control children and 49.3 % for high risk for autism children.

For data analysis, large areas of interest (AOIs) were defined. Gaze points on the left side of the screen were counted as looks at Image 1 and gaze points on the right side of the screen were counted as looks at Image 2. Looks near the center of the screen, at the gap between the two stimuli, were counted as looks away and were not included in the statistical analyses.

Since some of the data were sampled with settings specifying different rates that ranged from 16.5 to 20 ms, to be able to analyze all data uniformly we made the approximation that children kept looking at the same place during the 16–20 ms interval between two samples. Our time window of interest was of 1800 ms starting from noun onset. No other temporal or fixation area filters were used for data smoothing.

Accuracy

The accuracy scores were compared in a 3 (age: 18, 24, 36 months) × 2 (group: high-risk ASD vs. low-risk control) × 2 (word type: early vs. late) analysis of variance. This analysis revealed significant main effects of

---

[3] The accuracy measure used in this paper is computed using the formula (proportion of time looking at target)/(proportion of time looking at target + proportion of time looking at distractor). This formula by definition does not take into account time spent looking away or lost data for which the eye-tracker outputs an error code.
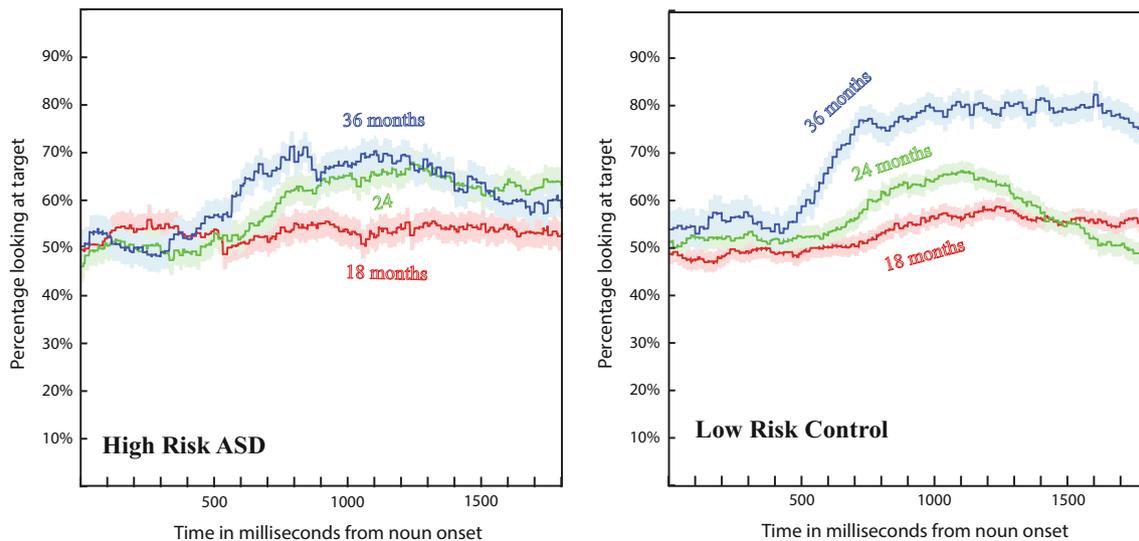
Fig. 2 Looks at the target image. The curves represent the percentage of trials in which the children were looking at the target image as opposed to the distractor image plotted as a function of time and starting with the target noun onset. The *curves* show an effect of age and differences between the two groups: high-risk ASD and low risk control. The *shaded* regions indicate one standard deviation error bars

Table 2 Accuracy and reaction time

Measure	Age	High risk ASD		Low risk control		t value	df	p value
		n	Mean (SE)	n	Mean (SE)			
Accuracy								
	18 months	24	0.53 (0.02)	49	0.54 (0.02)	−0.65	71	0.517
	24 months	20	0.60 (0.02)	36	0.58 (0.02)	1.02	54	0.313
	36 months	21	0.63 (0.02)	17	0.74 (0.02)	−4.21	36	<0.001
Reaction time								
	18 months	24	635.67 (49.67)	49	641.23 (30.81)	−0.1	71	0.921
	24 months	20	600.06 (37.63)	36	551.39 (23.76)	1.15	54	0.256
	36 months	21	476.27 (21.31)	17	496.06 (42.05)	−0.42	36	0.678

Italics represent statistically significant differences

age, $F(2, 322) = 30.5$, $p < 0.001$, and word type, $F(1, 322) = 103.8$, $p < 0.001$, as well as an age × group interaction, $F(2, 322) = 5.92$, $p = 0.003$. Follow-up tests showed that the high risk ASD and low risk control groups were similarly accurate for 18- and 24-month-olds, but the high risk ASD group was significantly less accurate than the low risk control group for 36-month-olds, $t(36) = -4.21$, $p < 0.001$ (see Table 2). This is in agreement with the difference on the offline measure of MSEL verbal DQ scores on which, at 36 months, high risk ASD children score significantly lower $t(33) = -3.36$, $p < 0.002$ (see Table 1). On our measure, the high risk ASD 36-month-olds had significantly lower scores than the low-risk control group of the same age, both for words expected to be acquired early (high risk ASD: M = 0.70, SD = 0.11; low risk control: M = 0.81, SD = 0.10),

$t(36) = -3.07$, $p = 0.004$, and for words expected to be acquired late (high risk ASD: M = 0.56, SD = 0.12; low risk control: M = 0.68, SD = 0.11), $t(36) = -3.16$, $p = 0.003$. Both groups had higher scores on words expected to be acquired early than on words expected to be acquired late.

Reaction Time

The reaction time scores were compared in a 3 (age: 18, 24, 36 months) × 2 (group: high-risk ASD vs. low-risk control) × 2 (word type: early vs. late) analysis of variance. This analysis revealed significant main effects of age, $F(2, 309) = 10.3$, $p < 0.001$, and word type, $F(1, 309) = 14.5$, $p < 0.001$. There was no significant main effect of group nor were there interaction effects,

suggesting that there were no differences in terms of the speed of processing between children at high risk for ASD and low risk controls.

Discussion

This study aimed to compare the receptive language abilities of children at high risk for ASD and low risk control children in order to shed light on the sources of receptive language deficits that are related to familial risk for ASD. We used an online eye gaze task that taps into vocabulary knowledge in different ways than offline measures of standardized tests do, by posing minimal response demands, by being better suited for children with potential social and pragmatic deficits, by measuring basic comprehension in the absence of a rich lexical representation and by offering information about the speed of language processing.

The results show both similarities and differences between children at high risk for ASD and low risk controls. The high risk ASD and low risk control groups were similar in their language processing speed as shown by the reaction time measure. The speed of processing increased with age for both groups, which is in line with the findings of Fernald et al. (2006) and words typically acquired earlier in childhood were processed more rapidly by children in both groups, which is in line with the findings of Venker et al. (2013). The two groups were also similar in their vocabulary knowledge at 18 and 24 months of age, as indicated by the accuracy measure. However, at 36 months the high risk ASD group performed significantly less accurately than low risk controls, suggesting that they comprehended fewer words at this age. This pattern of results confirms our third hypothesis, namely that the speed of processing for speech input does not seem to be compromised in children at high risk for ASD and that their receptive language difficulties might arise from deficits in other skills required for vocabulary acquisition.

One possibility is that the difference in word knowledge between the two groups at 36-months is due to different word learning strategies that reflect difficulties in processing social cues. Norbury et al. (2010) found that children with ASD are significantly worse than typically developing controls at using social cues to acquire new words. This might also be true of children at high risk for ASD, which might explain why they comprehend fewer nouns in our measure at 36 months than the low risk controls. Also, studies have shown that if they are to use social cues, children at high risk for ASD, as opposed to low risk controls, require rich and redundant social cues that combine gaze shifts, vocalizations and pointing in order to orient to a target (Presmanes et al. 2007; Stone

et al. 2007). Because of this, high risk ASD children might be learning less than low risk control children from listening to ambient conversations. Some of the words in our measure, for example "coffee," are words that one would not typically use in speech directed to infants, but that are often present in conversations that the child would hear. It is possible that high risk ASD children spend less time attending to other people's conversations and actions and using social cues to infer what they mean.

The difference in word knowledge between the two groups at 36 months could also be due to high risk ASD children engaging less proficiently in social and communicative interactions, and thus they might be missing on opportunities to learn new words. Toth et al. (2007) found that high risk ASD children (non-autistic siblings of children with ASD) have lower communication skills and use fewer words, gestures and responsive social smiles than typically developing controls and Bedford et al. (2013) found that high risk ASD children with unascertained outcome do not benefit from social feedback for the retention of words.

Another possibility is that high risk ASD children have a different language environment compared to low risk controls, due to the fact that they belong to a family with an older child with autism, which could have an impact on how parents are interacting with them. The eye-gaze receptive language measure can reflect differences in early language experience: for example, Fernald and Marchman (2012) used this measure to capture differences in language ability between children with high SES backgrounds and children with low SES backgrounds who differed in the amount and quality of the language input that they typically received. The measure proved sensitive to differences in ability that were a result of language experience; therefore, in the case of the groups studied here, this measure cannot elucidate whether the language deficits seen in children at high risk are due to ASD risk per se, through shared genetic inheritance, or to differences in experience with spoken language. Evidence against this interpretation however are the findings of Talbott et al. (2013), who have shown that mothers of non-diagnosed high risk ASD infants (many of whom overlap with our sample) gesture more frequently than mothers of low risk infants and that maternal gesture use promotes later language development. These findings seem to suggest that mothers are aware of their children's risk status and are actively promoting language development. Also, parents in our sample of participants have high educational backgrounds and high-income levels which is often correlated with high levels of parental verbal input (Hart and Risley 1995). Moreover, parents of high risk ASD children have likely participated in parent training as part of their older ASD child's intervention program.

The developmental pattern seen here, differences in word knowledge appearing at 36 months but not at 18 or 24 months, might be emerging as a combined effect of the nature of our measure and suboptimal word learning strategies employed by children at high risk for ASD. Our online gaze measure, while sensitive to processing speed, lacks sensitivity in terms of the richness of the lexical representation. It is possible that at 18 and 24 months the two groups appear to have the same word knowledge and the same speed in word processing but the word knowledge of low risk controls might be deeper, possibly as a result of their higher social engagement. Differences in the breadth of vocabulary knowledge would become visible only later, by 36 months, after children have gone through a period of rapid vocabulary acquisition. Further research using more complex versions of online gaze measures of receptive language (for example with variations in the number of distractor images and manipulations of how close in meaning the distractor nouns are to the target nouns, which would better capture the depth of vocabulary knowledge) is needed to elucidate the development of these differences seen at 36 months and the relationship between the processing speed and vocabulary knowledge.

There are several limitations to this study that will hopefully be addressed by future work. In order to conclusively elucidate how word learning is affected by poor social and communication skills and how this leads to poor receptive language in children at high risk for ASD, further work is needed to study both the relation between social and communication skills and receptive language skills in children at high risk for ASD and vocabulary acquisition in a word learning paradigm that would compare word-learning strategies for the two groups. Another limitation comes from the fact that this study employed cross-sectional data. A longitudinal analysis of the development of speech processing and word learning between 24 and 36 months could reveal more in terms of individual differences of the children in both groups. Research adopting individual differences or clustering approaches would also help understand whether sub-phenotypes characterized by language difficulties exist among high risk ASD children as suggested by the findings of Messinger et al. (2013). Another limitation is the interpretability and utility of this measure at 18 months, age at which we found that both typically developing children and children at high risk for ASD in our study perform at chance level. Future research is needed to test the differences between our modified paradigm and that used by Fernald et al. (2006) who have successfully used this measure with 15 month-olds.

It would also be valuable to directly compare children at high risk for ASD with no ASD outcome to children at high-risk with ASD outcome. The findings of Mitchell et al. (2006) suggest that children at risk for ASD with no ASD outcome share in part the language impairment of children with ASD outcome but to a lesser degree and less pervasively. This raises the question of whether the same mechanisms are causing high risk children with no ASD outcome and those with ASD outcome to have receptive language deficits.

In conclusion, the eye-tracking online eye gaze measure of receptive language offered some important insights into the nature of the language difficulties of the high risk ASD group: the two groups did not differ on any of the two measures at 18 or 24 months but children at high risk for ASD had significantly lower accuracy at 36 months than the low risk control group, although the groups did not differ on the reaction time measure at this age. This pattern of results suggests that the speech processing speed of high risk ASD children is not compromised and that they might differ from low risk control children in their word acquisition abilities by failing to form more robust lexical representations of words using social and communicative skills. These results also raise new questions about ASD-related language deficits, which can be further studied using variants of this online eye gaze measure of receptive language (with a different number of distractor images or a different choice of distractor nouns). For example: are there differences in the depth of vocabulary knowledge between the two groups? What is the relationship between the depth of vocabulary knowledge and the speed of processing? Is this different for the two groups? In a word-learning paradigm, how does social input affect the word knowledge and the speed of processing for the two groups? More detailed future studies, with a focus on the interplay between processing speed, depth as well as breadth of vocabulary knowledge and word acquisition strategies will hopefully further elucidate mechanisms underlying word learning and shed more light on the causes of ASD-related language deficits.

Appendix

See Table 3.

Table 3 Noun pairs associated with image pairs

Early-acquired		Late-acquired	
Target nouns	Distractor nouns	Target nouns	Distractor nouns
Bottle	Bowl	Weasel	Chipmunk
	Spoon		Fox
Doggy	Bird	Coffee	Quiche
	Cow		Crèmebrulee
Baby	Boy	Nail	Hammer
	Girl		Wrench
Car	Bus	Kiwi	Blackberry
	Boat		Kumquat
Shoe	Pants	Bow	Belt
	Hat		Box

References

Barbaro, J., & Dissanayake, C. (2012). Developmental profiles of infants and toddlers with ASD spectrum disorders identified prospectively in a community-based setting. *Journal of Autism and Developmental Disorders, 42*, 1939–1948. doi:10.1007/s10803-012-1441-z.

Bedford, R., Gliga, T., Frame, K., Hudry, K., Chandler, S., Johnson, M. H., & Charman, T. (2013). Failure to learn from feedback underlies word learning difficulties in toddlers at risk for ASD. *Journal of Child Language, 40*(1), 29–46. doi:10.1017/S0305000912000086.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers, 28*(1), 125–127.

Elsabbagh, M., & Johnson, M. H. (2010). Getting answers from babies about autism. *Trends in Cognitive Sciences, 14*(2), 81–87. doi:10.1016/j.tics.2009.12.005.

E-Prime Exstensions for Tobii [Computer software]. Sharpsburg, PA: Psychology Software Tools.

Fenson, L., Marchman, V. A., Thal, D., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates communicative development inventories: User's guide and technical manual* (2nd ed.). Baltimore: Brookes.

Ference, J., & Curtin, S. (2013). Attention to lexical stress and early vocabulary growth in 5-month-olds at risk for ASD spectrum disorder. *Journal of Experimental Child Psychology, 116*, 891–903. doi:10.1016/j.jecp.2013.08.006.

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development, 83*(1), 203–222. doi:10.1111/j.1467-8624.2011.01692.x.

Fernald, A., Marchman, V. A., & Hurtado, N. (2008). Input affects uptake: How early language experience influences processing efficiency and vocabulary learning. *7th IEEE international conference on development and learning, 2008. ICDL 2008* (pp. 37–42). doi:10.1109/DEVLRN.2008.4640802

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science, 16*(2), 234–248. doi:10.1111/desc.12019.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology, 42*, 98–116. doi:10.1037/0012-1649.42.1.98.

Fernald, A., Swingley, D., & Pinto, J. (2001). When half a word is enough: Infants can recognize spoken words using partial acoustic–phonetic information. *Child Development, 72*, 1003–1015.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008b). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. Fernandez, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97–135). Amsterdam: John Benjamins.

Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes Publishing.

Houston-Price, C., Mather, E., & Sakkalou, E. (2007). Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *Journal of Child Language, 34*, 701–724. doi:10.1017/S0305000907008124.

Hudry, K., Chandler, S., Bedford, R., Pasco, G., Gliga, T., Elsabbagh, M., et al. (2014). Early language profiles in infants at high-risk for autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44*, 154–167.

Landa, R., & Garrett-Mayer, E. (2006). Development in infants with ASD spectrum disorders: A prospective study. *Journal of Child Psychology and Psychiatry, 47*, 629–638. doi:10.1111/j.1469-7610.2006.01531.x.

Lord, C., Risi, S., Lambrecht, L., Cook, Edwin 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*(3), 205–223. doi:10.1023/A:1005592401947.

Luyster, R. J., Kadlec, M. B., Carter, A., & Tager-Flusberg, H. (2008). Language assessment and development in toddlers with ASD spectrum disorders. *Journal of ASD and Developmental Disorders, 38*, 1426–1438. doi:10.1007/s10803-007-0510-1.

Luyster, R. J., Seery, A., Talbott, M. R., & Tager-Flusberg, H. (2011). Identifying early-risk markers and developmental trajectories for language impairment in neurodevelopmental disorders. *Developmental Disabilities Research Reviews, 17*(2), 151–159. doi:10.1002/ddrr.1109.

Mathematica (Version 9.0) [Computer software]. Champaign, IL:Wolfram.

Marchman, V., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science, 11*(3), F9–F16. doi:10.1111/j.1467-7687.2008.00671.x.

Messinger, D., Young, G. S., Ozonoff, S., Dobkins, K., Carter, A., Zwaigenbaum, L., et al. (2013). Beyond ASD: A baby siblings research consortium study of high-risk children at three years of age. *Journal of the American Academy of Child and Adolescent Psychiatry, 52*(3), 300–308. doi:10.1016/j.jaac.2012.12.011.

Mitchell, S., Brian, J., Zwaigenbaum, L., Roberts, W., Szatmari, P., Smith, I., et al. (2006). Early language and communication development of infants later diagnosed with ASD spectrum disorder. *Journal of Developmental and Behavioral Pediatrics, 27*, S69–S78. doi:10.1097/DBP.0b013e3181df7f3c.

Mullen, E. M. (1995). *Mullen scales of early learning* (AGS ed.). Circle Pines, MN: American Guidance Service Inc.

Norbury, C. F., Griffiths, H., & Nation, K. (2010). Sound before meaning: Word learning in autistic disorders. *Neuropsychologia, 48*, 4012–4019. doi:10.1016/j.neuropsychologia.2010.10.015.

Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., et al. (2011). Recurrence risk for autism spectrum disorders: A Baby Siblings Research Consortium study. *Pediatrics, 128*(3), e488–e495.

Presmanes, A. G., Walden, T. A., Stone, W. L., & Yoder, P. J. (2007). Effects of different attentional cues on responding to joint

attention in younger siblings of children with ASD spectrum disorders. *Journal of ASD and Developmental Disorders, 37*(1), 133–144. doi:10.1007/s10803-006-0338-0.

Reynell, J. K., & Grubber, C. P. (1990). *Reynell developmental language scale*. Los Angeles: Western Psychological Association.

Rutter, M., Bailey, A., & Lord, C. (2003). *Social communication questionnaire*. Los Angeles: Western Psychological Services.

Sandin, S., Lichtenstein, P., Kuja-Halkola, R., Larsson, H., Hultman, C. M., & Reichenberg, A. (2014). The familial risk of autism. *JAMA, 311*(17), 1770–1777.

Stone, W. L., McMahon, C. R., Yoder, P. J., & Walden, T. A. (2007). Early social-communicative and cognitive development of younger siblings of children with ASD spectrum disorders. *Archives of Pediatrics and Adolescent Medicine, 161*, 384–390. doi:10.1001/archpedi.161.4.384.

Swingley, D., Pinto, J., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71*, 73–108. doi:10.1016/S0010-0277(99)00021-9.

Talbott, M., Nelson, C. A., & Tager-Flusberg, H. (2013). Maternal gesture use and language development in infant siblings of children with autism spectrum disorder. *Journal of Autism and Developmental Disorders,*. doi:10.1007/s1-803-013-1820-0.

Tobii Studio [Computer software]. Stockholm, Sweden:Tobii Technology.

Toth, K., Dawson, G., Meltzoff, A. N., Greenson, J., & Fein, D. (2007). Early social, imitation, play and language abilities of young non-autistic siblings of children with ASD. *Journal of Autism and Developmental Disorders, 37*, 145–157. doi:10.1007/s10803-006-0336-2.

Venker, C. E., Eernisse, E. R., Saffran, J. R., & Ellis Weismer, S. (2013). Individual differences in the real-time comprehension of children with ASD. *Autism Research,*. doi:10.1002/aur.1304.

Wiig, E. H., Secord, W., & Semel, E. (1992). *CELF-preschool: Clinical evaluation of language fundamentals—preschool version*. San Antonio, TX: Psychological Corporation.

Yirmiya, N., Gamliel, I., Shaked, M., & Sigman, M. (2007). Cognitive and verbal abilities of 24 to 36-month-old siblings of children with ASD. *Journal of Autism and Developmental Disorders, 37*, 218–229. doi:10.1007/s10803-006-0163-5.

Zimmerman, I., Steiner, V., & Pond, R. (2002). *Preschool Language Scale* (4th ed.). San Antonio, TX: The Psychological Corporation.

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of ASD in the first year of life. *International Journal of Developmental Neuroscience, 23*, 143–152. doi:10.1016/j.ijdevneu.2004.05.001.