Journal of Autism and Developmental Disorders
https://doi.org/10.1007/s10803-019-03924-2

ORIGINALPAPER

Does an Early Speech Preference Predict Linguistic and Social-Pragmatic Attention in Infants Displaying and Not Displaying Later ASD Symptoms?

Amy Yamashiro[1] · Suzanne Curtin[2] · Athena Vouloumanos[1]

Abstract

Human infants show a robust preference for speech over many other sounds, helping them learn language and interact with others. Lacking a preference for speech may underlie some language and social-pragmatic difficulties in children with ASD. But, it's unclear how an early speech preference supports later language and social-pragmatic abilities. We show that across infants displaying and not displaying later ASD symptoms, a greater speech preference at 9 months is related to increased attention to a person when they speak at 12 months, and better expressive language at 24 months, but is not related to later social-pragmatic attention or outcomes. Understanding how an early speech preference supports language outcomes could inform targeted and individualized interventions for children with ASD.

Keywords Speech preference · Linguistic attention · Social-pragmatic attention · Language outcomes · Autism spectrum disorder

From birth, human infants privilege speech over other sounds, listening longer to speech over altered or naturally occurring non-speech sounds (Colombo and Bundy 1981; Samples and Franklin 1978; Vouloumanos et al. 2010; Spence and DeCasper 1987; Vouloumanos and Werker 2004, 2007). This speech preference may help infants learn to communicate and interact with others (Curtin and Vouloumanos 2013; Vouloumanos and Curtin 2014), whereas lacking a speech preference may hinder infants' language and social-pragmatic skills. A lack of preference for speech over non-speech has been observed in children diagnosed with autism spectrum disorder (ASD; Klin 1991, 1992; Kuhl et al. 2005), who have difficulty interacting socially with others and often show deficits in language (American Psychiatric Association 2013; Center for Disease Control and Prevention 2018; Wodka et al. 2013) and this difference in speech preference may begin in infancy (Curtin and Vouloumanos 2013). However, it is unclear how an early speech preference supports later language and social-pragmatic outcomes. An early speech preference may allow infants to identify and attend to the relevant linguistic and social-pragmatic information from communicative interactions in their environment, and may help infants learn language and learn to interact with others (Chawarska et al. 2012; Yamashiro and Vouloumanos 2019). Here we asked, for infants who displayed symptoms of ASD at 24 months and infants who did not display ASD symptoms at 24 months, whether an early speech preference relates to how they attend to linguistic and social-pragmatic information in communicative interactions, and how speech preference predicts language and social-pragmatic outcomes over 1 year later.

Only hours after birth (Butterfield and Siperstein 1970; Spence and DeCasper 1987; Vouloumanos and Werker 2007) and throughout their first year of life (Colombo and Bundy 1981; Curtin and Vouloumanos 2013; Vouloumanos and Werker 2004), infants exhibit a robust perceptual bias for speech. Infants listen longer to speech compared to many non-speech sounds including filtered speech (Spence and DeCasper 1987), white noise (Butterfield and Siperstein 1970; Colombo and Bundy 1981), and synthetic sine-wave analogs of speech, which share many acoustic properties with speech like pitch contour, but lack the broad band

✉ Amy Yamashiro
amy.yamashiro@nyu.edu

1 Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, USA

2 Department of Psychology, University of Calgary, 2500 University Drive NW, Calgary, AB T2N 1N4, Canada

Published online: 21 February 2019

frequencies and harmonic structure that characterize human speech (Vouloumanos and Werker 2004, 2007). Infants also listen longer to speech compared to environmental sounds like running water and non-speech human sounds like laughter or coughing, suggesting that speech is preferred even to naturally occurring sounds and other human vocalizations (Shultz and Vouloumanos 2010). Listening to speech also elicits different neural responses in newborns and 3- and 4-month-olds than reversed speech, tones, or monkey calls (Dehaene-Lambertz 2000; Dehaene-Lambertz et al. 2002; Peña et al. 2003; Shultz et al. 2014). Thus, speech is an important and privileged signal for infants—infants' robust bias for speech in their first year of life may play a key role in orienting them to the relevant social and communicative signals of their own species.

Orienting to speech early in life helps infants learn from the people around them (reviewed in Vouloumanos and Waxman 2014). Between 3 and 12 months, infants use speech, but not non-speech, to categorize (Balaban and Waxman 1997; Ferry et al. 2010; Fulkerson and Waxman 2007), label (Mackenzie et al. 2011), and individuate (Xu 2002; Xu et al. 2005) objects. By 5 months, infants use speech to learn who to communicate with, matching speech, but not non-speech vocalizations like laughter, to human rather than monkey faces (Vouloumanos et al. 2009). By 6 months, infants recognize speech and eye gaze as signals of communication (Csibra and Gergely 2009, 2011) and understand that speech, but not non-speech, can transfer information to others (Cheung et al. 2012; Martin et al. 2012; Vouloumanos et al. 2014). In these studies, infants watched a scene in which a Communicator could no longer reach a preferred target object, so she vocalized to a Listener using speech or non-speech. The Listener, who could reach the objects, selected one of the two objects and presented it to the Communicator. Infants' look relatively longer when the Listener selects the incorrect non-target object compared to the correct target object after the Communicator speaks, but look equally at correct and incorrect responses after the Communicator produces coughs, emotional vocalizations, or humming, suggesting that infants understand that only speech, but not non-speech, can inform the Listener about the target object (Martin et al. 2012; Vouloumanos 2018; Vouloumanos et al. 2014; Yamashiro and Vouloumanos 2018). Within their first few years, infants use speech to learn how to communicate and interact with others.

Just as a speech preference helps infants learn from others, lacking a speech preference may underlie atypical language and social-pragmatic abilities. Between 2 and 6 years, children with ASD prefer non-speech to speech (Klin 1991, 1992), listening longer to a compilation of environmental noises and overlapping voices than to their mothers' speech (Klin 1991, 1992; Kuhl et al. 2005) and listening longer to a synthetic sine-wave analogs of speech than to infant-directed speech (Kuhl et al. 2005). Moreover, a lack of speech preference is related to social-pragmatic abilities and language abilities in children with ASD: a greater preference for non-speech compared to speech is related to showing more ASD symptoms, initiating joint attention less often, and saying fewer words (Kuhl et al. 2005). Further, listening longer to speech at 12 months is related to saying more words at 18 months, while listening longer to non-speech is not related to later language abilities (Vouloumanos and Curtin 2014). Without a preference for speech, children with ASD may not orient to important linguistic and social-pragmatic information in their environment, which could explain some of the language and social-pragmatic difficulties characteristic of this population.

This lack of a speech preference may begin in infancy. Though ASD cannot be reliably diagnosed until 2 years (e.g., Guthrie et al. 2013), prospective studies have examined earlier differences in the infant siblings of children diagnosed with ASD (Zwaigenbaum et al. 2007, 2013), who are at a 12-fold higher risk of being diagnosed themselves (Ozonoff et al. 2011). Studying this high-risk sample prospectively increases the odds that some of the infants will be diagnosed with ASD. It can also reveal early behavioral and biological markers of ASD, and identify when and how developmental trajectories diverge (Zwaigenbaum et al. 2007, 2013). Some studies have shown that high-risk infants also lack a preference for speech. While the low-risk infant siblings of neurotypical children listen longer to speech than non-speech at 12 months, high-risk infants listen equally to both sounds, and for high-risk infants, a greater preference for non-speech at 12 months is related to showing more autism-like behaviors at 18 months (Curtin and Vouloumanos 2013). This suggests that lacking a preference for speech may disrupt infants' ability to learn how to communicate and interact with others. Though the diagnostic outcomes of high-risk infants were not known in this study, it is possible that the differences in speech preference between the low- and high-risk infants were driven by the high-risk infants who later displayed symptoms of ASD.

Though a speech preference can predict later language abilities in infancy (Vouloumanos and Curtin 2014) and a lack of a preference can predict ASD symptoms in toddlers with ASD and high-risk infants (Curtin and Vouloumanos 2013; Kuhl et al. 2005), it is unclear why. It is possible that attending to speech allows infants to more easily identify and attend to the linguistic and social-pragmatic information in communicative interactions. By 12 months, infants can process communicative interactions in real time with adult-like accuracy (Thorgrimsson et al. 2014, 2015; Yamashiro and Vouloumanos 2018), and their linguistic and social-pragmatic attention during communicative interactions separately predicts language and social-pragmatic abilities 1 year later (Yamashiro and Vouloumanos 2019).

Specifically, for infants displaying and not displaying later ASD symptoms, a greater increase in their slope of looking at a person when she speaks (linguistic attention) at 12 months is related to producing more words at 24 months. And, for infants displaying later symptoms, not detecting an incongruity in a person's response to speech (social-pragmatic attention) at 12 months, is related to showing more ASD symptoms at 24 months (Yamashiro and Vouloumanos 2019). Thus, attending to speech early in infancy may be related to infants' attention to the relevant linguistic and social-pragmatic information in their environment as well as later language and social-pragmatic abilities.

In this prospective longitudinal study, we asked whether a preference for speech at 9 months in infants displaying and not displaying later ASD symptoms can predict their linguistic and social-pragmatic attention during communicative interactions at 12 months and language and social-pragmatic outcomes at 24 months. To measure infants' speech preference at 9 months, we examined infants' average looking time to a screen while listening to 3 trials each of speech and complex non-speech. To measure infants' linguistic and social-pragmatic attention at 12 months, we tracked infants' eye movements as they watched a video of a communicative interaction in which we manipulated a Listener's response to a Communicator's speech (Martin et al. 2012). All infants were evaluated for ASD symptoms and language abilities at 24 months.

We expected that on average, infants not displaying later ASD symptoms would listen longer to speech compared to non-speech, but infants displaying later symptoms would listen equally or longer to non-speech than speech, as in prior studies (Curtin and Vouloumanos 2013; Klin 1991, 1992; Kuhl et al. 2005). If infants' speech preference predicts later linguistic attention, we expected that a greater preference for speech at 9 months would predict a greater increase in their slope of looking at a Communicator's face when she spoke. Infants' looking at a speaker's face (including eyes and/or mouth) can reflect their attention to the linguistic information in the scene (Chawarska et al. 2012; Tenenbaum et al. 2015; Yamashiro and Vouloumanos 2019; Young et al. 2009). Moreover, measuring infants' slope of looking at a Communicator rather than calculating average overall looking times, as in prior studies (Chawarska et al. 2012; Tenenbaum et al. 2015; Vivanti et al. 2016; Young et al. 2009), reveals how the timing, rate, and trajectory (increasing or decreasing) of their looking changed in real time as a Communicator spoke. Overall looking times could obscure these informative patterns and do not capture how infants' attention changes in real time.

If infants' speech preference is related to later social-pragmatic attention, we expected infants' speech preference to predict the degree to which they detected an incongruity in a Listener's response to speech (social-pragmatic information). Recognizing how a person should respond to speech and noticing when the response is incongruent requires infants to attend to the social-pragmatic information within the communicative interaction (Yamashiro and Vouloumanos 2019). Finally, in line with previous studies with older infants (Curtin and Vouloumanos 2013; Vouloumanos and Curtin 2014), we expected that a greater preference for speech at 9 months would predict better expressive language and social-pragmatic outcomes at 24 months.

Method

Participants

Participants were 107 full-term 9-month-old infants (50 females) who had at least one older sibling, and who participated as part of a larger prospective longitudinal study. The sample consisted of 61 low-risk infants, with no immediate family members diagnosed with ASD (28 females; $M_{age}=9.10$ months; range 7.23–10.30), and 46 high-risk infants, with at least one older sibling diagnosed with ASD by a pediatrician or psychologist (22 females; $M_{age}=9.18$; range 7.83–9.90). Low-risk infants were recruited from maternity wards at local hospitals, through parent fairs, flyers, and advertisements. High-risk infants were recruited through local pediatric clinics and autism organizations. Infants were excluded if they presented with neurological disorder of known etiology, significant sensory or motor impairment, major physical abnormalities, and history of serious head injury and/or neurological disease. All participants were evaluated for potential ASD symptoms and likelihood of diagnosis at 24 months by a licensed clinical psychologist (see ASD symptoms and diagnosis below).

Data were included only for infants who contributed both speech preference data at 9 months and eye-tracking data at 12 months. The final sample consisted of 47 infants (18 females): 27 low-risk infants (11 females) and 20 high-risk infants (7 females). However, data collection for the observational measures of general cognitive development and language development (see below) varied slightly depending on whether each infant completed all tasks at each visit (see Table 1 for sample sizes for each measure). This sample size was justified by an a priori power analysis (GPOWER; Faul et al. 2007) using the effect size from a related study that found that for infants displaying and not displaying later ASD symptoms, slope of looking at a Communicator's predicted later language outcomes (Yamashiro et al. 2017). The power analysis suggested that we would need at least $n=40$ infants to detect a significant effect with 80% power at alpha level $p<.05$ based on an effect size of $f^2=0.21$.

Speech preference data from an additional 30 infants (15 low risk, 15 high risk) were excluded due to computer

Table 1 Observed means and standard deviations for looking times during the speech preference task at 9 months, looking behaviors to the Communicator and Listener, and general cognitive abilities at 12 months, and outcome variables at 24 months for infants displaying and not displaying later ASD symptoms

	No ASD symptoms			ASD symptoms		
	N	*M*	*SD*	*N*	*M*	*SD*
9 months						
Speech	34	13.5	6.1	13	12.2	5.0
Non-speech	34	10.8	5.6	13	14.8	6.2
Speech preference score	34	2.7	6.8	13	−2.6	4.3
12 months						
Communicator slope	34	0.07	0.12	13	0.07	0.12
Listener difference score	34	0.09	0.18	13	0.00	0.18
General cognitive abilities	34	98.8	11.7	11	94.7	14.5
24 months						
MB-CDI Words Produced	32	58.1	28.4	12	35.5	34.0
ADOS Social Affect Raw score	34	2.0	1.8	13	14.0	4.2
ADOS Restricted and Repetitive Behaviors Raw score	34	0.82	1.1	13	2.4	1.4

program error (9:3 low risk, 6 high risk) or experimenter error (21:12 low risk, 9 high risk). A computer programming error caused incorrect trial orders such that infants heard the same sound type in more than 2 consecutive trials. Experimenter errors occurred when errors in online coding caused test trials to end prematurely or to continue past the minimum look-away criteria of 2 consecutive seconds (see Procedure below). These errors resulted in an inconsistency in the contingency between infants' looks away from the screen and the end of the trial, making it difficult to assess infants' preference for either sound type.

Eye-tracking data from an additional 27 infants were excluded due to fussiness or crying (7:5 low risk, 2 high risk) or insufficient gaze tracking by the eye tracker with data loss or inattention during more than 50% of two or more trials (19:13 low risk, 6 high risk). Finally, data were excluded for 3 infants who received a clinical best estimate diagnosis of atypical development without ASD at 24 months (1 low risk, 2 high risk).

Parents gave informed consent on behalf of their infants and received $20 at each visit or $100 at the last visit (compensation varied by site following the guidelines of each institution's ethics review board), as well as a certificate and small gifts of toys or t-shirts. All procedures were approved independently by the University Committee on Activities Involving Human Subjects at New York University (IRB-FY 2016-170) and the Conjoint Health Research Ethics Board at the University of Calgary (REB15-1002).

Speech Preference Task at 9 months

Stimuli

Stimuli were speech and complex non-speech sounds that have been used in multiple prior studies and are robust in eliciting speech preferences across infants' first year: at birth (Vouloumanos and Werker 2007), 2–7 months (Vouloumanos and Werker 2004) and 12 months (Vouloumanos and Curtin 2014).

Speech Human speech consisted of monosyllabic nonsense words spoken by a female native English speaker. The 12 distinct tokens (six 'lif' tokens and six 'neem' tokens) varied in intonational contour (average minimum and maximum pitch: 197 Hz and 350 Hz) and duration (525–1155 ms).

Complex Non-speech Non-speech stimuli consisted of time varying sinusoidal waves tracking the main regions of significant energy, specifically the fundamental frequency and the first three formants of speech. These non-speech analogues reproduce the main spectral and temporal changes in natural speech, retaining the duration, pitch contour, amplitude envelope, relative formant amplitude, and relative intensity of speech counterparts. By tracking changes across time for the peak frequencies of their speech counterparts, complex non-speech analogues preserved the pitch contour and followed very closely the spectral and timing changes of natural speech, however complex non-speech analogues differed from the speech stimulus in: (1) voice quality (non-speech analogues have none), (2) naturalness or biological quality (non-speech analogues are artifacts), and (3) characteristics of the source (speech has one source, the vocal tract, while non-speech analogues have four, one per sinusoidal tone). For a complete description of stimulus creation, see Vouloumanos et al. 2001; Vouloumanos and Werker 2004.

Procedure

Testing was conducted in a sound-attenuated room using an infant-controlled sequential preferential looking procedure

(Cooper and Aslin 1990; Vouloumanos and Werker 2004). In this procedure, infants controlled the trial onset and offset by looking at or away from a central monitor displaying a bulls-eye. An online coder, who was blind to the type of experimental trial, coded when and for how long infants were looking at or away from the monitor. Infants sat on a parent's lap 35″ (89 cm) in front of a 30″ (76.25 cm) computer monitor. At the start of the experiment, infants' attention was drawn to the monitor by a flashing colored circle. Once the infant fixated on the monitor, a stationary bulls-eye appeared in tandem with one set of sounds, either speech or non-speech presented at an average amplitude of 60 dB (± 5 dB). Sounds played until the infant looked away from the monitor for 2 consecutive seconds or for a maximum for 40 s, at which time the sounds and the visual display ceased. Infants' attention was drawn back to the monitor by the flashing colored circle and the other sound set was then presented. Before and after testing, infants heard a 40 s clip of Bach's Concerto for Violin and Orchestra No 1 in A minor (BWV 1041-III. Allegro Assai). This familiarized infants with the procedure (e.g., Cooper and Aslin 1994; Vouloumanos and Werker 2004). During testing, three trials each of speech and non-speech were presented in pseudorandom order such that each pair of trials included one sound of each type for a total of 6 trials. To obtain precise measures of infants' looking time to the monitor during speech and complex non-speech analogues, we used frame-by-frame offline coding (30 frames per second) by offline coders blind to the type of experimental trial. An intraclass correlation (ICC) estimate based on a mean-rating, absolute-agreement, 1-way random effects model shows high reliability, $ICC = 0.97$, 95% CIs [0.96, 0.98] (Koo and Li 2016; Shrout and Fleiss 1979). Infants rarely looked at the screen for the maximum trial length: infants not displaying later symptoms of ASD looked at the screen for 37 s or more during 3% of test trials and infants displaying later symptoms of ASD, during 4% of test trials.

Eye-Tracking Task at 12 Months

Stimuli

At 12 months, infants watched a video of two actors interacting with two novel objects, based on a live action study (Martin et al. 2012; Vouloumanos et al. 2014). Each infant saw a target experimental outcome and a non-target experimental outcome (see Procedure below). In the target outcome, the target object was an orange, angular, hourglass shaped object with a purple base and block in the center and the non-target object was an inverted ring tower covered in pink tape. In the non-target outcome, the target object was a gray box with a red base and rod on top and the non-target object was a yellow cone with a green, coiled wire on top.

Procedure

Each experimental outcome included three trials: two familiarization trials and one test trial. During the familiarization trials, the Communicator was present in the back window with the top of her face and arms visible. Two novel objects were within her reach. She looked briefly at both objects and picked up the target object. The familiarization trial was presented two times, each lasting 4 s with an interstimulus interval of 430 ms (Fig. 1a). During the 6 s action phase of the test trial, both the Communicator and Listener were present. However, the Communicator could no longer reach the objects, because her arms were behind an opaque occluder. The Listener could reach both objects. The Communicator looked briefly at both objects, turned to the Listener, and produced a speech vocalization (e.g., "koba"). Then, the Listener selected either the target object (target experimental outcome) or non-target object (non-target experimental outcome) and lifted it just below the Communicator's face (Fig. 1b). After the action phase, the image then froze for 20 s (Fig. 1c). Infants within each risk group were randomly assigned to see either the target experimental outcome first or the non-target experimental outcome first.

We selected this communicative scenario to examine infants' attention to the linguistic and social-pragmatic information in a communicative scene as it has been validated in multiple prior studies with infants 6–12 months (Martin et al. 2012; Vouloumanos et al. 2014; Yamashiro and Vouloumanos 2018, 2019).

Apparatus

Infants were seated upright in a high chair approximately 60 cm from a 29 cm × 47 cm screen in a sound-attenuated room. At a viewing distance of 60 cm, the stimulus scene measured 27.8° vertical and 43.1° horizontal visual angle. Gaze was measured with the SensoMotoric Instruments (SMI) RED-m infrared eye tracker system. The system recorded pupil and corneal reflection, while sampling at 120 Hz. Calibration and stimulus presentation were controlled by SMI IView X™ (Version 2.8, 2014) and Experiment Center™ (Version 3.4, 2014). Before beginning data collection, the eye tracker software obtained a 2-point calibration using a telescoping bulls-eye or animated image of Elmo accompanied by infant friendly sounds. The calibration was repeated if we did not obtain an accurate calibration. Prior to the familiarization trials, participants' attention was attracted to the center of the screen using a telescoping bulls-eye or flashing circle. The eye tracker obtained a 4-point validation using a telescoping bulls-eye at the end of data collection to assess whether the accuracy of the tracking shifted from the initial calibration (Oakes 2010).

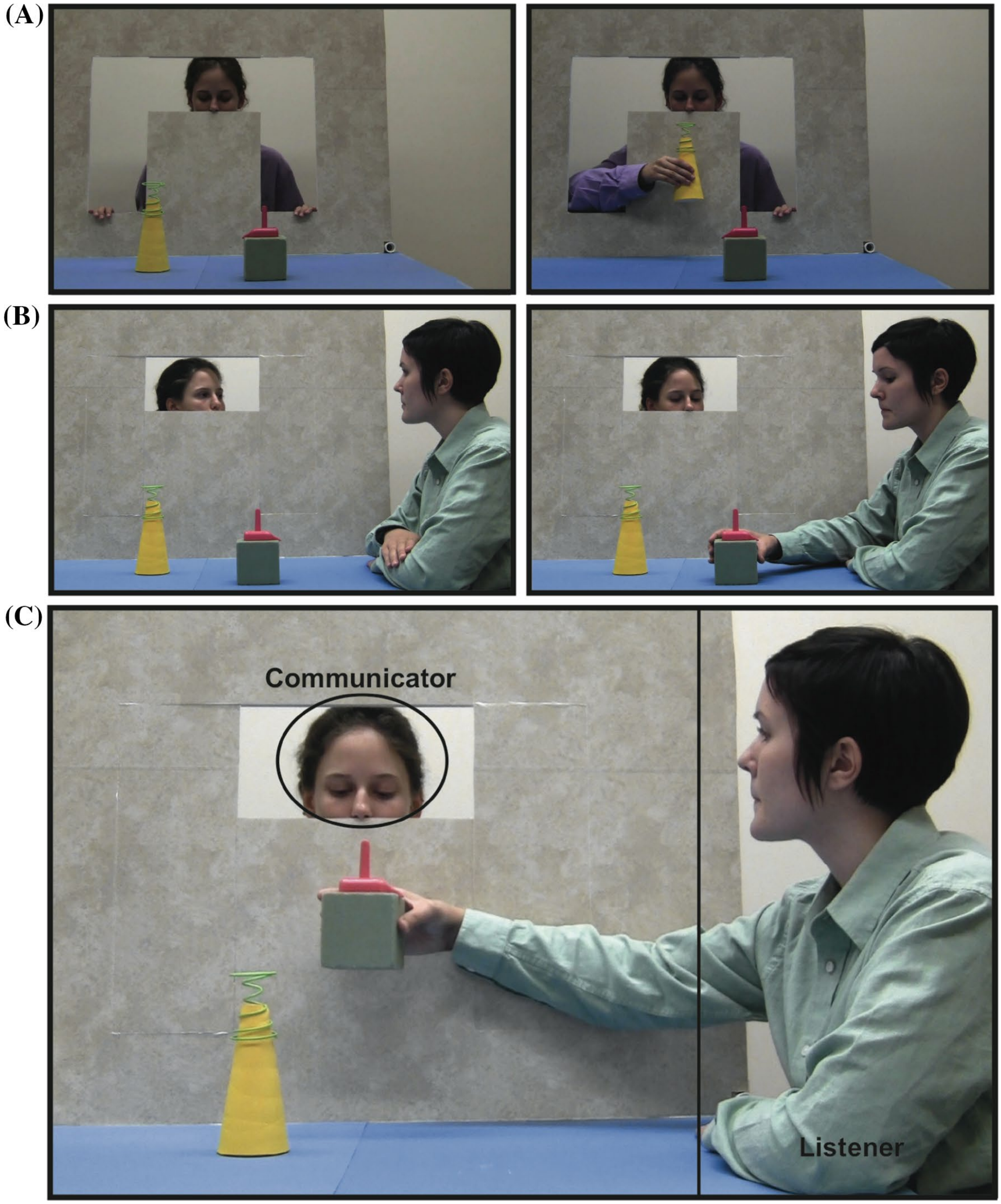

Fig. 1 Third-party communicative interaction. Infants saw a third-party communicative interaction in a sequence of three trials. **a** During the two familiarization trials, the Communicator looked at two novel objects, and grasped the target object. **b** During the action phase of the test trial, the Communicator could no longer reach the objects so she spoke to the Listener (left panel). The Listener selected either the target or non-target object (right panel). **c** In the still image phase, the final image of the test trial froze for the remainder of the test trial. Actor AOIs used in all trials are shown in **c**

Data Reduction and Analyses

Eye-tracking data were processed using SMI BeGaze™ Eye-tracking Analysis Software (Version 3.4, 2014). We defined multiple areas of interest (AOIs) including the Listener's face and body (657,090 pixels) and the Communicator's face (69,276 pixels; Fig. 1c). AOIs remained the same size for the duration of the trials. However, the shape of the Communicator AOI was adjusted to avoid overlap with an object on two occasions: when the Communicator lifted the object below her face in the familiarization trials and when the Listener lifted the object to the Communicator in the test trials. The eye-tracker calculated fixation lengths and locations by filtering raw looking data by predetermined criteria (80 ms, 100 pixels of dispersion) for each individual participant.

The quality of eye tracking data did not differ between infants displaying and not displaying later ASD symptoms, as both groups showed no difference in proportion of missing data due to data loss or inattention during the familiarization trials (no ASD symptoms: $M=0.25$, $SD=0.02$; later ASD symptoms: $M=0.23$, $SE=0.04$; $F(1,45)=0.42$, $p=.52$), nor during the action phase of the test trial (no ASD symptoms: $M=0.19$, $S=0.02$; later ASD symptoms: $M=0.17$, $SE=0.04$; $F(1,45)=0.16$, $p=.69$).

To quantify infants' linguistic and social-pragmatic attention at 12 months, we measured their looking behavior to the Communicator and Listener during the action phase of the communicative scene. We divided the action phase into 25 bins of 250 ms each, which would capture infants' typical response time of 200 ms in action processing tasks (Rosander and von Hofsten 2011). Then, to quantify infants' linguistic attention we estimated individual infants' slope of looking at the Communicator when she spoke (−250 to 750 ms; Pfister et al. 2013). To calculate slope of looking at the Communicator, we used piecewise linear regression on individual infants' proportion of looking at the Communicator out of total looking time to the scene (Cudeck and Klebe 2002; Yamashiro and Vouloumanos 2019). In a piecewise linear regression, a nonlinear trend is divided at fixed time points into connected linear segments (Cohen 2008; Cudeck and Klebe 2002). The regression estimates reflect individual infants' average linear change or slope of looking behavior in each segment. Because we are most interested in how looking to the Communicator changes when she spoke, in our analyses, we use the regression estimate that reflects each infants' average slope of looking between −250 and 750 ms, which is when the Communicator spoke (see Fig. 2 for a timeline of the actions). Thus, the slope of looking at the Communicator can be interpreted as infants' average increase in proportion looking at the Communicator for every unit increase in time (250 ms) when she spoke (Yamashiro and Vouloumanos 2019). Because the action phase of each experimental outcome was identical when the Communicator spoke, we include each infants' most positive, non-zero slope from either experimental outcome in our analyses. Using this single best slope, rather than an average of slopes, prevents possible artifacts like inattention or momentary loss of tracking during some trials from underestimating infants' slope of looking.

To quantify infants' social-pragmatic attention, we calculated individual infants' difference in mean proportion looking at the Listener out of total looking time to the scene when she selected the non-target object minus when she selected the target object (750–4000 ms; see Fig. 2). This score reflects the degree to which infants treated the Listener selecting the non-target object as incongruent relative to when the Listener selects the target object. A positive score reflects greater proportion looking at the Listener when she selected the non-target than the target object, suggesting infants detected an incongruity. A negative score or a score around zero would suggest that infants did not detect an incongruity in the Listener's response (Yamashiro and Vouloumanos 2019).

Observational Measures at 12 and 24 Months

General Cognitive Development

We assessed infants' general cognitive development at 12 months using the Mullen Scales of Early Learning (MSEL; Mullen 1995), a comprehensive assessment of cognitive development across several domains composed of 5 subscales including gross motor, fine motor, visual reception, receptive language, and expressive language. Participants receive a raw score for each subscale, which is converted into a T score, percentile rank, and age equivalent. The MSEL also provides an Early Learning Composite score, which is based on the T scores of the last four subscales and provides a measure of general cognitive ability. All subscales are standardized for infants and children aged 0–69 months, except for the gross motor scale, which is standardized for children aged 0–29 months.

Language Development

Infants' language abilities at 24 months were assessed by the MacArthur-Bates Communicative Development Inventories (MD-CDI; Fenson et al. 1993), a set of reliable parental report measures of expressive and receptive language skills in infants and children. We used the Words and Sentences form (normed for 16–30 months). The MB-CDI were scored using the Scoring Program for MB-CDI (vers. 4.1; Marchman 2013) which calculates percentile scores for each participants' expressive language compared to others of the same age and sex.

ASD Symptoms and Diagnosis

A licensed clinical psychologist with expertise in diagnosing and treating individuals with ASD evaluated all participants (high-risk and low-risk) for a comprehensive clinical best estimate diagnosis at 24 months using the Toddler Module of the Autism Diagnostic Observation Schedule (ADOS; Luyster et al. 2009) and their observations of children's behaviors and interactions during the MSEL. The clinical psychologists were research reliable on their administration

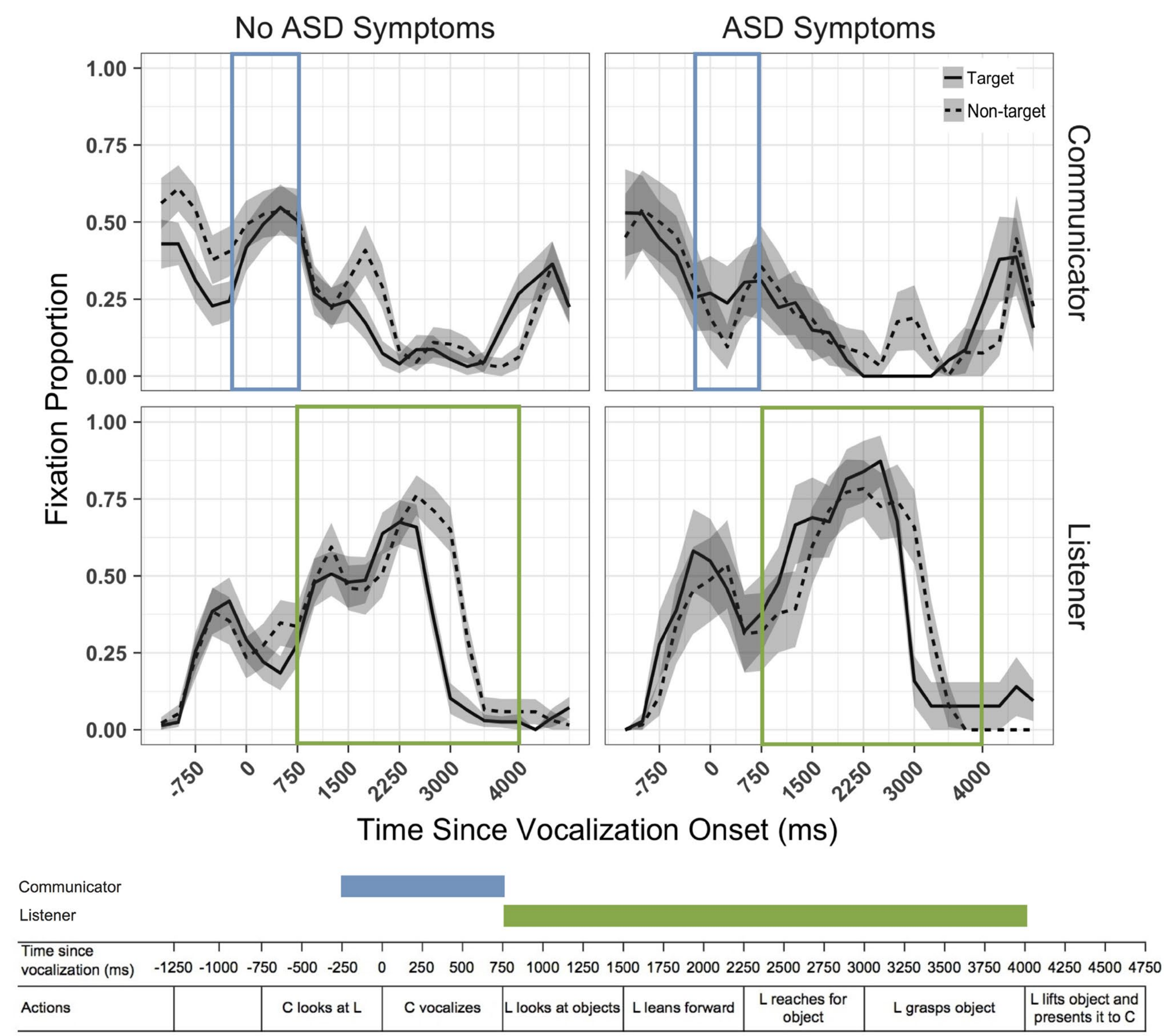

Fig. 2 Infants' raw looking behavior and timeline of actions. Time course of fixations ± SEM (ribbon) to the Communicator (top panels) and Listener (bottom panels) during the Target outcome (solid lines) and Non-target outcome (dashed lines) for infants not displaying later ASD symptoms (left panels) and infants displaying later ASD symptoms (right panels). Colored boxes indicate the time segments included in regression models. Below is a timeline of the start and end of each time segment aligned with the start and end of each action in the action segment of the test trial. "C" is Communicator and "L" is Listener

of the ADOS. The ADOS is a reliable diagnostic instrument for quantifying behaviors indicative of ASD in children; it assesses social interaction, communication, and restricted and repetitive behaviors through planned "presses" within semi-structured play-based interactions. The Toddler Module of the ADOS is intended for children under 30 months. The ADOS provides algorithm scores for social affect and for restricted and repetitive behaviors, which are summed together for an overall total score. Thirteen infants met the cutoff criteria for ASD on the ADOS and received a clinical best estimate diagnosis of ASD; we will refer to these infants as later displaying symptoms of ASD: 3 displayed symptoms in the mild-to-moderate range of concern (2 low risk, 1 high risk; 1 female, 2 male) and 10 displayed symptoms in the moderate-to-severe range of concern (2 low risk, 8 high risk; 10 male) at 24 months. Because of our limited sample, we included both male and female infants displaying later symptoms of ASD in our analyses. Thirty-four infants were assessed as typically developing and did not display symptoms of ASD in the mild-to-moderate or moderate-to-severe range on the ADOS or symptoms of atypical development (e.g., language and/or cognitive delay) at 24 months

(23 low-risk, 11 high-risk; 18 female, 16 male); we will refer to these infants as not displaying later symptoms of ASD.

Prevalence rates of ASD symptomology in our sample were higher than previously reported diagnostic rates for both low- and high-risk infants (American Psychiatric Association 2013; Center for Disease Control and Prevention 2018; Ozonoff et al. 2011). However, this sample is part of a larger longitudinal study in which the prevalence rates at 24 months are more similar to those previously reported in research settings (8% of low-risk participants and 26% of high-risk participants have received a clinical best estimate diagnosis of ASD).

Results

Overall Speech Preference at 9 Months

To examine whether infants displaying and not displaying later ASD symptoms show differences in their preference for speech over non-speech at 9 months, we ran a 2 (sound type: speech vs. non-speech) × 2 (symptom group: no ASD symptoms vs. ASD symptoms) mixed design analysis of variance (ANOVA) with sound type as a within-subjects factor and symptom group as a between-subjects factor. There were two outliers (1 not displaying ASD symptoms, 1 displaying ASD symptoms), whose average looking times during non-speech trials was more than 2 standard deviations above the mean. However, the results of the ANOVA did not change after outliers were removed. We report results that include all participants' data.

We found a significant interaction between sound type and symptom group ($F(1, 45) = 6.86$, $p = .01$, $\eta^2 = 0.04$) where infants not displaying later ASD symptoms listened longer to speech ($M = 13.3$, $SE = 1.2$) compared to non-speech ($M = 10.6$, $SE = 1.2$, $t(33) = 2.70$, $p = .02$). In contrast, infants displaying later ASD symptoms listened equally to speech ($M = 11.9$, $SE = 1.4$) and non-speech ($M = 14.5$, $SE = 1.4$, $t(12) = 2.64$, $p = .13$; see Fig. 3). The ANOVA and all graphs were run in R using ez vers. 4.40 and ggplot2 vers. 2.1.0 (Lawrence 2016; R Core Team 2013; Wickham 2009).

Overall Differences in Linguistic and Social-Pragmatic Attention at 12 Months

To examine whether infants displaying and not displaying later ASD symptoms showed overall differences in their linguistic and social-pragmatic attention at 12 months, we ran independent t-tests on infants' slope of looking at the Communicator when she spoke and difference score of looking at the Listener when she selected an object. However, we found no significant differences between the symptom groups for either measure ($ps > .10$).

Predicting Linguistic and Social-Pragmatic Attention at 12 Months

To examine whether individual infants' preference for speech over non-speech at 9 months predicted linguistic and social-pragmatic attention at 12 months, we calculated infants' speech preference score at 9 months by subtracting mean looking time during speech trials minus mean

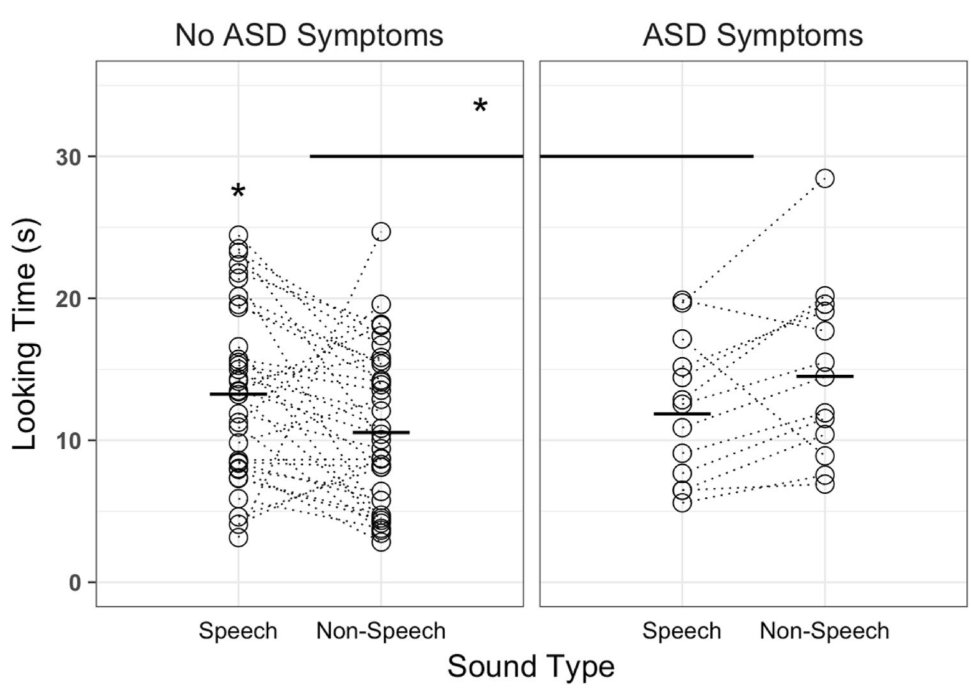

Fig. 3 Mean looking times during speech preference task at 9 months. Mean looking times during speech and non-speech trials during the speech preference task at 9 months for infants not displaying later ASD symptoms (left panel) and infants displaying later ASD symptoms (right panel). Circles represent individual infants' mean looking times. Dotted lines indicate individual infants' means for each sound type. Removing outliers did not change results. *$p < .05$

looking time during non-speech trials. A positive score reflects longer looking during speech than non-speech. A negative score reflects longer looking during non-speech than speech. We mean centered all continuous predictors and effects coded symptom group with infants not displaying later ASD symptoms as -1 and infants later displaying ASD symptoms as 1. All regression analyses were conducted in SPSS (vers. 22.0; IBM Corp. Released 2014).

We used separate regression models to test whether infants' speech preference score at 9 months predicted (1) linguistic attention: slope of looking at the Communicator when she spoke (−250 to 750 ms), and (2) social-pragmatic attention: difference in mean looking at the Listener when she selected the non-target object minus when she selected the target object (750 to 4000 ms) at 12 months. We also included symptom group as a predictor variable in each model and included general cognitive abilities at 12 months, as measured by the Early Learning Composite score from the MSEL (Mullen 1995), as a covariate in each model (see Table 1 for variable means and Table 2 for correlations among variables).

Diagnostic statistics of individual participants' influence on the regression equation such as Cook's Distance, DFFIT, and DF Beta, identified three outliers for the effect of infants' speech preference score on linguistic attention and two outliers for the effect of infants' speech preference score on social-pragmatic attention. Results of the regressions were unchanged after removing outliers, therefore we report results using all participants' data.

Predicting Linguistic Attention

Across symptom groups, infants' speech preference score at 9 months significantly predicted infants' most positive, non-zero slope of looking at the Communicator in either experimental outcome when she spoke to the Listener, where listening longer to speech than non-speech was related to a greater increase in looking at the Communicator when she spoke ($b = 0.01$, $t(42) = 2.12$, $p = .04$, $f^2 = 0.11$, 95% CIs [0.00, 0.01]; see Fig. 4).

Predicting Social-Pragmatic Attention

Infants' speech preference score at 9 months did not significantly predict infants' difference score (non-target–target) of looking at the Listener when she selected an object ($p = .50$).

Symptom group ($ps > .11$) and general cognitive abilities at 12 months ($ps > .35$) did not significantly predict infants' looking behaviors to the Communicator or the Listener in either model (see Table 3 for full model results).

Predicting Language Production at 24 Months

We ran a regression model to test whether individual infants' speech preference at 9 months predicted language production at 24 months quantified by standardized percentiles of the

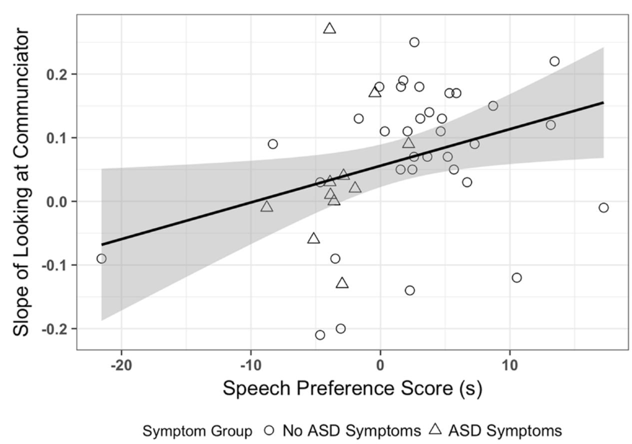

Fig. 4 Infants' speech preference scores predict linguistic attention. Regression plot with 95% confidence intervals (ribbon) of speech preference scores at 9 months on slope of looking at the Communicator when she spoke at 12 months for infants not displaying later ASD symptoms (circles) and infants displaying later ASD symptoms (triangles). Removing outliers did not change results

Table 2 Correlations between looking behaviors to Communicator and Listener at 12 months, words produced at 24 months, symptom group, speech preference scores at 9 months, and general cognitive abilities at 12 months for both infants displaying and not displaying later ASD symptoms

	N	1	2	3	4	5
1. Communicator slope	45	–				
2. Listener difference score	45	0.02	–			
3. MB-CDI Words Produced	42	0.24	0.38*	–		
4. Speech preference score	45	0.33*	0.21†	0.25†	–	
5. Symptom group	45	−0.12	−0.30*	−0.33*	−0.39**	–
6. General cognitive abilities	45	−0.06	0.00	0.10	−0.32*	−0.14

*$p < .05$, †$p < .10$

Table 3 Fixed effects of linear regression models for symptom group, speech preference score at 9 months, and general cognitive abilities at 12 months on looking behaviors to Communicator and Listener (top panel) and on MB-CDI Words Produced at 24 months for all infants and standardized ADOS Social Affect severity scores at 24 months for infants displaying later ASD symptoms (bottom panel)

	Communicator slope						Listener difference score					
	$R^2\Delta$	*b*	*SE*	*p*	95% CI		$R^2\Delta$	*b*	*SE*	*p*	95% CI	
Step 1												
Constant	0.11[†]	0.06	0.02	0.00	0.02	0.10	0.10	0.04	0.03	0.26	−0.03	0.10
Symptom group		0.00	0.02	0.95	−0.04	0.04		−0.05	0.03	0.11	−0.12	0.01
Speech preference score		**0.01**	**0.00**	**0.04**	**0.00**	**0.01**		0.00	0.00	0.50	−0.01	0.01
Step 2												
Constant	0.00	0.07	0.02	0.00	0.02	0.11	0.00	0.04	0.03	0.28	−0.03	0.10
Symptom group		0.00	0.02	0.86	−0.04	0.05		−0.05	0.04	0.13	−0.13	0.02
Speech preference score		**0.01**	**0.00**	**0.04**	**0.00**	**0.01**		0.00	0.01	0.55	−0.01	0.01
General cognitive abilities		0.00	0.00	0.68	0.00	0.00		0.00	0.00	0.97	−0.01	0.01
	Words produced						Social affect					
	$R^2\Delta$	*b*	*SE*	*p*	95% *CI*		$R^2\Delta$	*b*	*SE*	*p*	95% *CI*	
Step 1												
Constant	0.12[†]	47.07	5.78	0.00	35.37	58.76	0.06	6.91	0.67	0.00	5.40	8.42
Symptom group		−10.01	6.31	0.12	−22.78	2.76						
Speech preference score		0.73	0.99	0.46	−1.27	2.72		−0.20	0.26	0.46	−0.78	0.38
Step 2												
Constant	0.02	48.33	5.94	0.00	36.30	60.35	0.01	6.91	0.70	0.00	5.29	8.53
Symptom group		−7.60	6.81	0.27	−21.40	6.19						
Speech preference score		1.29	1.15	0.27	−1.04	3.63		−0.18	0.3	0.53	−0.82	0.46
General cognitive abilities		0.42	0.44	0.35	−0.48	1.32		0.01	0.05	0.79	−0.11	0.14

Significant effects ($p < .05$) are given in bold

MB-CDI (Fenson et al. 1993). Again, we included symptom group as a predictor variable and included general cognitive abilities at 12 months as a covariate (see Table 1 for variable means and Table 2 for correlations among variables).

There was a positive correlation between speech preference score at 9 months and expressive language at 24 months ($r = .25$, $p = .05$) where listening longer to speech than non-speech at 9 months was related to higher expressive language scores at 24 months (see Fig. 5). However, infants' speech preference score, symptom group, and general cognitive abilities did not significantly predict expressive language at 24 months in the regression model ($ps > 0.12$). Thus, there was a linear, positive relation between speech preference scores at 9 months and later expressive language, but individual infants' speech preference scores did not significantly predict later expressive language, adjusting for symptom group and general cognitive abilities (see Table 3 for full model results).

There was also a negative correlation between symptom group and speech preference score ($r = -.39$, $p = .02$), such that infants not displaying later ASD symptoms showed a greater preference for speech compared to infants displaying later ASD symptoms, and a negative correlation between

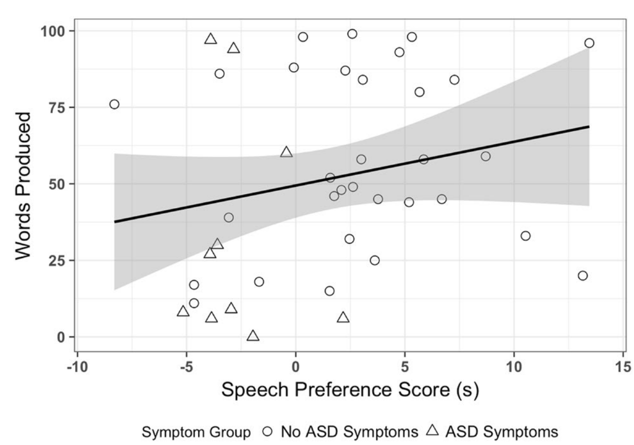

Fig. 5 Infants' speech preference scores predicting later language abilities. Regression plot with 95% confidence intervals (ribbon) of speech preference scores at 9 months on MB-CDI Words Produced percentiles at 24 months for infants not displaying later ASD symptoms (circles) and infants displaying later ASD symptoms (triangles). Removing outliers did not change results

symptom group and expressive language at 24 months ($r = -.33$, $p = .02$), where infants not displaying later ASD symptoms said more words at 24 months compared to infants displaying later ASD symptoms.

Predicting Social-Pragmatic Communication at 24 Months

We used a similar regression model to examine for infants displaying later ASD symptoms, whether speech preference at 9 months predicted later social-pragmatic communication at 24 months. In a regression model, we tested whether infants' speech preference scores at 9 months predicted social-pragmatic outcomes at 24 months quantified by standardized Social Affect severity scores from the ADOS Toddler Module (Esler et al. 2015; Luyster et al. 2009). General cognitive abilities at 12 months was included as a covariate (see Table 1 for variable means and Table 4 for correlations among variables). Neither speech preference scores nor general cognitive abilities significantly predicted ADOS scores at 24 months ($ps > .46$; see Table 3 for full model results).

Discussion

An early speech preference predicts infants' ability to communicate and interact with others (Curtin and Vouloumanos 2013; Klin 1991, 1992; Kuhl et al. 2005; Vouloumanos and Waxman 2014), possibly because it is related to how infants extract relevant linguistic and social-pragmatic information from their environment. We found that across both infants displaying and not displaying later ASD symptoms, an early preference for speech at 9 months predicts infants' tendency to attend to a person when they speak in communicative interactions (linguistic attention), and is related to later language abilities. Though on average, infants displaying later ASD symptoms do not prefer speech over non-speech as infants not displaying later ASD symptoms do, we found no evidence that infants' speech preference predicts their social-pragmatic attention or social-pragmatic outcomes. The correlational nature of our findings leads us to tentatively speculate that an early speech preference may be a mechanism for learning to identify and attend to linguistic information and is related to how infants learn language, but may not be related to infants' social-communicative abilities.

Table 4 Correlations between standardized ADOS Social Affect severity scores at 24 months, speech preference scores at 9 months, looking behavior at the Listener at 12 months, and general cognitive abilities at 12 months for infants displaying later ASD symptoms

	N	1	2	3
1. ADOS Social Affect score	11	–		
2. Speech preference score	11	−0.25	–	
3. Listener difference score	11	−0.29	−0.31	–
4. General cognitive abilities	11	0.15	−0.23	−0.04

Infants' early speech preference may be important for learning to attend to a person when she speaks and may allow infants to learn the meanings of novel words. Across symptom group, listening longer to speech than non-speech at 9 months was related to a faster increase in looking at the Communicator when she spoke at 12 months and was related to producing more words at 24 months. Infants use speech to identify potential communicative partners (Vouloumanos et al. 2009) and recognize speech as a communicative signal (Csibra and Gergely 2009, 2011). Furthermore, 12-month-olds who listen longer to speech produce more words at 18 months (Vouloumanos and Curtin 2014), and toddlers with ASD who listen longer to non-speech than speech produce fewer words (Kuhl et al. 2005), suggesting that attending to speech may be important for learning language and not attending to speech may hinder language learning. Our results corroborate and expand on previous findings by suggesting that attending to speech at 9 months is related to language outcomes at 24 months, and is also related to how infants displaying and not displaying later ASD symptoms identify and attend to a person when she speaks at 12 months, which is an indicator of infants' attention to the linguistic information in a communicative scene (Chawarska et al. 2012; Tenenbaum et al. 2015; Yamashiro and Vouloumanos 2019; Young et al. 2009). Thus, an early speech preference is related to how infants' extract and attend to the linguistic information from communicative interactions and may be important for learning language.

Contrary to our expectations, we did not find that an early speech preference predicted social-pragmatic attention or later social-pragmatic outcomes. Across groups, speech preference at 9 months did not significantly predict infants' ability to detect the incongruity in a Listener's response to speech at 12 months, and for infants displaying later ASD symptoms, speech preference did not predict social-pragmatic outcomes. This suggests that an early speech preference may be related to language, not social-pragmatic communication. Linguistic and social-pragmatic attention during communicative interactions may be separable constructs: infants' attention to a communicator when she speaks and their ability to evaluate a Listener's response were not correlated (see Table 2) and these attention abilities separately predict later language and social-pragmatic outcomes (Yamashiro and Vouloumanos 2019). Thus, attending to speech early in infancy may be a mechanism for learning to attend to linguistic, not social-pragmatic information.

Though we did not find that an early speech preference is directly related to social-pragmatic attention or outcomes, we found that overall, infants not displaying later ASD symptoms listened longer to speech than non-speech, while infants displaying later ASD symptoms did not, which is consistent with previous studies (Curtin and Vouloumanos 2013; Klin 1991, 1992; Kuhl et al. 2005; Vouloumanos and Curtin 2014). Thus, infants displaying later ASD symptoms lack a preference for speech as early as 9 months, which is over a year before the current age of a reliable diagnosis (e.g., Guthrie et al. 2013) and may be important for informing early intervention strategies. For example, targeting infants' early attention to speech could improve language abilities and, as improved language abilities can make ASD interventions more effective, strengthening speech preference in infancy could indirectly improve social-pragmatic outcomes in children with ASD (Gillespie-Lynch et al. 2012; Magiati et al. 2014).

A limitation of the current study is our small sample of infants displaying later ASD symptoms. The longitudinal nature of our study, which required families to attend multiple visits over 2 years, made it difficult to enroll a large sample of infants. Additionally, our strict inclusion criteria further limited our sample to infants who had both useable speech preference data at 9 months and eye-tracking data at 12 months. This inclusion criteria confirms that any relations between individual infants' speech preference, linguistic or social-pragmatic attention, and expressive language or social-pragmatic outcomes were not biased by missing data (Little and Rublin 1987; Sterne et al. 2009). It is also possible that we did not find a predictive relation between speech preference and social-pragmatic outcomes for infants displaying later ASD symptoms, as in previous studies (Curtin and Vouloumanos 2013; Kuhl et al. 2005) because our limited sample size did not provide enough power to detect the effect, as we only had 9% power to detect this effect at alpha level $p < .05$ with 11 infants displaying later ASD symptoms (2 infants were not included in the regression because they were missing data for general cognitive abilities; GPOWER; Faul et al. 2007). Future studies could confirm our findings with a larger sample of infants displaying later ASD symptoms.

Furthermore, to retain power for our analyses, we chose to collapse across low- and high-risk infants who did not display later ASD symptoms. All low- and high-risk infants who did not display later ASD symptoms received a clinical best estimate diagnosis of typical development at 24 months, did not meet criteria for ASD on the ADOS, and did not present with any language or cognitive delays. Though previous studies have found that low-risk infants listen longer to speech than non-speech and high-risk infants listen equally to both sounds, the infants in this prior study were not assessed for a later ASD diagnosis (Curtin and Vouloumanos 2013). It is possible that the difference in speech preference between low- and high-risk infants was driven by the high-risk infants who went on to display later ASD symptoms. Toddlers and children diagnosed with ASD also lack a preference for speech, unlike their neurotypical peers (Klin 1991, 1992; Kuhl et al. 2005). Future studies with larger samples could further examine potential differences in speech preference between low- and high-risk infants who did not display later ASD symptoms.

An additional question for future research is to examine how an early preference for other socially relevant signals like human faces predicts social-pragmatic attention and later social-pragmatic outcomes. An early preference for human faces may help infants attend to social-pragmatic information from others (Moore and Corkum 1998; Pascalis et al. 1998; Scaife and Bruner 1975; Tronick 1989). Not looking at faces may signal atypical social-pragmatic functioning in infants displaying later ASD symptoms, who look at faces less than infants not displaying later ASD symptoms (Chawarska et al. 2013) and show a decrease in looking at others' eyes over time (Jones and Klin 2013).

Despite our limited sample, our findings provide meaningful and interesting results showing that an early speech preference is related to infants' attention to the linguistic information in communicative interactions at 12 months, just as infants are beginning to produce their first words, and related to language outcomes over a year later. However, we found no evidence that an early speech preference is related to infants' social-pragmatic attention or later social-pragmatic outcomes. Understanding how an early speech preference supports linguistic attention may reveal why and how an early emerging speech preference in infancy is crucial for the language abilities central to our human experience and could have important implications for early diagnosis and intervention strategies for individuals with ASD.

Acknowledgments We'd like to thank members of the NYU Infant Cognition and Communication Lab, the members of the University of Calgary Speech Development Lab, and the parents and infants who participated.

Author Contributions SC and AV designed the study. AY collected and analyzed the data. All authors wrote, edited, and approved the final manuscript.

Funding This research was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number R01HD072018 awarded to AV and SC.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in the study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. Arlington, VA: APA

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology, 64*(1), 3–26. https://doi.org/10.1006/jecp.1996.2332.

Butterfield, E. C., & Siperstein, G. N. (1970). Influence of contingent auditory stimulation upon nonnutritional suckle. In J. F. Bosma (Ed.), *Third symposium on oral sensation and perception: The mouth of the infant* (pp. 313–334). Springfield, IL: Charles C. Thomas.

Center for Disease Control and Prevention. (2018). Prevalence of autism spectrum disorder among children aged 8 years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014.

Chawarska, K., Macari, S., & Shic, F. (2012). Context modulates attention to social scenes in toddlers with autism. *Journal of Child Psychology and Psychiatry, 53*(8), 903–913. https://doi.org/10.1111/j.1469-7610.2012.02538.x.

Chawarska, K., Macari, S., & Shic, F. (2013). Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum disorders. *Biological Psychiatry, 74*(3), 195–203. https://doi.org/10.1016/j.biopsych.2012.11.022.

Cheung, H., Xiao, W., & Lai, C. M. (2012). Twelve-month-olds' understanding of intention transfer through communication. *PLoS ONE, 7*(9), e46168. https://doi.org/10.1371/journal.pone.0046168.

Cohen, P. (2008). *Applied data analytic techniques for turning points research*. New York: Taylor & Francis Group, LLC.

Colombo, J., & Bundy, R. S. (1981). A method for the measurement of infant auditory selectivity. *Infant Behavior and Development, 4*(1), 219–223. https://doi.org/10.1016/S0163-6383(81)80025-2.

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*(5), 1584–1595. https://doi.org/10.1111/j.1467-8624.1990.tb02885.x.

Cooper, R. P., & Aslin, R. N. (1994). Developmental differences in infant attention to the spectral properties of infant-directed speech. *Child Development, 65*(6), 1663–1677. https://doi.org/10.1111/j.1467-8624.1994.tb00841.x.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005.

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 366*(1567), 1149–1157. https://doi.org/10.1098/rstb.2010.0319.

Cudeck, R., & Klebe, K. J. (2002). Multiphase mixed-effects models for repeated measures data. *Psychological Methods, 7*(1), 41–63. https://doi.org/10.1037//1082-989X.7.1.41.

Curtin, S., & Vouloumanos, A. (2013). Speech preference is associated with autistic-like behavior in 18-months-olds at risk for autism spectrum disorder. *Journal of Autism and Developmental Disorders, 43*(9), 2114–2120. https://doi.org/10.1007/s10803-013-1759-1.

Dehaene-Lambertz, G. (2000). Cerebral specialization for speech and non-speech stimuli in infants. *Journal of Cognitive Neuroscience, 12*(3), 449–460. https://doi.org/10.1162/089892900562264.

Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science, 298*(5600), 2013–2015. https://doi.org/10.1126/science.1077066.

Esler, A. N., Bal, V. H., Guthrie, W., Wetherby, A., Weismer, S. E., & Lord, C. (2015). The Autism Diagnostic Observation Schedule, Toddler Module: Standardized Severity Scores. *Journal of Autism and Developmental Disorders, 45*(9), 2704–2720. https://doi.org/10.1007/s10803-015-2432-7.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146.

Fenson, L., Bates, E., Dale, P. S., Marchman, V. A., Reznick, J. S., & Thal, D. J. (1993). *MacArthur-Bates Communicative Development Inventories: User's guide and technical manual*. Baltimore: Paul H. Brookes.

Ferry, A. L., Hespos, S. J., & Waxman, S. R. (2010). Categorization in 3- and 4-month-old infants: An advantage of words over tones. *Child Development, 81*(2), 472–479. https://doi.org/10.1111/j.1467-8624.2009.01408.x.

Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not Tones) facilitate object categorization: Evidence from 6- and 12-month-olds. *Cognition, 105*(1), 218–228. https://doi.org/10.1016/j.cognition.2006.09.005.

Gillespie-Lynch, K., Sepeta, L., Wang, Y., Marshall, S., Gomez, L., Sigman, M., & Hutman, T. (2012). Early childhood predictors of the social competence of adults with autism. *Journal of Autism and Developmental Disorders, 42*(2), 161–174. https://doi.org/10.1007/s10803-011-1222-0.

Guthrie, W., Swineford, L. B., Nottke, C., & Wetherby, A. M. (2013). Early diagnosis of autism spectrum disorder: Stability and change in clinical diagnosis and symptom presentation. *Journal of Child Psychology and Psychiatry, 54*(5), 582–590. https://doi.org/10.1111/jcpp.12008.

IBM Corp. (2017). *IBM SPSS statistics for Mac, Version 25.0*. Armonk, NY: IBM.

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2–6-month-old infants later diagnosed with autism. *Nature, 504*(7480), 427–431. https://doi.org/10.1038/nature12715.

Klin, A. (1991). Young autistic children's listening preferences in regard to speech: A possible characterization of the symptom of social withdrawal. *Journal of Autism and Developmental Disorders, 21*(1), 29–42. https://doi.org/10.1007/BF02206995.

Klin, A. (1992). Listening preferences in regard to speech in four children with developmental disabilities. *Journal of Child Psychology and Psychiatry, 33*(4), 763–769. https://doi.org/10.1111/j.1469-7610.1992.tb00911.x.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

Kuhl, P. K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological measures. *Developmental Science, 8*(1), F1–F12. https://doi.org/10.1111/j.1467-7687.2004.00384.x.

Lawrence, M. A. (2016). ez: Easy analysis and visualization of factorial experiments. R package version 4.4-0. https://CRAN.R-project.org/package=ez.

Little, R., & Rublin, D. (1987). *Statistical analysis with missing data* (p. 381). New York.: Wiley. https://doi.org/10.1002/9781119013563.

Luyster, R. J., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., … Lord, C. (2009). The autism diagnostic observation schedule—Toddler module: A new module of a standardized diagnostic measure for autism spectrum disorders. *Journal of Autism* *and Developmental Disorders, 39*(9), 1305–1320. https://doi.org/10.1007/s10803-009-0746-z.

Mackenzie, H., Graham, S. A., & Curtin, S. (2011). Twelve-month-olds privilege words over other linguistic sounds in an associative learning task. *Developmental Science, 14*(2), 249–255. https://doi.org/10.1111/j.1467-7687.2010.00975.x.

Magiati, I., Tay, X. W., & Howlin, P. (2014). Cognitive, language, social and behavioural outcomes in adults with autism spectrum disorders: A systematic review of longitudinal follow-up studies in adulthood. *Clinical Psychology Review*. https://doi.org/10.1016/j.cpr.2013.11.002.

Marchman, V. A. (2013). *Scoring Program for the MacArthur-Bates Communicative Development Inventories*. Retrieved from https://mb-cdi.stanford.edu/scoringdb_p.htm.

Martin, A., Onishi, K. H., & Vouloumanos, A. (2012). Understanding the abstract role of speech in communication at 12months. *Cognition, 123*(1), 50–60. https://doi.org/10.1016/j.cognition.2011.12.003.

Moore, C., & Corkum, V. (1998). Infant gaze following based on eye direction. *British Journal of Developmental Psychology, 16*(4), 495–503. https://doi.org/10.1111/j.2044-835X.1998.tb00767.x.

Mullen, E. M. (1995). *Mullen Scales of Early Learning*. Circle Piens, MN: American Guidance Services, Inc. (AGS ed.).

Oakes, L. M. (2010). Infancy guidelines for publishing eye-tracking data. *Infancy, 15*(1), 1–5. https://doi.org/10.1111/j.1532-7078.2010.00030.x.

Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., … Stone, W. L. (2011). Recurrence risk for autism spectrum disorders: A Baby Siblings Research Consortium study. *Pediatrics, 128*(3), e488–e495. https://doi.org/10.1542/peds.2010-2825.

Pascalis, O., de Haan, M., Nelson, C. A., & de Schonen, S. (1998). Long-term recognition memory for faces assessed by visual paired comparison in 3- and 6-month-old infants. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(1), 249–260. https://doi.org/10.1037/0278-7393.24.1.249.

Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., & Mehler, J. (2003). Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences of the United States of America, 100*(20), 11702–11705. https://doi.org/10.1073/pnas.1934290100.

Pfister, R., Schwarz, K., Carson, R., & Jancyzk, M. (2013). Easy methods for extracting individual regression slopes: Comparing SPSS, R, and Excel. *Tutorials in Quantitative Methods for Psychology, 9*(2), 72–78.

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Rosander, K., & von Hofsten, C. (2011). Predictive gaze shifts elicited during observed and performed actions in 10-month-old infants and adults. *Neuropsychologia, 49*(10), 2911–2917. https://doi.org/10.1016/j.neuropsychologia.2011.06.018.

Samples, J. M., & Franklin, B. (1978). Behavioral responses in 7 to 9 month old infants to speech and non-speech stimuli. *Journal of Auditory Research, 18*(2), 115–123.

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature, 253*(5489), 265–266. https://doi.org/10.1038/253265a0.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420.

Shultz, S., & Vouloumanos, A. (2010). Three-month-olds prefer speech to other naturally occurring signals. *Language Learning and Development, 6*(4), 241–257. https://doi.org/10.1080/15475440903507830.

Shultz, S., Vouloumanos, A., Bennett, R. H., & Pelphrey, K. (2014). Neural specialization for speech in the first months of life. *Developmental Science, 17*(5), 766–774. https://doi.org/10.1111/desc.12151.

Spence, M. J., & DeCasper, A. J. (1987). Prenatal experience with low-frequency maternal-voice sounds influence neonatal perception of maternal voice samples. *Infant Behavior and Development, 10*(2), 133–142. https://doi.org/10.1016/0163-6383(87)90028-2.

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., … Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *BMJ, 338*, b2393. https://doi.org/10.1136/bmj.b2393.

Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Shah, R. J., Malle, B. F., & Morgan, J. L. (2015). Attention to the mouth and gaze following in infancy predict language development. *Journal of Child Language, 42*(6), 1173–1190. https://doi.org/10.1017/S0305000914000725.

Thorgrimsson, G. B., Fawcett, C., & Liszkowski, U. (2014). Infants' expectations about gestures and actions in third-party interactions. *Frontiers in Psychology, 5*, 321. https://doi.org/10.3389/fpsyg.2014.00321.

Thorgrimsson, G. B., Fawcett, C., & Liszkowski, U. (2015). 1- and 2-year-olds' expectations about third-party communicative actions. *Infant Behavior and Development, 39*, 53–66. https://doi.org/10.1016/j.infbeh.2015.02.002.

Tronick, E. Z. (1989). Emotions and emotional communication in infants. *American Psychologist, 44*(2), 112–119. https://doi.org/10.1037/0003-066X.44.2.112.

Vivanti, G., Hocking, D. R., Fanning, P., & Dissanayake, C. (2016). Social affiliation motives modulate spontaneous learning in Williams syndrome but not in autism. *Molecular Autism, 7*(1), 40. https://doi.org/10.1186/s13229-016-0101-0.

Vouloumanos, A. (2018). Voulez-vous jouer avec moi? Twelve-month-olds understand that foreign languages can communicate. *Cognition, 173*, 87–92. https://doi.org/10.1016/j.cognition.2018.01.002.

Vouloumanos, A., & Curtin, S. (2014). Foundational tuning: How infants' attention to speech predicts language development. *Cognitive Science, 38*(8), 1675–1686. https://doi.org/10.1111/cogs.12128.

Vouloumanos, A., Druhen, M. J., Hauser, M. D., & Huizink, A. T. (2009). Five-month-old infants' identification of the sources of vocalizations. *Proceedings of the National Academy of Sciences of the United States of America, 106*(44), 18867–18872. https://doi.org/10.1073/pnas.0906049106.

Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The tuning of human neonates' preference for speech. *Child Development, 81*, 517–527.

Vouloumanos, A., Kiehl, K. A., Werker, J. F., & Liddle, P. F. (2001). Detection of sounds in the auditory stream: Event-related fMRI evidence for differential activation to speech and nonspeech. *Journal of Cognitive Neuroscience, 13*(7), 994–1005. https://doi.org/10.1162/089892901753165890.

Vouloumanos, A., Martin, A., & Onishi, K. H. (2014). Do 6-month-olds understand that speech can communicate? *Developmental Science, 17*(6), 872–879. https://doi.org/10.1111/desc.12170.

Vouloumanos, A., & Waxman, S. R. (2014). Listen up! Speech is for thinking during infancy. *Trends in Cognitive Sciences, 18*(12), 642–646. https://doi.org/10.1016/j.tics.2014.10.001.

Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: The privileged status of speech for young infants. *Developmental Science, 73*, 270–276.

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*. https://doi.org/10.1111/j.1467-7687.2007.00549.x.

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. New York: Springer. https://doi.org/10.1007/978-0-387-98141-3.

Wodka, E. L., Mathy, P., & Kalb, L. (2013). Predictors of phrase and fluent speech in children with autism and severe language delay. *Pediatrics, 131*(4), e1128–e1134.

Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition, 85*(3), 223–250. https://doi.org/10.1016/S0010-0277(02)00109-9.

Xu, F., Cote, M., & Baker, A. (2005). Labeling guides object individuation in 12-month-old infants. *Psychological Science, 16*(5), 372–377. https://doi.org/10.1111/j.0956-7976.2005.01543.x.

Yamashiro, A., & Vouloumanos, A. (2018). How do infants and adults process communicative events in real time? *Journal of Experimental Child Psychology, 173*, 268–283. https://doi.org/10.1016/j.jecp.2018.04.011.

Yamashiro, A., & Vouloumanos, A. (2019). Are linguistic and social-pragmatic attention separable in neurotypical infants and infants later diagnosed with ASD? *Developmental Psychology*. https://doi.org/10.1037/dev0000676.

Young, G. S., Merin, N., Rogers, S. J., & Ozonoff, S. (2009). Gaze behavior and affect at 6 months: Predicting clinical outcomes and language development in typically developing infants and infants at risk for autism. *Developmental Science, 12*(5), 798–814. https://doi.org/10.1111/j.1467-7687.2009.00833.x.

Zwaigenbaum, L., Bryson, S., & Garon, N. (2013). Early identification of autism spectrum disorders. *Behavioural Brain Research, 251*, 133–146. https://doi.org/10.1016/j.bbr.2013.04.004.

Zwaigenbaum, L., Thurm, A., Stone, W., Baranek, G., Bryson, S., Iverson, J., … Sigman, M. (2007). Studying the emergence of autism spectrum disorders in high-risk infants: Methodological and practical issues. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-006-0179-x.