Journal of Autism and Developmental Disorders
https://doi.org/10.1007/s10803-026-07516-9

ORIGINAL ARTICLE

Infant Attention to Eyes Versus Mouth Varies With Mouth Movement Speed, Family History of Autism and Language Development

Mary Agyapong[1] · Luke Mason[2,3] · Jannath Begum Ali[3] · Mark H. Johnson[3,4] · Tony Charman[1] · Emily J. H. Jones[3] · the STAARS Team

Received: 9 April 2024 / Accepted: 24 August 2026
© The Author(s) 2026

Abstract

Purpose Patterns of visual attention to faces are mixed in infants with a family history of autism relative to controls. Attention deficit hyperactivity disorder (ADHD) often co-occurs with autism and could represent one source of variance in findings. Heterogeneity in language skills may represent another source of variance. We investigated longitudinal patterns of attention to eyes and mouths in infants with and without an elevated familial likelihood (EL) of autism and/or ADHD, and examined associations with language development.

Methods The sample included 160 infants (EL-autism = 79, EL-ADHD = 31, EL-autism+ADHD = 21, Typical Likelihood (TL) of autism and/or ADHD = 29). We measured eye-to-mouth looking ratios of 5, 10 and 14 month olds who viewed pairs of faces with mouths moving at low (1 Hz), medium (6 Hz) and high (10 Hz/15 Hz) temporal frequency. Language skills were measured throughout infancy and at 36 months.

Results Eye looking (relative to mouth) decreased as a function of age and was reduced for speech-like frequencies (6 Hz) compared to slower frequencies. Infants with EL-autism looked more to the eyes than TL infants, but this effect was attenuated in the presence of familial ADHD. Controlling for 14 months word acquisition, higher receptive language at 36 months was associated with greater eye looking.

Conclusion Our findings confirm age-related changes in eye-to-mouth looking across development and suggest that variability in familial ADHD and heterogeneity in language skills influence attentional differences in infants with a family history of autism. Future family history studies should investigate this heterogeneity using neurophysiological methods.

Keywords Autism · Attention deficit hyperactivity disorder · Infancy · Language development · Visual attention · Eye-tracking

Introduction

Autism and attention deficit hyperactivity disorder (ADHD) are highly heritable neurodevelopmental conditions which frequently co-occur (Hollingdale et al., 2020; Taylor et al., 2013). Autism is characterised by challenges with social communication, the presence of restricted and repetitive behaviours and/or interests, and hypo/hypersensitivity to sensory stimuli (American Psychiatric Association [APA], 2022). Pragmatic, phonological, morphological and lexical/semantic language delays are often present in autism (Eigsti et al., 2011). ADHD is characterised by challenges with attentional control, hyperactivity and impulsivity, or a combination of these factors (APA, 2022). Language delays, whilst not a diagnostic criterion for ADHD, are often present (e.g., Gremillion & Martel, 2014; Kim & Kaiser, 2000).

✉ Mary Agyapong
mary.agyapong@kcl.ac.uk

1 Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience (IoPPN), King's College London, Denmark Hill, London SE5 8AF, UK

2 Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

3 Faculty of Sciences, Centre for Brain and Cognitive Development, Birkbeck College, University of London, London, UK

4 Department of Psychology, University of Cambridge, Cambridge, UK

Published online: 07 October 2026

Despite the common co-occurrence of autism and ADHD within individuals, each condition may have different routes to language delays, with language in autism affected by social difficulties and language in ADHD affected by impulsivity and difficulties with turn-taking (Geurts & Embrechts, 2008). Understanding the factors that cause deviation from a typical linguistic developmental trajectory in these conditions, and the mechanisms responsible, provides insight into their converging and diverging pathways.

Identifying the factors that underpin differences in language trajectories in autism and ADHD requires studies prior to diagnosis, because differences in language development emerge in infancy and toddlerhood (Johnson et al., 2015). Identification of putative markers of altered developmental processes, such as visual attention and language, allow us to map deviations from a typical developmental trajectory and identify potential targets for supportive interventions.

Attention to Faces

Eye-tracking is a promising method for studying visual attention in young infants, being inexpensive, non-invasive, and suitable in non-verbal samples (Katus et al., 2019). Given that social difficulties are a core diagnostic criterion for autism and are related to quality of life (Oakley et al., 2021), much eye tracking research has focused on measures of social attention. Findings are sometimes inconsistent for attention to faces over other body parts and non-social content. Where infants with an elevated likelihood (EL) for autism have demonstrated longer looking to faces in infancy, they have shown lower face recognition abilities at 3-years which may indicate face processing difficulties (de Klerk et al., 2014). There is some evidence for longer looking to faces in infants later diagnosed with autism when compared to typically developing (TD) infants, while other studies show shorter looking to faces or no group differences at all (Falck-Ytter et al., 2023; Johnson et al., 2015)—reflecting the phenotypic heterogeneity of autism and the methodological heterogeneity of autism research, e.g. Mastergeorge et al., 2021.

The extent to which there are different patterns of attention to internal facial features (such as eyes and mouths) in autism is also unclear. Before 5 months of age, TD infants look longer at the eyes than the mouth. Thereafter, they spend an increasing amount of time looking at the mouth, which remains their predominant focus up until 12 months of age (Lewkowicz & Hansen-Tift, 2012; Tsang et al., 2018). Some (generally earlier) studies report a decreased eye-to-mouth looking index (EMI) in autism (i.e. less attention to the eyes relative to the mouth; Jones et al., 2008; Klin et al., 2002). This has been associated with lower social communication skills (Chawarska et al., 2012); however, others report mixed or null results on attention to these features (Guillon et al., 2014; Papagiannopoulou et al., 2014). In some familial likelihood studies, infants with EL-autism show similar proportions of looking to the eyes and mouth compared to infants with typical-likelihood (TL) for autism (Elsabbagh et al., 2014; Shic et al., 2014; Young et al., 2009). However, there is a dearth of literature studying EMI in infants with EL-ADHD. It is likely that multiple factors lead to alterations in attention to eyes and mouths, including the nature of the stimuli, the degree of audio-visual synchrony between speech sounds and mouth movements, stages of development and language acquisition (Falck-Ytter et al., 2013), and culture (Ma et al., 2021).

Language Development and EMI

Given that TD infants begin to shift attention to mouths versus eyes over the same developmental timescale as emerging comprehension, language ability may be related to gaze patterns. Children with a family history of autism show language delays and lower receptive and expressive language scores compared to children with no family history of autism (Belteki et al., 2022; Hudry et al., 2014; Marrus et al., 2018). There is also evidence for early language delays in ADHD from retrospective studies with infants and toddlers who were later diagnosed with ADHD (Gurevitz et al., 2012), though it is unclear if this extends to infants with a family history of ADHD. Thus, studying variability in language skills within cohorts of infants enriched for neurodevelopmental outcomes could clarify associations between having a family history of autism and/or ADHD and heterogeneity in eyes versus mouth looking.

Indeed, preference for the mouth compared to the eyes has been associated with higher concurrent expressive language scores in autistic children (Habayeb et al., 2021) and has been associated with higher later expressive language abilities in prospective autism cohorts (Elsabbagh et al., 2014; Young et al., 2009). This association between EMI and expressive language was only observed in particular contexts, such as where infants had already acquired first words (Habayeb et al., 2021) and for 'mouths only' or 'peek-a-boo' stimulus videos (Elsabbagh et al., 2014). However, the same associations were not observed between EMI and receptive language in these studies (or between looking to the mouth and later receptive language in infants with a family history of autism; Chawarska et al., 2022). It is currently unclear how both infants' eye-mouth looking behaviour and having first words in infancy may relate to later language skills in infants with a family history of autism and/or ADHD.

Language-Relevant Temporal Frequencies

Falck-Ytter et al. (2013) concluded that eye and mouth looking in EL-autism infants is determined by a constellation of factors beyond diagnostic category, such as developmental stage, presence of social cues, competing non-social stimulus characteristics, multisensory integration of speech sounds and mouth movements, and genetics. It is unknown whether preferences for language-relevant frequencies of visual mouth movements also modulates eye and mouth looking. Studies on the speech modulation spectrum have shown that amplitude modulation sensitivity peaks at about 4–5 Hz for fluent speech—within the syllabic speech production range (i.e. theta, 4–8 Hz; Giraud & Poeppel, 2012; Poeppel & Assaneo, 2020). Theta-band neural oscillations are primarily responsible for processing syllables, and studies with adults show that the intelligibility of speech is diminished when theta frequency has been acoustically degraded within the speech signal (Ghitza, 2014; Luo & Poeppel, 2007). Theta frequency neural activations in newborns highlights its importance in speech perception from the prenatal period (e.g. Mariani et al., 2023). It is possible that as infants become more familiar with language, they are drawn to audiovisual cues but also perceptually tune into relevant frequencies of mouth movements to aid intelligibility.

There is evidence of cortical tracking in 5 and 8 months old typically developing infants during silent videos of syllables repeated at low temporal frequency. Infants showed neural entrainment to 2 Hz stimuli, and positive associations between 8 months 'phase angle' to theta (4 Hz) and 24 months expressive and receptive language were observed (Choisdealbha et al., 2024). This highlights the importance of attention to language-relevant temporal frequencies for later language development.

Evolutionary research suggests that rhesus monkeys demonstrate perceptual tuning to theta frequency at a behavioural level (Ghazanfar et al., 2013). Monkeys showed a visual preference for avatars of monkeys which opened and closed their mouths at theta frequency, compared to lower and higher frequencies (Ghazanfar et al., 2013). However, it is unclear if human infants show a visual preference for human mouths moving at theta frequency (relative to slower or faster frequencies) and if this associates with emerging language skills.

Present Study

While recent studies have highlighted varying patterns of attention to faces in infants with a family history of autism (Falck-Ytter et al., 2023), a gap in the literature remains here for infants with a family history of ADHD. Unlike naturalistic speech tasks, silent experimental stimulus videos where frequencies of mouth movements are carefully controlled (i.e. Ghazanfar et al., 2013) target attention to visual properties of speech-like frequencies. However, our current knowledge of infant visual attention to mouths moving at frequencies similar or different to syllabic speech remains unclear, and its association with language development in infants with a family history of autism and ADHD is not known. Taken together, these factors may elucidate possible effects on visual attention to facial features in the development of autism and ADHD.

Therefore, in this study we aimed to extend previous work by examining the relationship between visual attention to faces, temporal frequency of mouth movements, and language outcomes, through a prospective longitudinal study of infants with and without a family history of autism and/or ADHD. Firstly, we sought to investigate the developmental pattern of eye versus mouth looking behaviour across our sample at 5, 10 and 14 months, and we predicted a decrease in infants' EMI (i.e. less looking to the eyes relative to the mouth) over time. We also predicted that EMI would not vary by familial likelihood status of autism and/or ADHD. Next, we predicted that TL infants would have higher early receptive and expressive language scores at 14 months than infants with a family history of either condition. We also examined whether infant face scanning patterns are modulated by temporal frequency and language skills. We predicted that at 14 months, infants would show a smaller EMI for mid-frequency stimuli, and that frequency and having first words would be associated with a smaller EMI. We also predicted that higher expressive (but not receptive) language scores at 36 months would be associated with smaller EMI across infancy. If having first words is found to modulate concurrent EMI at 14 months, we would further explore if this associated with 36 months language and developmental EMI.

Methods

Participants

The total sample of 161 infants comprised 80 infants with EL-autism, 31 infants with EL-ADHD, 21 infants with EL of both autism and ADHD (EL-autism+ADHD) and 29 infants with TL for both conditions. Of these, one participant with EL-autism had no valid eye-tracking task trials after data quality checks (see below) and was therefore excluded from the analyses. The remaining 160 infants with valid eye-tracking data from at least one of the infant visits (see Table S1) were included in the current analysis: 79 EL-autism, 31 EL-ADHD, 21 EL-autism+ADHD, and 29 TL (77% were of White ethnicity, see Table 1 for participant

Table 1 Sample characteristics

Mean (SD)	N	EL-autism	EL-ADHD	EL-autism+ADHD	TL	Group differences
		N = 79	N = 31	N = 21	N = 29	
5 months						
n Infants [female:male]	93	42 [20:22]	14 [7:7]	12 [5:7]	25 [9:16]	$p = .78$
Age in months	93	5.29 (0.67)	5.21 (0.43)	5.33 (0.49)	5.36 (0.57)	$p = .89$
MSEL ELC [n infants]	93	[42] 84.14 (10.24)	[14] 84.71 (7.42)	[12] 85.42 (11.75)	[25] 84.88 (9.49)	$p = .98$
Valid EMI trials out of 3 [n infants]	93	[42] 2.55 (0.67)	[14] 2.79 (0.43)	[12] 2.75 (0.62)	[25] 2.72 (0.54)	$p = .46$
EMI	93	0.72 (0.25)	0.78 (0.24)	0.73 (0.31)	0.58 (0.29)	$p = .08$
10 months						
n Infants [female:male]	141	72 [37:35]	25 [12:13]	19 [7:12]	25 [11:14]	$p = .70$
Age in months	141	10.01 (0.52)	10.28 (0.89)	10.16 (0.50)	9.96 (0.61)	$p = .20$
MSEL ELC [n infants]	141	[72] 87.18 (14.52)	[25] 85.44 (15.79)	[19] 83.74 (16.14)	[25] 87.80 (11.76)	$p = .76$
Valid EMI trials out of 3 [n infants]	141	[72] 2.65 (0.63)	[25] 2.76 (0.60)	[19] 2.68 (0.58)	[25] 2.76 (0.44)	$p = .80$
EMI	141	0.52 (0.32)	0.56 (0.32)	0.48 (0.34)	0.48 (0.29)	$p = .82$
14 months						
n infants [female:male]	137	70 [33:37]	25 [8:17]	19 [7:12]	23 [10:13]	$p = .57$
Age in months	137	14.26 (0.61)	14.12 (0.83)	14.32 (0.67)	14.22 (0.67)	$p = .78$
MSEL ELC [n infants]	136	[70] 77.41 (11.44)	[24] 79.08 (11.12)	[19] 72.53 (14.50)	[23] 78.78 (11.99)	$p = .27$
CDI comprehension [n infants]	129	[69] 89.54 (71.24)	[23] 119.22 (74.25)	[17] 130.29 (139.11)	[20] 121.00 (150.09)	$p = .28$
CDI production	129	10.45 (13.70)	11.65 (12.96)	14.29 (20.95)	14.55 (22.05)	$p = .69$
Valid EMI trials out of 3 [n infants]	137	[70] 2.74 (0.56)	[25] 2.72 (0.46)	[19] 2.79 (0.54)	[23] 2.78 (0.52)	$p = .96$
EMI	137	0.44 (0.28)	0.39 (0.31)	0.33 (0.29)	0.24 (0.18)	$p = .030$[a]
36 months						
n infants [female:male]	119	61 [33:28]	23 [11:12]	16 [5:11]	19 [7:12]	$p = .31$
Age in months	119	37.21 (1.46)	37.35 (2.69)	37.19 (1.52)	36.79 (1.78)	$p = .78$
MSEL ELC [n infants]	119	[61] 108.10 (18.53)	[23] 118.39 (18.83)	[16] 105.93 (19.90)	[19] 129.05 (11.75)	$p < .001$[b]
MSEL receptive language	119	34.41 (4.66)	35.87 (6.06)	32.75 (6.03)	39.26 (2.64)	$p < .001$[c]
MSEL expressive language	119	33.26 (6.02)	34.70 (6.03)	29.13 (6.27)	37.53 (4.09)	$p < .001$[c,d]

Sex ratios across groups were tested using Chi-Squared Test for independence. Group differences in means were tested using one-way ANOVAs with group as a between-subject factor, followed by uncorrected post-hoc pairwise comparisons of means (significance level $p < .05$)

CDI communicative development index, *EL* elevated-likelihood, *EMI* eye-mouth index (collapsed across frequency), *ELC* early learning composite, *SD* standard deviation, *TL* typical-likelihood

[a]TL < EL-autism

[b]EL-autism, EL-autism+ADHD < EL-ADHD, TL

[c]EL-autism, EL-ADHD, EL-autism+ADHD < TL

[d]EL-autism+ADHD < EL-autism, EL-ADHD

characteristics). Infants were part of the longitudinal Studying Autism and ADHD Risk in Siblings study (STAARS), recruited from a volunteer database at the Birkbeck Centre for Brain and Cognitive Development, existing research networks, and advertisements including via social media and with autism-related charities. The children were recruited for the longitudinal study if, based on parent report, they had a first degree relative with: a community clinical diagnosis of autism; a community clinical diagnosis or probable research diagnosis of ADHD; or, typically developing older sibling(s) and no first-degree relatives with either diagnosis (see Supplementary Materials for full details). Inclusion criteria included full-term birth (gestational age > 36 weeks) and no known medical or developmental conditions.

Families enrolled in the study during one of the infant timepoints (5, 10 or 14 months) and were subsequently followed-up at the remaining infant timepoints with an outcome visit at 36 months. Infant visits were carried out at the Birkbeck Babylab, while 36 months data was collected at King's College London. Parents provided informed written consent. Ethical approval was granted by the National Research Ethics Service (13/LO/0751) and the Psychological Sciences Research Ethics Committee, Birkbeck, University of London.

For analysis each infant was assigned a binary rating, to account for the absence or presence of a first-degree relative with autism and ADHD. This approach allowed us to test the independent effect of an EL of autism or ADHD, and their interaction.

Stimuli

Infants took part in a battery of eye-tracking tasks at 5, 10, and 14 months. The 'frequency task' involved pairs of 10 s social or non-social silent videos. The non-social condition was only shown at 5 months, during which two videos of checkerboards were presented side by side, and the checkerboards alternated between black and white squares in every other frame at either low- (1 Hz), mid- (6 Hz) or high-frequency (10 Hz at 5 months timepoint, 15 Hz at 10 and 14 months timepoints). The present study only focused on the social condition; therefore, the non-social condition will not be further discussed. During the social condition, two videos of a woman's face were presented side by side, and the mouth flashed open and closed at low-, mid- or high-frequency; see Fig. 1. Infants viewed a total of 3 social trials at each of the 5, 10 and 14 months visits. Each social video in the respective pairs was presented at different frequencies (e.g. one trial would present 1 Hz in the video on the left-side, 6 Hz in the video on the right-side, etc.). Infants were presented with all possible frequency pairs, with each frequency shown at least once on each side of screen. We increased the high-frequency condition from 10 to 15 Hz for 10 and 14 months visits after piloting the task, as 10 Hz did not seem developmentally appropriate for capturing the older infants' attention.

Procedure

Stimuli were presented on a 23-inch flat screen monitor with a 16:9 aspect ratio. Gaze data were acquired with a Tobii TX-300 (sampled at 120 Hz) infra-red eye tracker. Infants sat in their parent's lap at distance of ~ 20° of visual angle (approximately 60 cm viewing distance). Parents were asked to refrain from directing their child's attention to specific stimuli on the screen while helping the child stay engaged. After calibration, frequency task trials were presented as part of a larger battery of short eye-tracking videos. The order of presentation for each trial was counterbalanced across participants.

Offline Data Processing

Gaze data were processed using a customised MATLAB (MathWorks, 2016) script. Raw data were assigned to areas of interest (AOIs) for the upper face (eyes) and lower face (mouth) in each stimulus within the pair for each frequency task trial, i.e. respectively for each video presented on the left- and right-side of the screen; see Fig. 1. To minimise issues with estimating gaze samples that arose with binocular gaze coordinates, the X and Y gaze co-ordinates of monocular samples were averaged for both eyes; off-screen gaze was considered as missing.

Trials where participants had < 25% valid samples (e.g. due to infant fussiness or skipped trials) were deemed invalid and excluded from further analysis. Infants showed a mean of > 2.5 valid trials out of 3 at each timepoint (see Table 1; see also Table S1 for summary of timepoints where infants contributed valid EMI data). To calculate EMI for each face stimulus i.e. respectively for left- and right-side of the screen, we used the formula: *time spent looking at the eyes/time spent looking to the eyes + the mouth*. EMI means and SDs are presented in Table 1 and Fig. S1.

Language and Developmental Measures

MacArthur-Bates Communicative Development Inventory (CDI; Fenson et al., 2007) words and gestures form is a parent/caregiver report checklist to assess infants and toddlers communication and early language skills. We used the total raw word comprehension ('understands' + 'understands and says'; i.e. receptive language skills) and production



Fig. 1 The Frequency Task. **a** Schematic of the 10 s videos of face pairs presented to infants. The mouths opened and closed at low (1 Hz), mid (6 Hz) or high (5 months: 10 Hz; 10 and 14 months: 15 Hz) frequencies. Infants were presented with all possible frequency pairs. **b** Upper face and lower face areas of interest (AOIs) used to calculate eye-mouth looking index (EMI) for each face stimulus i.e. respectively for left- and right-side of the screen

('understands and says'; i.e. expressive language skills). The CDI scores at 14 months were the focus of the present study, as infants have generally developed more language by this age than earlier timepoints, providing more variance in the scores (see Table 1).

Mullen Scales of Early Learning (MSEL; Mullen, 1995) was administered by a trained examiner at 5, 10, 14, and 36 months to measure developmental ability with a composite score. Similar to Habayeb et al. (2021), we used a binary score for the absence or presence of first words at 14 months on item 11 of the expressive language subscale (*Says first words:* 0, no words; > 0, has words). We also focused on 36 months raw MSEL receptive and expressive language subscale scores; see Table 1.

Statistical Analyses

Data were analysed using R version 3.6.2 (R Core Team, 2019). Across 5, 10 and 14 months timepoints, EMI data were non-normally distributed (Kolmogorov–Smirnov normality tests: $p < .001$). Therefore, we conducted longitudinal analyses with generalised linear mixed models (glmer; maximum likelihood (Laplace) to provide accurate estimations of parameters despite missing values; logit link function and bobyqa optimizer) using the lme4 package. This approach accounted for EMI taking values between 0 and 1. Glmer are also appropriate for analyses with missing data across trials and timepoints, as analyses are conducted on each trial (Brown, 2021).

An initial glmer was run to establish a base model that converged and whether variables included as controls (i.e. side of presentation, trial number, sex) significantly contributed to the model. This initial glmer included EMI (across 5, 10 and 14 months) as the outcome variable, and the following fixed factors: Familial likelihood (autism, ADHD, and their interaction), frequency (low, mid, high), side of presentation (left, right; given differences in lateral looking behaviour in autistic infants; Donati et al., 2020) and trial number (1, 2, 3). Sex was also included as a fixed effect (male, female) and always retained in language analyses given literature suggesting that girls show higher scores on early language development measures compared to boys (Wallentin, 2009). Age (in days; centred around the mean to aid convergence) was included as a fixed and random effect to account for minor variations. Slopes were allowed to vary by participant (random effect). Where glmer was subsequently used to test longitudinal hypotheses, separate glmer were run by adding appropriate predictors and interaction terms to the base model.

In separate glmer, interaction terms for age by frequency were included to test developmental effects, and for familial likelihood by age to test familial likelihood effects across development.

For 14 months language analyses, to test infants' receptive and expressive language skills on the CDI by familial likelihood status we used separate Analysis of Covariance (ANCOVA) models, controlling for sex and age in days. To test effects of MSEL first words acquisition (no words, has words) and frequency on EMI at 14 months, we ran a multiple regression (robust standard errors clustered by participant ID) including sex and age (as well as side of presentation and trial number if they were significant in the base glmer), and an interaction between first words and frequency. Group differences were assessed by subsequently adding familial likelihood into the multiple regression model.

To test effects of language outcomes, EMI (across 5, 10 and 14 months) was the outcome variable,[1] and 36 months raw MSEL expressive and receptive language scores (centred around the mean for model convergence) were included as fixed effects in separate glmer. If first word status was found to predict EMI at 14 months then longitudinal effects would be explored by including interaction terms between 14 months first words and 36 months MSEL outcomes in the glmer.

To aid interpretation of significant effects in glmer models, estimates were back-transformed into predicted probabilities using the formula: *exp(logit)/[1 + exp(logit)]*, and their 95% confidence intervals were calculated. Similar to raw EMI scores, probabilities > 0.5 indicates infants were more likely to look to the eyes; while probabilities < 0.5 indicates infants were more likely to look to the mouth. Tukey's method for pairwise comparisons was applied for post-hoc analyses.

Results

Preliminary Analyses

Sex was not a significant predictor ($p = .95$) and therefore removed from the initial glmer. The resulting base glmer (model 0) was of the form: glmer (EMI ~ autism likelihood * ADHD likelihood + frequency + age + side of presentation + trial number + (1 + age | participant ID)). The effect of side of presentation was significant (right-side: estimate = −0.26, SE = 0.12, $z(1925) = -2.24$, $p = .025$), as was trial number 3 compared to trial number 1 (estimate = −0.41, SE = 0.16, $z(1925) = -2.60$, $p = .009$; but not trial

[1] EMI was entered as an outcome variable when testing the effects of language outcomes (as opposed to 36 month MSEL expressive or receptive language scores) in favour of model convergence. The association being tested remains unchanged—i.e. whether there is an association between language abilities and infant EMI.

Table 2 Selected output from glmer

Model number	Fixed effect	Estimate (logit)	Standard error	*z*	*p*	Predicted probability	95% CI_{prob}
1	Age	− 1.05	0.12	− 8.88	< .001(*)	0.26	0.22, 0.31
	Frequency (low)	0.36	0.15	2.32	.021(*)	0.59	0.51, 0.66
	Interactions						
	Age * Frequency (low)	− 0.06	0.15	− 0.42	.68	0.48	0.41, 0.56
	Age * Frequency (high)	0.18	0.15	1.22	.22	0.55	0.47, 0.62
2	Age	− 1.13	0.17	− 6.70	< .001(*)	0.24	0.19, 0.31
	Autism Likelihood	0.78	0.37	2.10	.036(*)	0.68	0.51, 0.82
	ADHD Likelihood	0.81	0.45	1.82	.068	0.69	0.48, 0.84
	Interactions						
	Autism Likelihood * ADHD Likelihood	− 1.24	0.61	− 2.04	.042(*)	0.22	0.08, 0.49
	Age * Autism Likelihood * ADHD Likelihood	− 0.08	0.35	− 0.23	.82	0.48	0.31, 0.65
4a	Sex	− 0.07	0.37	− 0.19	.85	0.48	0.31, 0.66
	14 months words	− 0.59	0.37	− 1.58	.11	0.36	0.21, 0.54
	MSEL 36 months Expressive	0.32	0.33	0.98	.33	0.58	0.42, 0.72
	Interactions						
	14 months words * MSEL 36 months Expressive	− 0.57	0.38	− 1.50	.13	0.36	0.21, 0.54
5a	Sex	− 0.25	0.36	− 0.68	.49	0.44	0.28, 0.61
	14 months words	− 0.74	0.38	− 1.95	.051(*)	0.32	0.18, 0.50
	MSEL 36 months Receptive	0.65	0.32	2.01	.044(*)	0.66	0.50, 0.78
	Interactions						
	14 months words * MSEL 36 months receptive	− 0.62	0.37	− 1.68	.09	0.35	0.21, 0.53

Estimates which are significantly different from 0 are denoted with (*). Estimates are provided in the logit scale; the formula exp(logit)/[1 + exp(logit)] applies a backward transformation, converting these values into predicted probabilities

Model 1: glmer (EMI ~ autism likelihood * ADHD likelihood + frequency * age + side of presentation + trial number + (1 + age | participant ID))

Model 2: glmer (EMI ~ autism likelihood * ADHD likelihood * age + frequency + side of presentation + trial number + (1 + age | participant ID))

Model 4a: glmer (EMI ~ autism likelihood * ADHD likelihood + frequency + MSEL Expressive Language 36 months * 14 months words + age + sex + side of presentation + trial number + (1 + age | participant ID)

Model 5a: glmer (EMI ~ autism likelihood * ADHD likelihood + frequency + MSEL Receptive Language 36 months * 14 months words + age + sex + side of presentation + trial number + (1 + age | participant ID)

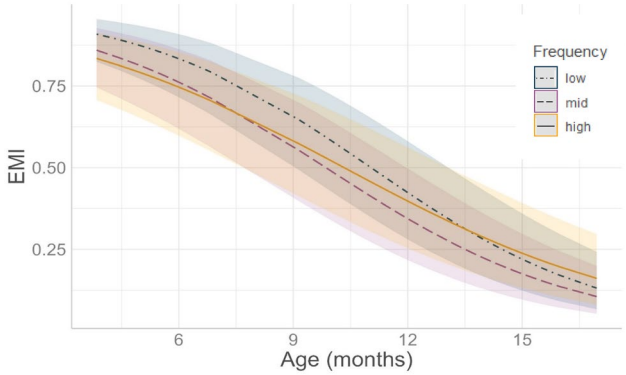


Fig. 2 Predicted probabilities of eye-mouth index (EMI) across age and temporal frequencies (modelled using 'ggpredict' function). Coloured ribbons represent 95% confidence intervals of the mean

number 2 compared to trial number 1, p = .40). Side of presentation and trial number were retained in subsequent glmer and multiple regression models; selected results are summarised in Table 2.

Developmental EMI

In the model with an interaction between age and frequency (model 1), looking to the eyes decreased as age increased (p < .001, predicted probability = 0.26, 95% CI_{prob} = 0.22, 0.31). Infants looked more at the eyes relative to the mouth for low- compared to mid-frequency stimuli (p = .021, predicted probability = 0.59, 95% CI_{prob} = 0.51, 0.66), but showed no difference in EMI between mid- and high-frequency stimuli (p = .25). There was no significant interaction between age in days and frequency such that there were no developmental changes in the frequency effect on eye/mouth looking (ps > .22), see Table 2 and Fig. 2.

Familial Likelihood and Developmental EMI

In the model including an interaction between familial likelihood and age (model 2), the effect of autism likelihood on eye-mouth looking was significant (p = .036, predicted

probability = 0.68, 95% CI_{prob} = 0.51, 0.82) while ADHD likelihood was non-significant but had a similar predicted probability to autism likelihood (p = .07, predicted probability = 0.69, 95% CI_{prob} = 0.48, 0.84). The interaction between autism and ADHD familial likelihood was significant (p = .042, predicted probability = 0.22, 95% CI_{prob} = 0.08, 0.49); the greater eye to mouth looking in infants with a family history of autism was attenuated where there was also a family history of ADHD. Pairwise comparisons also showed that infants with EL-autism had a larger EMI than the TL infants (OR = 0.46, SE = 0.17, p = .035). There were no significant differences between any other contrasts (ps > .07); see Tables 2, S2 and Fig. 3. Age remained a significant predictor of eye-mouth looking in the model (p < .001, predicted probability = 0.24, 95% CI_{prob} = 0.19, 0.31) but did not share any two- or three-way interactions with autism- or ADHD-likelihood (ps > .19; see Fig. S2).

14 Months Language and EMI

There were no main effects of familial likelihood on 14 months language comprehension (ANCOVA 1: F(3, 119) = 2.04, p = .11) or production (ANCOVA 2: F(3, 119) = 0.76, p = .52); see Fig. S3. Girls comprehended significantly more words than boys (F(1, 119) = 4.22, p = .042, η_p^2 = 0.03) and this effect was approaching significance for language production (F(1, 119) = 3.66, p = .058, η_p^2 = 0.03). Age was not a significant covariate for receptive language (p = .13), but older children produced more words than younger children (F(1, 119) = 5.08, p = .026, η_p^2 = 0.04).

We used a multiple regression (robust standard errors) to investigate if MSEL first words status and frequency predicted EMI at 14 months.

The overall model was significant (Model 3: Adjusted R^2 = 0.03; F(10, 135) = 2.03, p = .034). Infants who had acquired words by 14 months had a smaller concurrent EMI compared to those with no words, with medium-to-large effect (β = −0.14, SE = 0.06, t(59.15) = −2.36, p = .021, d = −0.61). However, there was no effect of frequency (ps > .42) or first words by frequency interaction (ps > .61); see Fig. S4.

Familial likelihood was subsequently added into the model. The overall model was significant (Model 3a: Adjusted R^2 = 0.08; F(13, 135) = 3.25, p < .001) and a significantly better fit compared to when familial likelihood was not included (Wald test: χ^2 (3) = 18.61, p < .001). First words status remained a significant predictor of 14 months EMI (β = −0.14, t(58.43) = −2.47, p = .016, d = −0.65). There were effects of autism-likelihood (β = 0.20, t(36.12) = 4.13, p < .001, d = 1.37), and ADHD likelihood (β = 0.16, t(42.99) = 2.17, p = .035, d = 0.66) such that infants with a family history of either autism or ADHD had a larger 14 months EMI. The autism-likelihood by ADHD-likelihood interaction was also significant with medium-to-large effect (β = −0.31, t(68.09) = −3.01, p = .004, d = −0.73). Pairwise comparisons showed that 14 months EMI was significantly larger in infants with EL-autism (t(720) = −4.13, SE = 0.05, p < .001, d = −0.31) and EL-ADHD (t(720) = −2.17, SE = 0.08, p = .030, d = −0.16) compared to TL infants. The EL-autism group also had a larger EMI compared to the EL-autism+ADHD group (t(720) = 2.04, SE = 0.07, p = .042, d = 0.14). No other contrasts reached significance (ps > .21); see Table S2. With the exception of right-side of presentation (β = −0.04, t(129.16) = −2.36, p = .020, d = −0.42), all other simple effects and interactions did not reach significance (ps > .17).

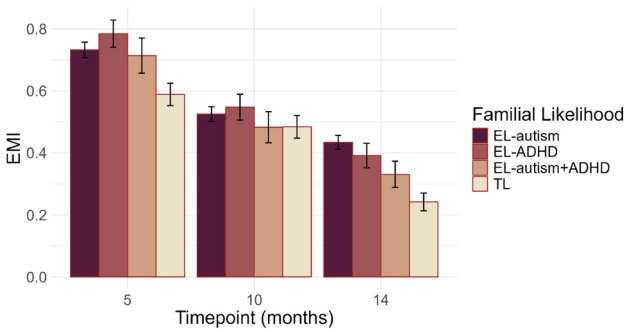


Fig. 3 Eye-mouth index (EMI) by timepoint and familial likelihood status (EL-autism, elevated likelihood for autism; EL-ADHD, elevated likelihood for attention deficit hyperactivity disorder; EL-autism+ADHD, elevated likelihood for autism and attention deficit hyperactivity disorder; TL, typical likelihood). Data were collapsed across temporal frequencies. Error bars represent standard errors of the mean

36 Months Language Outcomes and Developmental EMI

We then tested associations between 36 months language outcomes and the variance in developmental EMI.

In the model including 36 months expressive language (Model 4), there was no association between 36 months expressive language (p = .38) and developmental EMI; as was the finding in the model including 36 months receptive language (Model 5: p = .45).

Given the significant association between having first words and smaller EMI at 14 months (Model 3, Model 3a), we subsequently investigated the longitudinal relationship between developmental EMI, 14 months first words status and 36 months language outcomes. Interaction terms between 14 months first words and 36 months MSEL expressive and receptive language were separately added to the previous models.

In the 36 months expressive language model including 14 months first word status (Model 4a), having first words

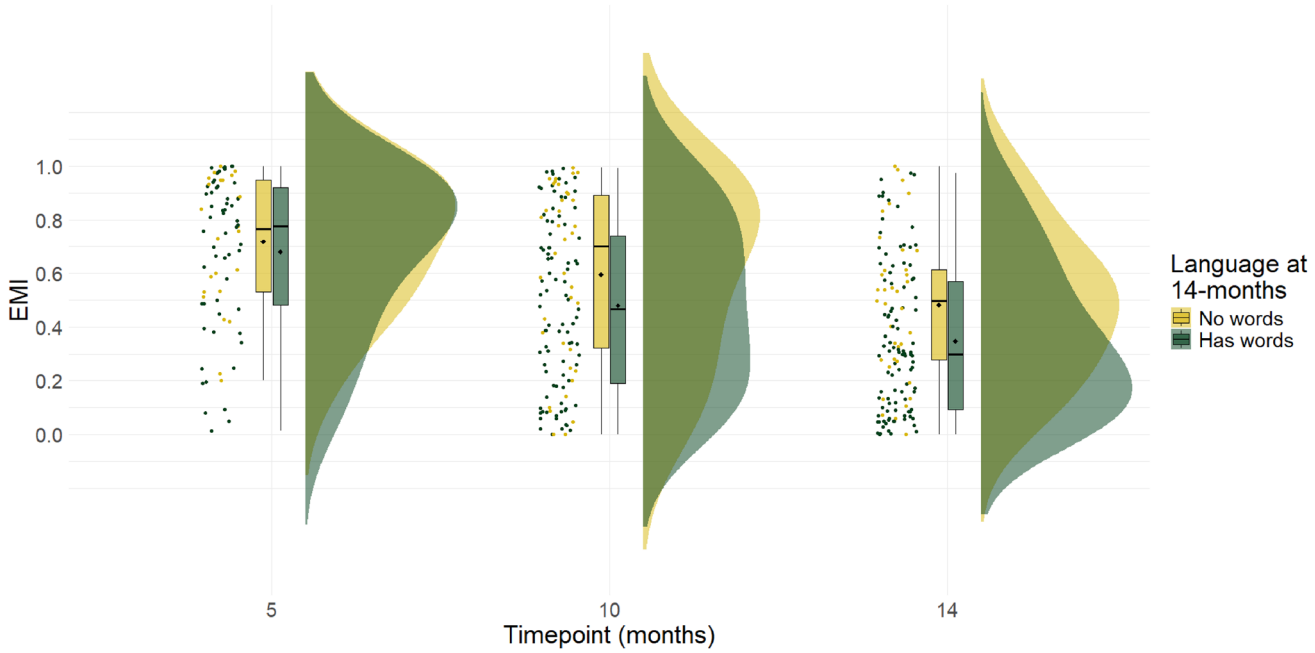


Fig. 4 Raincloud plots of eye-mouth index (EMI) by timepoint and 14 months first words status. Data were collapsed across temporal frequencies. Each coloured dot represents one participant, and the black dot on boxplots represents the mean. Infants were categorised as having acquired language at 14 months according to item 11 on the Expressive Language subscale (EL11) of the Mullen Scales of Early Learning (MSEL; No words: EL11 = 0; Has words: EL11 > 0)

was not associated with developmental EMI (p = .11), nor was the interaction between first words status and 36 months expressive language (p = .13).

However, in the 36 months receptive language model including 14 months first word status (Model 5a), the effect of first words was marginally significant (p = .051, predicted probability = 0.32, 95% CI_{prob} = 0.18, 0.50), as having words at 14 months was associated with having a smaller developmental EMI than infants without words at 14 months; see Fig. 4. The effect of MSEL 36 months receptive language was significant (p = .044, predicted probability = 0.66, 95% CI_{prob} = 0.50, 0.78); see Fig. S5 and Table 2. Higher receptive language scores at 36 months was associated with more looking to the *eyes* than the mouth across 5–14 months. The interaction term between 14 months word status and 36 months MSEL receptive language did not reach significance (p = .09).

Discussion

We investigated visual attention to faces and language outcomes in infants with a family history of autism and/or ADHD. EMI decreased with age across all infants, consistent with increasing focus on the mouth over the age range that infants develop their receptive and expressive language skills. Indeed, infants who had acquired words by 14 months showed greater concurrent mouth looking, consistent with previous work (Habayeb et al., 2021). Infants also showed more looking at the eyes relative to the mouth for low- compared to mid-frequency stimuli, suggesting some modulation by the nature of facial movement patterns. We found that both a family history of autism or a family history of ADHD was associated with more looking at eyes relative to mouth compared to infants without a family history of autism, though the latter effect was weaker and only significant at 14 months when concurrent language skills were accounted for. Further, a family history of ADHD appeared to moderate the relationship between EL-autism and EMI such that the pattern of greater eye looking in infants with a family history of autism was attenuated. Finally, higher receptive language scores at 36 months was associated with more looking to the *eyes* than the mouth when controlling for having words at 14 months. Taken together, our findings suggest that infants with a family history of autism show different face scanning patterns to those without a family history of autism in the first year of life. However, these effects may vary by the presence of ADHD in the family and language outcomes, potentially explaining some heterogeneity in previous reports.

EMI Across Infancy

As predicted, we found that there was a developmental effect on EMI, with EMI decreasing from 5 to 14 months. This is consistent with previous studies (e.g. Frank et al.,

2012; Tsang et al., 2018), showing a general shift from looking more to the eyes and towards the mouth in the first year of life. While some studies found a gradual re-orienting back towards the eyes from about 12 months (Lewkowicz & Hansen-Tift, 2012), we did not observe this in our sample. Indeed, some studies (e.g. Morin-Lessard et al., 2019) have found that increased attention to the mouth relative to the eyes remains up to age 5 in monolingual and bilingual children, supporting the notion that attention to the mouth (particularly when it is moving) is adaptive for language development (Falck-Ytter & von Hofsten, 2011).

EMI and Family History of Autism and/or ADHD

We found that infants with EL-autism looked more to the eyes relative to the mouth compared to TL infants. Contextual differences may contribute to the contrast with previous findings, where infants with EL-autism showed no differences in EMI compared to TL peers (e.g. Elsabbagh et al., 2014; Shic et al., 2014). For example, Shic et al. (2014) played background music while presenting selected stimuli to increase overall attention, whereas we reduced potential competition of input to other sensory modalities by presenting silent videos during this task. In Elsabbagh et al. (2014), the 'mouth only' condition is somewhat similar to the present study, with only the mouth moving in the videos. However, we presented pairs of stimuli (not a singular face) where the mouths were moving at varying frequencies and employed a more widely used calculation for EMI. Here, the mouths could be considered as more salient than other regions. Increased attention to the eyes relative to the mouth is typically seen earlier in infancy, with infants gradually directing their attention to the most socially relevant areas (Frank et al., 2012). Our findings may therefore reflect a developmental delay in the shift between eyes and mouth for infants with EL-autism. Future replication studies of existing experimental protocols could help strengthen previous developmental EMI findings; while novel eye-tracking research could investigate dynamic fixation patterns in infants and toddlers with EL-autism, to gain knowledge of continuous changes (i.e. by millisecond) in face scanning pathways.

We found a significant interaction between having a family history of autism and a family history of ADHD on EMI, such that the overall increase in eye looking in infants with a family history of autism was attenuated by the additional presence of ADHD. However, the effect of ADHD alone had a similar estimated probability as the effect of autism and was in the same direction (increased EMI). Further, at 14 months, we also found an association between a family history of ADHD and more looking to the eyes, and that EMI was significantly larger in infants with EL-autism when compared to EL-autism+ADHD. The development of face scanning behaviour in infants with EL-autism+ADHD may therefore be distinct from that in EL-autism, where having a family history of ADHD attenuates the relationship between EL-autism and EMI. It would be beneficial for ADHD to be mapped in cohorts with a family history of autism (Johnson et al., 2015), as ADHD familial history may affect the presence of observed effects. Future genetic research could explore non-additive effects of having a family history of autism and ADHD in relation to EMI.

Visual Attention to Theta Frequency and 14 Months Language

We found that across infancy, infants looked more at the eyes relative to the mouth for low- compared to mid-frequency stimuli. This suggests infants show some preference for mouths moving at the frequency of syllabic speech, providing partial support for perceptual tuning to theta (Ghazanfar et al., 2013). However, this was not an exclusive preference, as the effect of frequency on EMI did not hold for mid- versus high-frequency stimuli. Early infant studies using checkerboard stimuli have shown frequency stimulation preference is associated with arousal (Karmel et al., 2010). Further, recent research suggests that infants show highest levels of attention when at an intermediate level of arousal, reflecting an inverted-U shape relationship between attention and arousal (Kleberg et al., 2019). Future research could explore longitudinal effects of arousal on EMI by analysing physiological responses to mouths moving at low-, mid- and high-frequency across infancy.

Having words at 14 months was associated with more mouth looking concurrently and across infancy, compared to infants who did not have first words. However, this did not interact with frequency. Our findings can be understood in the light of attention to faces being experience dependent (Chawarska et al., 2012); and therefore, having first words possibly invites more scenarios to engage in refining and imitating word shapes from other speakers. Additionally, theta frequencies are one part of the multi-level frequency structure of language. Syllables occur in the theta range but are nested within lower (e.g. delta frequency—important for prosody) and higher frequency levels (e.g. gamma frequency—denotes phonemic units; Gross et al., 2013; Meyer, 2018). Therefore, it is possible that frequencies nested within other units of language are more salient than the frequency of syllabic speech alone at 14 months. Our study presented moving mouths and not speaking faces as such. Future studies could measure EEG during naturalistic speech videos, to examine neural responses to the different frequencies present in speech and if this associates with infant language development.

When controlling for sex and age, we observed no familial likelihood differences in 14 months CDI language production and comprehension. It is possible that differences in the analysis of language scores compared to previous studies may contribute to our findings (e.g. comprehension-production difference scores on the CDI in Hudry et al., 2014; compared to separate models in the present study).

Developmental EMI and 36 Months Language Outcomes

Initially, we did not observe an association between 36 months expressive or receptive language and EMI across 5–14 months. We subsequently found significant effects of 36 months language when controlling for whether infants could say their first words at 14 months; however, this was in the opposite direction to our predictions. When controlling for infants having first words at 14 months, higher *receptive* language at 36 months was associated with preference for the *eyes* relative to the mouth in infancy. Habayeb et al. (2021) found smaller EMI associated with higher concurrent expressive language in infants who had first words at 14 months—however, here we analysed language and longitudinal EMI. Our findings also contrast with previous literature which suggests that greater attention to the mouth in infancy associates with (a) better later receptive language in infants without a family history of autism (but not in infants with a family history of autism; Chawarska et al., 2022) and (b) better later expressive language in infants with and without a family history of autism (Young et al., 2009). One interpretation for this unexpected result is offered in recent work by Viktorsson et al. (2023). They suggest that gaze following (a necessary component of joint attention) plays a key role in infants' word learning—which may explain the positive correlation between 5 months preference for eyes and 14 months receptive language skills they observed in their study. Additionally, previous studies have found that attention to the eyes was positively correlated with interpersonal skills in TL and EL-autism infants (Key & Stone, 2012; Riddiford et al., 2022). Therefore, our present findings further highlight the interdependence of attending to both the eyes and the mouth in language development; while attending to the mouth is valuable for infants' emerging language skills, language abilities in toddlerhood are also associated with engaging with the eyes. Future work could explore the interplay between attention to eyes/mouth and language development in older children with a family history of autism and/or ADHD.

Limitations

Our results shed light on the associations between infants' EMI, a family history of autism and/or ADHD and language skills. Although we gain some insight into the different mechanisms at play here, we cannot infer causality from our analyses. While the family history design is informative for studying the development of conditions with some familial/genetic influence, the present study was underpowered to investigate autism or ADHD outcomes. Indeed, our experimental task featured videos where only the mouths were moving by design, which allowed for good control over the various frequencies of interest while allowing us to explore how infants direct their attention to faces. However, infants' attention here may differ from attention to naturalistic stimuli, as movements were isolated to the mouth in the present study, while naturalistic interactions may involve movements from other parts of the face and head (Esteve-Gibert & Guellaï, 2018). The stimuli in the present study did not emulate human speech, nor was there a stream of audio like in naturalistic speech tasks. Although previous work has identified some group differences across stimulus conditions (i.e. still, moving or speaking faces; Shic et al., 2014), recent evidence suggests that individual differences in EMI are largely independent of exact stimulus properties for 5-month-olds and toddlers (Constantino et al., 2017; Viktorsson et al., 2023). Finally, we did not control for arousal levels in the present study, or its association with frequency stimulation preference (Karmel et al., 2010).

Summary

In conclusion, using a prospective autism and/or ADHD longitudinal cohort, we studied infants' attention to the eyes and mouth using an experimental eye tracking task. We investigated whether their looking patterns were modulated by temporal frequency, and associations with later language skills. Infants generally showed smaller eye-mouth looking ratios over the first 14 months of life, and the greater eye to mouth looking in infants with a family history of autism was attenuated where there was also a family history of ADHD. Infants who had acquired words by 14 months looked more at the mouth, and higher 36 months receptive language scores was associated with a preference for the eyes across 5–14 months when controlling for first word status at 14 months. This highlights the importance of both eye and mouth looking in early development for social communication contexts and language skills.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10803-026-07516-9.

Acknowledgments We would like to warmly thank all the families that took part in this study. We would also like to thank the researchers who helped with data collection and recruitment: Kim Davies, Janice Fernandes and Natalie Vaz; and the placement students who helped in data collection and entry: Francesca Conti, Meg Jackson and Zoë Freeman. The STAARS team consists of: Tessel Bazelmans, Leila Dafner, Mutluhan Ersoy, Teodora Gliga, Amy Goodwin, Rianne Haartsen, Hanna Halkola, Alexandra Hendry, Rebecca Holman, Sarah Kalwarowsky, Anna Kolesnik, Sarah Lloyd-Fox, Nisha Narvekar, Greg Pasco, Laura Pirazzoli and Chloë Taylor. Thanks to Teresa Del Bianco and Virginia Carter Leno for helpful discussions on methods of analysis. For the purposes of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Accepted Author Manuscript version arising from this submission.

Author Contributions M.A. drafted the manuscript, which was refined and approved by all authors before submission. Conceptualization: M.H.J, T.C., E.J.H.J.; Methodology: E.J.H.J., M.A.; Data acquisition: J.B-A, M.A., STAARS Team; Formal analysis of data: M.A., L.M.; Original draft manuscript preparation: M.A.; Critical review and editing of manuscript: all authors.

Funding M.A. was supported by the Economic and Social Research Council (ESRC) London Interdisciplinary Social Science Doctoral Training Partnership (LISS DTP; Studentship Project Ref: 1917731). This work was also supported by awards from the Medical Research Council (MR/K021389/1), MQ (MQ14PP_83), EU Autism Innovative Medicine Studies-2-Trials (EU-AIMS/AIMS-2-Trials) programmes – funded by Innovative Medicines Initiative (IMI) Joint Undertaking (115300 and No. 777394, European Union's FP7 and Horizon 2020). This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme, with in-kind contributions from the European Federation of Pharmaceutical Industries and Associations (EFPIA) companies and funding from Autism Speaks, Autistica and SFARI. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results. Any views expressed are those of the author(s) and not necessarily those of the funders. This paper represents independent research part funded by the NIHR Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Data Availability Data are available following a review of requests subject to BASIS data sharing policies and ethical permissions as indicated here: https://www.basisnetwork.org/collaboration-and-project-affiliation/.

Declarations

Competing interests M.H.J. has received royalties from Wiley-Blackwell, OUP, and MIT Press. T.C. has served as a paid consultant to F. Hoffmann-La Roche Ltd and Servier, and has received royalties from Sage Publications and Guildford Publications. The remaining authors have declared that they have no competing or potential conflicts of interest.

Ethical Approval Ethical approval was granted by the National Research Ethics Service (13/LO/0751) and the Psychological Sciences Research Ethics Committee, Birkbeck, University of London.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders: DSM-5-TR* (5th edition, text revision). American Psychiatric Association Publishing.

Belteki, Z., Lumbreras, R., Fico, K., Haman, E., & Junge, C. (2022). The vocabulary of infants with an elevated likelihood and diagnosis of autism spectrum disorder: A systematic review and meta-analysis of infant language studies using the CDI and MSEL. *International Journal of Environmental Research and Public Health,19*(3), 1469.

Brown, V. A. (2021). An introduction to linear mixed-effects modeling in R. *Advances in Methods and Practices in Psychological Science,4*(1), 2515245920960351.

Chawarska, K., Lewkowicz, D., Feiner, H., Macari, S., & Vernetti, A. (2022). Attention to audiovisual speech does not facilitate language acquisition in infants with familial history of autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines,63*(12), 1466–1476.

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.

Choisdealbha, Á. N., Attaheri, A., Rocha, S., Mead, N., Olawole-Scott, H., Alfaro e Oliveira, M., Brough, C., Brusini, P., Gibbon, S., & Boutris, P. (2024). Cortical tracking of visual rhythmic speech by 5-and 8-month-old infants: Individual differences in phase angle relate to language outcomes up to 2 years. *Developmental Science,27*(4), Article e13502.

Constantino, J. N., Kennon-McGill, S., Weichselbaum, C., Marrus, N., Haider, A., Glowinski, A. L., Gillespie, S., Klaiman, C., Klin, A., & Jones, W. (2017). Infant viewing of social scenes is under genetic control and is atypical in autism. *Nature,547*(7663), 340–344.

de Klerk, C. C., Gliga, T., Charman, T., Johnson, M. H., Basis Team. (2014). Face engagement during infancy predicts later face recognition ability in younger siblings of children with autism. *Developmental Science,17*(4), 596–611.

Donati, G., Davis, R., & Forrester, G. S. (2020). Gaze behaviour to lateral face stimuli in infants who do and do not receive an ASD diagnosis. *Scientific Reports,10*(1), Article 13185.

Eigsti, I. M., de Marchena, A. B., Schuh, J. M., & Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. *Research in Autism Spectrum Disorders,5*(2), 681–691.

Elsabbagh, M., Bedford, R., Senju, A., Charman, T., Pickles, A., Johnson, M. H., Team B. (2014). What you see is what you get: contextual modulation of face scanning in typical and atypical development. *Social Cognitive and Affective Neuroscience,9*(4), 538–543.

Esteve-Gibert, N., & Guellaï, B. (2018). Prosody in the auditory and visual domains: A developmental perspective. *Frontiers in Psychology,9*, 338.

Falck-Ytter, T., Bölte, S., & Gredebäck, G. (2013). Eye tracking in early autism research. *Journal of Neurodevelopmental Disorders,5*(1), 28.

Falck-Ytter, T., Kleberg, J. L., Portugal, A. M., & Thorup, E. (2023). Social attention: Developmental foundations and relevance for autism spectrum disorder. *Biological Psychiatry,94*(1), 8–17.

Falck-Ytter, T., & von Hofsten, C. (2011). How special is social looking in ASD: A review. *Progress in Brain Research,189*, 209–222.

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

Frank, M. C., Vul, E., & Saxe, R. (2012). Measuring the development of social attention using free-viewing. *Infancy,17*(4), 355–375.

Geurts, H. M., & Embrechts, M. (2008). Language profiles in ASD, SLI, and ADHD. *Journal of Autism and Developmental Disorders,38*, 1931–1943.

Ghazanfar, A. A., Morrill, R. J., & Kayser, C. (2013). Monkeys are perceptually tuned to facial expressions that exhibit a theta-like speech rhythm. *Proceedings of the National Academy of Sciences of the United States of America,110*(5), 1959–1963.

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience,15*(4), 511.

Gremillion, M. L., & Martel, M. M. (2014). Merely misunderstood? Receptive, expressive, and pragmatic language in young children with disruptive behavior disorders. *Journal of Clinical Child & Adolescent Psychology,43*(5), 765–776.

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biology,11*(12), Article e1001752.

Guillon, Q., Hadjikhani, N., Baduel, S., & Rogé, B. (2014). Visual social attention in autism spectrum disorder: Insights from eye tracking studies. *Neuroscience & Biobehavioral Reviews,42*, 279–297.

Gurevitz, M., Geva, R., Varon, M., & Leitner, Y. (2012). Early markers in infants and toddlers for development of ADHD. *Journal of Attention Disorders,18*(1), 14–22.

Habayeb, S., Tsang, T., Saulnier, C., Klaiman, C., Jones, W., Klin, A., & Edwards, L. A. (2021). Visual traces of language acquisition in toddlers with autism spectrum disorder during the second year of life. *Journal of Autism and Developmental Disorders,51*, 2519–2530.

Hollingdale, J., Woodhouse, E., Young, S., Fridman, A., & Mandy, W. (2020). Autistic spectrum disorder symptoms in children and adolescents with attention-deficit/hyperactivity disorder: A meta-analytical review. *Psychological Medicine,50*(13), 2240–2253.

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

Johnson, M. H., Gliga, T., Jones, E., & Charman, T. (2015). Annual research review: Infant development, autism, and ADHD–early pathways to emerging disorders. *Journal of Child Psychology and Psychiatry,56*(3), 228–247.

Jones, W., Carr, K., & Klin, A. (2008). Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Archives of General Psychiatry,65*(8), 946–954.

Karmel, B. Z., Gardner, J. M., Meade, L. S., Cohen, I. L., London, E., Flory, M. J., Lennon, E. M., Miroshnichenko, I., Rabinowitz, S., Parab, S., Barone, A., & Harin, A. (2010). Early medical and behavioral characteristics of NICU infants later classified with ASD. *Pediatrics,126*(3), 457–467.

Katus, L., Hayes, N. J., Mason, L., Blasi, A., McCann, S., Darboe, M. K., de Haan, M., Moore, S. E., Lloyd-Fox, S., & Elwell, C. E. (2019). Implementing neuroimaging and eye tracking methods to assess neurocognitive development of young infants in low-and middle-income countries. *Gates Open Research,3*, 1113.

Key, A. P., & Stone, W. L. (2012). Same but different: 9-month-old infants at average and high risk for autism look at the same facial features but process them using different brain mechanisms. *Autism Research,5*(4), 253–266.

Kim, O. H., & Kaiser, A. P. (2000). Language characteristics of children with ADHD. *Communication Disorders Quarterly,21*(3), 154–165.

Kleberg, J. L., Del Bianco, T., & Falck-Ytter, T. (2019). How infants' arousal influences their visual search. *Child Development,90*(4), 1413–1423.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry,59*(9), 809–816.

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences of the United States of America,109*(5), 1431–1436.

Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron,54*(6), 1001–1010.

Ma, X. E., Gu, H., & Zhao, J. (2021). Atypical gaze patterns to facial feature areas in autism spectrum disorders reveal age and culture effects: A meta-analysis of eye-tracking studies. *Autism Research,14*(12), 2625–2639.

Mariani, B., Nicoletti, G., Barzon, G., Ortiz Barajas, M. C., Shukla, M., Guevara, R., Suweis, S. S., & Gervain, J. (2023). Prenatal experience with language shapes the brain. *Science Advances,9*(47), 3524.

Marrus, N., Hall, L., Paterson, S., Elison, J. T., Wolff, J. J., Swanson, M., Parish-Morris, J., Eggebrecht, A. T., Pruett, J. R., Jr., & Hazlett, H. (2018). Language delay aggregates in toddler siblings of children with autism spectrum disorder. *Journal of Neurodevelopmental Disorders,10*(1), 29.

Mastergeorge, A. M., Kahathuduwa, C., & Blume, J. (2021). Eye-tracking in infants and young children at risk for autism spectrum disorder: A systematic review of visual stimuli in experimental paradigms. *Journal of Autism and Developmental Disorders,51*, 2578–2599.

Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: State of the art and emerging mechanisms. *European Journal of Neuroscience,48*(7), 2609–2621.

Morin-Lessard, E., Poulin-Dubois, D., Segalowitz, N., & Byers-Heinlein, K. (2019). Selective attention to the mouth of talking faces in monolinguals and bilinguals aged 5 months to 5 years. *Developmental Psychology,55*(8), 1640.

Mullen, E. M. (1995). *Mullen scales of early learning*. AGS.

Oakley, B. F., Tillmann, J., Ahmad, J., Crawley, D., San José Cáceres, A., Holt, R., Charman, T., Banaschewski, T., Buitelaar, J., Simonoff, E., Murphy, D., & Loth, E. (2021). How do core autism traits and associated symptoms relate to quality of life? Findings from the Longitudinal European Autism Project. *Autism,25*(2), 389–404.

Papagiannopoulou, E. A., Chitty, K. M., Hermens, D. F., Hickie, I. B., & Lagopoulos, J. (2014). A systematic review and meta-analysis of eye-tracking studies in children with autism spectrum disorders. *Social Neuroscience,9*(6), 610–632.

Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience,21*(6), 322–334.

R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Riddiford, J. A., Enticott, P. G., Lavale, A., & Gurvich, C. (2022). Gaze and social functioning associations in autism spectrum disorder: A systematic review and meta-analysis. *Autism Research,15*(8), 1380–1446.

Shic, F., Macari, S., & Chawarska, K. (2014). Speech disturbs face scanning in 6-month-old infants who develop autism spectrum disorder. *Biological Psychiatry,75*(3), 231–237.

Taylor, M., Charman, T., Robinson, E., Plomin, R., Happé, F., Asherson, P., & Ronald, A. (2013). Developmental associations between traits of autism spectrum disorder and attention deficit hyperactivity disorder: A genetically informative, longitudinal twin study. *Psychological Medicine,43*(8), 1735–1746.

The MathWorks Inc. (2016). *Optimization toolbox version: 9.0* (R2016a). The MathWorks Inc. https://www.mathworks.com

Tsang, T., Atagi, N., & Johnson, S. P. (2018). Selective attention to the mouth is associated with expressive language skills in monolingual and bilingual infants. *Journal of Experimental Child Psychology,169*, 93–109.

Viktorsson, C., Portugal, A. M., Li, D., Rudling, M., Siqueiros Sanchez, M., Tammimies, K., Taylor, M. J., Ronald, A., & Falck-Ytter, T. (2023). Preferential looking to eyes versus mouth in early infancy: Heritability and link to concurrent and later development. *Journal of Child Psychology and Psychiatry,64*(2), 311–319.

Wallentin, M. (2009). Putative sex differences in verbal abilities and language cortex: A critical review. *Brain and Language,108*(3), 175–183.

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.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.