ASD Toddlers Exhibit Impaired Development of Neural Systems That Respond to and Guide Mother-Child Interactions

Yaqiong Xiao (✉ yaq.xiao@gmail.com)

University of California, San Diego <https://orcid.org/0000-0001-8719-5418>

Teresa Wen

University of California, San Diego

Lauren Kupis

University of Miami

Lisa Eyler

University of California, San Diego

Disha Goel

University of California, San Diego

Michael Lombardo

University of Cyprus

Karen Pierce

University of California, San Diego

Eric Courchesne

University of California San Diego <https://orcid.org/0000-0002-3772-5799>

Article

Keywords: Motherese, Social Mother-child Interactions, Similarity Network Fusion, Neural-clinical Clusters, Neural Responses

Posted Date: November 20th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-93365/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Nature Human Behaviour on January 3rd, 2022. See the published version at <https://doi.org/10.1038/s41562-021-01237-y>.

Abstract

Motherese is an experience-expectant, human-specific and innate form of parent speech that enhances social and language learning, and affect and emotion development in infants. An early sign of ASD is the child's lack of responding to motherese and reduced social mother-child interactions. To learn why, we devised a novel experiment quantifying (a) neural responses to motherese and other emotion speech with sleep fMRI and (b) active behavioral preference for motherese with eye tracking in ASD and TD toddlers. We combined the power of diverse neural and clinical data types using Similarity Network Fusion to reveal four neural-clinical clusters. The ASD cluster with the weakest neural responses to motherese and the poorest social and language abilities had the lowest eye tracking attention to motherese, while the TD cluster with the strongest neural response to motherese showed the opposite effects. We conclude that the ASD child's impairment in engaging in social mother-child interactions is due to impaired development of innate neural systems that normally respond to and guide behavior that maintains mother-child interactions.

Introduction

Social and language development are inextricably linked in typical infants¹ as a result of experience during mother-infant interactions. In these interactions, it is theorized that motherese or "infant-directed speech" is an experience-expectant form of speech that enhances social and language learning as well as affect and emotion development in babies and infants². It has the human-unique and special characteristics of higher pitch, slower tempo, and exaggerated speech contours accompanied by heightened positive affect. The prevailing view, then, is that motherese utterances create and maintain a mother-infant social and language interaction which promotes learning. Implicit is that the infant's neural response to motherese mediates this. Motherese occurs in all cultures and is thought to have a genetic basis that evolutionarily emerged in humans³⁻⁵. Typically developing (TD) infants attend to and prefer "motherese" over other forms of adult speech^{2,6-9}, and a small number of behavioral and neuroimaging studies suggest TD infants may process motherese differently from non-speech sounds¹⁰⁻¹⁴. However, the general theory that neural responsiveness to motherese speech in TD infants relates to variation in behavioral attention towards motherese and social and language learning and development remains largely untested.

Impairments in social and language learning and development are early-age signs of autism spectrum disorder (ASD), and the causes are essential to understand¹⁵⁻¹⁷. While behavioral literature has thoroughly described mother-infant joint social interaction deficits in ASD^{18,19}, there is remarkably scant evidence about how these deficits arise and whether a key causal factor may be abnormally reduced neural and behavioral responsiveness of ASD infants to motherese. There are no studies that characterize impairments in neural processing of motherese in ASD. If neural responses to motherese in ASD infants are reduced, then behavioral responses to motherese should likewise be reduced. This would

then disconnect the innate mother-infant interaction loop and result in reduced social and language learning and ability.

Here we hypothesize that ASD infants and toddlers exhibit impaired development of innately-driven neural mechanisms that respond to motherese, which underlies reduced behavioral preference for and attentiveness to motherese. To investigate this hypothesis, we measured functional magnetic resonance imaging (fMRI) responses to motherese speech as well as to non-motherese speech that has less affective prosody; assessed social and language development; and quantified gaze-contingent behavioral responses to females speaking motherese versus non-speech computer sounds and images. We first tested how TD and ASD differ in their neural response to speech, and also examined how TD and ASD brain response relates to social and language development as well as behavioral responses to motherese. Our prediction was that individual differences in early-age social and language development would be associated with a child's neural responses. Thus, for example, ASD infants and toddlers with the poorest neural responses to motherese, were predicted to have the most severe social symptoms, poorest language outcomes, and most impairment of behavioral preference and attention towards motherese. The converse was predicted for TD toddlers with the strongest neural responses to motherese; as a strong test of this, we included TD toddlers with a wide range of language abilities. Therefore, we expected across-subject heterogeneity as well as within-subject concordance across multiple modalities, including neural, symptom, and behavioral data.

Importantly, our assessment of the covariation of brain and behavioral traits included an examination of whether there were homogeneous subgroups within our sample of ASD and TD toddlers that differed with respect to attentional preference for motherese speech. We used Similarity Network Fusion (SNF), a novel unbiased, data-driven approach^{20,21}, not previously used in the ASD field, to identify patterns of neural and clinical characteristics that underlie greater and lesser eye tracking behavioral preference for and attention to females speaking motherese. The SNF approach integrates multiple and very different kinds of measures (e.g., fMRI, clinical) from individuals to identify clusters of subjects whose measures are maximally similar to each other and different from subjects in other clusters. In this approach, rather than determining diagnostic categories *a priori* through clinical measures, the similarity in patterns of different kinds of measures drives clusters independent of diagnosis and in this way reveals bio-behavioral dimensionality. The target outcome of interest, in this case, i.e., a child's preference and interest in motherese utterances and other affective prosody stimuli, can then be objectively and independently examined and tested for each neural-clinical cluster/subgroup.

Results

Similar activation patterns in TD adults and TD toddlers. We collected a total of 200 fMRI datasets from 71 toddlers and 41 fMRI datasets from 14 adults, separately (**Supplementary Table 1**). We first analyzed task activation data collected from sleeping TD toddlers and awake TD adults. For each language paradigm, whole-brain activation maps showed similar patterns in sleeping TD toddlers and those in passively listening adults (**Fig. 1a**), though, as expected, the percent signal changes in bilateral temporal

regions of interest (ROIs)—from meta-analytic activation map in Neurosynth (<https://neurosynth.org/>) with the term “language”—were significantly lower in sleeping TD toddlers than awake TD adults (**ROIs in Fig. 2; plots of toddler vs. adult differences in Supplementary Fig. 1**). These results validate that activation during sleep in TD toddlers reflects similar language patterns to those in awake adults albeit with reduced activation strength.

In addition, we evaluated the test-retest reliability of brain activation to language paradigms with varying levels of affective prosody within individual toddlers across time. The test-retest scans were divided into two groups based on intervals between initial and retest scans: short-term retest (1–4 months after initial scans) and long-term retest (12–15 months after initial scans). The overall test-retest reliability (initial scans versus all retest scans) was quantified with intraclass correlation coefficients, which showed moderate to good reliability for Karen language and Motherese but poor reliability for Story language, especially in the right temporal region (**Fig. 1b**).

Reduced neural response to speech with varying prosodic levels in ASD. We found robust and significant activation in temporal language regions in TD toddlers but reduced activation in ASD toddlers across all levels of affective prosody (**Fig. 3a**). Although there were no significant differences in whole-brain activation between TD and ASD toddlers, a ROI-based analysis demonstrated medium or large effect sizes (Cohen's d) for group comparisons across the three language paradigms in both left and right temporal ROIs (**Fig. 3b and 3c**), such that activation was lower in ASD than TD.

Correlations of neural response to speech with a toddler's social and communication abilities.

We further investigated correlations between a toddler's neural response to speech and his/her social and communication abilities. The results of mixed effects models showed significant correlations between fMRI activation and social and communication scores across subjects and language paradigms (left temporal ROI – communication scores: $p = .021$; left temporal ROI – social scores: $p = .009$; right temporal ROI – communication scores: $p = .013$; right temporal ROI – social scores: $p = .002$) (**Supplementary Table 2; Supplementary Fig. 2**).

Clustering results. We found five fMRI-clinical phenotypically distinct clusters, including four TD and ASD clusters, and one small mixed cluster of 2 TD and 3 ASD (**Fig. 4a**). At the individual subject level, 90% of TD toddlers fell into two clusters, TD Cluster 1 and Cluster 2; and 80% of ASD toddlers also fell into two clusters, ASD Cluster 3 and Cluster 4. Within each of four TD and ASD clusters, toddlers share similar fMRI-clinical attributes (**Fig. 4b and 4c; for plots of all five clusters, see Supplementary Fig. 3**). Across Clusters 1 to 4, right temporal responses to Motherese ranged from very high to virtually absent and, in parallel social, communication, language, cognitive and symptom severity scores also neatly ranged from high to very low. The very highest Motherese neural responses occur in those with the highest social and cognitive abilities, and the very lowest Motherese neural responses occur in those with the very lowest social and cognitive abilities and most severe ASD social symptoms. The other prosodic language

paradigms show slightly different patterns between clusters. Relative to Motherese, there is less distinction between the two TD clusters for Story language and less distinction between the two ASD clusters for Karen language. Thus, in the usual TD vs. ASD analysis that does not consider within-diagnosis heterogeneity, it is unsurprising that the TD vs ASD effect sizes were higher for the Story and Karen language than for the Motherese paradigm, which shows more differentiated activation between phenotypic subgroups.

Motherese eye-tracking results. As compared to TD toddlers, ASD toddlers showed substantially and significantly reduced percentage fixation towards motherese, preferring the computer “techno” sounds (**Fig. 5a**). We examined the association of the four neural-clinical clusters and gaze preference for motherese and found that toddlers in clusters predominantly including TD subjects, i.e., Clusters 1 and 2, showed significantly higher percentage fixation towards motherese as compared to toddlers in clusters dominated by ASD subjects, i.e., Clusters 4 (**Fig. 5b and 5c; for all five clusters, see Supplementary Fig. 4**). ASD Cluster 3, with better cognitive and developmental skills, in turn, had higher percentage fixation towards motherese than ASD Cluster 4 (66.6% versus 41.4%); there was a large effect size (Cohen’s $d = 1.01$) in motherese fixation between ASD toddlers in Cluster 3 and those in Cluster 4.

Discussion

Our within-subject fMRI, clinical and eye-tracking design provides unique evidence for the long-standing behavior-based theory that a toddler’s increased neural responses may mediate behavioral responses to motherese utterances and lead to increased social and language abilities at young ages. Using data-driven network fusion and clustering approaches, we disentangled TD and ASD heterogeneity into four main subgroups and showed that toddlers with greater neural responses to motherese and other affective prosody stimuli have greater attention towards and preference for motherese utterances and better social and language skills than those with lesser neural responsiveness. This effect was evident across TD and ASD toddlers, indicating that social preference and language development are intertwined across a wide spectrum of social and language ability and disability. Further, clustering results suggest high response TD and low response ASD toddlers stand at opposite ends of the neural-motherese-social-language spectrum, and overall, these results indicate that the biology and behavior of TD as well as ASD are dimensional.

Our study also points to the early-age neural origins of the core social deficits that first emerge in ASD infants and toddlers. As such, the absence of normal neural response to motherese and other affective language stimuli at the early age of clinical onset in ASD may be a biomarker of foundational dysregulation of social-emotional neural development that could underpin development of associated social cognitive and behavioral functions. Indeed, one of the first early signs of ASD in babies and infants is a sharply reduced or absent behavioral response to mother’s speech^{22,23}. The present study, conducted in sleeping ASD infants, provides evidence for the early-age neural basis of reduced behavioral preferences for motherese utterances. It is robustly evident in the ASD Cluster 4 with reduced neural responses to motherese, reduced behavioral attention to motherese and poor social and language

abilities, and, because the reduced neural response is observed during natural sleep, it cannot be attributed to attention, arousal, momentary distractions, or competing motivations. We literally image a near-failure of this neural cortical response in superior temporal regions. Such a near-failure in the awake ASD baby or infant will necessarily disconnect the foundational mother-infant loop that innately drives experience-expectant growth of social, language, and cognitive abilities in infants. This foundational dysfunction undermines not only the early experience-expectant mother-infant learning loop, but likely undermines later behavioral therapy efforts to socially engage ASD toddlers. Many early interventions that show success for some individuals hinge critically upon the idea of changing this attribute of early ASD development^{24–26}. The hope is that early intervention will increase engagement between the child and the social world and enable experience-expectant neuroplasticity to divert a child towards more typical developmental trajectories.

Motherese is the tool that tunes the human-special mother-infant interactive loop and is thought to have a genetic basis that evolutionarily emerged in humans; motherese speech is the same in all humans and cultures^{3–5}. As such, the neural, behavioral and experience-expectant responses in the infant are likely to also be genetically driven and evolutionarily emergent in human infants. Thus, the markedly deficits in neural and behavioral responses we identify in ASD toddlers are very likely to have a genetic basis, a hypothesis that should be pursued with brain-genetic studies. ASD is highly genetic with heritability of at least 81%²⁷. It is a prenatal multistage, multi-process disorder that begins in the first trimester with disruption of proliferation and neurogenesis and continues throughout second and third trimesters with disorder of neurite outgrowth, synaptogenesis and neural network function^{28,29}. Social and language impairments are the consequences. Indeed, genetic dysregulation in ASD is highly correlated with ASD social symptoms in infants and toddlers³⁰. We also identified gene co-expression networks in ASD that are associated with neural hypoactivation to language stimuli and with abnormal cortical growth in ASD toddlers who had very poor language development^{31,32}. Identified genes and networks enrich language-relevant, ASD-associated, human-specific, and prenatal genes, including genes involved in cell proliferation and excitatory neuron development. These several studies indicate that prenatal genetic dysregulations in ASD are a cause for early-age impairment in social and language development, with reduced neural responses to motherese speech being a pivotal impediment to socially engaging with mother. It is of major importance that future work directly test this critical theory.

Understanding early-age clinical and neural ASD heterogeneity is a major challenge. Heterogeneity among typically developing toddlers is equally important to address but is often overlooked leading to weakened power to detect and characterize differences among ASD and TD toddlers. Here we demonstrate the power of the unbiased data-driven SNF/clustering method used for the first time in the ASD field to resolve the neural bases of early-age social and language heterogeneity in both ASD and TD and revealed four main distinct subgroups. We used this method to additionally resolve, in a purely data-driven approach, heterogeneity among TD. The relevance of the neural-clinical subgroups to the behavior of interest—attention to and preference for motherese utterances—was then quantitatively measured and validated using gaze contingent eye-tracking assessments of toddlers choosing to view a female telling a

story in motherese or computer “techno” sounds and images. The results show that social and language ability and behavioral preference for motherese are linked to how strongly temporal cortex responds to motherese and other prosodic speech in toddlers across the neurodevelopmental spectrum from typical to language and social impaired.

Lastly, another exciting finding is that our purely unbiased data-driven SNF/clustering method appears to have replicated the same two ASD subgroups we previously identified using a subjective stratification approach^{31,33} in a new sample of toddlers using new paradigms. Previously, we arbitrarily stratified ASD based on a child being above (“ASD Good”) or below (ASD Poor”) one standard deviation on the Mullen expressive and receptive language scales. The same two ASD subgroups emerged from the purely unbiased SNF/clustering method while also resolving TD toddler heterogeneity. Therefore, these may be robust and reliable ASD subgroups which are both etiologically and biologically meaningful. As such, future work should investigate how they may open early-age diagnostic, prognostic and/or treatment avenues for biomarker discovery.

In conclusion, in a one of a kind fMRI study of ASD, we resolve ASD and TD neural-social-language-behavioral heterogeneity into four main discrete subgroups and show robust and systematic differences in how the ASD and TD toddler brain responds to varying levels of emotional prosodic speech including motherese. These neural differences underlie distinctions in behavioral preference for social-emotional motherese utterances and relate to social and communication developmental differences. These findings support the longstanding behavior-based theory that motherese is an essential and innately driven stimulus that evokes neural responses driving infants to engage with mother in social and language learning. Neural responsiveness may lead to such learning, while sharply reduced or absent neural responsiveness may impede or preclude it. It suggests neural and behavioral deficits together may be a biomarker of foundational dysregulation of social-emotional neural development and learning. As such, different ASD neural-clinical-behavioral subgroups were identified that may benefit from different treatment approaches.

Methods

Participants. Toddlers were recruited through community referral and a population-based screening method in collaboration with pediatricians via the 1-Year Well-Baby Check-Up Approach^{15,17}. All toddlers participated in clinical assessments, including the ADOS³⁴, Mullen Scales of Early Learning³⁵, and Vineland Adaptive Behavior Scales³⁶. Toddlers who received their initial diagnostic and clinical evaluations at < 36 months were invited to return for repeat evaluations until they reached 48 months. Clinical scores at the outcome visit were used as a best estimate of a child’s abilities (**Supplementary Table 3**). Clinical testing occurred at the University of California, San Diego Autism Center of Excellence. Adult participants were recruited by word of mouth. This study was approved by the University of California, San Diego Institutional Review Board. Informed consent was obtained from parents or guardians of toddlers and from adult participants.

Clinical scores and fMRI scans were collected from 71 toddlers (41 ASD/30 TD). Scans were conducted during natural sleep, which has been proven to yield robust activation in ASD and TD toddlers^{37–40}. Toddlers were considered TD if their diagnosis at outcome was TD and their Mullen Early Learning Composite scores fell within 2 standard deviations of the group mean. This allows us to examine activation patterns along a continuum of language and cognitive abilities in TD children. A subset of toddlers (4 ASD/6 TD) had test-retest fMRI scans collected at intervals ranging from 1 – 15 months after the initial scan. fMRI scans were also obtained from 14 TD adults (6M/8F, 20 – 37 years old).

Language paradigms. We presented three types of language stimuli with varying levels of emotional valence, including the “Story” language paradigm we used in earlier studies^{31,33,38,41,42} (i.e., low emotional valence); a newly-created “Karen” language paradigm (i.e., moderate emotional valence); and a Motherese paradigm (i.e., high emotional valence). Stimuli were created by recording female voices reading nursery stories or age-appropriate phrases either using neutral or infant-directed utterances (e.g., Motherese^{2,8}). Language paradigms were presented in a block design (20s stimulus/20s rest).

The Story language paradigm has been used previously to identify ASD language-related brain signatures^{38,39} and to develop predictors of language outcome among ASD toddlers^{31,33}. It consists of three types of speech stimuli (simple forward speech, complex forward speech, and backward speech; 9 cycles of speech + rest; 6min 25s). Forward and backward speech stimuli were combined, and our main contrast of interest was all speech versus rest^{31,33}.

To expand our understanding of neurofunctional language and affect development in ASD, we developed two new social orienting language paradigms. The Karen language paradigm utilizes 18 different nursery story stimuli (12min 5s) with moderate levels of emotion and prosody. The Motherese paradigm includes 12 phrases (8min 5s) recorded using high-pitched, intonational, lyrical and sing-songy speech characteristic of motherese.

Emotionality level testing. Two computer-based surveys were administered to TD adults to test emotionality levels of language paradigms.

Each fMRI paradigm consists of unique language segments, i.e., 2 Story language, 18 Karen language, and 12 Motherese segments. For survey 1, each unique segment was presented in random order (same order for each participant). Subjects were instructed to listen to each segment and respond using a Likert scale of 1 – 5 with a rating of 1 indicating the least amount of emotionality and a rating of 5 indicating the most. TD adults (n=19) rated Story language as the least emotional, followed by Karen language, and Motherese segments as the most emotional (**Supplementary Fig. 5a**).

Survey 2 consisted of 18 trials, each containing a Story language segment, a Karen language segment, and a Motherese segment. Presenting all three stimulus types allowed for evaluation of differences in emotionality across all language paradigms. TD adults (n=15) then rated each segment using a Likert scale of 1 through 3 with a rating of 1 indicating the least amount of emotionality, 2 – some emotionality,

and 3 – very emotional. Participants rated Story language as the least emotional, followed by Karen language, and Motherese segments as the most emotional (**Supplementary Fig. 5b**).

fMRI data acquisition. All fMRI data were collected in a 3T GE scanner at the University of California, San Diego Center for Functional MRI. Functional images were acquired with a multi-echo EPI protocol (echo times (TE) = 15ms, 28ms, 42ms, 56ms; TR = 2500ms; flip angle = 78°; matrix size = 64 × 64; slice thickness = 4 mm; field of view (FOV) = 256 mm; 34 slices). Structural images were acquired using a T1-weighted MPRAGE sequence (FOV = 256 mm; TE = 3.172ms; TR= 8.142ms; Flip angle = 12°).

Imaging data preprocessing. Functional data were preprocessed using the ME-ICA analysis pipeline “meica.py”^{43,44} implemented in AFNI⁴⁵ and Python. First, the first 4 volumes of each run were discarded to allow for magnetization to reach steady state. Next, motion correction parameters were calculated based on the first TE images (TE = 15 ms) using a rigid-body alignment procedure. Slice timing correction was implemented for functional images of each TE, which then were normalized to an age-matched infant template⁴⁶. The time series of four TEs were combined into a single time series⁴⁷. Both principal and independent component analyses were applied to denoise the data through isolation of thermal (i.e., random) noise from structured signals (i.e., BOLD and non-BOLD signals) and separation of BOLD and non-BOLD signals. Only the BOLD-like components were retained in the preprocessed images, which were then spatially smoothed with a 8 mm FWHM Gaussian kernel.

Head motion was quantified via framewise displacement (FD)⁴⁸. For adults and sleeping toddlers, head motion was minimal (mean FD < 0.1 mm). Group differences between ASD and TD toddlers in mean FD were only seen in the Motherese paradigm; there were no group differences between adults and toddler groups (**Supplementary Table 4**).

Whole-brain analyses. First-level and second-level whole-brain activation analyses were conducted with the general linear model (GLM) in SPM12 (<https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>). Events in first-level models were based on the canonical hemodynamic response function and its temporal derivative. Given that scans were collected at multiple time points, including retest scans, we ran second-level whole-brain analyses with mixed effects models using 3dMVM program in AFNI:

$$brain\ activation = \beta_0 + \beta_1 * group + \beta_2 * age + \beta_3 * gender + \beta_4 * mean\ FD + \varepsilon$$

In mixed effects models, brain activation to each language paradigm (i.e., speech versus rest contrast) served as a dependent variable. Individual subjects were treated as a random effect, which allows for fixed effects (i.e., age, gender, and mean FD) to vary for each subject.

Using a similar approach, we conducted whole-brain analyses with adult data for each language paradigm. However, only within-group tests were performed as all adult participants had typical development.

Resulting activation maps were corrected for multiple comparisons with the family-wise error (FWE) approach using 3dClustSim program in AFNI (voxel wise $p = 0.005$ and cluster size > 138 voxels for adults and cluster size > 186 voxels for toddlers). This spatial cluster correction took into account spatial autocorrelation by using the ‘-acf’ option in 3dClustSim.

Calculation of percent signal changes in temporal language cortex. Two language-relevant ROIs from the meta-analytic activation map in Neurosynth (<https://neurosynth.org/>) with the term “language”, including left and right temporal regions, were used for ROI analysis (**Fig. 2**). These ROIs were identical to those used in previous papers^{31,33}. Given that a toddler template was used for toddler samples, ROIs were co-registered to the toddler template using FSL’s flirt function^{49,50}. For each language paradigm, percent signal changes were calculated with first-level models in speech versus rest contrast for all toddlers and adults.

Examination of Stability and Validation of fMRI Activation in Toddlers: Toddler-Adult, Test-Retest, and Sleep-Awake. Given the challenges relating to implementing sleep imaging with toddlers, test-retest is rarely examined, but essential for determining the rigor of this approach. Additional key questions surround the degree to which functional activation patterns vary along the dimension of sleep and awake states, and developmental periods such as between toddlers and adults or in individuals across time. Here we took steps towards filling these gaps and tested: 1) whether brain activation to language stimuli in sleeping TD toddlers is similar to that in passively listening adults; 2) whether brain activation patterns are stable and reproducible in TD and ASD individuals across time.

These questions were addressed by comparing percent signal change values between TD adults and TD toddlers, and by computing intraclass correlation coefficients of brain activation within individuals who were scanned multiple times at intervals of 1–15 months, respectively.

Group differences in temporal cortex activation between ASD and TD. We compared percent signal change values between ASD and TD toddlers in a priori temporal ROIs relevant to language processing. Given the repeated time points, mixed effect models with lmer function (lme4 package) in R⁵¹ were used, in which age, gender, and mean FD were included as fixed effects while subjects served as a random effect.

Brain–behavior correlation analysis. Using similar mixed effects models as aforementioned, but with social or communication scores as a predictor of interest (age, gender, and mean FD as control variables/fixed effects, subjects as a random effect), we investigated the relevance of brain activation to a child’s social and communication abilities assessed by the Vineland Adaptive Behavior Scales³⁶.

Clustering analysis using Similarity Network Fusion. SNF is a novel approach for capturing heterogeneity in multiple types of patient data and forming clusters or subgroups. The method reduces noise by aggregating across multiple types of data, detects common and complementary signals from different types of data, and reveals the importance of each data type to patient similarity. Validation and clinical values emanate from follow-up analyses testing whether different clusters predict important “held-back”

clinical variables of interest. In the original description, five SNF clusters formed from multiple types of heterogeneous genomic and genetic variables were associated with significantly different patient survival times²⁰, a finding substantially superior to other single modality cluster approaches.

To identify clusters of clinical features of ASD and TD toddlers linked with patterns of fMRI activation to speech with varying emotionality levels, we used SNF²⁰ to integrate fMRI brain activation in three language paradigms and clinical measures and then used Louvain algorithm⁵² to detect clusters of the similarity network. For this analysis, we included a subset of 52 of the 71 toddlers who had successful scans of all three language paradigms. The analysis was performed with 6 ROI variables (left and right temporal activation for each of the three language paradigms) and 14 clinical variables (i.e., 3 ADOS variables, 6 Vineland variables, and 5 Mullen variables) in R with SNFtool package. First, ROI and clinical data were normalized separately. Next, pairwise distance matrices between subjects were calculated for ROI or clinical data. Affinity matrices (networks) were computed based on distance matrices. Each affinity matrix is equivalent to a similarity network where nodes are samples (e.g., subjects) and weighted edges represent pairwise sample similarities. Network fusion that iteratively updates every network was then performed, making two networks more similar to each other with every iteration. After a few iterations, two networks converged to a single network. We constructed the network with the strongest 15% connecting partners of each subject and ran the clustering analysis with the Louvain community algorithm. The clusters were visualized with Cytoscape⁵³.

Motherese eye-tracking task. We used a novel eye tracking Motherese task that utilized gaze contingent technology wherein a toddler's gaze activates what he/she sees and hears. Toddlers can choose to watch a movie depicting an actress telling a story using motherese speech or computer "techno" sounds and images. Motherese utterances used in this task were among those included in fMRI experiments. Fifty-four toddlers had moderate or good eye-tracking performance and total looking time > 50% and were therefore included in the analysis. For 38 toddlers, eye-tracking data were collected prior to fMRI while 16 toddlers completed the task after fMRI.

Eye tracking was conducted using Tobii software (Tobii Studio; Tobii Pro Lab), and fixation data were collected using a velocity threshold of 0.42 pixels/ms (Tobii Studio Tobii Fixation Filter) or 0.03 degrees/ms (Tobii Pro Lab Tobii IV-T Fixation Filter). Preference for Motherese was characterized by comparing percent fixation duration within Motherese versus computer "Techno" sounds and images, we tested Motherese preferences between ASD and TD and across clusters identified by clustering analysis.

Declarations

Acknowledgements

The authors sincerely thank the parents and children in San Diego who participated in our research, without whom this would not be possible. We are also fortunate to work with wonderful pediatricians and family practice physicians spanning a range of medical groups including UCSD, Sharp Rees-Stealy,

Scripps, Rady-Children's Primary Care Medical Group, Chula Vista Pediatrics, Graybill Medical Group, Grossmont Pediatrics, Linda Vista Health Care Center, Mills Pediatrics, North County Health Services, San Diego Family Care, and Sea Breeze Pediatrics. We are grateful for their support. This work was supported by NIDCD grant 1R01DC016385 awarded to Eric Courchesne and Karen Pierce; NIMH grants R01MH118879 and R01MH104446 awarded to Karen Pierce; and 755816 European Research Council awarded to Michael V. Lombardo and Eric Courchesne.

Author contributions

E.C., K.P., and L.E. conceived the idea and designed the study. L.P., D.G., T.W., Y.X., L.E., and E.C. collected the data. Y.X. conceived and performed all analyses. E.C., K.P., and L.E. obtained grant funding. Y.X., and E.C. wrote the manuscript. All authors contributed to editing the manuscript.

Competing Interests

The authors declare no competing interests.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Ramírez-Esparza, N., García-Sierra, A. & Kuhl, P. K. Look Who's Talking NOW! Parentese Speech, Social Context, and Language Development Across Time. *Front. Psychol.* **8**, 1008 (2017).
2. Saint-Georges, C. *et al.* Motherese in Interaction: At the Cross-Road of Emotion and Cognition? (A Systematic Review). *PLoS One* **8**, 1–17 (2013).
3. Kuhl, P. K. *et al.* Cross-language analysis of phonetic units in language addressed to infants. *Science* **277**, 684–686 (1997).
4. Grieser, D. A. L. & Kuhl, P. K. Maternal Speech to Infants in a Tonal Language: Support for Universal Prosodic Features in Motherese. *Dev. Psychol.* **24**, 14–20 (1988).
5. Falk, D. Prelinguistic evolution in early hominins: Whence motherese? 491–541 (2004).
6. Cooper, R. P. & Aslin, R. N. Preference for Infant-Directed Speech in the First Month after Birth. *Child Dev.* **61**, 1584 (1990).
7. Fernald, A. Four-month-old infants prefer to listen to motherese. *Infant Behav. Dev.* **8**, 181–195 (1985).
8. Kuhl, P. K., Coffey-Corina, S., Padden, D. & Dawson, G. Links between social and linguistic processing of speech in preschool children with autism: behavioral and electrophysiological measures. *Dev. Sci.* **8**, F1–F12 (2005).

9. Pegg, J. E., Werker, J. F. & McLeod, P. J. Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behav. Dev.* **15**, 325–345 (1992).
10. Saito, Y. *et al.* Frontal cerebral blood flow change associated with infant-directed speech. *Arch. Dis. Child. Fetal Neonatal Ed.* **92**, F113–F116 (2007).
11. Santesso, D. L., Schmidt, L. A. & Trainor, L. J. Frontal brain electrical activity (EEG) and heart rate in response to affective infant-directed (ID) speech in 9-month-old infants. *Brain Cogn.* **65**, 14–21 (2007).
12. Sulpizio, S. *et al.* fNIRS reveals enhanced brain activation to female (versus male) infant directed speech (relative to adult directed speech) in Young Human Infants. *Infant Behav. Dev.* **52**, 89–96 (2018).
13. Zangl, R. & Mills, D. L. Increased brain activity to infant-directed speech in 6- and 13-month-old infants. *Infancy* **11**, 31–62 (2007).
14. Zhang, Y. *et al.* Neural coding of formant-exaggerated speech in the infant brain. *Dev. Sci.* **14**, 566–581 (2011).
15. Pierce, K. *et al.* Detecting, studying, and treating autism early: The one-year well-baby check-up approach. *J. Pediatr.* **159**, 458–465.e6. (2011).
16. Pierce, K., Courchesne, E. & Bacon, E. To Screen or Not to Screen Universally for Autism is not the Question: Why the Task Force Got It Wrong. *J. Pediatr.* **176**, 182–194 (2016).
17. Pierce, K. *et al.* Evaluation of the Diagnostic Stability of the Early Autism Spectrum Disorder Phenotype in the General Population Starting at 12 Months. *JAMA Pediatr.* **173**, 578–587 (2019).
18. Bacon, E. C. *et al.* Rethinking the idea of late autism spectrum disorder onset. *Dev. Psychopathol.* **30**, 553–569 (2018).
19. Bruinsma, Y., Koegel, R. L. & Koegel, L. K. Joint attention and children with autism: A review of the literature. *Ment. Retard. Dev. Disabil. Res. Rev.* **10**, 169–175 (2004).
20. Wang, B. *et al.* Similarity network fusion for aggregating data types on a genomic scale. *Nat. Methods* **11**, 333–337 (2014).
21. Pai, S. & Bader, G. D. Patient Similarity Networks for Precision Medicine. *Journal of Molecular Biology* **430**, 2924–2938 (2018).
22. Klin, A. Listening Preferences in Regard to Speech in Four Children with Developmental Disabilities. *J. Child Psychol. Psychiatry* **33**, 763–769 (1992).
23. Klin, A. Young autistic children's listening preferences in regard to speech: A possible characterization of the symptom of social withdrawal. *J. Autism Dev. Disord.* **21**, 29–42 (1991).
24. Bacon, E. C. *et al.* Measuring Outcome in an Early Intervention Program for Toddlers with Autism Spectrum Disorder: Use of a Curriculum-Based Assessment. *Autism Res. Treat.* **2014**, 964704 (2014).
25. Dawson, G. *et al.* Randomized, controlled trial of an intervention for toddlers with autism: The early start Denver model. *Pediatrics* **125**, e17–e23 (2010).

26. Kasari, C., Freeman, S. & Paparella, T. Joint attention and symbolic play in young children with autism: A randomized controlled intervention study. *J. Child Psychol. Psychiatry Allied Discip.* **47**, 611–620 (2006).
27. Sandin, S. *et al.* The heritability of autism spectrum disorder. *JAMA - J. Am. Med. Assoc.* **318**, 1182–1184 (2017).
28. Courchesne, E. *et al.* The ASD Living Biology: from cell proliferation to clinical phenotype. *Mol. Psychiatry* **24**, 88–107 (2019).
29. Courchesne, E., Gazestani, V. H. & Lewis, N. E. Prenatal Origins of ASD: The When, What, and How of ASD Development. *Trends Neurosci.* **43**, 326–342 (2020).
30. Gazestani, V. H. *et al.* A perturbed gene network containing PI3K–AKT, RAS–ERK and WNT– β -catenin pathways in leukocytes is linked to ASD genetics and symptom severity. *Nat. Neurosci.* **22**, 1624–1634 (2019).
31. Lombardo, M. V. *et al.* Large-scale associations between the leukocyte transcriptome and BOLD responses to speech differ in autism early language outcome subtypes. *Nat. Neurosci.* **21**, 1680–1688 (2018).
32. Lombardo, M. V. *et al.* Atypical genomic patterning of the cerebral cortex in autism with poor early language outcome. *bioRxiv* (2020). doi:10.1101/2020.08.18.253443
33. Lombardo, M. V. *et al.* Different functional neural substrates for good and poor language outcome in autism. *Neuron* **86**, 267–277 (2015).
34. Lord, C., Elsabbagh, M., Baird, G. & Veenstra-Vanderweele, J. Autism spectrum disorder. *Lancet* **392**, 508–520 (2018).
35. Mullen, E. M. *Mullen Scales of Early Learning*. (American Guidance Service, 1995).
36. Sparrow, S., Cicchetti, D. & Balla, D. *Vineland-II Scales Of Adaptive Behavior: Survey Form Manual*. (American Guidance Service, 2005).
37. Dehaene-Lambertz, G., Dehaene, S. & Hertz-Pannier, L. Functional neuroimaging of speech perception in infants. *Science* **298**, 2013–2015 (2002).
38. Eyler, L. T., Pierce, K. & Courchesne, E. A failure of left temporal cortex to specialize for language is an early emerging and fundamental property of autism. *Brain* **135**, 949–960 (2012).
39. Redcay, E. & Courchesne, E. Deviant Functional Magnetic Resonance Imaging Patterns of Brain Activity to Speech in 2–3-Year-Old Children with Autism Spectrum Disorder. *Biol. Psychiatry* **64**, 589–598 (2008).
40. Redcay, E., Kennedy, D. P. & Courchesne, E. fMRI during natural sleep as a method to study brain function during early childhood. *Neuroimage* **38**, 696–707 (2007).
41. Pierce, K. & Redcay, E. Fusiform Function in Children with an Autism Spectrum Disorder Is a Matter of ‘Who’. *Biol. Psychiatry* **64**, 552–560 (2008).
42. Courchesne, E. *et al.* Mapping Early Brain Development in Autism. *Neuron* **56**, 399–413 (2007).

43. Kundu, P., Inati, S. J., Evans, J. W., Luh, W. M. & Bandettini, P. A. Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. *Neuroimage* **60**, 1759–1770 (2012).
44. Kundu, P. *et al.* Integrated strategy for improving functional connectivity mapping using multiecho fMRI. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 16187–16192 (2013).
45. Cox, R. W. AFNI: Software for Analysis and Visualization of Functional Magnetic Resonance Neuroimages. *Comput. Biomed. Res.* **29**, 162–173 (1996).
46. Shi, F. *et al.* Infant Brain Atlases from Neonates to 1- and 2-Year-Olds. *PLoS One* **6**, e18746 (2011).
47. Kundu, P. *et al.* Multi-echo fMRI: A review of applications in fMRI denoising and analysis of BOLD signals. *Neuroimage* **154**, 59–80 (2017).
48. Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L. & Petersen, S. E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* **59**, 2142–2154 (2012).
49. Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *Neuroimage* **17**, 825–841 (2002).
50. Jenkinson, M. & Smith, S. A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* **5**, 143–156 (2001).
51. Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
52. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, P10008 (2008).
53. Shannon, P. *et al.* Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

Figures

Figure 1

Stability and reliability of brain activation in toddlers. a, fMRI language activation patterns are similar in typically developing (TD) adults and sleeping toddlers across three language paradigms. All t-maps were corrected for multiple comparisons (FWE corrected) with voxel-wise $p = 0.005$ and minimal cluster size of 142 voxels for TD adult group and 190 voxels for TD toddler group. b, The line plots show brain response to three language paradigms within individuals across time, including initial scan, short-term retest (with intervals of 1–4 months), and long-term retest (with intervals of 12–15 months). Intraclass correlation coefficients (ICC) were calculated between initial scans and all retest scans (including short- and long-term retest scans), and the results suggest moderate to good reliability for Karen language and Motherese but poor reliability for Story language, especially in the right temporal region.

Figure 1

Stability and reliability of brain activation in toddlers. a, fMRI language activation patterns are similar in typically developing (TD) adults and sleeping toddlers across three language paradigms. All t-maps were corrected for multiple comparisons (FWE corrected) with voxel-wise $p = 0.005$ and minimal cluster size of 142 voxels for TD adult group and 190 voxels for TD toddler group. b, The line plots show brain response to three language paradigms within individuals across time, including initial scan, short-term retest (with intervals of 1–4 months), and long-term retest (with intervals of 12–15 months). Intraclass correlation coefficients (ICC) were calculated between initial scans and all retest scans (including short- and long-term retest scans), and the results suggest moderate to good reliability for Karen language and Motherese but poor reliability for Story language, especially in the right temporal region.

Figure 2

Neurosynth language regions of interest (ROIs). Two ROIs within the Neurosynth “Language” meta-analysis map (<https://neurosynth.org/>) in left and right temporal regions.

Figure 2

Neurosynth language regions of interest (ROIs). Two ROIs within the Neurosynth “Language” meta-analysis map (<https://neurosynth.org/>) in left and right temporal regions.

Figure 3

Reduced brain activation in ASD toddlers as compared to TD toddlers. a, Whole-brain activation to language paradigms with varying levels of prosody in TD and ASD toddlers. b, Percent signal changes in left and right temporal regions across three language paradigms. The boxes show interquartile range

(IQR; first quartile, Q1; third quartile, Q3); the vertical line inside the box represents the median; the whiskers indicate $Q1 - (1.5 \times IQR)$ or $Q3 + (1.5 \times IQR)$. c, The heatmap matrix shows standardized effect sizes (Cohen's d) for each pairwise group comparison (TD vs. ASD). Cohen's d value is shown in each cell and the standard effect size is also indicated by the color of the cell. The asterisk indicates significant results of two-sample t-tests between groups (TD vs. ASD). * $p < .05$, ** $p < .005$.

Figure 3

Reduced brain activation in ASD toddlers as compared to TD toddlers. a, Whole-brain activation to language paradigms with varying levels of prosody in TD and ASD toddlers. b, Percent signal changes in left and right temporal regions across three language paradigms. The boxes show interquartile range (IQR; first quartile, Q1; third quartile, Q3); the vertical line inside the box represents the median; the whiskers indicate $Q1 - (1.5 \times IQR)$ or $Q3 + (1.5 \times IQR)$. c, The heatmap matrix shows standardized effect sizes (Cohen's d) for each pairwise group comparison (TD vs. ASD). Cohen's d value is shown in each cell and the standard effect size is also indicated by the color of the cell. The asterisk indicates significant results of two-sample t-tests between groups (TD vs. ASD). * $p < .05$, ** $p < .005$.

Figure 4

TD and ASD subgroups with distinct fMRI-clinical patterns. **a**, Similarity Network Fusion and Louvain algorithm revealed 5 fMRI-clinical distinct subgroups (green, TD; red, ASD), including two TD clusters (Cluster 1 and 2), two ASD clusters (Cluster 3 and 4), and one mixed cluster (Cluster 5) with 2 TD and 3 ASD toddlers. **b**, Reduced brain activation to three language paradigms in ASD clusters as compared to TD clusters. **c**, Overall, severe symptoms and social and language deficits in ASD toddlers but Cluster 4 appears to be more severe and have a poorer prognosis as compared to Cluster 3. The boxes show interquartile range (IQR; first quartile, Q1; third quartile, Q3); the vertical line inside the box represents the median; the whiskers indicate $Q1 - (1.5 \times IQR)$ or $Q3 + (1.5 \times IQR)$. For brain activation and clinical scores in 5 subgroups, see Supplementary Fig. 3. Abbreviations: ASD, autism spectrum disorder; TD, typical development; ABC, adaptive behavior composite; EL, expressive language; RL, expressive language; ELC, early learning composite; ADOS, Autism Diagnostic Observation Schedule; SA, social affect; RRB, restricted and repetitive behavior.

Figure 4

TD and ASD subgroups with distinct fMRI-clinical patterns. **a**, Similarity Network Fusion and Louvain algorithm revealed 5 fMRI-clinical distinct subgroups (green, TD; red, ASD), including two TD clusters (Cluster 1 and 2), two ASD clusters (Cluster 3 and 4), and one mixed cluster (Cluster 5) with 2 TD and 3 ASD toddlers. **b**, Reduced brain activation to three language paradigms in ASD clusters as compared to TD clusters. **c**, Overall, severe symptoms and social and language deficits in ASD toddlers but Cluster 4 appears to be more severe and have a poorer prognosis as compared to Cluster 3. The boxes show interquartile range (IQR; first quartile, Q1; third quartile, Q3); the vertical line inside the box represents the median; the whiskers indicate $Q1 - (1.5 \times IQR)$ or $Q3 + (1.5 \times IQR)$. For brain activation and clinical scores in 5 subgroups, see Supplementary Fig. 3. Abbreviations: ASD, autism spectrum disorder; TD, typical development; ABC, adaptive behavior composite; EL, expressive language; RL, expressive language; ELC, early learning composite; ADOS, Autism Diagnostic Observation Schedule; SA, social affect; RRB, restricted and repetitive behavior.

Figure 5

Preference for motherese vs. computer “techno” images and sounds in TD and ASD as well as 4 clusters. a, Typical developing (TD) toddlers showed significant stronger preference for motherese versus computer “techno” images and sounds than ASD toddlers (d , Cohen's d ; white cross, group mean). b, Toddlers in TD clusters (Clusters 1 and 2) had significantly higher percentage fixation towards motherese versus computer “techno” images and sounds than ASD toddlers (Clusters 4). The white cross indicates group mean. For percentage fixation towards Motherese (% Motherese) in 5 clusters, see Supplementary Fig. 4. c, The heatmap matrix shows standardized effect sizes (Cohen's d) for each pairwise group comparison between clusters. Cohen's d value is shown in each cell and the standard effect size is also indicated by the color of the cell. The asterisk indicates significant results of two-sample t-tests. * $p < .05$, ** $p < .005$.

Figure 5

Preference for motherese vs. computer “techno” images and sounds in TD and ASD as well as 4 clusters. a, Typical developing (TD) toddlers showed significant stronger preference for motherese versus computer “techno” images and sounds than ASD toddlers (d , Cohen's d ; white cross, group mean). b, Toddlers in TD clusters (Clusters 1 and 2) had significantly higher percentage fixation towards motherese

versus computer “techno” images and sounds than ASD toddlers (Clusters 4). The white cross indicates group mean. For percentage fixation towards Motherese (% Motherese) in 5 clusters, see Supplementary Fig. 4. c, The heatmap matrix shows standardized effect sizes (Cohen’s d) for each pairwise group comparison between clusters. Cohen’s d value is shown in each cell and the standard effect size is also indicated by the color of the cell. The asterisk indicates significant results of two-sample t-tests. *p <.05, **p < .005.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementaryinformationfinal.docx](#)
- [supplementaryinformationfinal.docx](#)