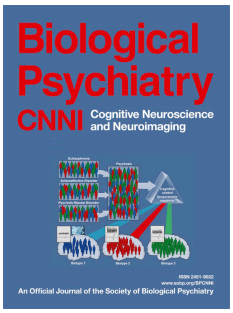


Aberrant emotional prosody circuitry predicts social communication impairments in children with autism

Simon Leipold, Daniel A. Abrams, Shelby Karraker, Jennifer M. Phillips, Vinod Menon

PII: S2451-9022(22)00245-2

DOI: https://doi.org/10.1016/j.bpsc.2022.09.016

Reference: BPSC 1010

To appear in: *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*

Received Date: 13 April 2022

Revised Date: 25 August 2022

Accepted Date: 30 September 2022

Please cite this article as: Leipold S., Abrams D.A., Karraker S., Phillips J.M. & Menon V., Aberrant emotional prosody circuitry predicts social communication impairments in children with autism, *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* (2022), doi: https://doi.org/10.1016/j.bpsc.2022.09.016.

Aberrant emotional prosody circuitry predicts social communication impairments in children with autism

Simon Leipold [1, $], *Daniel A. Abrams* [1, $], *Shelby Karraker* [1], *Jennifer M. Phillips* [1], *Vinod Menon* [1, 2, 3]

[$] These authors contributed equally.

[1] Department of Psychiatry and Behavioral Sciences

[2] Department of Neurology and Neurological Sciences

[3] Stanford Neurosciences Institute

Stanford University

Short title: Emotional prosody circuitry in children with autism

Corresponding authors:

Simon Leipold, Ph.D. and Vinod Menon, Ph.D.
Stanford Cognitive and Systems Neuroscience Laboratory
Department of Psychiatry and Behavioral Sciences
401 Quarry Rd.
Stanford, CA 94305
USA

Emails: simon.leipold@donders.ru.nl; menon@stanford.edu

Abstract

Background: Emotional prosody provides acoustical cues that reflect a communication partner's emotional state and is crucial for successful social interactions. Many children with autism have deficits in recognizing emotions from voices, however the neural basis for these impairments is unknown. Here we examine brain circuit features underlying emotional prosody processing deficits and their relation to clinical symptoms of autism.

Methods: We used an event-related fMRI task to measure neural activity and connectivity during processing of sad and happy emotional prosody and neutral speech in 22 children with autism and 21 matched control children (7-12 years old). We employed functional connectivity analyses to test competing theoretical accounts which attribute emotional prosody impairments to either sensory processing deficits in auditory cortex or theory of mind deficits instantiated in temporoparietal junction (TPJ).

Results: Children with autism showed specific behavioral impairments for recognizing emotions from voices. They also showed aberrant functional connectivity between voice-sensitive auditory cortex and bilateral TPJ during emotional prosody processing. Neural activity in bilateral TPJ during processing of both sad and happy emotional prosody stimuli was associated with social communication impairments in children with autism. In contrast, activity and decoding of emotional prosody in auditory cortex was comparable between autism and control groups and did not predict social communication impairments.

Conclusions: Our findings support a social-cognitive deficit model of autism by identifying a role for TPJ dysfunction during emotional prosody processing. Our study underscores the importance of "tuning in" to vocal-emotional cues for building social connections in children with autism.

Introduction

Understanding the emotional state of a communication partner is a crucial component of meaningful human interactions (1-3). The emotional state of a speaker is provided by emotional prosody and is expressed with distinct patterns in a speaker's tone of voice, intonation, emphasis, rhythm, and speech rate (4-6). Learning to identify these acoustical features and map them to a speaker's mental state is an important aspect of children's social and neurodevelopment (7). Children with autism spectrum disorders (ASD) have pronounced deficits in social interaction and communication (8). Difficulties in the perception and production of prosody have been identified as a prominent component of the autism behavioral phenotype since Kanner's initial description of childhood autism (9-13), and subsequent studies have identified deficits in interpreting emotional prosody cues in ASD (14-24). However, it is unclear why individuals with ASD may have deficits in emotional prosody processing. This question is critical as emotional prosody represents a unique interface between language, emotion, and social function, and understanding why children struggle with prosody perception may provide novel insights into improving verbal communication in affected individuals.

Two hypotheses have emerged to explain the deficits in interpreting vocal-emotional cues in ASD. The first hypothesis proposes that children with ASD have a sensory processing deficit in which abnormal sensory processing of emotional prosody cues precludes access to downstream systems that evaluate and categorize these acoustical features (25, 26). Sensory processing of prosodic stimuli is thought to be performed by auditory cortical brain regions in superior temporal cortex in both adults (5, 27) and children (28). Thus, the sensory deficit model predicts aberrant brain function at the level of auditory cortical regions that subserve acoustical analyses of vocal-emotional cues (29, 30). A second hypothesis posits that acoustical processing of emotional prosody cues is largely intact in individuals with ASD, however impairments in social cognitive and evaluative processes preclude the accurate interpretation of emotional information from these vocal cues. For example, one prominent theory proposes that the core social deficits of ASD are caused by a deficit in theory of mind, the ability to reason about other people's mental states and emotions (31-33). A key brain region that has consistently been implicated in the context of theory of mind processing is bilateral temporoparietal junction (TPJ) (34-38), a central hub of the "social brain network" (39). Thus, the social cognition model would predict that aberrant TPJ function contributes to deficits in inferring mental states based on vocal-emotional cues (40, 41). A significant gap in our knowledge of social communication deficits in children with ASD is that most studies have focused on visual stimuli (42, 43) with relatively few studies conducted in the auditory modality, particularly in children. Understanding the relative contribution of auditory sensory and social cognitive systems to emotional prosody perception deficits in children with ASD

would provide important information regarding the nature of this communication impairment and may provide insight into its remediation (44, 45).

Little is known regarding the brain basis of emotional prosody processing deficits in children with ASD. Examining this question in children is particularly important given that remediation efforts are generally more beneficial in children compared to adults (46). Previous studies in adults with ASD have failed to identify significant group differences in neural measures during emotional prosody processing when contrasted with neutral speech (47-49). In the only controlled study of emotional prosody in children with ASD, results showed that angry compared to neutral speech elicited greater activation in a distributed brain network in children with ASD than in typically developing (TD) children (50). However, this sample of children with ASD had lower cognitive and language abilities than TD children (50), and therefore it is not clear whether the observed neural effects reflect deficits in emotional prosody perception or general cognitive impairments. Thus, it remains unknown whether children with ASD have differential brain activation, connectivity, and decoding of emotional prosody stimuli compared to TD children.

Theoretical models have proposed a link between emotional prosody recognition and social abilities (2, 3). However, it is unknown whether neural processing of emotional prosody in children with ASD is related to individual differences in clinical measures of social communication, a key diagnostic symptom domain of ASD. Identifying sources of heterogeneity in key symptom domains remains an important challenge for autism research that has not been addressed in the emotional prosody literature.

To address these questions, we used event-related fMRI to measure neural responses in children with ASD and closely matched TD children aged 7 to 12 years old while they listened to emotional prosody and neutral speech (**Figure 1**). We had three major goals: First, we investigated whether neural responses and neural decoding of emotional prosody decoding are altered in children with ASD in brain regions serving auditory function, including voice-sensitive auditory cortex. Second, we investigated task-related functional connectivity during emotional prosody processing to examine the integrity of brain circuitry linking voice-sensitive auditory regions with social brain regions, including the TPJ. Third, we assessed whether heterogeneity in neural activation measured during emotional prosody processing is related to individual differences in key clinical domains of social communication and interaction impairments.

Methods and Materials

Participants

The Stanford University Institutional Review Board approved the study protocol. Parental consent and children's assent were obtained for all evaluation procedures, and participants were paid for their participation in the study. Our sample included n = 43 children between 7 and 12 years of age (22 children with ASD, 21 TD children). This age range is the youngest that can be practically included in multi-run task-based fMRI studies in ASD. Children in the ASD group were diagnosed based on an classification algorithm (51) that combines information from both the module 3 of the Autism Diagnostic Observation Schedule-Second Edition (ADOS-2; 52) and the Autism Diagnostic Interview-Revised (ADI-R; 53). Children with ASD and TD children underwent extensive neuropsychological testing to assess their cognitive abilities, language function, and emotion recognition abilities. Groups were matched for age, sex, cognitive and language abilities, and motion during functional imaging. Demographic and neuropsychological characteristics are given in **Table 1**. Details are provided in the **Supplementary Material**.

fMRI stimuli

The stimuli presented during fMRI scanning consisted of acoustic sentences spoken in emotional and neutral prosody as well as nonspeech environmental sounds. Emotional and neutral prosodic stimuli were recorded in a recording studio by a professional actress who produced vocal samples of two sentences, “a bag is in the room” (sentence #1) and “my spoon is on the table” (sentence #2), using sad, happy, and neutral emotions (**Figure 1B**). Emotional and neutral prosody stimuli for the fMRI experiment were selected based on results from a behavioral experiment conducted in an independent cohort of 27 school-age TD children who did not participate in the fMRI study and 9 adults (see **Figure 1C** and the **Supplementary Material**). Stimuli can be downloaded from the Open Science Framework (https://dx.doi.org/10.17605/OSF.IO/TYFXS).

fMRI task

Stimuli were presented in 10 separate fMRI runs, each lasting ~3.5 minutes. One run consisted of 39 trials of acoustic sentence stimuli spoken in sad (high and low intensity), happy (high and low intensity), and neutral prosody, as well as environmental sounds and catch trials. Across the 10 runs, a total of 60 exemplars per stimulus condition were presented to each participant. Acoustic stimuli were presented during silent intervals between volume acquisitions to eliminate the effects of scanner noise on auditory perception (**Figure 1D**). Functional images were subjected to preprocessing procedures, which included realignment,

slice timing correction, spatial normalization, and smoothing. Details are provided in the **Supplementary Material**.

Statistical analyses

The statistical analyses had four aims (**Figure 1E**). First, we compared neural decoding of emotional prosody between children with ASD and TD children in anatomically distinct subdivisions of auditory cortex, building on our previous work on emotional prosody decoding in TD children (28). Second, we probed differential functional connectivity of voice-sensitive auditory cortex between children with ASD and TD children. Third, we used voxel-wise general linear model (GLM) analysis to compare neural activation in response to vocal-emotional stimuli between children with ASD and TD children. And fourth, we investigated relationships between neural measures of emotional prosody processing and key clinical measures of social communication function in children with ASD.

Voxel-wise group comparison of fMRI activation

The goal of the voxel-wise analysis of fMRI activation was to identify brain regions that showed differential activity levels in children with ASD and TD children in response to emotional prosody stimuli, neutral prosody, and environmental sounds. A second-level analysis used a two-sample *t*-test to compare activation in children with ASD and TD children on the contrasts of interest [neutral prosody > environmental sounds], [sad prosody > neutral speech], and [happy prosody > neutral speech]. Details are provided in the **Supplementary Material**.

Multivariate and univariate analyses of auditory cortical activation during emotional prosody

To examine decoding and activation during emotional prosody within specific subregions of auditory cortex, regions of interest (ROIs) encompassing bilateral superior temporal auditory areas were constructed. These regions were defined as three ROIs along the anterior-posterior axis of auditory cortex in both the supratemporal plane, including Heschl's gyrus, planum temporale, planum polare, and the superior temporal sulcus (STS), including posterior STS (pSTS), middle STS (mSTS), and anterior STS (aSTS) (28, 54) (see **Supplementary Material).**

An ROI-based multivariate pattern classification method was used to examine whether children with ASD and TD children differentially discriminate emotional and neutral prosody stimuli in auditory cortical areas. Additionally, mean signal levels in response to emotional prosody stimuli were computed for auditory

cortical ROIs and compared between children with ASD and TD children. Details are reported in the **Supplementary Material**.

Modulation of voice-sensitive auditory cortex functional circuitry during emotional prosody

We examined functional connectivity of voice-sensitive auditory cortex during emotional prosody processing using a generalized psychophysiological interaction (gPPI) model (55) with the goal of identifying group differences between children with ASD and TD children in modulations of functional connectivity in response to emotional prosody. A second-level analysis used a two-sample *t*-test to compare functional connectivity of voice-sensitive auditory cortex in children with ASD and TD children on the contrasts of interest [sad prosody > neutral speech] and [happy prosody > neutral speech]. Details are provided in the **Supplementary Materials**.

Association between neural and clinical measures of autism symptom severity

Whole-brain regression analysis was used to examine the relationship between neural activation to emotional prosody and social communication impairments in children with ASD, as quantified using the ADI-R communication and language subscale (53). This ADI-R subscale was chosen based on prior evidence showing associations between social information processing and social communication deficits in adults with ASD (40). Details are provided in the **Supplementary Material**.

Results

Emotional face and prosody recognition accuracy in children with ASD and TD children

We first assessed behavioral emotion recognition abilities in children with ASD and TDs using vocal and facial emotional cues. Emotion recognition was assessed using the Diagnostic Analysis System of Nonverbal Accuracy 2 (DANVA2; 56, 57) (see **Supplementary Material**). Children with ASD showed lower accuracy compared to TD children in recognizing emotions across voices and faces ($F(1,38) = 5.12$, $p = 0.03$) (**Figure 2**). Follow-up analyses revealed that children with ASD had lower accuracies than TD children in recognizing emotions from voices ($t(26.75) = 2.35$, $p = 0.03$), but not faces ($t(32.6) = -1.39$, $p = 0.17$). The group difference in emotion recognition from voices had a moderate to large effect size (Hedges $g = 0.74$). These results suggest selective deficits for recognizing emotions from vocal cues in children with ASD.

Neural activation during emotional prosody in auditory cortex

We next assessed neural activation during emotional prosody processing in children with ASD and TD children within distinct subregions of auditory cortex (28). Results showed no group differences in auditory cortical activation to emotional prosody using either frequentist (all $p > 0.05$) or Bayesian statistical analyses (**Figure S1B and Supplementary Results**). See the **Supplementary Results** for additional analysis of group differences in activation at the whole-brain level.

Multivariate decoding of emotional prosody in auditory cortex

As multivariate decoding provides a more sensitive measure of differential response to stimuli (58), we assessed whether decoding of emotional prosody differed between children with ASD and TD children within distinct subregions of auditory cortex (28, 59). Results from ROI-based multivariate pattern analysis showed striking similarity of emotional prosody decoding for children with ASD and TD children across all investigated auditory cortical ROIs as revealed by both frequentist (all $p > 0.05$) and Bayesian statistical analyses (**Figure S1C** and **Supplementary Results**). Notably, children with ASD showed the same pattern of emotional prosody decoding in auditory cortex as previously reported in TD children (28) with mSTS and pSTS showing the highest classification accuracies for differentiating emotional prosody from neutral speech (**Figure S1C)**. Across both groups, all auditory ROIs decoded emotional prosody from neutral speech above chance level for both sad and happy emotions (all $p < 0.05$).

Functional connectivity of voice-sensitive auditory cortex during emotional prosody processing

We next examined the integrity of voice-sensitive functional circuitry in children with ASD. We used whole-brain gPPI analysis (55) to examine group differences between TD and ASD children with regards to functional connectivity of voice-sensitive auditory cortex during the processing of emotional prosody stimuli.

For the [sad prosody > neutral speech] contrast, group comparisons of functional connectivity revealed a pronounced pattern of hyperconnectivity in children with ASD. Specifically, results revealed greater task-based functional connectivity between voice-sensitive cortex and bilateral TPJ in children with ASD compared to TD children (left TPJ: 1520 mm^3, $p_{FWE} < 0.01$; right TPJ: 1320 mm^3, $p_{FWE} < 0.01$; See **Table S1** for Montreal Neurological Institute [MNI] coordinates of all significant clusters). Children with ASD further revealed hyperconnectivity between voice-sensitive cortex and right fusiform gyrus (944 mm^3, $p_{FWE} < 0.01$), primary visual cortex (V1; 800 mm^3, $p_{FWE} < 0.01$), and the left middle temporal gyrus (792 mm^3, $p_{FWE} < 0.01$). For the [happy prosody > neutral speech] contrast, children with ASD again revealed hyperconnectivity between voice-sensitive cortex and left middle temporal gyrus compared to TD children (1216 mm^3, $p_{FWE} <$

0.01; **Figure 3**). Across both sad and happy prosody, no brain areas showed reduced emotional prosody-related connectivity with voice-sensitive cortex in children with ASD compared to TD children. These results demonstrate hyperconnectivity of voice-sensitive auditory cortex during sad prosody processing in children with ASD.

Association between TPJ activation and clinical assessments of communication and language impairments in children with ASD

Our final goal was to determine whether neural measures of emotional prosody processing are associated with clinical measures of social communication impairments in children with ASD. We performed a whole-brain linear regression on neural activation using participants' social communication impairment scores as a regressor (**Figure 4**). For both [sad prosody > neutral speech] and [happy prosody > neutral speech] contrasts, results from this analysis showed a striking pattern of brain-behavior relations: Children with ASD with more severe social communication impairments, reflected by greater scores on the ADI-R, showed reduced activation during sad prosody in bilateral TPJ (left TPJ: 1432 mm^3, $p_{FWE} < 0.01$; right TPJ: 1128 mm^3, $p_{FWE} < 0.01$). Additional clusters where activation was negatively associated with social communication impairments were found in the right lateral occipital cortex (3280 mm^3, $p_{FWE} < 0.01$) and in prefrontal cortex (2688 mm^3, $p_{FWE} < 0.01$). Similarly, during processing of happy prosody, children with more severe social communication impairments showed reduced activation in bilateral TPJ (left TPJ: 5168 mm^3, $p_{FWE} < 0.01$; right TPJ+: 12896 mm^3, $p_{FWE} < 0.01$). Additional clusters were widely distributed over the cortex and subcortical areas (see **Table S3** for a full list and MNI coordinates of significant clusters).

Discussion

Understanding a speaker's emotional state is critical for navigating the social world and many children with ASD have deficits in interpreting the vocal-acoustical cues that signal a speaker's emotions, known as emotional prosody. In contrast to studies using emotional faces, investigations of the neural processing of emotional prosody in individuals with ASD, particularly in children, have been few and limited in scope (50). Results from the current study revealed behavioral deficits for recognizing vocal emotions in children with ASD compared to TD children who were matched for sex, age, cognitive abilities, and language function. Behavioral deficits in recognizing emotions from voices in children with ASD were accompanied by aberrant functional coupling between voice-sensitive auditory areas and bilateral TPJ, a brain region associated with theory of mind (34) and creation of a social context for behavior (60). Crucially, neural activation during emotional prosody processing in bilateral TPJ predicted the severity of social communication impairments in children with ASD, with reduced brain activity in more severely affected children. In contrast, children with ASD showed accurate neural decoding of emotional prosody stimuli across auditory cortex subdivisions, and decoding accuracies were comparable to decoding accuracies measured in TD children (28). Converging results from brain connectivity and social communication brain-behavior analyses support the hypothesis that TPJ dysfunction in ASD is not only associated with deficits in explicit mentalizing and social cognitive tasks, most of which have involved visual stimuli, but also extends to fine-grained analysis of the acoustical features that signal the emotional state of a speaker.

Aberrant functional coupling between voice-sensitive auditory cortex and the TPJ during emotional prosody processing in children with ASD

Results from whole-brain functional circuit analyses revealed hyperconnectivity of voice-sensitive auditory cortex with bilateral TPJ in children with ASD while processing sad prosody compared to neutral vocal stimuli. This result has several important implications for understanding the neurobiological basis of social information processing in ASD. First, previous studies in individuals with ASD that have highlighted TPJ dysfunction have focused primarily on visual processing tasks that examine complex aspects of social cognition (40, 61, 62). These previous findings suggest that TPJ dysfunction in ASD is associated with active inferential cognitive processes associated with explicit theory of mind tasks. Results from the current study suggest that TPJ circuit dysfunction extends beyond these explicit mentalizing tasks and impact the passive processing of sensory-level cues that signal mental states in a communication partner. Therefore, our findings suggest a broader role for TPJ circuitry in social impairments in ASD compared to previous studies.

Our findings also provide important new information regarding voice processing circuitry in children with ASD more generally. An influential model of voice processing circuitry in the brain posits that the STS serves as a hub for the voice processing network (63), and empirical evidence has supported many of the predictions of this model in TD individuals (54, 64-66). Importantly, investigations into the brain circuitry underlying voice processing in children with ASD have shown that aberrant STS connectivity with other key brain systems appears to play a central role in voice-related processing deficits in these individuals (67, 68). Results from the current study add to this literature by showing that aberrant functional connectivity, specifically hyperconnectivity, between voice-sensitive STS and the TPJ, a core node of the social brain, is a neural signature underlying impaired processing of vocal emotions in children with ASD. Findings support a model of voice processing in ASD that places dysfunction of voice-sensitive STS circuitry at the center of multiple aspects of voice perception deficits in affected individuals.

TPJ activation is related to social communication symptom severity in children with ASD

Our whole-brain voxel-wise regression analysis revealed a striking association between activation within bilateral TPJ during emotional prosody perception and symptom severity for social communication and social interaction impairments in children with ASD. Specifically, reduced activation in the TPJ during emotional prosody processing was correlated with more severe social communication deficits in these children.

Results are consistent with previous studies in adults (40) and adolescents with ASD (41) that revealed a similar relationship between reduced activation of the TPJ during social cognitive theory of mind tasks and symptom severity for social communication deficits. Findings from the current study add to this literature by showing that, in contrast with the active inferential processes highlighted in previous studies (40, 41), the link between TPJ activity and autism clinical symptom severity extends to social stimuli presented through the auditory modality and passive listening tasks in children with ASD. Importantly, results showing a relation between bilateral TPJ activation and symptom severity for social communication provide complementary evidence for TPJ dysfunction in children with ASD relative to our finding of group-level STS-TPJ hyperconnectivity during emotional prosody processing (**Figure 3**). Together, results provide converging evidence that upregulation of TPJ activity and normalization of TPJ functional circuitry may serve as neurobiological targets for remediation of social communication deficits in individuals with ASD (69).

Implications for sensory and social cognitive theories of ASD

Processing of the subtle acoustical cues in speech that signal a speaker's emotions may be viewed as a crucial nexus of sensory and social cognitive processing that requires the mapping of brief and fine-grained pitch and timbral features onto representations of distinct emotional state information that is essential for successful and meaningful human interaction. Importantly, autism is closely associated with both auditory sensory and social cognitive deficits, and therefore the current study provides an opportunity to adjudicate models of autism that support primary contributions of sensory and social cognitive deficits in ASD. Results from the current study showing aberrant functional coupling and activity patterns in the TPJ during emotional prosody processing are consistent with social cognitive accounts of ASD, which have consistently implicated the TPJ as a key brain region underlying deficits for theory of mind-related processes (40, 42, 61). Importantly, results from our decoding and activation analyses showed strikingly similar sensory neural function in ASD and TD groups across a wide range of auditory cortical regions in superior temporal cortex during emotional prosody perception. Our findings are consistent with prior evidence examining auditory social information processing in children and adolescents with ASD (50, 67), and in adults with ASD (47, 48), which identified only subtle differences, or no differences at all, in neural function at the level of auditory cortex when compared to TD individuals. Our findings do not support a sensory deficit model of autism (25, 26) in the context of emotional prosody processing and suggest that auditory sensory processing of vocal-emotional cues is not a primary deficit associated with impaired emotional prosody perception. Rather, findings suggest a crucial role for the TPJ in social cognitive impairments in ASD which encompass sensory elements that contribute to social communication and cognition.

Task-based neuroimaging in clinical pediatric samples

Reproducibility in neuroimaging research represents a challenge for the study of clinical pediatric populations whose data is more difficult to acquire compared to adult participants (70). While the sample size used in the current study (N=43) is modest in comparison to recent task-based brain imaging studies of adult populations, an important consideration is that deep sampling in individual participants improves the precision of the measurements and consequently increases replicability (71). Therefore, a primary goal for our data acquisition was to obtain a large amount (7-10 runs; 60 stimuli/condition) of high-quality brain imaging data per participant to enable robust within-participant sampling.

To keep fMRI runs short and maximize participant engagement, we employed a simple implicit fMRI task with two basic emotions. However, an important future direction for this work is to use prosodic processing tasks with greater complexity to investigate different emotions and processing modes in children with ASD (49).

287 ***Conclusions***

288 We identified neural alterations underlying impaired emotional prosody processing in school-age children
289 with ASD, which were predictive of pervasive social communication and interaction difficulties. Children
290 with ASD showed atypical modulation of functional connectivity between voice-sensitive auditory cortex
291 and bilateral TPJ, which represents a hub of the social brain network. Reduced activation of bilateral TPJ
292 during emotional prosody processing was associated with greater social communication and interaction
293 impairments. These findings support social-cognitive accounts of social deficits in ASD and highlight the
294 importance of "tuning in" to vocal-emotional cues for building social connections in children with autism.

Acknowledgments

We thank all the children and their parents who participated in our study and the staff at the Stanford Lucas Center for Imaging for assistance with data collection. We thank Dawlat El-Said and Carlo de los Angeles for assistance with data analysis and K. D'Arcey for help with stimulus production. This work was supported by National Institutes of Health Grants: K01MH102428 (to D.A.A.), R21DC017950 and R21DC017950-S1 (to D.A.A. and V.M.), and R01MH084164 (to V.M.); a NARSAD Young Investigator Grant from the Brain and Behavior Research Foundation (to D.A.A.); the Singer Foundation; the Simons Foundation/SFARI: 308939 (to V.M.); and the Swiss National Science Foundation: P2ZHP1_187704 (to S.L.).

Financial Disclosures

All authors report no biomedical financial interests or potential conflicts of interest.

References

1. Pell MD, Kotz SA (2021): Comment: The Next Frontier: Prosody Research Gets Interpersonal. *Emotion Review*. 13:51-56.

2. Keltner D, Haidt J (1999): Social Functions of Emotions at Four Levels of Analysis. *Cognition & Emotion*. 13:505-521.

3. Van Kleef GA (2009): How Emotions Regulate Social Life: The Emotions as Social Information (EASI) Model. *Current Directions in Psychological Science*. 18:184-188.

4. Hammerschmidt K, Jürgens U (2007): Acoustical Correlates of Affective Prosody. *Journal of Voice*. 21:531-540.

5. Schirmer A, Kotz SA (2006): Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*. 10:24-30.

6. Banse R, Scherer KR (1996): Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*. 70:614-636.

7. Morningstar M, Nelson EE, Dirks MA (2018): Maturation of vocal emotion recognition: Insights from the developmental and neuroimaging literature. *Neuroscience & Biobehavioral Reviews*. 90:221-230.

8. American Psychiatric Association (2013): Neurodevelopmental Disorders. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*: American Psychiatric Publishing.

9. Kanner L (1943): Autistic disturbances of affective contact. *Nervous Child*. 2:217-250.

10. Baltaxe CAM, Simmons JQ (1985): Prosodic Development in Normal and Autistic Children. In: Schopler E, Mesibov GB, editors. *Communication Problems in Autism*. Boston, MA: Springer US, pp 95-125.

11. McCann J, Peppé S (2003): Prosody in autism spectrum disorders: a critical review. *International Journal of Language & Communication Disorders*. 38:325-350.

12. Shriberg LD, Paul R, McSweeny JL, Klin A, Cohen DJ, Volkmar FR (2001): Speech and Prosody Characteristics of Adolescents and Adults With High-Functioning Autism and Asperger Syndrome. *Journal of Speech, Language, and Hearing Research*. 44:1097-1115.

13. Paul R, Augustyn A, Klin A, Volkmar FR (2005): Perception and Production of Prosody by Speakers with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*. 35:205-220.

14. Mazefsky CA, Oswald DP (2007): Emotion Perception in Asperger's Syndrome and High-functioning Autism: The Importance of Diagnostic Criteria and Cue Intensity. *Journal of Autism and Developmental Disorders*. 37:1086-1095.

15. Rutherford MD, Baron-Cohen S, Wheelwright S (2002): Reading the Mind in the Voice: A Study with Normal Adults and Adults with Asperger Syndrome and High Functioning Autism. *Journal of Autism and Developmental Disorders*. 32:189-194.

16. Golan O, Baron-Cohen S, Hill JJ, Rutherford MD (2007): The 'Reading the Mind in the Voice' Test-Revised: A Study of Complex Emotion Recognition in Adults with and Without Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders*. 37:1096-1106.

17. Lindner JL, Rosén LA (2006): Decoding of Emotion through Facial Expression, Prosody and Verbal Content in Children and Adolescents with Asperger's Syndrome. *Journal of Autism and Developmental Disorders*. 36:769-777.

18. Oerlemans AM, Droste K, van Steijn DJ, de Sonneville LMJ, Buitelaar JK, Rommelse NNJ (2013): Co-segregation of Social Cognition, Executive Function and Local Processing Style in Children with ASD, their Siblings and Normal Controls. *Journal of Autism and Developmental Disorders*. 43:2764-2778.

19. Taylor LJ, Maybery MT, Grayndler L, Whitehouse AJO (2015): Evidence for shared deficits in identifying emotions from faces and from voices in autism spectrum disorders and specific language impairment. *International Journal of Language & Communication Disorders*. 50:452-466.

20. Oerlemans AM, van der Meer JMJ, van Steijn DJ, de Ruiter SW, de Bruijn YGE, de Sonneville LMJ, et al. (2014): Recognition of facial emotion and affective prosody in children with ASD (+ADHD) and their unaffected siblings. *Eur Child Adolesc Psychiatry*. 23:257-271.

21. Fridenson-Hayo S, Berggren S, Lassalle A, Tal S, Pigat D, Bölte S, et al. (2016): Basic and complex emotion recognition in children with autism: cross-cultural findings. *Molecular Autism*. 7:52.

22. McCann J, Peppé S, Gibbon FE, O'Hare A, Rutherford M (2007): Prosody and its relationship to language in school-aged children with high-functioning autism. *International Journal of Language & Communication Disorders*. 42:682-702.

23. Peppé S, McCann J, Gibbon F, O'Hare A, Rutherford M (2007): Receptive and Expressive Prosodic Ability in Children With High-Functioning Autism. *Journal of Speech, Language, and Hearing Research*. 50:1015-1028.

24. Scheerer NE, Shafai F, Stevenson RA, Iarocci G (2020): Affective Prosody Perception and the Relation to Social Competence in Autistic and Typically Developing Children. *J Abnorm Child Psychol*. 48:965-975.

25. Robertson CE, Baron-Cohen S (2017): Sensory perception in autism. *Nature Reviews Neuroscience*. 18:671-684.

26. Baum SH, Stevenson RA, Wallace MT (2015): Behavioral, perceptual, and neural alterations in sensory and multisensory function in autism spectrum disorder. *Progress in Neurobiology*. 134:140-160.

27. Brück C, Kreifelts B, Wildgruber D (2011): Emotional voices in context: A neurobiological model of multimodal affective information processing. *Physics of Life Reviews*. 8:383-403.

28. Leipold S, Abrams DA, Karraker S, Menon V (2022): Neural decoding of emotional prosody in voice-sensitive auditory cortex predicts social communication abilities in children. *Cerebral Cortex*.

29. Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B (2000): Voice-selective areas in human auditory cortex. *Nature*. 403:309-312.

30. Gervais H, Belin P, Boddaert N, Leboyer M, Coez A, Sfaello I, et al. (2004): Abnormal cortical voice processing in autism. *Nature Neuroscience*. 7:801-801.

31. Baron-Cohen S, Leslie AM, Frith U (1985): Does the autistic child have a "theory of mind" ? *Cognition*. 21:37-46.

32. Happé FGE (1993): Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*. 48:101-119.

33. Premack D, Woodruff G (1978): Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*. 1:515-526.

34. Saxe R, Kanwisher N (2003): People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind". *NeuroImage*. 19:1835-1842.

35. Saxe R, Powell LJ (2006): It's the Thought That Counts: Specific Brain Regions for One Component of Theory of Mind. *Psychological Science*. 17:692-699.

36. Schurz M, Radua J, Aichhorn M, Richlan F, Perner J (2014): Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews*. 42:9-34.

37. Samson D, Apperly IA, Chiavarino C, Humphreys GW (2004): Left temporoparietal junction is necessary for representing someone else's belief. *Nature Neuroscience*. 7:499-500.

38. Mars RB, Sallet J, Schuffelgen U, Jbabdi S, Toni I, Rushworth MFS (2012): Connectivity-Based Subdivisions of the Human Right "Temporoparietal Junction Area": Evidence for Different Areas Participating in Different Cortical Networks. *Cerebral Cortex*. 22:1894-1903.

39. Adolphs R (2009): The Social Brain: Neural Basis of Social Knowledge. *Annual Review of Psychology*. 60:693-716.

40. Lombardo MV, Chakrabarti B, Bullmore ET, Baron-Cohen S (2011): Specialization of right temporo-parietal junction for mentalizing and its relation to social impairments in autism. *NeuroImage*. 56:1832-1838.

41. O'Nions E, Sebastian CL, McCrory E, Chantiluke K, Happé F, Viding E (2014): Neural bases of Theory of Mind in children with autism spectrum disorders and children with conduct problems and callous-unemotional traits. *Developmental Science*. 17:786-796.

42. Pelphrey KA, Morris JP, McCarthy G, LaBar KS (2007): Perception of dynamic changes in facial affect and identity in autism. *Social Cognitive and Affective Neuroscience*. 2:140-149.

43. Pelphrey KA, Shultz S, Hudac CM, Wyk BCV (2011): Research Review: Constraining heterogeneity: the social brain and its development in autism spectrum disorder. *Journal of Child Psychology and Psychiatry*. 52:631-644.

44. Gengoux GW, Abrams DA, Schuck R, Millan ME, Libove R, Ardel CM, et al. (2019): A Pivotal Response Treatment Package for Children With Autism Spectrum Disorder: An RCT. *Pediatrics*. 144.

45. Koegel RL, Koegel LK (2006): *Pivotal response treatments for autism: Communication, social, & academic development*. Baltimore, MD, US: Paul H Brookes Publishing.

46. Volkmar FR (2014): Editorial: The Importance of Early Intervention. *Journal of Autism and Developmental Disorders*. 44:2979-2980.

47. Charpentier J, Latinus M, Andersson F, Saby A, Cottier J-P, Bonnet-Brilhault F, et al. (2020): Brain correlates of emotional prosodic change detection in autism spectrum disorder. *NeuroImage: Clinical*. 28:102512.

48. Gebauer L, Skewes J, Hørlyck L, Vuust P (2014): Atypical perception of affective prosody in Autism Spectrum Disorder. *NeuroImage: Clinical*. 6:370-378.

49. Rosenblau G, Kliemann D, Dziobek I, Heekeren HR (2017): Emotional prosody processing in autism spectrum disorder. *Social Cognitive and Affective Neuroscience*. 12:224-239.

50. Eigsti I-M, Schuh J, Mencl E, Schultz RT, Paul R (2012): The neural underpinnings of prosody in autism. *Child Neuropsychology*. 18:600-617.

51. Risi S, Lord C, Gotham K, Corsello C, Chrysler C, Szatmari P, et al. (2006): Combining Information From Multiple Sources in the Diagnosis of Autism Spectrum Disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*. 45:1094-1103.

52. Lord C, Risi S, Lambrecht L, Cook JEH, Leventhal BL, DiLavore PC, et al. (2000): The Autism Diagnostic Observation Schedule—Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *Journal of Autism and Developmental Disorders*. 30:205-223.

53. Lord C, Rutter M, Le Couteur A (1994): Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*. 24:659-685.

54. Abrams DA, Kochalka J, Bhide S, Ryali S, Menon V (2020): Intrinsic functional architecture of the human speech processing network. *Cortex*. 129:41-56.

55. McLaren DG, Ries ML, Xu G, Johnson SC (2012): A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *NeuroImage*. 61:1277-1286.

56. Baum KM, Nowicki S (1998): Perception of Emotion: Measuring Decoding Accuracy of Adult Prosodic Cues Varying in Intensity. *Journal of Nonverbal Behavior*. 22:89-107.

57. Nowicki S, Duke MP (1994): Individual differences in the nonverbal communication of affect: The diagnostic analysis of nonverbal accuracy scale. *Journal of Nonverbal Behavior*. 18:9-35.

58. Haynes JD (2015): A primer on pattern-based approaches to fMRI: Principles, pitfalls, and perspectives. *Neuron*. 87:257-270.

59. Ethofer T, Van De Ville D, Scherer K, Vuilleumier P (2009): Decoding of Emotional Information in Voice-Sensitive Cortices. *Current Biology*. 19:1028-1033.

60. Carter RM, Huettel SA (2013): A nexus model of the temporal–parietal junction. *Trends in Cognitive Sciences*. 17:328-336.

61. Murdaugh DL, Nadendla KD, Kana RK (2014): Differential role of temporoparietal junction and medial prefrontal cortex in causal inference in autism: An independent component analysis. *Neuroscience Letters*. 568:50-55.

62. Kana RK, Libero LE, Hu CP, Deshpande HD, Colburn JS (2014): Functional Brain Networks and White Matter Underlying Theory-of-Mind in Autism. *Social Cognitive and Affective Neuroscience*. 9:98-105.

63. Belin P, Fecteau S, Bédard C (2004): Thinking the voice: neural correlates of voice perception. *Trends in Cognitive Sciences*. 8:129-135.

64. Abrams DA, Chen T, Odriozola P, Cheng KM, Baker AE, Padmanabhan A, et al. (2016): Neural circuits underlying mother's voice perception predict social communication abilities in children. *Proceedings of the National Academy of Sciences of the United States of America*. 113:6295-6300.

65. Steiner F, Bobin M, Frühholz S (2021): Auditory cortical micro-networks show differential connectivity during voice and speech processing in humans. *Communications Biology*. 4:801.

66. Ethofer T, Anders S, Erb M, Herbert C, Wiethoff S, Kissler J, et al. (2006): Cerebral pathways in processing of affective prosody: A dynamic causal modeling study. *NeuroImage*. 30:580-587.

67. Abrams DA, Padmanabhan A, Chen T, Odriozola P, Baker AE, Kochalka J, et al. (2019): Impaired voice processing in reward and salience circuits predicts social communication in children with autism. *eLife*. 8.

68. Abrams DA, Lynch CJ, Cheng KM, Phillips J, Supekar K, Ryali S, et al. (2013): Underconnectivity between voice-selective cortex and reward circuitry in children with autism. *Proceedings of the National Academy of Sciences of the United States of America*. 110:12060-12065.

69. Luckhardt C, Schütz M, Mühlherr A, Mössinger H, Boxhoorn S, Dempfle A, et al. (2021): Phase-IIa randomized, double-blind, sham-controlled, parallel group trial on anodal transcranial direct current stimulation (tDCS) over the left and right tempo-parietal junction in autism spectrum disorder—StimAT: study protocol for a clinical trial. *Trials*. 22:248.

70. Yuan W, Altaye M, Ret J, Schmithorst V, Byars AW, Plante E, et al. (2009): Quantification of head motion in children during various fMRI language tasks. *Human Brain Mapping*. 30:1481-1489.

71. Nee DE (2019): fMRI replicability depends upon sufficient individual-level data. *Communications Biology*. 2:1-4.

72. Vul E, Harris C, Winkielman P, Pashler H (2009): Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on Psychological Science*. 4:274-290.

73. Constantino JN, Gruber CP (2012): *Social Responsiveness Scale: SRS-2*. Western Psychological Services.

Figure Legends

Figure 1. Overview of participants, stimuli, experimental design, analysis strategy, and behavioral data.
(A) Our sample included children with autism spectrum disorders (ASD) and typically developing (TD) children between 7 and 12 years of age. **(B)** fMRI stimuli consisted of acoustic sentences spoken in emotional and neutral prosody. Spectrograms of sentence #2, "my spoon is on the table", spoken in sad prosody (left panel), in neutral prosody (middle panel), and in happy prosody (right panel). **(C)** Stimuli were selected based on results from a behavioral experiment conducted in an independent cohort of 27 school-age TD children and nine adults, who provided ratings on a 5-point scale ("how sad or happy is this voice?"). Stimuli that were consistently rated "1" and "5" were identified as the high intensity "sad" and "happy" stimuli for the fMRI experiment; stimuli rated "2" and "4" were identified as the low intensity "sad" and "happy" stimuli, and the stimulus consistently rated "3" was identified as the "neutral" control stimulus. **(D)** A sparse sampling fMRI scanning protocol with a repetition time (TR) larger than the acquisition time (TA) was used to present acoustic stimuli during silent intervals between volume acquisitions to eliminate the effects of scanner noise on auditory perception. **(E)** Schematic of the analyses employed in the study. (i) Neural decoding of emotional prosody within auditory cortical regions of interest, (ii) neural connectivity during emotional prosody, and (iii) neural activation during emotional prosody were compared between children with ASD and TD children. (iv) Relationships between neural measures and clinical measures of social function were assessed within the group of children with ASD. **(F)** ASD (green) and TD (orange) groups were closely matched regarding measures of intelligence, as quantified by the Wechsler Abbreviated Scale of Intelligence (WASI), language abilities measured with the Wechsler Individual Achievement Test, Second Edition (WIAT-II), and scanner motion.

Figure 2. Reduced emotion recognition accuracy in children with ASD.

Selective impairments in recognizing emotion for voices on the Diagnostic Analysis System of Nonverbal Accuracy, Second Edition (DANVA2), a standardized emotion recognition test.

* = $p < 0.05$.

Figure 3. Functional connectivity of voice-sensitive auditory cortex.

Functional connectivity during emotional prosody processing in children with autism spectrum disorders (ASD) and typically developing (TD) children was examined using a generalized psychophysiological interaction (gPPI) model. **(A)** The gPPI analysis was seeded in bilateral voice-sensitive auditory cortex. **(B)** gPPI connectivity was calculated from voice-sensitive auditory cortex to every gray-matter voxel in the brain. **(C)** Group comparisons of emotional prosody-related connectivity revealed overconnectivity of voice-sensitive cortex to bilateral temporoparietal junction (TPJ) in children with ASD compared to TD children during sad prosody. For sad prosody, voice-sensitive cortex was also functionally hyperconnected with left middle temporal gyrus (MTG), primary visual cortex (V1), and right fusiform gyrus in children with ASD. For happy prosody, voice-sensitive cortex was functionally hyperconnected with left MTG in children with ASD. *** = $p < 0.001$.

Figure 4. Neural activation during emotional prosody and its relationship to social communication impairments in children with ASD.

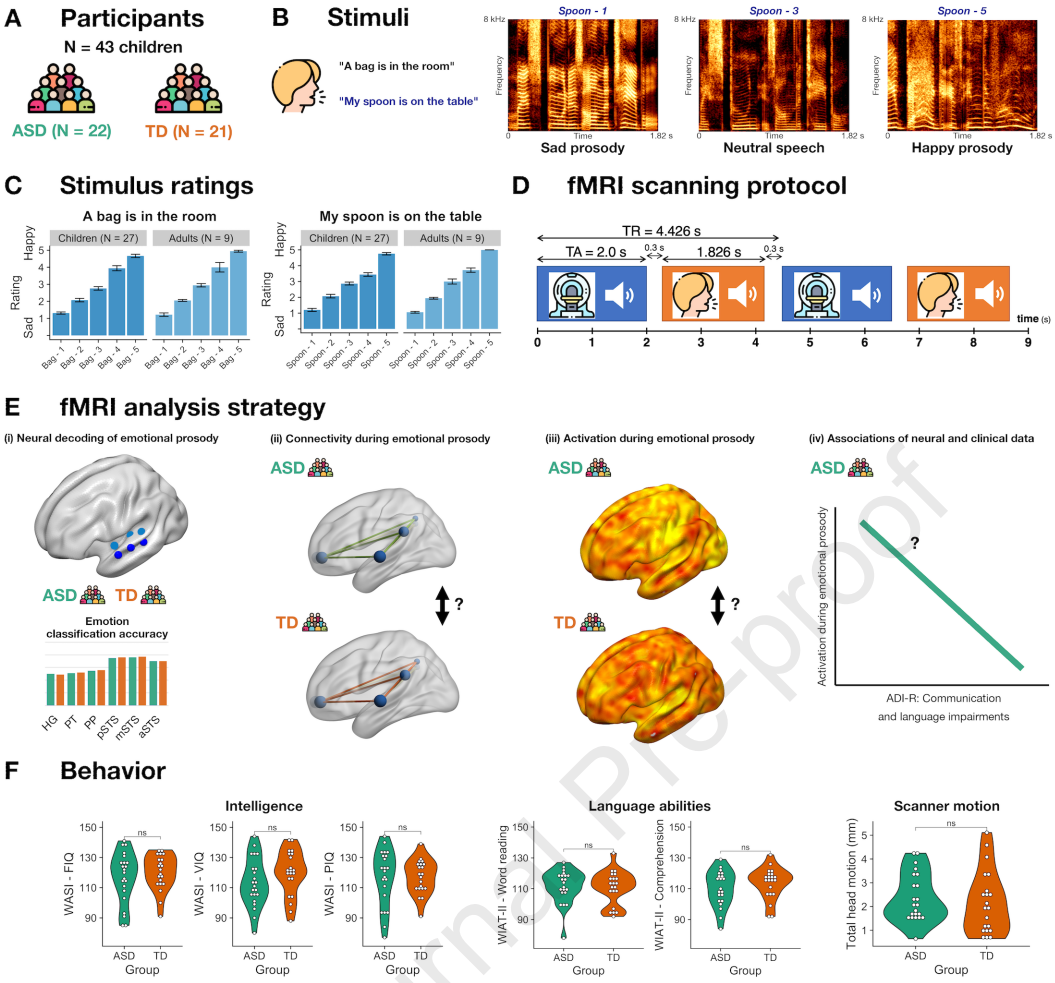
(A) During processing of sad prosody, children with ASD who had lower social communication impairments showed greater activation in both left and right temporoparietal junction (TPJ). **(B)** During processing of happy prosody, children with lower social communication impairments showed greater activation in left and right TPJ. Additional smaller clusters were widely distributed over the cortex and subcortical areas. Significant correlations are inherent to all scatterplots because they are based on results from the whole-brain general linear model analysis (72); however, the results provide important information regarding the distributions and covariation of activity strength in response to prosodic stimuli and social communication scores. See **Table S3** for a full list and Montreal Neurological Institute (MNI) coordinates of significant clusters.

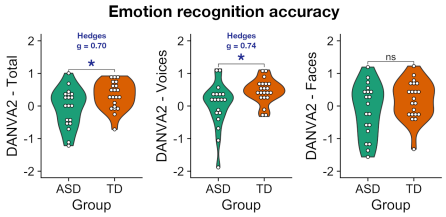
Table 1. Participants' demographic, neuropsychological, and clinical characteristics.

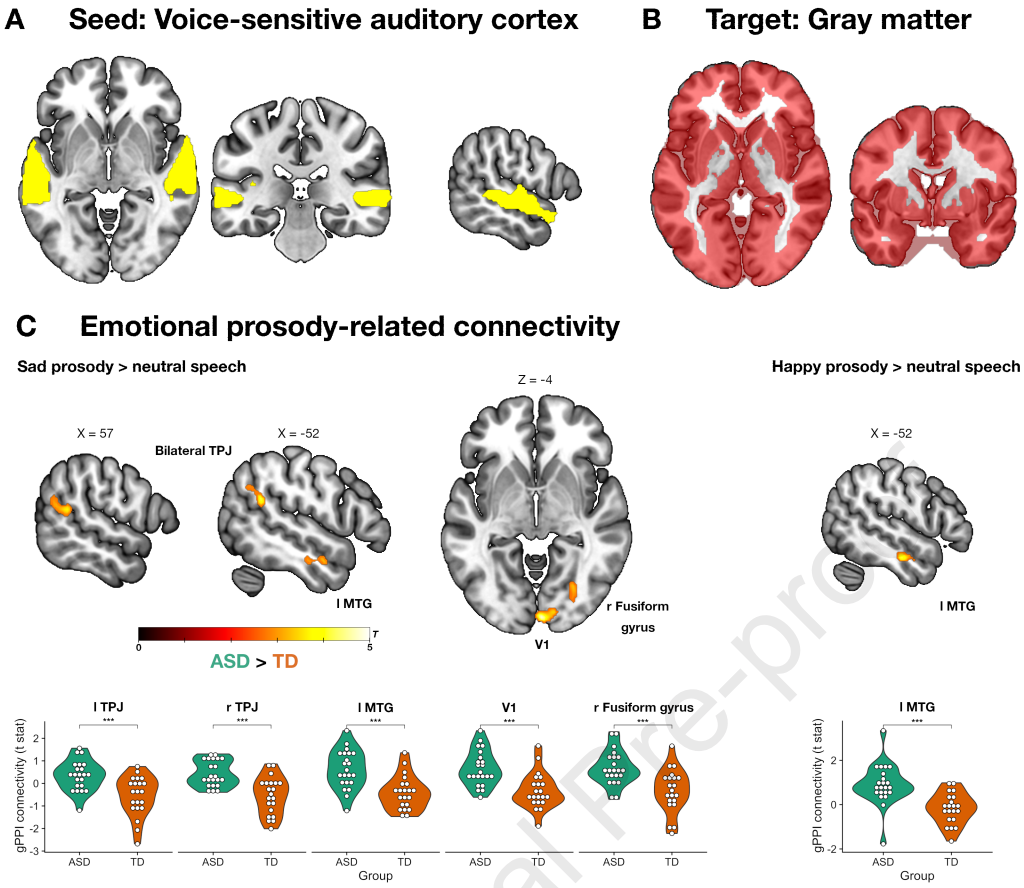
Continuous measures are given as mean ± standard deviation. *P* values derived from two-sample Welch's *t*-test for continuous variables and derived from chi-squared test for categorical variables.

Characteristic	ASD	TD	*P*
Number of participants	22	21	
Sex (female / male)	6 / 16	8 / 13	0.67
Age (years)	10.73 ± 1.66	10.71 ± 1.38	0.97
WASI: Full-scale IQ	116.91 ± 16.87	120.29 ± 11.68	0.45
WASI: Verbal IQ	113.55 ± 16.11	119.71 ± 14.94	0.20
WASI: Performance IQ	116.86 ± 18.65	117.14 ± 11.19	0.95
WIAT-II: Word reading	111.57 ± 10.95	110.76 ± 10.81	0.81
WIAT-II: Reading comprehension	110.24 ± 11.48	113.48 ± 10.16	0.34
DANVA2: Receptive tests *z*-scored accuracy *	-0.05 ± 0.60	0.32 ± 0.43	**0.03**
DANVA2: Faces *z*-scored accuracy $	-0.13 ± 0.81	0.19 ± 0.59	0.17
DANVA2: Voices *z*-scored accuracy +	0.03 ± 0.70	0.45 ± 0.37	**0.03**
SRS-2: Total standard *t*-score	82.05 ± 15.31	47.19 ± 8.08	**<0.001**
ADI-R: Communication and language	11.32 ± 6.33		
ADI-R: Reciprocal social interaction	10.77 ± 6.70		
ADI-R: Restricted and repetitive behaviors	4.82 ± 2.91		
ADOS-2: Social affect	9.95 ± 3.12		
fMRI total head motion (mm) %	2.37 ± 1.02	2.18 ± 1.31	0.59
fMRI mean volume-to-volume motion (mm)	0.15 ± 0.07	0.14 ± 0.09	0.72
fMRI max. volume-to-volume motion (mm)	0.93 ± 0.48	0.88 ± 0.58	0.76

* *z*-scored emotion recognition accuracy averaged across four subtests: adult facial expressions, adult paralanguage, child facial expressions, and child paralanguage. $ *z*-scored emotion recognition accuracy averaged across adult and child facial expressions. + *z*-scored emotion recognition accuracy averaged across adult and child paralanguage. % sum of six motion parameter ranges estimated during realignment. Abbreviations: ADI-R = Autism Diagnostic Interview-Revised (53); ADOS-2 = Autism Diagnostic Observation Schedule-Second Edition (52); DANVA2 = Diagnostic Analysis System of Nonverbal Accuracy, Second Edition (56); IQ = intelligence quotient; max = maximum; SRS-2 = Social Responsiveness Scale, Second Edition (73); WASI = Wechsler Abbreviated Scale of Intelligence; WIAT-II = Wechsler Individual Achievement Test, Second Edition.







Relation between TPJ activation and social communication impairments