

Emotional prosody processing in autism spectrum disorder

Gabriela Rosenblau,^{1,2,3} Dorit Kliemann,^{1,2,4,5} Isabel Dziobek,^{1,6} and Hauke R. Heekeren^{1,2,7}

¹Cluster of Excellence ‘Languages of Emotion’, Freie Universität Berlin, Berlin 14195, Germany, ²Department of Education and Psychology, Freie Universität Berlin, Berlin 14195, Germany, ³Yale Child Study Center, Yale University, 230 S. Frontage Road, New Haven, CT 06519, USA, ⁴McGovern Institute for Brain Research, Massachusetts Institute of Technology, 43 Vassar Street, Cambridge, MA 02139, USA, ⁵Department of Neurology, Massachusetts General Hospital/Harvard Medical School, 149 Thirteenth Street, Charlestown, MA 02129, USA, ⁶Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Unter den Linden 6, Berlin 10099, Germany and ⁷Dahlem Institute for Neuroimaging of Emotion, Freie Universität, Berlin, Germany

Correspondence should be addressed to Gabriela Rosenblau, Center for Translational Developmental Neuroscience, Yale Child Study Center, Yale University, 230 S. Frontage Rd., New Haven, CT 06519, USA. E-mail: gabriela.rosenblau@yale.edu

Abstract

Individuals with Autism Spectrum Disorder (ASD) are characterized by severe deficits in social communication, whereby the nature of their impairments in emotional prosody processing have yet to be specified. Here, we investigated emotional prosody processing in individuals with ASD and controls with novel, lifelike behavioral and neuroimaging paradigms. Compared to controls, individuals with ASD showed reduced emotional prosody recognition accuracy on a behavioral task. On the neural level, individuals with ASD displayed reduced activity of the STS, insula and amygdala for complex vs basic emotions compared to controls. Moreover, the coupling between the STS and amygdala for complex vs basic emotions was reduced in the ASD group. Finally, groups differed with respect to the relationship between brain activity and behavioral performance. Brain activity during emotional prosody processing was more strongly related to prosody recognition accuracy in ASD participants. In contrast, the coupling between STS and anterior cingulate cortex (ACC) activity predicted behavioral task performance more strongly in the control group. These results provide evidence for aberrant emotional prosody processing of individuals with ASD. They suggest that the differences in the relationship between the neural and behavioral level of individuals with ASD may account for their observed deficits in social communication.

Key words: emotion; prosody; behavioral assessment; fMRI; autism

Introduction

Noticing a person’s negative undertone to a seemingly neutral comment is crucial for choosing an adequate response. Emotional prosody, i.e. tone of voice, conveys important information about the speaker’s communicative intention and is processed mainly implicitly (i.e. in the absence of explicit verbal cues) (Wildgruber et al., 2006). In contrast to basic emotions

(e.g. happy, angry) that involve universal, highly stereotypical physiological reactions (Ekman and Friesen, 1971; Ekman, 1992; Zinck and Newen, 2008), understanding complex emotions (e.g. gratitude or jealousy) requires successful decoding and integration of contextual, social information (Zinck and Newen, 2008).

How do humans extract emotional meaning from prosody? Across various tasks, emotional prosody processing has been shown to involve activity of the right superior temporal sulcus

Received: 2 November 2015; Revised: 26 July 2016; Accepted: 12 August 2016

© The Author (2016). Published by Oxford University Press. For Permissions, please email: journals.permissions@oup.com

(STS) and the bilateral inferior frontal gyrus (IFG) (Schirmer and Kotz, 2006; Wildgruber et al., 2006). A current prosody processing model proposes that the right STS is involved in extracting acoustic information, which is subsequently evaluated within the bilateral IFG (Ethofer et al., 2006; Wildgruber et al., 2006). IFG, amygdala and the ventral striatum are also involved in processing the emotional salience of auditory stimuli (Schirmer and Kotz, 2006). It is, however, an open question how the interplay between these regions differentiates intact from impaired emotional prosody processing.

Autism Spectrum Disorder (ASD) has been associated with both impairments in emotional prosody production and processing (Tager-Flusberg, 1981; Baltaxe and D'Angiola, 1992; McCann and Peppe, 2003). However, empirical research investigating prosody processing in autism produced mixed results. Some studies reported aberrant prosody processing of basic and complex emotions in individuals with ASD compared to controls (Hobson et al., 1988; Baron-Cohen et al., 1993; Loveland et al., 1995; Deruelle et al., 2004; Golan et al., 2007; Kuchinke et al., 2011), whereas other studies did not find such group differences (Loveland et al., 1997; Boucher et al., 2000; Chevallier et al., 2011). These inconsistencies likely reflect substantial differences in methodology between studies (McCann and Peppe, 2003). Studies investigating emotional prosody processing with abstract, non-word stimuli (Brennand et al., 2011), a limited number of mostly basic emotions (Boucher et al., 2000; Paul et al., 2005), including one or two speakers and two answer options (Chevallier et al., 2011), may lack the sensitivity to detect subtle impairments in prosody processing of high-functioning individuals with ASD.

Furthermore, neural processing of emotional prosody in ASD has remained an under-researched topic with inconclusive results. There is the notion that individuals with ASD show increased and more widely spread neural activity during prosody processing compared to controls (Wang et al., 2006; Eigsti et al., 2012; Gebauer et al., 2014). With respect to the visual domain, research has shown that emotion recognition impairments of individuals with ASD are linked to dysfunctional activity of the social perception system including the amygdala, the posterior STS and the fusiform gyrus (Baron-Cohen et al., 1999a; Castelli et al., 2002; Pelphrey et al., 2011; Kliemann et al., 2012; Rosenblau et al., 2016).

The primary aim of this study was to corroborate previous reports of aberrant emotional prosody processing in individuals with ASD. Our results may also help to identify how the interplay of brain regions involved in prosody processing relates to prosody recognition performance and thus to intact vs impaired prosody processing. These insights help to further specify models of emotional prosody-processing. Given the striking social deficits of individuals with ASD in naturalistic settings (Dziobek et al., 2006; Rosenblau et al., 2016), we investigated emotional prosody processing with naturalistic behavioral and neuroimaging tasks. Our study overcomes important drawbacks of previous studies: most previous studies included a very limited number of mostly basic emotions, few speakers and abstract stimuli, which may lack sensitivity to detect impairments in prosody processing of high functioning individuals with autism.

We developed behavioral and fMRI tasks, which comprise a variety of complex emotions, speakers, as well as implicit and explicit task conditions. To approximate the communication challenges individuals face in real life, audio stimuli consisted of semantically neutral, short sentences spoken with either emotional or neutral prosody. In accordance with previous

studies (Bach et al., 2008), we assessed implicit emotional prosody processing with a gender discrimination task, asking participants to determine the speaker's gender rather than the emotion conveyed in the spoken sentences, while in the scanner. In the explicit emotional prosody tasks, participants were asked to label the emotion conveyed in the speaker's tone of voice. We expected individuals with ASD to score lower than controls on the explicit behavioral prosody recognition task and their emotion recognition deficit to be reflected in aberrant activity and effective connectivity of core prosody processing regions, such as the STS, IFG and amygdala.

Materials and methods

Procedure

The study consisted of a behavioral and an fMRI experiment (average time interval between the sessions was 18 days (SD = 15 days)). Participants were invited to participate in both, if they met MRI inclusion criteria. The behavioral session took place in testing rooms at Freie Universität Berlin, Germany. Participants completed the behavioral prosody task online through the project's website under the supervision of trained experimenters. The fMRI experiment was scheduled at the DINE (Dahlem Institute for Neuroimaging of Emotion, Freie Universität Berlin, Germany; <http://www.loe.fuberlin.de/dine/index.html>). All participants received payment for participation and gave written informed consent in accordance with the requirements of the German Society for Psychology ethics committee (DGPs).

Participants

Behavioral experiment. Twenty-seven adults with ASD (18 male, mean age = 33, age range: 19–47) and 22 control participants (16 male, mean age = 32, age range: 20–46) with no reported history of psychiatric or neurological disorders were matched according to gender, age and verbal IQ as measured with a vocabulary test [Mehrfachwahl-Wortschatz Test (MWT); Lehrl, 1989; Table 1]. All participants were right handed and had normal or corrected-to-normal vision. ASD participants were recruited through the autism outpatient clinic for adults of the Charité—University Medicine Berlin, Germany or were referred to us by specialized clinicians. ASD participants were diagnosed according to the DSM-IV criteria for Asperger syndrome and autism without intellectual disabilities (American Psychiatric Association, 2000). Diagnoses were confirmed by at least one of the two gold-standard diagnostic instruments: the Autism Diagnostic Observation Schedule (ADOS) (Lord et al., 2002) and the Autism Diagnostic Interview—Revised (ADI-R; (Lord et al., 1994), if parental informants were available ($n = 15$)). For 12 participants, the diagnostic methods included both ADOS and ADI-R. Additionally, the diagnosis of Asperger syndrome was confirmed with the Asperger Syndrome and High-Functioning Autism Diagnostic Interview (ASDI) (Gillberg et al., 2001).

fMRI experiment. Seven of the 27 ASD participants met exclusion criteria for participation in the fMRI experiment (claustrophobia: $N = 2$; no normal or corrected to normal vision $N = 1$, no current health insurance: $N = 1$; psychotropic medication: $N = 3$). Two of the 22 controls chose not to participate in the fMRI experiment (one male and one female), and one female only participated in the fMRI experiment. The fMRI sample thus

Table 1. Demographical and symptom characteristics

	Total sample						P	fMRI sample						
	Controls			ASD				Controls			ASD			
	N	M	SD	N	M	SD		N	M	SD	N	M	SD	P
Sex: N(F/M)	6/16	–	–	9/18			0.760	6/15			6/14			0.595
Age		31.8	8.5		33.1	8.7	0.600		31.9	9.3		31.8	9.3	0.970
MWT-IQ		108.6	13.2		112.9	16.7	0.330		108.3	13.6		113	17.3	0.335
ADOS	–	–	–	24	10.5	3.4	–				19	10.4	3.5	–

Means (M), standard deviations (SD) and sample size (N) of group characteristics. P-values: two-tailed significance-value for F- and χ^2 -tests in ASD vs Controls; Abbreviations: ASD: Autism Spectrum Disorders; F: female; M: male; MWT: Mehrfachwahl-Wortschatz Test; not applicable (–); ADOS: Autism Diagnostic Observation Schedule; fMRI: functional magnetic resonance imaging.

comprised 20 ASD and 21 control participants matched for age, gender and IQ (Table 1). All participants were right-handed.

Tasks and materials

Behavioral prosody task. The newly developed task comprised 25 semantically neutral sentences (e.g. ‘They were all invited to the meeting’) spoken by a total of 16 professional actors [6 male, varying age (20–50 years)]. All sentences (mean length = 5.1 seconds, SD = 0.9) were spoken with emotional prosody. In sum, the task covered four basic (angry, sad, happy, surprised) and 21 complex emotions (interested, frustrated, curious, passionate, contemptuous, furious, confident, proud, desperate, relieved, offended, concerned, troubled, expectant, confused, hurt, bored, in love, enthusiastic, lyrical and shocked). After listening to the audio excerpt, participants were asked to select the correct emotion label out of four different options and drag and drop it into the target panel (see Figure 1A for an example). Distractor labels consisted of (i) two emotions of the same valence, with one resembling the correct option more closely with respect to emotional arousal than the other one and (ii) one emotion of opposite valence (e.g. target emotion: angry, same valence distractors: desperate and embarrassed, opposite valence distractor: enthusiastic). Participants read introduction slides before completing the task (approximate total task duration: 15 min). Throughout the entire task, participants used the mouse to navigate through introduction screens and solve the 25 task items. There was no time limit to solve each item, but participants were instructed to perform as fast and as accurately as possible. No trial and thus no target emotion was repeated. Also no feedback was provided about whether the items had been solved correctly or not. Items were presented in randomized order across participants. The prosody task was designed and programmed as a web-based application in cooperation with a digital agency (gosub communications GmbH, www.gosub.de). Please refer to the supplemental section for detailed information about the chosen emotions, stimuli and task validation procedure.

fMRI prosody task. In the block-design fMRI task, participants were presented with semantically neutral sentences (mean length: 2.9 s, SD = 0.01) spoken with emotional or neutral prosody by 10 different actors (5 male). The task was presented using Presentation (Version 14.1, Neurobehavioral Systems Inc., Albany, CA) in two runs of 10 min 34 s each. Participants had to either indicate the speaker’s gender (implicit condition) or the correct emotion label from two options (explicit condition) (Figure 2A). To make a choice, they had to press a button with

either index or middle finger of their right hand. The position of the correct option and distractor on the screen (left or right) were counterbalanced (see example blocks for each condition in Table 2). Each fMRI task block (30s) started with a cue screen (2 s), which indicated the condition (‘gender’ for implicit blocks; ‘emotion’ for explicit blocks). The cue was followed by four audio trials (4 s each), interleaved with four answer screens (3 s each). Note that we simplified the explicit emotion recognition condition by reducing the number of target emotions (6 basic and 6 complex emotions) and answering options relative to the behavioral prosody task (4 basic and 21 complex emotions). Based on the ratings obtained by (Hepach et al., 2011), the six basic emotions (happy, surprised, fearful, sad, disgusted and angry) were matched for valence (Wilcoxon signed-ranks: $P = 0.75$) and arousal (Wilcoxon signed-ranks: $P = 0.92$) with six complex emotions (jealous, grateful, contemptuous, shocked, concerned, disappointed). In all explicit task blocks (neutral, basic and complex emotions) participants were asked to select the correct emotion label from two options. We limited the number of options to two (from the previous 4 in the behavioral task) to reduce task demands and thus possible load-related between group differences in Blood Oxygen Level Dependent (BOLD) signal change. One of the options was the correct emotion label. The other option, the distractor, was randomly chosen from five different emotion labels (4 of the same valence, differing in how much they resembled the valence and arousal of the correct label, and 1 emotion label of opposite valence). Eight blocks contained audios with neutral prosody (4 in the implicit and 4 in the explicit task condition, 32 audio stimuli in total) and 24 blocks contained audios with emotional prosody (12 in the implicit and 12 in the explicit condition, 96 audio stimuli in total). To increase design efficiency, task block should contain similar emotions, which would elicit similar neural responses. Given that several studies report different activation patterns for stimuli of positive vs negative valence (Viinikainen et al., 2012), we presented positive and negative emotions in separate blocks.

Out of the 12 blocks per condition, 4 blocks contained positive emotions (2 blocks basic and 2 blocks complex positive emotions) and 8 blocks contained negative emotions (4 blocks basic and 4 blocks complex negative emotions). Blocks of audios were counterbalanced with respect to the type of emotion and speaker’s gender across runs and conditions. There was no overlap between sentences used in the behavioral and fMRI task. The average duration of audio stimuli in the fMRI task was 2.9 s (SD = 0.75 s, range: 2–4 s). Mean duration of basic and complex emotional prosody audios did not differ [$t(94) = 0.14$; $P = 0.84$].

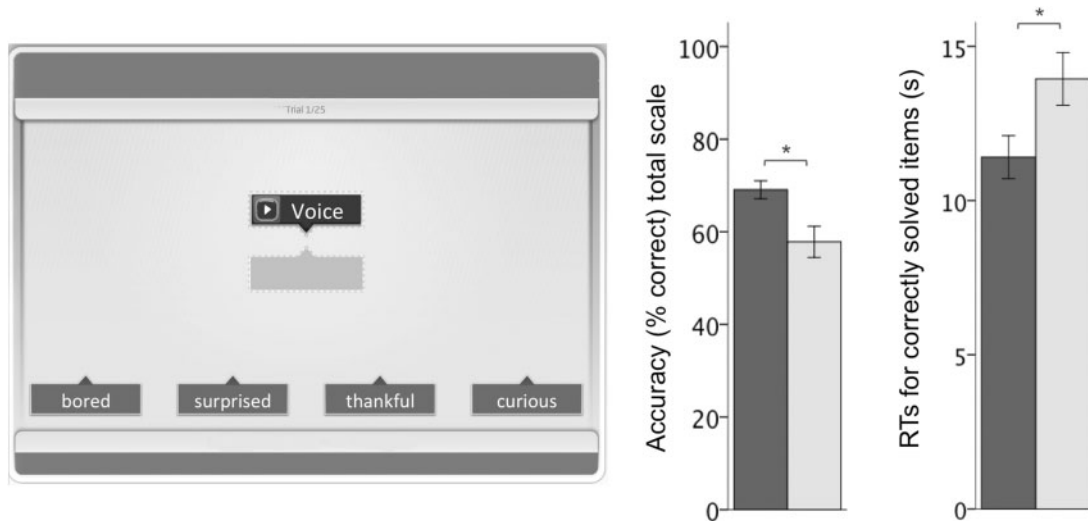


Fig. 1. Behavioral emotional prosody task. (A) Example item. Participants heard semantically neutral sentences that contained emotional prosody and were subsequently asked to label the emotional prosody from four different options. (B) Mean accuracy scores and reaction times for correctly solved items in Controls and ASD participants. Dark and light grey bars illustrate mean task performance of controls and ASD participants, respectively. * Significant difference between controls and ASD groups ($P < 0.05$). ASD: Autism Spectrum Disorders.

FMRI data acquisition

MRI data were acquired on a 3 Tesla scanner (Tim Trio; Siemens, Erlangen, Germany) using a 12-channel head coil. Functional data were acquired using an echo-planar T2*-weighted gradient echo pulse sequence (TR=2000 ms, TE=30 ms, flip angle=70, 64×64 matrix, field of view=192 mm, voxel size= $3 \times 3 \times 3$ mm³). A total of 37 axial slices (3 mm thick, no gap) were sampled for whole-brain coverage. Functional imaging data were acquired in two separate 310-volume runs of 10 min 34 s each. Both runs were preceded by two dummy volumes to allow for T1 equilibration. For each participant, a high-resolution T1-weighted anatomical whole brain scan was acquired in the same scanning session, which was later used for registration of the fMRI data (256×256 matrix, voxel size= $1 \times 1 \times 1$ mm³).

FMRI data analysis

FMRI's Software Library (FSL, version 4.1.8; Oxford Centre of fMRI of the Brain, www.fmrib.ox.ac.uk/fsl) (Smith et al., 2004) was used for fMRI data analysis on the High-Performance Computing system at Freie Universität Berlin (<http://www.zedat.fu-berlin.de/HPC>).

Preprocessing. fMRI data were preprocessed and analyzed using FEAT (FMRI Expert Analysis Tool) within the FSL toolbox. After brain extraction, slice timing, and motion correction, volumes were spatially smoothed using an 8-mm full width at half maximum (FWHM) Gaussian kernel. Low frequency artifacts were subsequently removed with a high-pass temporal filter (Gaussian-weighted straight line fitting, $\sigma = 100$ s). Functional data were first registered to individuals' T1-weighted structural image and then registered to standard space using the FMRI's Linear Image Registration Tool (FLIRT) (Jenkinson and Smith, 2001).

fMRI single-subject analysis. We modeled the time series individually for each participant and run including ten epoch regressors [representing the factor levels for the three factors

emotion complexity (complex and basic prosody), valence (positive, negative and neutral prosody) and condition (implicit and explicit condition)], as well as one regressor for all button presses that occurred during the experiment. Additionally, we included six regressors modeling head movement parameters. There were no differences between groups in the total amount of motion between functional volumes [mean relative displacement: $t(39) = 1.21$, $P = 0.236$; see Supplementary Figure S1 in the supplemental section]. The regressors of interest were then convolved with a Gamma hemodynamic response function (HRF). Contrast images were computed for each condition, run, and participant. They were spatially normalized, transformed into standard space and then submitted to a second-order within-subject fixed-effects analysis across the two runs.

FMRI group analysis. All reported group analyses were higher-level mixed-effects analyses using the FMRI Local Analysis of Mixed Effects tool provided by FSL (FLAME, stage 1 & 2). The models included age and IQ as regressors of no interest. Additionally, we added a gender regressor. Given the growing literature on gender differences in ASD, we explored if any group effects were additionally modulated by gender. We report clusters of maximally activated voxels that survived family wise error (FWE) cluster correction for multiple comparisons at a statistical threshold of $P < 0.05$ and a z-value > 2.3 . Given our *a priori* hypothesis regarding group differences in amygdala activity, we performed separate region of interest (ROI) analyses using an anatomically defined ROI of the bilateral amygdala. These analyses were also corrected for multiple comparisons at a statistical threshold of $P < 0.05$ and a z-value > 2.3 .

Common emotional prosody network. To investigate which regions are involved in emotional prosody processing across groups, we computed a conjunction map of the overlap between activation in the control and ASD group for the contrast emotional vs neutral prosody (Nichols et al., 2005). We additionally report changes in neural activity for emotional vs. neutral prosody separately for each group in Table 3. Subsequently, we

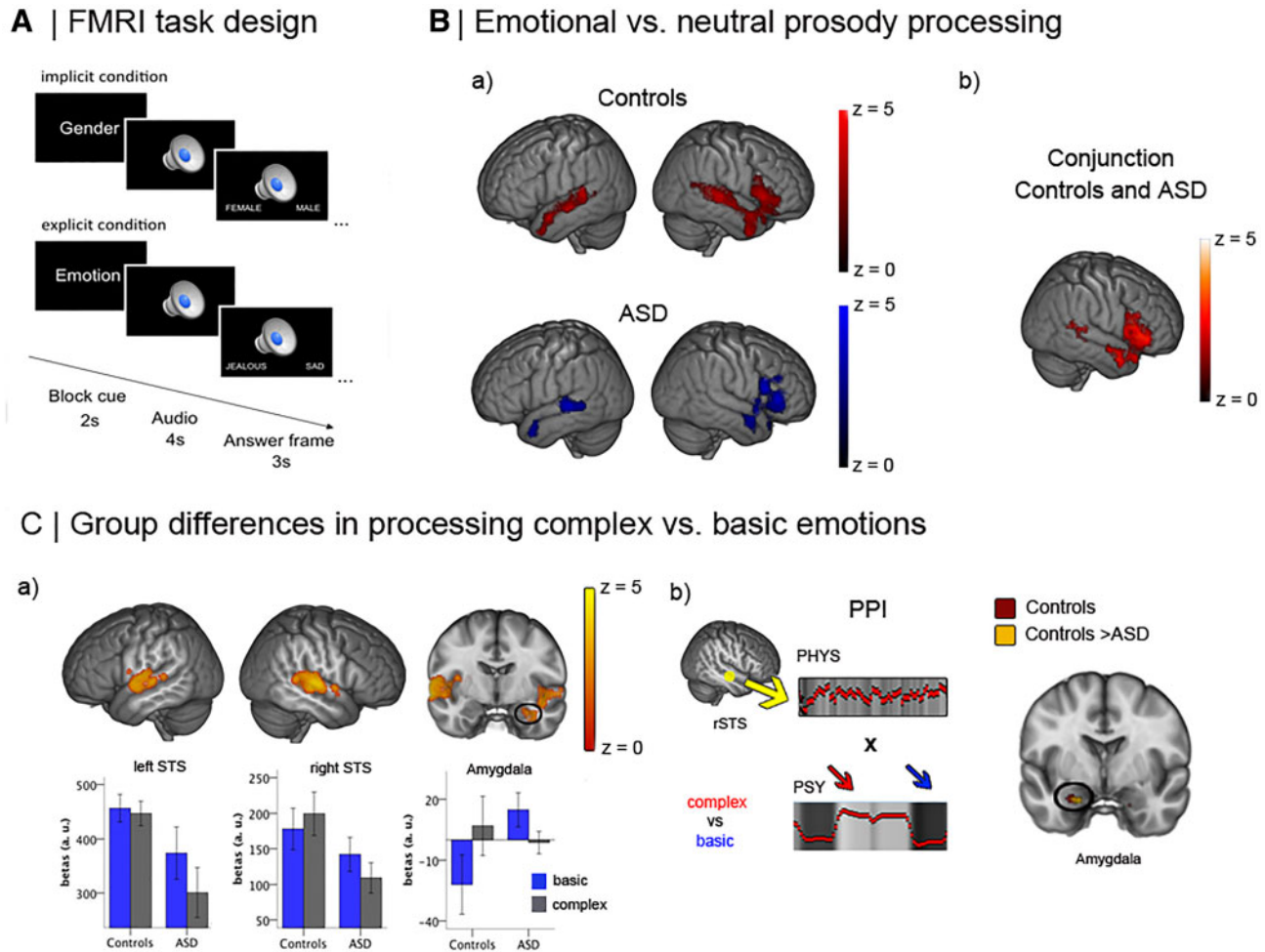


Fig. 2. fMRI emotional prosody task. (A) The task comprised blocks of semantically neutral sentences spoken with (basic or complex) emotional prosody or with neutral prosody. Participants either indicated the speaker’s gender (implicit condition) or the correct emotion label from two options (explicit condition). (B) Brain regions showing significantly greater activation during emotional compared to neutral prosody processing (a) in controls and in ASD participants, (b) in both groups. (C) (a) Brain regions showing significantly greater activation during complex compared to basic emotional prosody processing in controls compared to individuals with ASD. Parameter estimates extracted from the amygdala and STS are illustrated in bar graphs (blue color: basic emotions, grey color: complex emotions). Error bars indicate standard error of mean. All clusters are significant at $P < 0.05$ and $z = 2.3$ family wise error (FWE) cluster corrected for multiple comparisons. (b) Effective connectivity between the right STS (seed region in yellow) and left Amygdala in controls (red) is greater (yellow) than in individuals with ASD. The psychophysiological interaction (PPI) is the interaction between the physiological regressor (PHYS: is the extracted time course from the STS seed region) and the psychological regressor (PSY: complex vs basic emotional prosody). All clusters are significant at $P < 0.05$ and $z = 2.3$ family wise error (FWE) cluster corrected for multiple comparisons. Abbreviations: Autism Spectrum Disorders (ASD); Blood Oxygen Level Dependent signal (BOLD signal); Inferior Frontal Gyrus (IFG); Superior Temporal Sulcus (STS) a.u. = arbitrary units.

Table 2. Example blocks for the explicit emotion recognition condition in the fMRI task

Trial	Basic negative emotions		Complex negative emotions		Neutral	
	Correct label	Distractor	Correct label	Distractor	Correct label	Distractor
1	Angry	Enraged	Concerned	Compassionate	Neutral	Contemptuous
2	Sad	Compassionate	Disappointed	Doubtful	Neutral	Hurt
3	Disgusted	Interested	Contemptuous	Embarrassed	Neutral	Doubtful
4	Sad	Shocked	Jealous	Doubtful	Neutral	Guilty

performed whole brain analyses to investigate *group* differences (controls vs ASD) in emotional prosody processing and whether the emotional prosody network was distinctly modulated by *condition* (implicit vs explicit) and *emotion complexity* (complex vs basic) in controls vs ASD participants. For the sake of completeness, we report significant clusters of activation for these contrasts for each group separately in Table 3.

Psychophysiological interaction. To investigate group differences in effective connectivity of brain regions during prosody processing, we conducted a psychophysiological interaction (PPI) analysis following the guidelines by O’Reilly et al. (2012). The PPI analysis reveals how the coupling between a seed region and any other voxel in the brain changes with task condition (Friston et al., 1997; Rogers et al., 2007; O’Reilly et al., 2012).

Table 3. Emotional prosody recognition performance

	Controls			ASD			P
	N	M	SD	N	M	SD	
<i>Behavioral task</i>							
	22			27			
Accuracy		0.69	0.10		0.58	0.18	0.006**
RT total scale (s)		11.4	3.26		13.9	4.44	0.030*
<i>fMRI task</i>							
	21			20			
Accuracy basic emotions		0.79	0.16		0.79	0.13	0.904
Accuracy complex emotions		0.64	0.15		0.60	0.14	0.330
RT basic emotions (s)		1.48	0.23		1.57	0.31	0.334
RT social emotions (s)		1.78	0.30		1.76	0.30	0.895

Means (M), standard deviations (SD) and sample size (N) of group characteristics. P-values: two-tailed significance-value for independent sample t-tests in ASD vs Controls.

*Significant difference between controls and ASD ($P < 0.05$).

**Significant difference between ASD and controls ($P < .01$). Abbreviations: ASD: Autism Spectrum Disorder; Reaction times for correctly solved items (RT); seconds (s).

Specifically, we sought to identify group differences in the coupling of brain regions when processing emotional vs neutral prosody and complex vs basic emotional prosody.

The PPI represents the interaction between task condition (e.g. emotional vs neutral prosody) and the correlation of activity in two or more brain regions. External effects, such as the main effect of condition (e.g. brain activity for emotional vs neutral prosody), are regressed out in the PPI approach. We selected the right STS as the seed region for the PPI based on a previous study (Ethofer et al., 2006) that identified the right STS as the input region of the prosody processing network. The seed ROI was defined by drawing a 10 mm sphere around the peak-activated voxel of the STS cluster (MNI coordinates: 52, -18, -10) in the conjunction map. The conjunction represents the overlap of activation for emotional vs neutral prosody across groups, and is therefore not biased by group differences. On the single-subject level, the PPI model included four main regressors and additional nuisance regressors as described in the preprocessing section. The physiological regressor was the demeaned time course from the seed ROI (right STS). The psychological regressor contrasted the experimental conditions (e.g. emotional vs neutral prosody). A third regressor represented the added effect of both task conditions (e.g. emotional and neutral prosody). Finally, the PPI regressor was the vector product of the physiological and psychological regressors. On the group level, we investigated differences in effective connectivity between controls and individuals with ASD.

Brain behavior relationship. To investigate whether brain activity during prosody processing correlated with prosody recognition accuracy, we added accuracy scores from the independent behavioral prosody task as a covariate into the fMRI group analysis. We investigated whether activity for emotional vs neutral prosody and complex vs basic emotional prosody were modulated by prosody recognition accuracy.

Furthermore, we investigated whether the coupling between brain regions during emotional prosody processing predicted prosody recognition accuracy on the independent behavioral task. For this analysis, we added performance on the behavioral prosody recognition task as a covariate into the PPI group analyses.

Results

Behavioral results: emotional prosody recognition

Performance measures for both tasks comprised accuracy scores (percentages of correct answers) and reaction times (time to choose the correct emotion label) for correctly solved items.

Behavioral prosody task. To avoid the repetition of basic emotion in the task, the majority of items conveyed complex emotions (21 out of 25 task items). Due to the different numbers of included basic and complex emotions, we refrained from analyzing group differences in basic emotion recognition and from comparing basic and complex emotion recognition in the behavioral task. Independent sample t-tests revealed that controls were more accurate and faster than individuals with ASD [accuracy: $t(41) = 2.72$, $P = 0.006$; RT: $t(47) = -2.23$, $P = 0.03$ (homogeneity of variance is not met); see Figure 1B]. In the ASD group, accuracy scores correlated negatively with autism symptomatology, as measured by the ADOS [$r(22) = -0.448$, $P = 0.028$] and the ASDI [$r(22) = -0.478$, $P = 0.018$], indicating that more severely affected individuals scored lower on the task (see Supplementary Figure S2 in the supplemental section). Furthermore, task accuracy was positively correlated with verbal IQ in the control group [$r(20) = 0.497$, $P = 0.019$] but not in the ASD group [$r(25) = -0.100$, $P = 0.619$]. The difference between the correlations is significant (Fisher's $Z = 2.10$; $P < 0.05$).

fMRI task. The number of blocks containing basic and complex emotions in the fMRI prosody task was equal, and thus we compared emotion recognition behavior of complex vs basic emotions by adding the within-subject factor emotion complexity to the analysis. Repeated measures ANOVAs with the within subject factor complexity (complex vs basic emotions) and the between subject factor group (Controls vs ASD) were performed for accuracy rates and RT separately. Over all participants, basic emotions were recognized faster and more accurately than complex emotions [accuracy: $F(1, 39) = 47.9$, $P < 0.01$, $\eta_p^2 = 0.551$; RT: $F(1, 39) = 55.93$, $P < 0.01$, $\eta_p^2 = 0.589$]. The groups showed comparable emotion recognition performance for basic and complex emotions [accuracy: $F(1, 39) = 0.43$, $P = 0.516$; RT: $F(1, 39) = 0.18$, $P = 0.667$]. Furthermore, there was no significant

Table 4. Significant activations in the contrasts of interest in Controls and in individuals with ASD

	Side	Cluster size (Voxel)	Peak voxel MNI Coordinates (mm)			Peak Z score
			x	y	z	
<i>Emotional prosody > neutral prosody in controls</i>						
Cluster 1		6068				
Superior temporal gyrus	R		48	-18	0	6.08
Superior temporal sulcus	R		52	-26	4	5.96
			56	-18	0	5.91
			48	-22	0	5.9
			66	-30	6	5.48
			58	-24	-2	5.04
Inferior frontal gyrus	R		56	24	0	4.56
Cluster 2		3715				
Heschl's gyrus	L		-42	-26	2	5.7
			-48	-14	-2	5.64
			-46	-18	0	5.64
Superior temporal sulcus	L		-46	-24	-4	5.23
			-48	-12	-6	5.21
Superior temporal gyrus	L		-64	-28	6	5.2
<i>Emotional prosody > neutral prosody in ASD</i>						
Cluster 1		3841				
Inferior frontal gyrus	R		46	30	2	4.46
			58	34	8	4.42
			52	32	4	4.23
Superior temporal sulcus	R		50	4	-22	4.44
			58	20	20	4.34
			52	4	-18	4.27
Cluster 2		2017				
Superior temporal gyrus	L		-66	-34	6	4.73
			-68	-26	4	4.1
			-60	-22	4	3.5
Superior temporal sulcus	L		-54	-18	-14	3.56
Temporal pole	L		-38	4	-32	3.53
Heschl's gyrus	L		-46	-4	-6	3.5
Cluster 3		1715				
Intracalcarine cortex	L		16	-72	6	3.75
			0	-88	2	3.34
			16	-72	14	3.33
Cuneal cortex	L		-10	-76	20	3.46
Lingual gyrus	L		-18	-60	-2	3.3
Occipital pole	L		14	-86	16	3.23
<i>Explicit > implicit emotional prosody in controls</i>						
Cluster 1		4859				
Superior temporal sulcus	R		56	-14	-8	4.86
			58	-8	-4	4.26
			54	-32	0	4.17
			60	-16	-2	4.09
Inferior frontal gyrus	R		54	28	-4	4.36
			56	28	0	4.09
Cluster 2		1842				
Superior temporal sulcus	L		-52	2	-20	5.29
			-52	-2	-20	4.91
Superior temporal gyrus	L		-50	-10	-2	3.81
			-48	-38	6	3.74
			-60	-26	8	3.69
			-64	-26	6	3.69
<i>Explicit > implicit emotional prosody in ASD</i>						
Cluster 1		4859				
Superior temporal sulcus	R		56	-14	-8	4.86
			58	-8	-4	4.26
			54	-32	0	4.17
			60	-16	-2	4.09
Inferior frontal gyrus	R		54	28	-4	4.36

(Continued)

Table 4.. (Continued)

	Side	Cluster size (Voxel)	Peak voxel MNI Coordinates (mm)			Peak Z score
			x	y	z	
Cluster 2		1842	56	28	0	4.09
Superior temporal sulcus	L		-52	2	-20	5.29
			-52	-2	-20	4.91
Superior temporal gyrus	L		-50	-10	-2	3.81
			-48	-38	6	3.74
			-60	-26	8	3.69
			-64	-26	6	3.69
<i>Explicit > implicit emotional prosody in ASD</i>						
Cluster 1		1573				
Superior temporal sulcus	R		54	-16	-8	4.02
			52	-20	-2	3.95
			52	-16	-4	3.95
Planum polare	R		50	-10	-4	3.81
	R		42	-20	-4	3.48
<i>Implicit > explicit emotional prosody in controls</i>						
Cluster 1		42364				
Precuneus cortex	L		-2	-48	38	7.37
Lateral occipital cortex	L		-44	-70	40	6.96
Frontal pole	L		-22	36	36	6.94
Medial prefrontal cortex	L/R		14	50	-4	6.76
Middle frontal gyrus	R		24	32	36	6.49
Anterior cingulate cortex	L/R		2	46	-8	6.33
<i>Implicit > explicit emotional prosody in controls</i>						
Cluster 1		42364				
Precuneus cortex	L		-2	-48	38	7.37
Lateral occipital cortex	L		-44	-70	40	6.96
Frontal pole	L		-22	36	36	6.94
Medial prefrontal cortex	L/R		14	50	-4	6.76
Middle frontal gyrus	R		24	32	36	6.49
Anterior cingulate cortex	L/R		2	46	-8	6.33
<i>Implicit > explicit emotional prosody in ASD</i>						
Cluster 1		19119				
Parietal operculum cortex	L		-50	-26	24	6.48
			-48	-26	20	5.7
Posterior cingulate cortex	L/R		-4	-46	38	6.2
Precuneus cortex	L/R		-4	-48	42	5.84
Precentral gyrus	L/R		-2	-30	48	5.32
Supplementary motor cortex	L/R		10	-4	48	5.12
Cluster 2		4584				
Frontal pole	R		12	62	0	5.3
Paracingulate gyrus	L/R		6	42	-8	4.73
			10	48	-4	4.46
Parahippocampal gyrus	L		-26	-26	-18	4.6
			-26	-26	-22	4.52
Cluster 3		2342				
Insular cortex	R		40	-14	12	4.95
			40	-16	16	4.62
			36	-10	2	4.48
			36	-20	2	4.32
			38	-14	2	4.3
Heschl's Gyrus	R		46	-12	10	4.67
<i>Basic > complex emotional prosody in ASD</i>						
Cluster 1		6690				
Superior Temporal gyrus	L		-64	-6	-2	8.36
			-60	-10	-6	7.86
Superior temporal sulcus			-64	-8	-8	7.86
			-56	-22	0	6.66
Cluster 2		5752				
Planum temporale	R		42	-28	8	7.62
			42	-32	12	7.12

(Continued)

Table 4.. (Continued)

	Side	Cluster size (Voxel)	Peak voxel MNI Coordinates (mm)			Peak Z score
			x	y	z	
Superior Temporal gyrus	R		38	-32	8	6.7
			38	-28	8	6.67
			62	-12	-2	7.41
			64	-28	6	7.23

All reported clusters are family wise error cluster corrected for multiple comparisons (FWE) at a statistical threshold of $P < 0.05$ and a z-value of 2.3.

group by complexity interaction for accuracy rates [$F(1, 39) = 0.61$, $P = 0.441$] and RT [$F(1, 39) = 2.05$, $P = 0.161$; see also Table 4]. In the implicit task condition, participants had to correctly label the gender of the speaker. Participants accuracy overall conditions was greater than 95%. There was no between group difference in either accuracy (Controls: 95%, $SD = 6$; ASD group: 96%, $SD = 6$) or reaction times (Controls: 0.9 s, $SD = 0.2$; ASD group: 0.9 s, $SD = 0.2$).

FMRI results

Emotional prosody processing network. Contrasting emotional with neutral prosody revealed a previously described fronto-temporal network including the STS and IFG in both groups (Figure 2B-a and Table 3) whereby ASD and controls showed overlapping neural activity in the right temporal pole, STS and IFG (Figure 2B-b and Table 5). There were no between group differences in overall emotional prosody processing.

Effects of condition and emotion complexity on emotional prosody processing. With regards to emotion complexity, we did not find regions that showed stronger activity for complex vs basic emotions in either group. We did, however, find a significant group by complexity interaction. Compared with the ASD group, controls showed a significantly greater increase in activity of bilateral fronto-temporal regions including the STS, insular cortex, superior temporal gyrus (STG) and right amygdala for complex vs basic emotions (Figure 2C-a and Table 5). The ASD group recruited temporal regions, such as the STS, more when processing basic emotions. We found a significant main effect of condition (basic vs complex) in the ASD group only (Table 3). Interestingly, we also found significant gender differences in the ASD group. Female ASD participants showed more activity of right temporal regions such as the STG and STS for complex vs basic emotions compared to males (Table 5).

In both groups, explicit vs implicit emotional prosody processing yielded increased activity of prosody processing regions such as the STS (Table 3). Implicit vs explicit prosody processing recruited cortical midline regions, such as the PCC and the frontal pole in both groups (Table 3). There was also a significant condition by group interaction. Controls showed increased activity of occipital and prefrontal regions compared to the ASD group (Table 5).

Effective connectivity between brain regions during emotional prosody processing. The PPI analysis did not reveal between group differences in processing emotional vs neutral prosody. We did, however, find between-group differences in effective

connectivity for complex vs basic emotional prosody. STS and amygdala (peak voxel: -20, -6, -22) (Figure 2C-b and Table 5).

Relationship between neural processing of emotional prosody and behavioral performance. We found group differences in the relationship between brain activity for emotional vs. neutral stimuli and prosody recognition performance on the behavioral task. Brain activity in a wide network of frontal and temporal regions, including the STG and the superior frontal gyrus, correlated more strongly with prosody recognition performance in individuals with ASD compared to controls (Figure 3A; Table 6). There were no significant correlations between brain activity for complex vs. basic emotions and behavioral task performance in either group.

When investigating the relationship between effective connectivity and behavioral task performance, we found the opposite group difference. Higher coupling between the right STS and anterior cingulate cortex (ACC, MNI coordinates: 0, 48, -4) for emotional vs neutral prosody predicted prosody recognition accuracy in controls compared to individuals with ASD (Figure 3B). Furthermore, we found similar group differences when investigating the relationship between effective connectivity for complex vs basic emotional prosody and behavioral performance. The coupling between right STS, fusiform cortex (FC) and precentral gyrus (PG) was stronger correlated with task accuracy in controls than in the ASD group (Table 7).

Discussion

The aim of the current study was to investigate differences in emotional prosody processing between individuals with ASD and healthy controls in behavior and brain function. In the behavioral experiment, we found that the ASD group was slower and less accurate in recognizing emotional prosody than controls. Symptom severity was negatively correlated with accurate recognition of emotional prosody. More impaired individuals scored lower on the task. The fMRI experiment, replicated the well-established emotional prosody network, including the STS and IFG, overall participants. Complex vs basic emotional prosody elicited less activity of core prosody processing regions, such as the STS and amygdala, in individuals with ASD compared to controls. Also, the STS and amygdala were less functionally connected in individuals with ASD. Importantly, the relationship between behavioral performance and neural processing of emotional prosody differed between groups. In the ASD group, brain activity in a wider spread network of cortical regions was more strongly related to behavioral task accuracy. In controls, on the other hand, the magnitude of effective

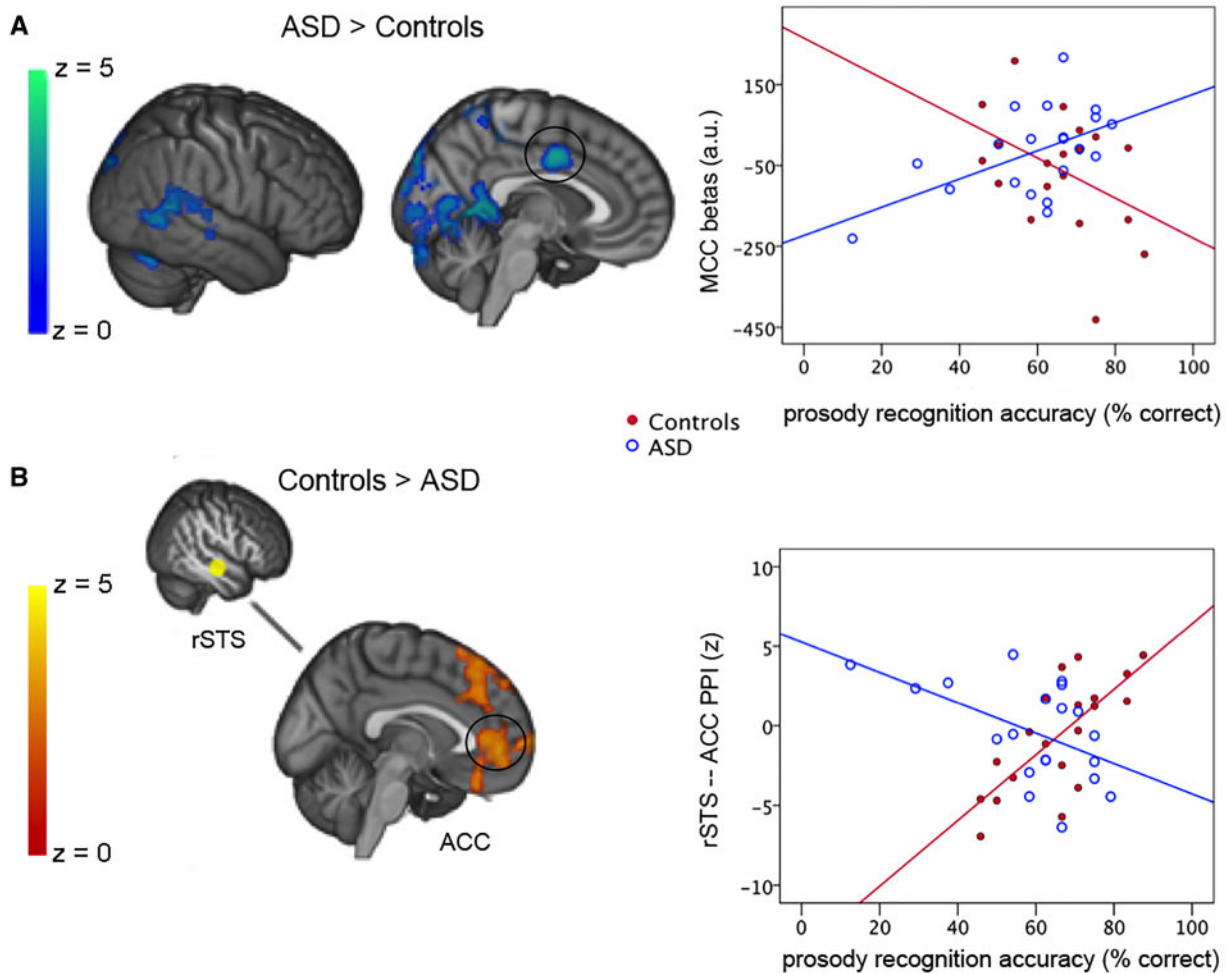


Fig. 3. Brain behavior relationship. (A) Stronger correlation between brain activity during emotional vs neutral prosody processing and accuracy on the behavioral prosody recognition task in ASD participants compared to controls. Correlation plot illustrates the relationship between parameter estimates extracted from the MCC and task accuracy in controls (red) and in individuals with ASD (blue). (B) Stronger correlation between rSTS—ACC effective connectivity and accuracy on the behavioral prosody recognition task in controls compared to individuals with ASD. Correlation plot illustrates the relationship between effective connectivity and accuracy in controls (red) and in individuals with ASD (blue). All clusters are significant at $P < 0.05$ and $z = 2.3$ family wise error (FWE) cluster corrected for multiple comparisons. Abbreviations: Autism Spectrum Disorders (ASD); Middle Cingulate Cortex (MCC); Anterior Cingulate Cortex (ACC); right Superior Temporal Sulcus (rSTS).

connectivity between STS and ACC during emotional prosody processing more strongly predicted behavioral accuracy.

Processing emotional prosody robustly activated the well-replicated prosody network both in the control and ASD group (Schirmer and Kotz, 2006; Wildgruber et al., 2006). Furthermore, both groups showed overlapping clusters of activation in the right IFG and STS for emotional vs neutral prosody. The right STS and right IFG have been more strongly implicated in emotional prosody processing than their contralateral homologues (Ross, 1981; Schirmer and Kotz, 2006). There were no between-group differences in overall prosody processing. We did, however, find group differences in processing complex vs basic emotions and implicit vs explicit prosody processing.

Individuals with ASD displayed reduced activity in bilateral temporal regions, such as the superior temporal gyrus, temporal pole and right STS for complex vs. basic emotions. These regions have been extensively implicated in auditory processing (Belin et al., 2000), in particular in processing emotional prosody (Wildgruber et al., 2005). This interaction effect in temporal regions, such as the STS, is due to the fact that individuals with ASD engage these regions more when processing basic vs

complex emotions. Previous research has shown that the STS does not distinguish between social and nonsocial information in individuals with ASD (Pelphrey et al., 2011). In this study, both basic and complex emotions represent social stimuli. Basic emotions, however, are less socially motivated; accurately recognizing basic emotions relies more on decoding physiological states than interpersonal relations (Ekman, 1992). Greater activity in the STS when processing basic emotions could mean that they are more salient. This may also explain greater processing accuracy of basic emotions in ASD.

Furthermore, ASD participants exhibited reduced activity of the bilateral insula and right amygdala, regions associated with emotion processing (Pessoa and Adolphs, 2010). Groups further differed in the magnitude of effective connectivity between the right STS and left amygdala for complex vs basic emotions. Typically developing controls exhibited a stronger coupling between STS and amygdala than individuals with ASD. Our results are in line with previous studies, which showed reduced functional connectivity of STS and amygdala in ASD in both the visual and auditory modality (Kriegstein and Giraud, 2004; Monk et al., 2010). Both, the amygdala and STS, have been implicated in social perception across modalities (Pelphrey and Carter,

Table 5. Significant activations in the contrasts of interest over all participants and between group differences

	Side	Cluster size (Voxel)	Peak voxel MNI coordinates (mm)			Peak Z score
			x	y	z	
<i>Emotional vs neutral prosody</i>						
Cluster 1		15077				
Lingual gyrus	R		20	-56	-4	4.79
Planum temporale	L		-64	-22	12	4.67
Superior temporal gyrus	L		-62	-22	6	4.57
Planum temporale	L		-20	-58	-6	4.55
			-64	-26	16	4.43
Angular gyrus	L		-38	-48	20	4.44
Cluster 2		1523				
Middle cingulate cortex	L/R		10	-2	40	4.33
			-10	-10	36	3.66
Posterior cingulate cortex	L/R		-12	-26	44	3.72
			-12	-30	44	3.7
Supplementary motor cortex	L/R		-10	4	40	3.49
Superior frontal gyrus	L/R		-8	-36	50	3.38
<i>Complex > basic emotional prosody in controls > ASD</i>						
Cluster 1		3169				
Planum temporale	R		38	-34	10	6.27
			42	-30	10	6.14
			52	-18	6	4.25
Parietal operculum cortex			44	-34	24	3.9
Planum polare	R		46	2	-14	3.78
			44	-12	-4	3.56
Superior temporal sulcus	R		56	-38	10	2.94
Amygdala	R		26	-4	-22	2.74
Cluster 2		2284				
Heschl's gyrus			-44	-22	2	4.39
Superior temporal sulcus	L		-64	-8	-6	4.36
Superior temporal gyrus	L		-64	-6	0	4.16
Planum temporale	L		-50	-22	2	3.75
Central opercular cortex	L		-58	-8	8	3.59
			-56	-4	6	3.48
Insular cortex	L		-38	-4	10	3.08
<i>Gender differences in complex > basic emotional prosody in ASD</i>						
Cluster 1		2107				
Planum temporale	R		58	-18	6	3.97
Supramarginal gyrus	R		46	-26	36	3.42
Superior temporal sulcus	R		64	-30	4	3.35
Superior temporal gyrus	R		66	-32	8	3.33
Planum temporale	R		38	-34	14	3.31
			52	-30	10	3.22
<i>Implicit > explicit emotional prosody in controls > ASD</i>						
Cluster 1		1430				
Angular gyrus	L		-44	-68	44	4.7
			-46	-68	40	4.69
			-60	-56	26	4.56
Lateral occipital cortex	L		-48	-68	44	4.48
			-56	-62	34	3.94
			-58	-64	24	3.81
Cluster 2		1218				
Frontal pole	L		-16	58	4	4.07
			-24	48	-10	3.57
			-26	48	-4	3.45
			-22	68	-8	3.07
Anterior cingulate cortex	L/R		-6	46	22	3.58
	L/R		6	42	20	3.33

(Continued)

Table 5.. (Continued)

	Side	Cluster size (Voxel)	Peak voxel MNI coordinates (mm)			Peak Z score
			x	y	z	
Cluster 3		1038				
Superior frontal gyrus	L		-24	36	52	4.17
			-20	28	30	4.08
			-20	30	26	4
			-20	28	22	3.9
			-18	36	54	3.76
			-14	34	56	3.51

All reported clusters are family wise error cluster corrected for multiple comparisons (FWE) at a statistical threshold of $P < 0.05$ and a z-value of 2.3.

Table 6. Relationship between neural processing of emotional prosody and behavioral performance in ASD vs Controls

	Side	Cluster size (Voxel)	Peak voxel MNI coordinates (mm)			Peak Z score
			x	y	z	
<i>Emotional prosody > Neutral prosody over all participants (conjunction analysis)</i>						
Cluster 1		4208				
Superior temporal sulcus	R		52	-18	-10	3.78
Inferior frontal gyrus	R		52	16	12	3.20
Temporal pole	R		58	8	-22	2.88
Temporo-parietal junction	R		56	-44	20	2.91
<i>Psychophysiological interaction: complex > basic emotional prosody in controls > ASD</i>						
Amygdala	L	13	-20	-6	-22	2.75

All reported clusters are family-wise error cluster corrected for multiple comparisons (FWE) at a statistical threshold of $p < .05$ and a z-value of 2.3.

Table 7. Relationship between effective connectivity between brain regions and behavioral performance in Controls vs ASD

	Side	Cluster size (Voxel)	Peak voxel MNI coordinates (mm)			Peak Z score
			x	y	z	
<i>Psychophysiological interaction with rSTS as seed region for complex vs basic emotional prosody</i>						
Cluster 1		1580				
Temporal occipital fusiform cortex	L		-28	-58	-6	3.73
			-24	-62	-10	3.47
Temporal fusiform cortex			-28	-38	-14	3.67
Lateral occipital cortex	L		-50	-70	-4	3.65
			-50	-70	-8	3.52
			-50	-72	6	3.46
Cluster 2		819				
Precentral gyrus			-8	-16	70	3.52
			-6	-12	70	3.45
			-4	-16	68	3.43
			-16	-22	70	3.33
			-12	-16	72	3.3
			2	-18	68	3.25

All reported clusters are family wise error cluster corrected for multiple comparisons (FWE) at a statistical threshold of $P < 0.05$ and a z-value of 2.3.

2008), which precedes and supports later developing mentalizing abilities (Allison et al., 2000; Adolphs et al., 2005).

The social perception deficits of individuals with ASD concern both visual and auditory modalities and persist from early childhood (Chawarska et al., 2010; Chawarska et al., 2012, 2013) throughout adulthood (Rutherford et al., 2002). In the visual domain, the amygdala and the posterior STS extending into the temporoparietal junction (TPJ) have been tightly linked to aberrant social perception of individuals with ASD

(Critchley et al., 2000; Castelli et al., 2002; Dziobek et al., 2010; Kliemann et al., 2012), in particular to their deficits in inferring others' intentions (Baron-Cohen et al., 1999a; Lombardo et al., 2011; Pelphrey et al., 2011; Pantelis et al., 2015). In contrast, very little is known about auditory social information processing of individuals with ASD. Our findings indicate that the amygdala and STS underlie the social information processing deficits of individuals with ASD also in the auditory modality.

In contrast to previous studies (Takahashi *et al.*, 2004; Alba-Ferrara *et al.*, 2011), we did not find increased activity of core mentalizing regions such as the ACC in controls for complex vs basic emotional prosody. The lack of a modulation by emotion complexity in typically developing controls suggests basic and complex emotions might be comparably salient and thus elicit similar activity of prosody processing regions.

An exploratory analysis of gender differences for emotional prosody processing revealed that females with ASD exhibit greater STS activity when processing complex vs basic emotions compared to males. These differences in neural processing could be linked to previously observed gender differences in autism symptomatology (Van Wijngaarden-Cremers *et al.*, 2014). However, we did not find gender differences on the behavioral level. Given the limited sample size of individuals with ASD, larger-scale studies are needed to explore gender differences in emotional prosody processing in greater detail.

In line with previous studies (Grandjean *et al.*, 2005; Sander *et al.*, 2005; Bach *et al.*, 2008; Fruhholz *et al.*, 2012), we found a modulation of the emotion prosody network by task condition (implicit vs explicit). Explicit evaluation of emotional prosody produced increased activity of the STS and IFG, regions assigned to the core prosody network in both groups. In accordance with previous studies, our results thus provide evidence of greater involvement of the core prosody regions (STS and IFG) in directing attention to emotional prosody (explicit condition) vs away from it (implicit condition) (Buchanan *et al.*, 2000; Wildgruber *et al.*, 2005; Bach *et al.*, 2008; Ethofer *et al.*, 2009). Implicit compared to explicit emotional prosody processing yielded activity of cortical midline regions, such as PCC, in both groups. Thus, in accordance with the literature, our study suggests that implicit and explicit prosody processing are mediated by distinct neural networks (Bach *et al.*, 2008; Fruhholz *et al.*, 2012). Furthermore, controls showed greater activity in the angular gyrus, and prefrontal regions such as the ACC, for implicit vs explicit prosody processing than individuals with ASD. The angular gyrus has been implicated in processing semantic information, fact retrieval, shifting attention to relevant tasks and is believed to represent a cross-modal hub, which integrates these multiple cognitive processes across sensory modalities (Seghier, 2013). Increased activity of this region in the control group relative to the ASD group might thus indicate a higher degree of cross modal integration of relevant information during implicit processing of emotional prosody in controls vs individuals with ASD.

We found significant group differences in emotional prosody recognition on the behavioral task. Individuals with ASD showed lower performance on the newly developed prosody recognition task compared to controls. Accuracy rates were negatively correlated with symptom severity in individuals with ASD, with more impaired individuals scoring lower. Along with basic emotional expressions, the newly developed task covers a wide range of complex emotions portrayed by a large number of male and female speakers. The higher degree of complexity and ecological validity of the task most likely increased its sensitivity to the subtle impairments of our sample of high-functioning ASD participants. Our results are in line with studies showing emotion recognition difficulties from voices of individuals with ASD (Hobson, 1986; Hobson *et al.*, 1988; Baron-Cohen *et al.*, 1993). Given that the recognition of complex emotions may involve mental state processing (Hoffman, 2000; de Vignemont and Singer, 2006; Decety and Jackson, 2006), the impaired recognition of complex emotions in individuals with ASD likely reflects their core deficit in understanding others'

mental states (Baron-Cohen *et al.*, 2001). In the simpler fMRI version of the task, which comprised a more limited number of speakers and emotions (six basic and six complex emotions) with only two answer options, we did not find behavioral between-group differences. Similarly, some studies that also used a more limited number of speakers, emotions or answer options report no differences in emotional prosody recognition between individuals with ASD and controls (Loveland *et al.*, 1997; Boucher *et al.*, 2000; Chevallier *et al.*, 2011). Our study thus stresses the importance of using more naturalistic tasks than previously done to sensitively assess the subtle social cognitive impairments of high-functioning individuals with ASD.

Finally, we took the first step towards establishing a neurocognitive model of prosody processing in ASD by investigating the relationship between neural processing of emotional prosody and prosody recognition performance on an independent task.

We found significant group differences in the relationship between behavioral and neural prosody processing. In typically developing individuals the coupling between STS and ACC during emotional prosody processing was a stronger predictor of task accuracy than in individuals with ASD. While the STS is involved in assessing the social salience of nonverbal stimuli (Allison *et al.*, 2000; Pelphrey *et al.*, 2011), the ACC is more strongly implicated in the explicit evaluation of emotions (Bush *et al.*, 2000; Bach *et al.*, 2008). A higher connectivity between the two regions may facilitate emotion detection in the auditory modality and thus increase emotion recognition accuracy. Moreover, increased connectivity between the STS, FC and PG while processing complex vs basic emotions, was also more strongly related to prosody recognition in controls compared to individuals with ASD. The relationship between task-based functional connectivity of emotion processing regions and emotion recognition accuracy has been very little explored. A recent study that investigated the relationship between resting state functional connectivity and emotion recognition found that the intrinsic connectivity between STS and prefrontal regions was more predictive of emotion recognition in typically developing individuals than in individuals with ASD (Alaerts *et al.*, 2014)). Reduced connectivity of the STS and prefrontal regions during emotion processing could account for the emotion recognition deficits of individuals with ASD. In contrast, higher activity of a wide-spread network of cortical regions including the STG and PCC was more strongly related to performance accuracy in the ASD than in the control group. ASD participants, however, were overall less accurate on the task. This indicates that the neural processes supporting accurate emotional prosody recognition in typically developing individuals differ from those in individuals with ASD.

Control participants' verbal IQ was positively correlated with emotional prosody recognition performance on the behavioral task. This was not the case for ASD participants, suggesting that their deficits in emotional prosody processing may be independent of verbal IQ. The IQ measure used in this study, however, provides a partial picture of an individual's verbal competence. Future studies should exhaustively explore the potential relationship between language and emotional prosody processing by including a more general IQ test with more fine-grained assessments of verbal and pragmatic language skills. Another limitation to the current study is the lack of an implicit behavioral prosody processing task. Future studies should explore the relationship between implicit and explicit prosody processing with comparable performance based tasks.

In sum, our study provides new insights into typical and atypical prosody processing that most likely have important implications for typical and impaired social communication in real life. We found significant differences between typically developing individuals and individuals with ASD on the behavioral and neural level as well regarding the relationship between behavioral and neural processing of emotional prosody.

Acknowledgements

We gratefully acknowledge the actors, who participated in the audio recordings and the members of the Computer and Media Service team (CMS) at the Humboldt University for their support with the production and postproduction of the audio stimuli. We also thank our student research assistants for their help with the stimulus production and data acquisition. The High Performance Computing Center at the Freie Universität Berlin (ZEDAT) is acknowledged for providing us with computational resources. In addition, we would like to thank our participants on the autism spectrum for their participation in the study.

Funding

The study was funded by a grant from the German Research Foundation (DFG; EXC 302).

Supplementary data

Supplementary data are available at SCAN online.

Conflict of interest. None declared.

References

- Adolphs, R., Gosselin, F., Buchanan, T.W., Tranel, D., Schyns, P., Damasio, A.R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, **433**(7021), 68–72.
- Alaerts, K., Woolley, D.G., Steyaert, J., Di Martino, A., Swinnen, S.P., Wenderoth, N. (2014). Underconnectivity of the superior temporal sulcus predicts emotion recognition deficits in autism. *Social Cognitive and Affective Neuroscience*, **9**(10), 1589–600.
- Alba-Ferrara, L., Hausmann, M., Mitchell, R.L., Weis, S. (2011). The neural correlates of emotional prosody comprehension: disentangling simple from complex emotion. *PLoS One*, **6**(12), e28701.
- Allison, T., Puce, A., McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences*, **4**(7), 267–78.
- American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Association.
- Bach, D.R., Grandjean, D., Sander, D., Herdener, M., Strik, W.K., Seifritz, E. (2008). The effect of appraisal level on processing of emotional prosody in meaningless speech. *Neuroimage*, **42**(2), 919–27.
- Baltaxe, C.A., D'Angiola, N. (1992). Cohesion in the discourse interaction of autistic, specifically language-impaired, and normal children. *Journal of Autism Developmental Disorders*, **22**(1), 1–21.
- Baron-Cohen, S., Harrison, J., Goldstein, L.H., Wyke, M. (1993). Coloured speech perception: is synaesthesia what happens when modularity breaks down?. *Perception*, **22**(4), 419–26.
- Baron-Cohen, S., Ring, H.A., Wheelwright, S., et al. (1999a). Social intelligence in the normal and autistic brain: an fMRI study. *European Journal of Neurosciences*, **11**(6), 1891–8.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., Plumb, I. (2001). The “Reading the Mind in the Eyes” Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, **42**(2), 241–51.
- Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, **403**(6767), 309–12.
- Boucher, J., Lewis, V., Collis, G.M. (2000). Voice processing abilities in children with autism, children with specific language impairments, and young typically developing children. *Journal of Child Psychology and Psychiatry*, **41**(7), 847–57.
- Brennand, R., Schepman, A., Rodway, P. (2011). Vocal emotion perception in pseudo-sentences by secondary-school children with Autism Spectrum Disorder. *Research in Autism Spectrum Disorders*, **5**(4), 1567–73.
- Buchanan, T.W., Lutz, K., Mirzazade, S., et al. (2000). Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Brain Research. Cognitive Brain Research*, **9**(3), 227–38.
- Bush, G., Luu, P., Posner, M.I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, **4**(6), 215–22.
- Castelli, F., Frith, C., Happe, F., Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, **125**(Pt 8), 1839–49.
- Chawarska, K., Macari, S., Shic, F. (2012). Context modulates attention to social scenes in toddlers with autism. *Journal of Child Psychology and Psychiatry*, **53**(8), 903–13.
- Chawarska, K., Macari, S., Shic, F. (2013). Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum disorders. *Biological Psychiatry*, **74**(3), 195–203.
- Chawarska, K., Volkmar, F., Klin, A. (2010). Limited attentional bias for faces in toddlers with autism spectrum disorders. *Archives of General Psychiatry*, **67**(2), 178–85.
- Chevallier, C., Noveck, I., Happe, F., Wilson, D. (2011). What's in a voice? Prosody as a test case for the Theory of Mind account of autism. *Neuropsychologia*, **49**(3), 507–17.
- Critchley, H.D., Daly, E.M., Bullmore, E.T., et al. (2000). The functional neuroanatomy of social behaviour: changes in cerebral blood flow when people with autistic disorder process facial expressions. *Brain*, **123** (Pt 11), 2203–12.
- de Vignemont, F., Singer, T. (2006). The empathic brain: how, when and why? *Trends in Cognitive Sciences*, **10**(10), 435–41.
- Decety, J., Jackson, P.L. (2006). A social-neuroscience perspective on empathy. *Current Directions in Psychological Science*, **15**(2), 54–8.
- Deruelle, C., Rondan, C., Gepner, B., Tardif, C. (2004). Spatial frequency and face processing in children with autism and Asperger syndrome. *Journal of Autism Developmental Disorders*, **34**(2), 199–210.
- Dziobek, I., Bahnemann, M., Convit, A., Heekeren, H.R. (2010). The role of the fusiform-amygdala system in the pathophysiology of autism. *Archives of General Psychiatry*, **67**(4), 397–405.
- Dziobek, I., Fleck, S., Kalbe, E., et al. (2006). Introducing MASC: a movie for the assessment of social cognition. *Journal of Autism Developmental Disorders*, **36**(5), 623–36.
- Eigsti, I.M., Schuh, J., Mencl, E., Schultz, R.T., Paul, R. (2012). The neural underpinnings of prosody in autism. *Child Neuropsychology*, **18**(6), 600–17.

- Ekman, P. (1992). Are there basic emotions? *Psychological Review*, *99*(3), 550–3.
- Ekman, P., Friesen, W.V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124–9.
- Ethofer, T., Anders, S., Erb, M., et al. (2006). Cerebral pathways in processing of affective prosody: a dynamic causal modeling study. *Neuroimage*, *30*(2), 580–7.
- Ethofer, T., Kreifelts, B., Wiethoff, S., et al. (2009). Differential influences of emotion, task, and novelty on brain regions underlying the processing of speech melody. *Journal of Cognitive Neurosciences*, *21*(7), 1255–68.
- Friston, K.J., Buechel, C., Fink, G.R., Morris, J., Rolls, E., Dolan, R.J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage*, *6*(3), 218–29.
- Fruhholz, S., Ceravolo, L., Grandjean, D. (2012). Specific brain networks during explicit and implicit decoding of emotional prosody. *Cerebral Cortex*, *22*(5), 1107–17.
- Gebauer, L., Skewes, J., Hørlyck, L., Vuust, P. (2014). Atypical perception of affective prosody in Autism Spectrum Disorder. *NeuroImage: Clinical*, *6*, 370–8.
- Gillberg, C., Rastam, M., Wentz, E. (2001). The Asperger Syndrome (and high-functioning autism) Diagnostic Interview (ASDI): a preliminary study of a new structured clinical interview. *Autism*, *5*(1), 57–66.
- Golan, O., Baron-Cohen, S., Hill, J.J., Rutherford, M.D. (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 Developmental Disorders*, *37*(6), 1096–106.
- Grandjean, D., Sander, D., Pourtois, G., et al. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature Neuroscience*, *8*(2), 145–6.
- Hepach, R., Kliemann, D., Gruneisen, S., Heekeren, H.R., Dziobek, I. (2011). Conceptualizing emotions along the dimensions of valence, arousal, and communicative frequency—implications for social-cognitive tests and training tools. *Frontiers in Psychology*, *2*, 266.
- Hobson, R.P. (1986). The autistic child's appraisal of expressions of emotion: a further study. *Journal of Child Psychology and Psychiatry*, *27*(5), 671–80.
- Hobson, R.P., Ouston, J., Lee, A. (1988). Emotion recognition in autism: coordinating faces and voices. *Psychological Medicine*, *18*(4), 911–23.
- Hoffman, M.L. (2000). *Empathy and Moral Development: Implications for Caring and Justice*, Cambridge, UK/New York: Cambridge University Press.
- Jenkinson, M., Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*(2), 143–56.
- Kliemann, D., Dziobek, I., Hatri, A., Baudewig, J., Heekeren, H.R. (2012). The role of the amygdala in atypical gaze on emotional faces in autism spectrum disorders. *Journal of Neuroscience*, *32*(28), 9469–76.
- Kriegstein, K.V., Giraud, A.L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage*, *22*(2), 948–55.
- Kuchinke, L., Schneider, D., Kotz, S.A., Jacobs, A.M. (2011). Spontaneous but not explicit processing of positive sentences impaired in Asperger's syndrome: pupillometric evidence. *Neuropsychologia*, *49*(3), 331–8.
- Lehr, S. (1989) *Mehrfachwahl-Wortschatz Intelligenztest: MWT-B [Multiple-Choice Vocabulary Test]*. Nürnberg, Germany: Perimed-Fachbuch-Verlag-Ges.
- Lombardo, M.V., Chakrabarti, B., Bullmore, E.T., Consortium, M.A., Baron-Cohen, S. (2011). Specialization of right temporo-parietal junction for mentalizing and its relation to social impairments in autism. *Neuroimage*, *56*(3), 1832–8.
- Lord, C., Rutter, M., Dilavore, P., Risi, S. (2002) *Autism Diagnostic Observation Schedule*. Los Angeles, CA: Western Psychological Services.
- 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 Developmental Disorders*, *24*(5), 659–85.
- Loveland, K.A., Tunali-Kotoski, B., Chen, Y.R., et al. (1997). Emotion recognition in autism: verbal and nonverbal information. *Developmental Psychopathology*, *9*(3), 579–93.
- Loveland, K.A., Tunalikhotoski, B., Chen, R., Brelsford, K.A., Ortegón, J., Pearson, D.A. (1995). Intermodal perception of affect in persons with autism or down-syndrome. *Development and Psychopathology*, *7*(3), 409–18.
- McCann, J., Peppe, S. (2003). Prosody in autism spectrum disorders: a critical review. *International Journal of Language and Communication Disorders*, *38*(4), 325–50.
- Monk, C.S., Weng, S.J., Wiggins, J.L., et al. (2010). Neural circuitry of emotional face processing in autism spectrum disorders. *Journal of Psychiatry & Neuroscience* *35*(2), 105–14.
- Nichols, T., Brett, M., Andersson, J., Wager, T., Poline, J.B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage*, *25*(3), 653–60.
- O'Reilly, J.X., Woolrich, M.W., Behrens, T.E., Smith, S.M., Johansen-Berg, H. (2012). Tools of the trade: psychophysiological interactions and functional connectivity. *Social Cognitive & Affective Neuroscience*, *7*(5), 604–9.
- Pantelis, P.C., Byrge, L., Tyszka, J.M., Adolphs, R., Kennedy, D.P. (2015). A specific hypoactivation of right temporo-parietal junction/posterior superior temporal sulcus in response to socially awkward situations in autism. *Social Cognitive and Affective Neuroscience*, *10*(10), 1348–56.
- Paul, R., Augustyn, A., Klin, A., Volkmar, F.R. (2005). Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism Developmental Disorders*, *35*(2), 205–20.
- Pelphrey, K.A., Carter, E.J. (2008). Charting the typical and atypical development of the social brain. *Developmental Psychopathology*, *20*(4), 1081–102.
- Pelphrey, K.A., Shultz, S., Hudac, C.M., Vander Wyk, B.C. (2011). Research review: Constraining heterogeneity: the social brain and its development in autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, *52*(6), 631–44.
- Pessoa, L., Adolphs, R. (2010). Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. *Nature Reviews Neuroscience*, *11*(11), 773–83.
- Rogers, B.P., Morgan, V.L., Newton, A.T., Gore, J.C. (2007). Assessing functional connectivity in the human brain by fMRI. *Journal of Magnetic Resonance Imaging*, *25*(10), 1347–57.
- Rosenblau, G., Kliemann, D., Lemme, B., Walter, H., Heekeren, H.R., Dziobek, I. (2016). The role of the amygdala in naturalistic mentalising in typical development and in autism spectrum disorder. *British Journal of Psychiatry*, *208*(6), 556–64.
- Ross, E.D. (1981). The aprosodias. Functional-anatomic organization of the affective components of language in the right hemisphere. *Archives of Neurology*, *38*(9), 561–9.
- Rutherford, M.D., 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**(3), 189–94.
- Sander, D., Grandjean, D., Pourtois, G., et al. (2005). Emotion and attention interactions in social cognition: brain regions involved in processing anger prosody. *Neuroimage*, **28**(4), 848–58.
- Schirmer, A., Kotz, S.A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Science*, **10**(1), 24–30.
- Seghier, M.L. (2013). The angular gyrus: multiple functions and multiple subdivisions. *Neuroscientist*, **19**(1), 43–61.
- Smith, S.M., Jenkinson, M., Woolrich, M.W., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, **23** (suppl 1), S208–19.
- Tager-Flusberg, H. (1981). On the nature of linguistic functioning in early infantile autism. *Journal of Autism Developmental Disorders*, **11**(1), 45–56.
- Takahashi, H., Yahata, N., Koeda, M., Matsuda, T., Asai, K., Okubo, Y. (2004). Brain activation associated with evaluative processes of guilt and embarrassment: an fMRI study. *Neuroimage*, **23**(3), 967–74.
- Van Wijngaarden-Cremers, P.J.M., van Eeten, E., Groen, W.B., Van Deurzen, P.A., Oosterling, I.J., Van der Gaag, R.J. (2014). Gender and age differences in the core triad of impairments in autism spectrum disorders: a systematic review and meta-analysis. *Journal of Autism and Developmental Disorders*, **44**(3), 627–35.
- Viinikainen, M., Katsyri, J., Sams, M. (2012). Representation of perceived sound valence in the human brain. *Human Brain Mapping*, **33**(10), 2295–305.
- Wang, A.T., Lee, S.S., Sigman, M., Dapretto, M. (2006). Neural basis of irony comprehension in children with autism: the role of prosody and context. *Brain*, **129**(Pt 4), 932–43.
- Wildgruber, D., Ackermann, H., Kreifelts, B., Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Progress in Brain Research*, **156**, 249–68.
- Wildgruber, D., Riecker, A., Hertrich, I., et al. (2005). Identification of emotional intonation evaluated by fMRI. *Neuroimage*, **24**(4), 1233–41.
- Zinck, A., Newen, A. (2008). Classifying emotion: a developmental account. *Synthese*, **161**(1), 1–25.