Atypical perceptual and neural processing of emotional prosodic changes in children with autism spectrum disorders

R. Lindström [a,*], T. Lepistö-Paisley [a,c], T. Makkonen [a], O. Reinvall [b,c], T. Nieminen-von Wendt [d], R. Alén [e], T. Kujala [a]

[a] Cognitive Brain Research Unit, Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Finland
[b] Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Finland
[c] Department of Pediatric Neurology, Helsinki University Hospital, Helsinki, Finland
[d] Neuropsychiatric Rehabilitation and Medical Centre NeuroMental, Helsinki, Finland
[e] Department of Child Neurology, Central Finland Central Hospital, Jyväskylä, Finland

ARTICLE INFO

HIGHLIGHTS

- Anomalous neural prosody discrimination in children with autism spectrum disorder (ASD).
- Impaired orienting to prosodic changes in children with ASD.
- Sluggish perceptual prosody discrimination in children with ASD.

ABSTRACT

Objective: The present study explored the processing of emotional speech prosody in school-aged children with autism spectrum disorders (ASD) but without marked language impairments (children with ASD [no LI]).

Methods: The mismatch negativity (MMN)/the late discriminative negativity (LDN), reflecting pre-attentive auditory discrimination processes, and the P3a, indexing involuntary orienting to attention-catching changes, were recorded to natural word stimuli uttered with different emotional connotations (neutral, sad, scornful and commanding). Perceptual prosody discrimination was addressed with a behavioral sound-discrimination test.

Results: Overall, children with ASD (no LI) were slower in behaviorally discriminating prosodic features of speech stimuli than typically developed control children. Further, smaller standard-stimulus event related potentials (ERPs) and MMN/LDNs were found in children with ASD (no LI) than in controls. In addition, the amplitude of the P3a was diminished and differentially distributed on the scalp in children with ASD (no LI) than in control children.

Conclusions: Processing of words and changes in emotional speech prosody is impaired at various levels of information processing in school-aged children with ASD (no LI).

Significance: The results suggest that low-level speech sound discrimination and orienting deficits might contribute to emotional speech prosody processing impairments observed in ASD.

© 2018 International Federation of Clinical Neurophysiology. Published by Elsevier B.V. All rights reserved.

1. Introduction

Deficient social communication skills, narrow interests, and repetitive behavior are the main diagnostic features of autism spectrum disorders (ASD) (APA, 2013). Some individuals with ASD show significant delays and abnormalities in their language development (later referred to as children with ASD (LI) in this article), whereas some individuals with ASD show rather typical formal language development (later referred to as children with ASD (no LI)) (WHO, 1993; Rapin and Dunn, 2003; Gillberg and Coleman, 2000). However, individuals with ASD (no LI) may have deficits in semantic-pragmatic language skills (Gillberg and Coleman, 2000).

* Corresponding author at: Cognitive Brain Research Unit, P.O. Box 9, FIN-00014, University of Helsinki, Finland.
  E-mail address: riikka.h.lindstrom@helsinki.fi (R. Lindström).

Semantic and pragmatic information, as well as information on the speaker's intentions and emotions, are conveyed by speech prosody (for reviews, see Wagner and Watson, 2010; Witteman et al., 2012). The f0, intensity, and duration changes in speech mainly carry the prosodic information (Wagner and Watson, 2010; Witteman et al., 2012). It has been shown that emotional speech prosody activates the auditory cortices irrespective of listeners' attention, suggesting that the early phases of emotional prosody detection are pre-conscious (Ethofer et al., 2006; Grandjean et al., 2005; for reviews, see Kotz and Paulmann, 2010; Brück et al., 2011). At the later stages, the acoustic cues are integrated, and finally, the emotional information carried by the vocalizations is evaluated (Schirmer and Kotz, 2006; Kotz and Paulmann, 2010; Brück et al., 2011).

Even though atypical prosody production is often documented in ASD (Shriberg et al., 2001; for a review, McCann and Peppé, 2003), behavioral studies of speech prosody comprehension in ASD show conflicting results. Some studies suggest that individuals with ASD have difficulties understanding emotional prosody (Golan et al., 2007; Peppé et al., 2007; Rutherford et al., 2002; Lindner and Rosén, 2006; Chevallier et al., 2011; McCann and Peppé, 2003). Children with ASD were less accurate in matching vocal emotional expressions with facial emotional expressions than their typically developed peers (Linder and Rosén, 2006). Also, they had difficulties judging, whether the speaker was liking or disliking food based on the speech prosody (Peppé et al., 2007). Consistently, adults with ASD have been found to score lower than the control group when matching spoken phrases presented with different emotional prosody with written labels of emotions (Rutherford et al., 2002; Golan et al. 2007).

In contrast, some studies suggest rather typical emotional prosody comprehension in ASD. Children with ASD were shown to have no difficulties in naming vocally expressed emotions from spoken words (Boucher et al., 2000). Also, children and adolescents with ASD had no deficits in recognizing emotions from spoken sentences or pseudo language utterances (Heikkinen et al., 2010; Brennand et al. 2011; Le Sourn-Bissaoui et al., 2013; Grossman et al., 2010; Chevallier et al., 2011). However, adolescents with ASD had difficulties comprehending the emotional vocal cues from sentences when being under a high cognitive load (Chevallier et al., 2011).

The variation in the instructions and cognitive task demands might contribute to the above mentioned conflicting behavioral results on emotional speech prosody comprehension in ASD. Therefore, the auditory event related potentials (ERPs), including components that are elicited task-independently, could serve as a suitable tool for investigating prosody processing in ASD (Taylor and Baldeweg, 2002). Repetitive speech sounds elicit the P1, N2, and N4 deflections that reflect the detection and encoding of speech in children (Čeponienė et al., 2008). It was suggested that the P1 reflects the detection of physical stimulus features of sounds, whereas the N2 and N4 reflect more complex sound analysis (Čeponienė et al., 2001, 2005, 2008, such as the "speechness" of the stimuli (Čeponienė et al., 2008). Previous studies have reported diminished N4 amplitudes both in children with ASD (LI) (Lepistö et al., 2005) and in children with ASD (no LI) (Lepistö et al., 2006) for repetitive vowels. Also, diminished ERPs to a repetitive word stimulus were found in children with ASD (LI) (Lindström et al., 2016).

The mismatch negativity (MMN), in turn, reflects pre-conscious auditory discrimination and it is elicited by any discriminable change in physical or even abstract properties of sounds in a sound sequence (Kujala and Näätänen, 2010). The MMN amplitude and latency reflect sound discrimination accuracy; MMNs with large amplitudes and short latencies are associated with more accurate and speeded behavioral sound discrimination (Kujala and Näätänen, 2010). In children, another change-related ERP deflec-

tion, the late discriminative negativity (LDN), can follow the MMN within 400–600 ms from the deviant stimulus onset (Korpilahti et al., 1995, 2001; Čeponienė et al., 1998; for reviews, see Cheour et al., 2001; Wetzel et al., 2014). However, the brain processes underlying the LDN elicitation are still poorly understood (Wetzel et al., 2014). Distracting sound changes elicit a positive deflection called the P3a that reflects involuntary attention switching processes (Escera et al., 2000; Escera and Corral, 2007). The P3a amplitude indexes the magnitude of the sound change: more distracting sounds elicit P3a with a larger amplitude than minor sound changes (Escera et al., 2000).

ERP studies on natural speech prosody processing in ASD are scarce. Diminished MMN for scornful prosodic sound changes and prolonged MMN latency for commanding deviants, suggesting aberrant neural discrimination of emotional prosody, were found in adults with ASD (no LI) (Kujala et al., 2005). Lindström et al. (2016), in turn, investigated children with ASD (LI) with the same word stimuli used in Kujala et al. (2005) study by presenting the children a repetitive neutral word stimulus that was occasionally replaced by scornfully, commandingly, or sadly uttered word. They showed a diminished MMN/LDN for the scornful prosodic change, peaking at about 500 ms from the deviant stimulus onset, in these children. Thus, the results of Lindström et al. (2016) and Kujala et al. (2005) suggest hyposensitive neural discrimination of emotional prosodic speech changes in ASD. However, Korpilahti et al. (2007) reported enhanced MMNs to an angrily uttered deviant word stimulus presented occasionally among tenderly-uttered words in school-aged children with ASD (no LI), suggesting hyperactive neural responsiveness for this prosodic change.

In Lindström et al. (2016) study, the P3a elicited by the scornful deviant was diminished in amplitude in children with ASD (LI), suggesting impaired orienting to emotional speech sound changes in the ASD group. However, to our knowledge, it has so far not been determined whether involuntary orienting to emotional prosodic changes in natural speech, as reflected by the P3a, is abnormal in children with ASD (no LI).

The aim of the current study was to explore how school-age children with ASD (no LI) detect and encode physical stimulus features of naturally-spoken words and how they behaviorally and neurally discriminate and involuntarily orient to prosodic changes in these words uttered with different emotional connotations. We used the same ERP paradigm that was previously applied in investigating children with ASD (LI) (Lindström et al., 2016). Based on the studies by Chevallier et al. (2011) and Peppé et al. (2007), the children with ASD were hypothesized to have lower hit rates and slower reaction times for prosodic changes in the sound discrimination test than the control participants. Based on our previous results (Lindström et al., 2016), it was hypothesized that the participants with ASD would show ERP responses diminished in amplitude to a repetitive word stimulus. Further, it was hypothesized that they would show diminished MMN/LDNs to the sad and scornful prosodic deviants including minor acoustic changes in the stimuli. However, based on Korpilahti et al. (2007), we expected an enhanced MMN/LDN to the commanding deviant. Finally, based on Lindström et al. (2016), the prosodic changes were expected to elicit diminished P3a in children with ASD (no LI).

2. Methods

2.1. Participants

16 children with ASD (no LI) and 16 control children were recruited for the experiment. However, as a result of noisy EEG signal, the data of one participant with ASD were rejected. Due to

a technical error, the behavioral data of one participant with ASD (no LI) was not available. Also, the behavioral data of one participant with ASD (no LI) were rejected as an outlier (Participants' hit rate being between 5 and 24%, whereas the ASD (no LI) group average was between 93 and 98%). Altogether, the ERP data consisted of 15 participants with ASD (no LI) and 16 controls, and the behavioral data consisted of 13 participants with ASD (no LI) and 16 controls.

Altogether, 15 children with ASD (no LI) fulfilling the ICD-10 (WHO, 1993) and DSM-IV (APA, 1994) criteria for Asperger syndrome (n = 13) or DSM-5 (APA, 2013) criteria for ASD (n = 2) participated the experiment (all boys; two left handed; mean age 10.4 years, range 8.2–12.4; monolingual Finnish speakers). Children with ASD (no LI) were recruited from the Helsinki University Central Hospital (HUCH), the Neuropsychiatric Rehabilitation and Medical Centre NeuroMental, and the Central Hospital of Central Finland. They all had been diagnosed by experienced clinicians utilizing all information available from multidisciplinary teams. Additional diagnostic information was collected with parental Autism Diagnostic Interview-Revised (Lord et al., 1994; Rutter et al., 2003). Based on the patient medical records available all the children with ASD (no LI) had no other major neurological or genetic disorders, and their language development was not delayed. One child with ASD (no LI) was reported to previously having had a mild unilateral hearing deficit. However, the hearing of this child was measured with audiometer before the experiment, and as it was found to be normal, he was included in the study. One child with ASD (no LI) was taking psychostimulant medication and took five days' break from medication before the EEG experiment. Based on parental reports all the other children with ASD (no LI) had no hearing problems and were unmedicated.

The control group consisted of 16 participants (all boys; two left handed; mean age 10.1 years, range 7.5–11.8; monolingual Finnish speakers). They were recruited from elementary schools or among participants of a previous EEG study at the Cognitive Brain Research Unit (CBRU), University of Helsinki. Parental reports showed that the control children had no past or present hearing problems or neurological disorders, or any language, learning, or psychiatric problems. Furthermore, their relatives did not have autism spectrum disorders or developmental/psychiatric disorders. They were all unmedicated.

Verbal and non-verbal cognitive performance of children with ASD (no LI) had been assessed in the hospital or as a part of a previous study protocol at the CBRU with the Finnish version of the Wechsler Intelligence Scale for Children III (WISC-III; Wechsler, 1991) or the Wechsler Intelligence Scale for Preschoolers III (n = 2) (WPPSI-III; Wecshler, 2009). WISC-III (n = 14) and the WPPSI-III (n = 2) were used to assess the control children. The mean of the performance IQ (PIQ) was 98 (range 75–133, sd 12.89) and the mean verbal IQ (VIQ) 108 (range 88–137, sd 14.72) in children with ASD. The mean PIQ was 108 (range 85–136, sd 12.9) and the mean VIQ 116 (range 83–144, sd 15.3) in control participants. The independent sample t-tests showed that there were no statistically significant age or VIQ differences between the groups. Although the participants with ASD (no LI) had a PIQ in the normal range (>70), a marginally significant PIQ difference between the groups was found (t(29) = −1.90, p = .066). However, regression analyses suggested that the PIQ had no significant effect on the ERPs (measured from the frontal region of interest [ROI] including the electrodes AF3, F1, FC1 AF4, Afz, Fz, F2, FC2, and FCz) or reaction times (RTs) in any condition.

2.2. Stimuli and procedure

The stimuli, created by Leinonen et al. (1997), consisted of Finnish female name "Saara" that was uttered with neutral, commanding, sad, and scornful emotional connotations by a female speaker. They were previously used to study adults with ASD by Kujala et al. (2005) and typically developed children and children with ASD (LI) by Lindström et al. (2012, 2016). The stimulus length varied as follows: neutral stimulus 577 ms, commanding 538 ms, sad 775 ms, and scornful 828 ms. The peak loudness of the stimuli varied randomly within 5 dB (Leinonen et al., 1997), but the stimuli were not manipulated in other ways. The f0 of the standard stimulus changed as follows: start 158 Hz, peak 207 Hz (from 454 ms stimulus onset), end 195 Hz. In the deviant stimuli the f0 changed as follows: commanding: start 230 Hz, peak 235 Hz (114 ms), end 139 Hz; scornful: start 191 Hz, peak 191 Hz (224 ms), end 98 Hz; and sad: start 172 Hz, peak 172 Hz (163 ms), end 99 Hz. In the standard stimulus the intensity changed as follows: start 59 dB, peak 80 dB (427 ms), end 53 dB, and in the deviant stimuli as follows: commanding: start 66 dB, peak 83 dB (120 ms), end 55 dB; scornful: start 53 dB, peak 79 dB (270 ms), end 46 dB, and sad: start 59 dB, peak 81 dB (206 ms), end 50 dB.

Before the experiment, the children's consent to participate in the study was obtained and a written informed consent was signed by a parent. The experiment was accepted by the HUCH and Central Hospital of Central Finland Ethical Committees. The experiment followed the Declaration of Helsinki.

Electroencephalogram (EEG) was recorded in an electrically and acoustically shielded room. The participants were presented with eight blocks of stimuli (268 stimuli in each) at 56 dB (SPL; measured at the approximate location of the head of a participant) via loudspeakers (OWI-202 [OWI Inc. CA., USA]). The oddball paradigm was used: a repetitive neutral stimulus (79%) was infrequently replaced by a deviant (commanding, sad, scornful; 7% of each) stimulus (the stimulus onset asynchrony [SOA] 1300 ms). A deviant stimulus was always followed by at least two standard stimuli. The stimulus sequences for each block were fixed and the order of the blocks were randomized.

During the EEG recordings, participants sat in an armchair and watched silent film from the screen that was located in front of the participant. The stimuli were presented from loudspeakers that were located on the left and right side of the screen. The distance between the loudspeakers was 108 cm and the distance between the loudspeaker and the participant's head was 157 cm. The parents were in the experimental chamber with the participants if the participants wanted to. The participants were video-monitored continuously during the whole experiment. The duration of the EEG-experiment was about an hour.

2.3. Behavioral task

The same behavioral prosody discrimination test was used as in Lindström et al. (2012). The behavioral test was performed after the EEG-experiment. The behavioral test consisted of three blocks of stimulus pairs (40 pairs in each block) that were presented via loudspeakers at 56 dB (SPL). The stimulus pairs were either identical (50%; two neutral stimuli) or different (50%; a neutral stimulus followed by one of the deviants) (Table 1). The within-pair stimulus onset asynchrony (SOA) was 1077 ms and the between-pair SOA 3900 ms. The stimulus sequence for each block was fixed but the order of the stimulus pairs was randomized. The order of the blocks was balanced between the participants. Participant's task was to answer with a response button if the sounds were the same or if they were different. To make sure that the participants understood the task, 12 training trials were presented before the behavioral test.

2.4. ERP recordings and analysis

The Biosemi Active Two Mk2 with a 64-channel active electrode set-up (BioSemi B.V.) was used to record the continuous EEG

Table 1
Results of the behavioral stimulus discrimination task.

Stimulus	ASD (no LI)		Control	
	Hit rate (% ± SD)	Reaction time (ms ± SD)	Hit rate (% ± SD)	Reaction time (ms ± SD)
Neutral	97.5 (3.2)	1184.0 (210.3)	97.9 (3.5)	1039.7 (127.6)
Scornful	97.5 (5.2)	1232.6 (132.8)	98.1 (3.5)	1087.4 (180.0)
Commanding	93.3 (8.6)	1087.3 (200.8)	98.4 (2.3)	940.0 (145.2)
Sad	97.7 (4.1)	1213.4 (137.7)	96.9 (5.5)	1059.2 (174.8)

Mean hit rates in percentage values and reaction times in milliseconds (standard deviations in brackets) in control children and children with ASD (no LI).

(DC-104 Hz; sampling rate 512 Hz). Supplementary electrodes were placed at the left and right mastoids and to the tip of the nose (the off-line reference electrode). The eye movements were recorded with electrodes that were placed above and at the outer corner of the left eye.

The continuous EEG was processed with the EEGLAB toolbox (Delorme and Makeig, 2004). First, it was down-sampled to 256 Hz. Then, the data were off-line high-pass filtered (1 Hz) and manually checked, and all the noisy channels were marked. The EEG data were divided into epochs that were 1100 ms long, containing 100 ms pre-stimulus baseline. The epochs were then baseline-corrected. All epochs surpassing the voltage changes ±300 µV at any electrode were removed from the data. Blink artifacts were removed with fast ICA (Hyvärinen and Oja, 2000) algorithm calculated from all the non-noisy channels. One blink component per participant was removed with ICA. The EEG data with ICA corrections were then compared with the original EEG data, to be certain that only the eye blinks were removed from the data. The data were then low-pass filtered (30 Hz) and epochs with amplitudes exceeding ±150 µV were removed. Channels with data indicating unstable signal behavior were interpolated from the neighboring channels.

Finally, the ERPs were created by averaging the data for the standard and deviant stimuli and for each condition separately. To calculate the difference waves, the ERPs elicited by the standard stimuli were subtracted from the ERPs elicited by the deviant stimuli. Accepted deviant trials for the children with ASD (no LI) were 132 (range 96–148) and for the control children 140 (range 113–149).

The same ERP deflections were chosen for the analysis as in Lindström et al. (2016). The mean amplitudes for the standard stimulus ERPs were calculated as follows. First, the group-average peak latencies of each standard stimulus ERPs were determined from the Fz electrode for control and ASD groups separately. Then, the mean amplitudes were determined by integrating the ERP signal over a ±25-ms time period centered to the grand-mean peak latency of each standard stimulus response. For the latency analysis, the individual-participant peak latencies of each

standard stimulus ERPs were identified at the Fz electrode, from the same time windows that were used in Lindström et al. (2016) (see Table 2).

Similarly, the mean amplitudes of the MMN/LDN and P3a were calculated from the standard-deviant difference waves. The MMN/LDNs were identified from the Fz electrode and the P3a from the Cz electrode. For the MMN/LDN, the mean amplitudes were determined by averaging the EEG signal over a ±25-ms time centered at the grand-mean peak latency of the component and for the P3a over a ±50-ms time period centered at the grand mean peak latency of the component. For the latency analysis, the individual peak latencies of the MMN/LDN and P3a were determined at the Fz (MMN/LDN) or Cz (P3a) electrodes. Table 2 presents the time windows that were used for the MMN/LDN and P3a peak latency identification.

2.5. Statistical analyses

For the behavioral task, the button presses which appeared before the 200 ms from the presentation of the second stimulus of the stimulus pair and occurred after the presentation of the next stimulus pair were excluded in the analysis. The normality of the RTs and hit rates were analyzed with Shapiro-Wilk test of normality. The hit rates were analyzed with the Mann-Whitney U tests, and the reaction times with Repeated Measures Analyses of Variance (rANOVA) (Group × Stimuli).

To ensure that the statistical analyses were applied to a real and reliable ERP responses (not noise), the statistical significance of each brain response was tested by comparing the mean amplitudes to zero at either the Fz electrode (for standard stimulus ERPs and the MMN/LDN) or at the Cz electrode (for the P3a) with one-sample t-tests (Table 2). With this approach it was possible to exclude the insignificant responses from the analysis beforehands, and thereby avoid comparing, firstly, insignificant ERP amplitudes with each others, and secondly, the scalp distributions of the signal (a significant response) and noise (an insignificant response). The ERPs that were elicited statistically significantly in at least one of the groups were chosen for further amplitude analysis (Table 2).

Table 2
The mean amplitudes and latencies (standard deviations in brackets) of the standard stimulus ERPs, the MMn/LDN and the P3a, with the corresponding t-values, degrees of freedom, and p-values of t-tests.

Stimulus type	Response	ASD					Control				
		Amplitude µV (sd)	t	df	p-value	Latency ms (sd)	Amplitude µV (sd)	t	df	p-value	Latency ms (sd)
Neutral	1st	2.8 (1.9)	5.5	14	.000	186 (44)	3.5 (1.9)	7.1	15	.000	192 (39)
	2nd	−1.1 (2.3)	−1.8	14	.096	362 (21)	−1.4 (1.7)	−3.3	15	.004	374 (11)
	3rd	0.3 (2.5)	.56	14	.583	483 (48)	0.2 (1.4)	.5	15	.633	498 (32)
	4th	−3.2 (2.1)	−6.1	14	.000	684 (56)	−4.8 (1.8)	−10.7	15	.000	716 (36)
Scornful	MMN/LDN	−2.5 (3.0)	−3.2	14	.006	502 (45)	−3.5 (3.3)	−4.2	15	.001	480 (40)
	P3a	1.7 (3.1)	2.2	14	.050	769 (60)	1.9 (2.9)	2.7	15	.022	751 (85)
Commanding	MMN/LDN	−1.1 (4.1)	−1.0	14	.332	478 (59)	−1.4 (2.8)	−2.0	15	.066	500 (66)
	P3a	1.1 (2.4)	1.8	14	.094	760 (90)	1.3 (2.2)	2.7	15	.038	752 (58)
Sad	MMN/LDN	−5 (3.1)	−6.3	14	.000	444 (28)	−4.3 (2.6)	4.3	15	.001	453 (25)
	P3a	2.0 (1.8)	4.3	14	.001	734 (60)	2.6 (2.3)	4.5	15	.000	765 (56)
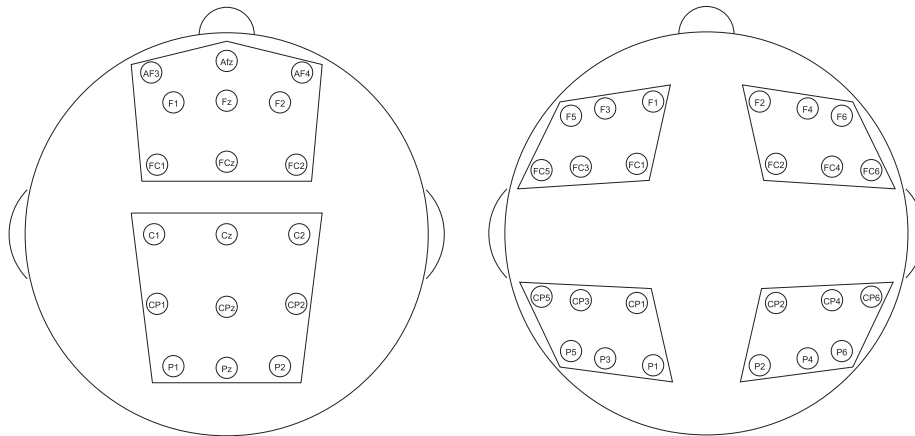
Fig. 1. Electrodes used in data analysis in the study. The electrodes were fitted into cap using cartesian XYZ electrode position coordinates of standard Biosemi 64 channels headcap (www.biosemi.com). The images are projections in XY plane. The figure on the left shows the frontal (AF3, AFz, AF4, F1, Fz, F2, FC1, FCz, FC2) and the centroparietal (C1, Cz, C2, CP1, CPz, CP2, P1, Pz, P2) ROIs that were used in group amplitude comparisons. The figure on the right shows the left frontal (F5, F3, F1, FC5, FC3, FC1), the right frontal (F2, F4, F6, FC2, FC4, FC6), the left centroparietal (CP5, CP3, CP1, P5, P3, P1) and the right centroparietal (CP2, CP4, CP6, P2, P4, P6) ROIs that were used in the scalp distribution comparisons.

To compare the ERP amplitude differences between the groups, two regions of interest (ROIs), the frontal and the centroparietal ROI, were calculated based on Kujala et al. (2010) study by averaging the ERP data over nine electrodes (Please, see Fig. 1 for more detailed information about the electrodes chosen for the analysis). Based on Kujala et al. (2010), a priori planned one-way ANOVAs were conducted to each response and ROI separately for standard stimulus responses. Similarly, based on Kujala et al. (2010), for the MMN/LDN and P3a, Group × Deviant rANOVAs were conducted to each ROI separately.

Scalp distribution comparisons were analyzed for those responses that were significantly elicited in both groups. Four ROIs were calculated by averaging the ERP data over six electrodes based on Kujala et al. (2010): the left frontal ROI, the right frontal ROI, the left parietal ROI and the right parietal ROI (Fig. 1). For standard stimulus responses, a priori planned three-way ANOVA [Group × Anterior-Posterior × Laterality] was conducted to each response separately. For the MMN/LDN and P3a, a three-way rANOVA [Group × Deviant × Anterior-Posterior × Laterality] was conducted to each response separately. For all ANOVAs/rANOVAs, significant interactions were analyzed with Sidak correction, to correct multiple comparisons within each ANOVA. However, to avoid increasing the Type II error and reducing power (Nakagawa, 2004; Perneger, 1998; Rothman, 1990), multiple comparisons were not corrected between the ANOVAs.

For all amplitude analyses, the normality of the data was analyzed with Shapiro-Wilk test of normality. The between group amplitude differences were analyzed with the Mann-Whitney U tests, if the data did not meet the normality assumption. Multivariate analysis of variance (MANOVA) was used instead of rANOVA, if the assumption of sphericity was violated. Between-group response latencies were investigated with a one-way ANOVA or Mann-Whitney U tests for the stimuli that were statistically significant in both groups. All statistical analyses were made with IBM SPSS statistics 24.

3. Results

3.1. Behavioral results

There were no hit rate differences between the groups. However, a significant main effect for Group (F[1,27] = 7.79, p = .010; partial ETA^2 = .22) was found for the reaction times: the RTs were slower in the participants with ASD than in the control children (Table 1).

3.2. Standard stimulus responses

The standard stimulus ERP consisted of four peaks (Fig. 2, Table 2). The 1st and 4th peaks were statistically significant in both groups and the 2nd peak in the control group (Table 2). At the frontal ROI the 4th peak was smaller in the children with ASD than in the controls (F [1,29] = 4.8, p = .037; partial ETA^2 = .14) (Table 2, Fig. 2). There were no standard stimulus ERP latency differences between the groups.

3.3. Deviant stimulus responses

3.3.1. The MMN/LDN

Within 500 ms from the deviant stimulus onset, a prominent negative deflection was elicited for all the deviant stimuli (Figs. 3 and 4; Supplementary Figs. S1–S9). Both the MMN and LDN could contribute to this deflection, and therefore we call this deflection as an MMN/LDN, consistently with our previous study using the same stimuli (Lindström et al., 2016). The amplitude of the MMN/LDN was statistically significant for scornful and sad stimuli in the ASD group (Table 2). In controls, it was significant for the scornful and sad stimuli, and marginally significant (p < .07) for the commanding stimuli (Table 2).

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.clinph.2018.08.018.

Significant Group × deviant interaction was found in the centro-parietal ROI (Wilks's lambda = .77, F[2,28] = 4.16, p = .026, partial ETA^2 = .23). Further, post-hoc analyses showed that the MMN/LDN amplitude for the scornful deviant was diminished in children with ASD (no LI) (t(29) = 2,1, p = .045, d′ = .75). There were no MMN/LDN latency differences between the groups.

3.3.2. The P3a

A positive deflection was elicited by all deviant stimuli at about 760 ms from the deviant stimulus onset, which presumably is the P3a (see also, Lindström et al., 2016) (Figs. 3 and 4; Supplementary Figs. S1–S9). The amplitude of the P3a was statistically significant for all the deviant stimuli in the control children. However, in children with ASD (no LI) it differed from zero for the sad stimuli and
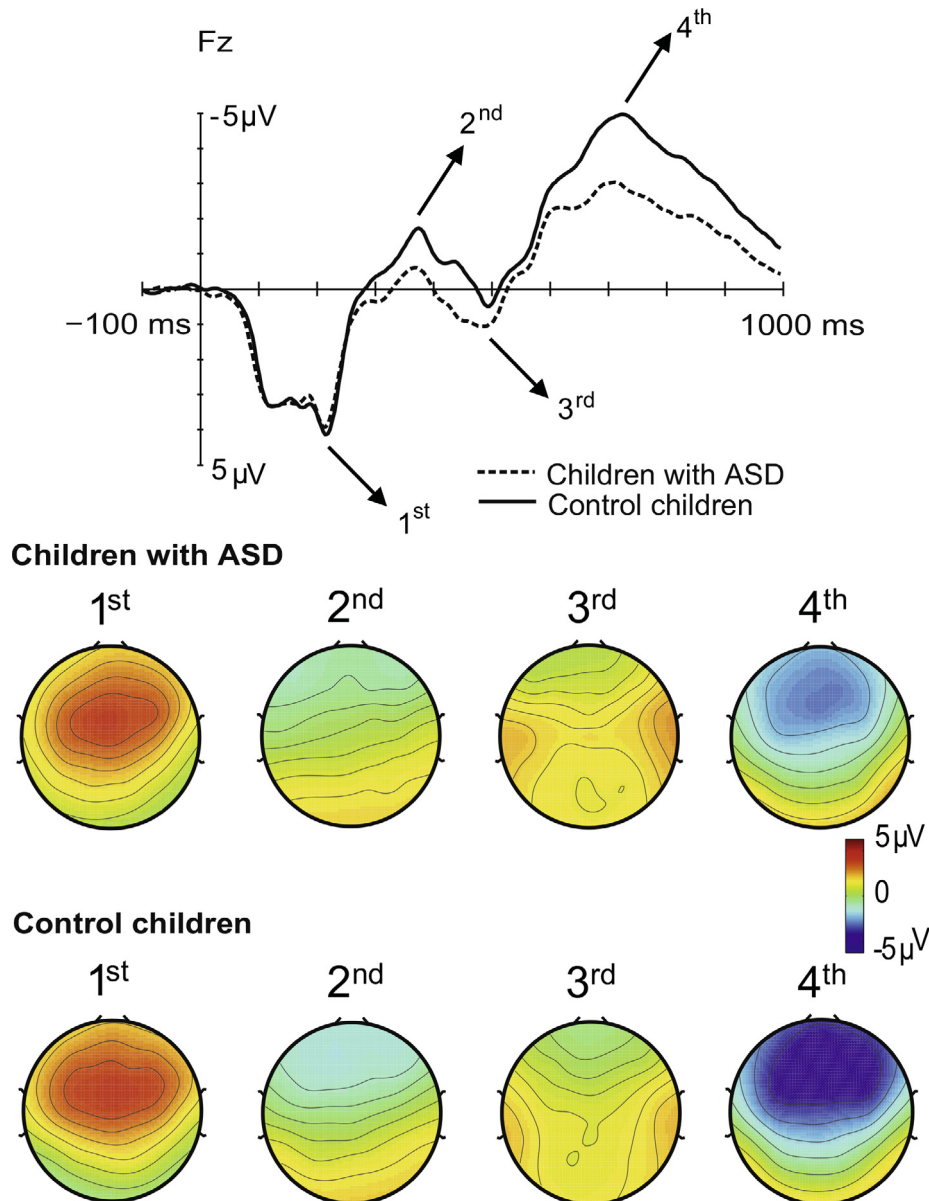
Fig. 2. Responses to the repetitive word stimuli at the Fz electrode. Four deflections chosen for analysis are marked with arrows. Respective topographic maps of 1st, 2nd, 3rd, and 4th deflections are presented below the waveforms. The stimulus onset is at 0 ms.

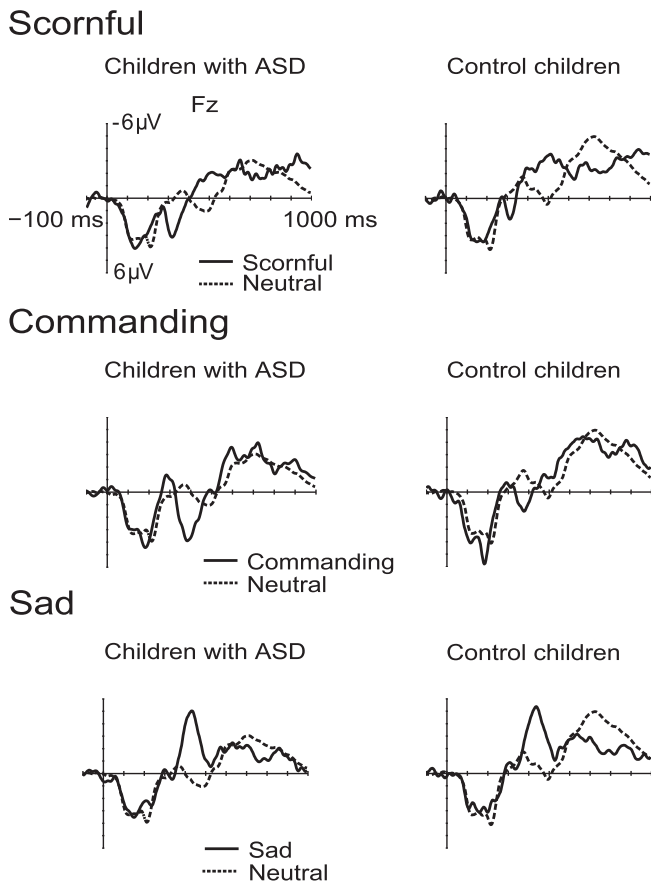was marginally significant for the scornful stimuli (p = .05), but insignificant for the commanding stimuli (Table 2).

Significant Group × Anterior-Posterior interaction effect was found for sad and scornful deviants (F[1,29] = 10.16, p = .003, partial ETA^2 = .26). Pos-hoc comparisons indicated that children with ASD (no LI) had a diminished P3a at the right and left frontal ROIs (t(29) = 2.1, p = .037, d' = .81) compared to the controls. Also, the P3a was larger at the frontal electrodes than in the parietal electrodes (t(29) = 4.3, p < .001, d' = 1.05) in control children, but no such amplitude distribution effect was found for the ASD group. The latencies of the P3a elicited by the scornful or sad stimuli did not differ between groups.

4. Discussion

The present study determined speech encoding and discrimination of prosodic speech features, as well as orienting to prosodic changes in speech, in children with ASD (no LI). ERPs elicited by natural words that were uttered with neutral voice or with commanding, sad, or scornful prosody were compared between typically developed children and children with ASD (no LI). Functions at the perceptual level were compared with a behavioral sound discrimination task, using the same stimuli, between typically developed children and children with ASD (no LI). The standard-stimulus ERPs (4th peak), the MMN/LDN and the P3a were diminished in the children with ASD (no LI). Furthermore, the scalp distribution of the P3a differed between the groups. The RTs were found to be slower in the participants with ASD (no LI) than in the control children for all the stimulus contrasts in the discrimination test. These results suggest impaired processing of words and prosodic speech features in neural speech encoding, discrimination and involuntary orienting level and slower perceptual prosody discrimination in children with ASD (no LI).

Our behavioral results showed that children with ASD were slower in discriminating prosodic features of speech stimuli than their controls. However, no Group x Stimulus interaction was observed, suggesting that the participants with ASD were reacting

Scornful

Children with ASD

Fz

-6μV

−100 ms   1000 ms

6μV   —— Scornful
                        ····· Neutral

Commanding

Children with ASD   Control children

—— Commanding
····· Neutral

Sad

Children with ASD   Control children

—— Sad
····· Neutral

Fig. 3. Dotted lines represent the grand-average ERPs to neutral stimulus and the solid lines the grand-average ERPs to the scornful, commanding, and sad stimuli at the Fz electrode. The stimulus onset is at 0 ms.

slowly irrespective of the emotional prosodic features. These results are consistent with Chevallier et al. (2011) showing slower RTs in adolescents with ASD in a vocal emotional identification task, and previous studies reporting slower than usual RTs both in children and adults with ASD to speech duration changes relevant for prosody perception (Lepistö et al., 2006; Lepistö et al., 2007).

Our result of the diminished 4th peak of the standard stimulus ERP in children with ASD suggests aberrant speech sound encoding in children with ASD (no LI). Further, they are consistent with our previous findings showing impaired encoding of the same repetitive word stimuli in children with ASD (LI) (Lindström et al., 2016) and with those of Lepistö et al. (2005, 2006) showing diminished N4 both in children with ASD (LI) and in children with ASD (no LI). In addition, in Lindström et al. (2016) study, the amplitude of both the 3rd and the 4th peak of the ERP to standard word stimuli were diminished in children with ASD (LI), whereas no between-group amplitude differences were observed in the present study for the 3rd peak of the standard stimulus ERP. The present results, and those of Lindström et al. (2016), suggest impaired encoding of words in both groups of children with ASD. However, these deficits might be more pronounced in children with ASD (LI) (Lindström et al., 2016), than in the children with ASD (no LI).

Diminished MMN/LDN amplitude for the scornful deviant was found in the present study in children with ASD (no LI), suggesting anomalous cortical discrimination of prosody in these children. These results are consistent with previous studies showing diminished MMN or MMN/LDN to the same scornful deviant stimulus both in children with ASD (LI) (Lindström et al., 2016) and adults

with ASD (Kujala et al., 2005). The scornful vs. neutral stimulus contrast was possibly harder to discriminate than the other stimulus contrasts. This is supported by the fact that the reaction times for the scornful stimuli were the longest ones in the behavioral discrimination test. Therefore, these three studies provide converging results suggesting that subtle prosodic sound features are particularly hard to discriminate for individuals with ASD.

Based on Korpilahti et al. (2007), who showed enhanced MMNs to the angry prosodic deviant in children with ASD (no LI), enhanced brain reactions were expected to the commanding deviant in children with ASD (no LI). However, we found no significant MMN/LDN amplitude differences between the groups for this deviant, consistent with Lindström et al. (2016) study. These opposite results of Korpilahti et al. (2007) vs. the present ones, Lindström et al. (2016), and Kujala et al. (2005) are compatible with the suggestion that ASD is characterized by both hypo- and hypersensitive sound processing (for reviews, see O'Connor et al., 2012; Kujala et al., 2013). Possibly, the stimuli of Korpilahti et al. (2007) included more differences in acoustical and emotional aspects than the stimuli used in the present study and in Lindström et al. (2016), and Kujala et al. (2005), leading to enhanced change-related responses. This conclusion is supported by the notion that the MMN amplitudes in Korpilahti et al. (2007) were larger (varying between -2.8 μV and -3.6 μV in children with ASD, and between -1.9 μV and -2.7 μV in control children) compared with the MMN/LDN elicited by the commanding deviant in the present study (-1.1 μV in children with ASD (no LI) and -1.4 μV in controls).

Our results showed that the P3a was diminished at the frontal scalp areas for the scornful and sad prosodic changes in children with ASD (no LI), consistent with previous studies showing diminished P3a amplitudes for speech sound changes both in children with ASD (no LI) (Lepistö et al., 2006) and children with ASD (LI) (Lindström et al., 2016; Čeponienė et al., 2003; Lepistö et al., 2005, 2008). These impairments in involuntary attention shifting towards prosodic speech features may affect the social-communication development of children with ASD, and further, contribute to the well-known social attention deficits observed in ASD (for a review, see Chevallier et al., 2012).

Further, the P3a was differently distributed in the children with ASD (no LI) than in the control children. In controls, a larger P3a was elicited at the frontal electrodes than at the parietal electrodes, whereas, no P3a amplitude distribution effect was found in the ASD group, indicating atypical P3a generator sources in ASD. Possibly, frontal P3a generator activation (Escera et al., 2000) is decreased in children with ASD as compared to typically developed children. Consistent with this interpretation, atypical frontal-lobe metabolism has been found in ASD (George et al., 1992; Zilbovicius et al., 1995).

The current stimuli were natural and thus acoustically highly variable (for further details, please see Lindström et al., 2012). Therefore, it is not possible to determine to what extent ERPs elicited by the deviant stimuli reflect the emotional category changes or the acoustical differences between the stimuli. However, as the speech prosody is conveyed via several acoustical speech features (Banse and Scherer, 1996; Wiethoff et al., 2008), controlling further the acoustical differences between the stimuli would have affected their emotional content (Wiethoff et al., 2008), and made the stimuli unnatural and ecologically invalid. The present stimuli and experimental paradigm served as valid setting to study several levels of prosody processing in ASD (no LI) and comparison of the results with those found in adults with ASD (no LI) (Kujala et al., 2005) and in children with ASD (LI) (Lindström et al., 2016). The results obtained in these studies suggest an extensive atypical processing pattern in encoding, discriminating, orienting, and reacting to natural speech prosody in ASD, which might
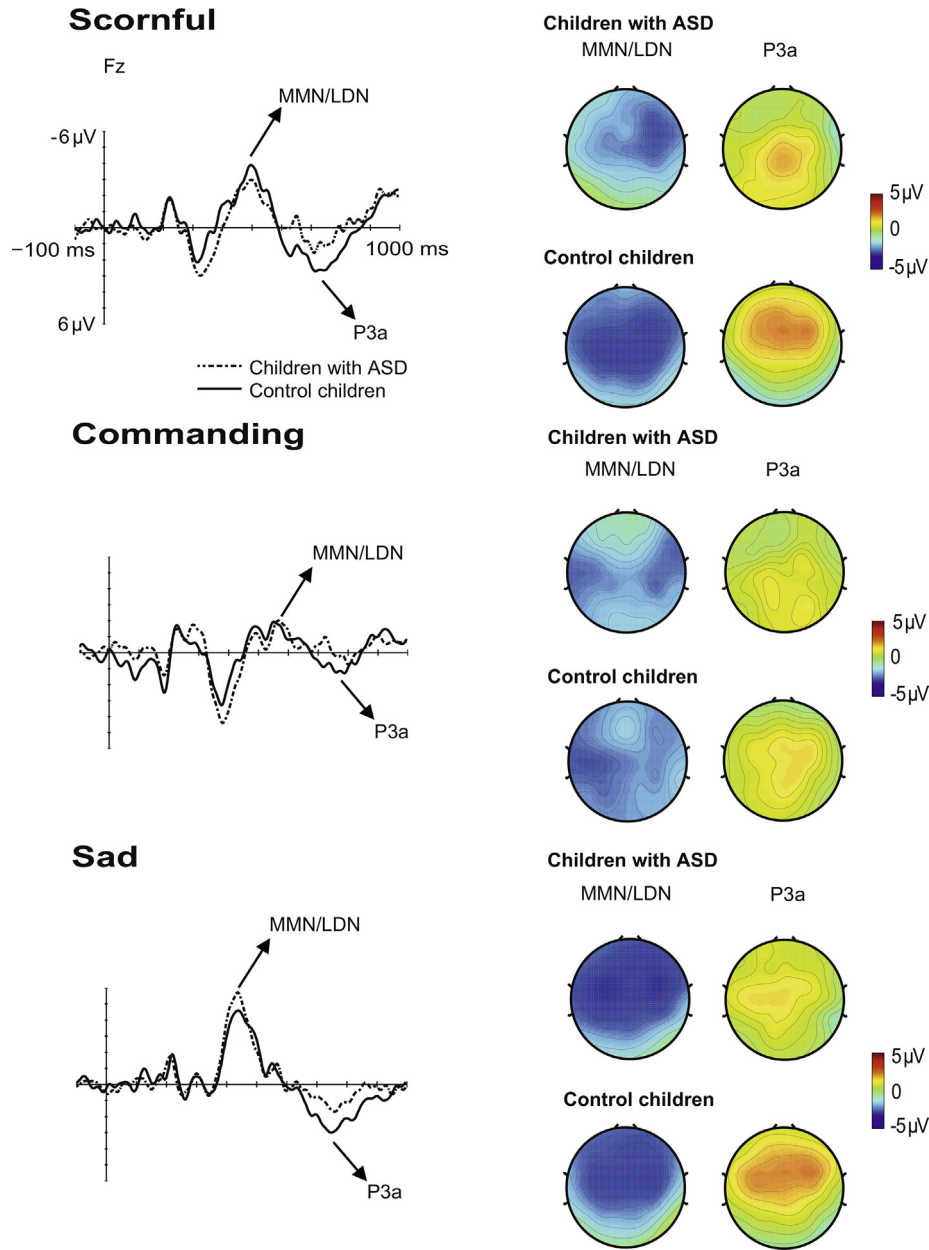
Fig. 4. The difference waves elicited by the deviant sounds at the Fz electrode. The stimulus onset is at 0 ms. Topographic maps of the MMN/LDN and the P3a are presented beside the waveforms.

underlie the observed difficulties in attributing other persons' emotional state based on his tone of voice in ASD.

4.1. Conclusions

Taken together, our results suggest impaired speech encoding, as reflected by diminished standard stimulus ERP, and aberrant neural discrimination of prosodic features, as reflected by diminished MMN/LDN, in children with ASD (no LI). Further, involuntary attention switching towards prosodic speech sound changes was found to be altered in children with ASD (no LI), as suggested by diminished and atypically distributed P3a. These neural speech sound processing deficits might contribute to the aberrant emotional prosody comprehension observed in ASD. Consistent with this, children with ASD (no LI) were found to be slower than normal in perceptually discriminating prosodic changes in speech. Our results support the hypothesis suggesting that emotional

speech prosody processing impairments observed in ASD have low-level neurofunctional origins.

Conflict of Interest

The authors have no potential conflict of interests to be disclosed.

References

American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5. 5th ed. Washington, DC: American Psychiatric Publishing; 2013.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. Washington, DC: American Psychiatric Association; 1994.

Banse R, Scherer K. Acoustic profiles in vocal emotion expression. J Pers Soc Psychol 1996;70:614–36.

Boucher J, Lewis V, Collis GM. Voice processing abilities in children with autism, children with specific language impairments, and young typically developing children. J Child Psychol Psychiatry 2000;41:847–57.

Brennand R, Schepman A, Rodway P. Vocal emotion perception in pseudo-sentences by secondary-school children with autism spectrum disorder. Res Autism Spectr Disord Disorders 2011;5:1567–73.

Brük C, Kreifelts B, Wildgruber D. Emotional voices in context: A neurobiological model of multimodal affective information processing. Phys Life Rev 2011;8:383–403.

Čeponienė R, Torki M, Alku P, Koyama A, Townsend J. Event-related potentials reflect spectral differences in speech and non-speech stimuli in children and adults. Clin Neurophysiol 2008;119:1560–77.

Čeponienė R, Alku P, Westerfield M, Torki M, Townsend J. ERPs differentiate syllable and nonphonetic sound processing in children and adults. Psychophysiol 2005;42:391–406.

Čeponienė R, Lepistö T, Shestakova A, Vanhala R, Alku P, Näätänen R, et al. Speech-sound-selective auditory impairment in children with autism: They can perceive but do not attend. Proc Natl Acad Sci USA 2003;100:5567–72.

Čeponienė R, Shestakova A, Balan P, Alku P, Yiaguchi K, Näätänen R. Children's auditory event-related potentials index sound complexity and "speechness". Int J Neurosci 2001;109:245–60.

Čeponienė R, Cheour M, Näätänen R. Interstimulus interval and auditory event-related potentials in children: evidence for multiple generators. Electroencephalog Clin Neurophysiol 1998;108:345–54.

Cheour M, Korpilahti P, Martynova O, Lang AH. Mismatch negativity and late discriminative negativity in investigating speech perception and learning in children and infants. Audiol Neurotol 2001;6:2–11.

Chevallier C, Kohls G, Troiani V, Brodkin E, Schultz R. The social motivation theory of autism. Trends Cogn Sci 2012;16:231–9.

Chevallier C, Noveck I, Happé F, Wilson D. What's in a voice? Prosody as a test case for the theory of mind account of autism. Neuropsychologia 2011;49:507–17.

Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. J Neurosci Methods 2004;134:9–21.

Escera C, Corral MJ. Role of mismatch negativity and novelty-P3 in involuntary auditory attention. J Psychophysiol 2007;21:251–64.

Escera C, Alho K, Schröger E, Winkler I. Involuntary attention and distractibility as evaluated with event-related brain potentials. Audiol Neurootol 2000;5:151–66.

Ethofer T, Anders S, Wiethoff S, Erb M, Herbert C, Saur R, et al. Effects of prosodic emotional intensity on activation of associative auditory cortex. NeuroReport 2006;17:249–53.

George MS, Costa DC, Kouris K, Ring HA, Ell PJ. Cerebral blood flow abnormalities in adults with infantile autism. J Nerv Ment Dis 1992;180:413–7.

Gillberg C, Coleman M. The biology of the autistic syndromes. London: Cambridge University Press; 2000.

Golan O, Baron-Cohen S, Hill J, Rutherford M. The 'Reading the mind in the voice' test-revised: A study of complex emotion recognition in adults with and without autism spectrum conditions. J Autism Dev Disord 2007;37:1096–106.

Grandjean D, Sander D, Pourtois G, Schwartz S, Seghier ML, Scherer KR, et al. The voices of wrath: brain responses to angry prosody in meaningless speech. Nat Neurosci 2005;8:145–6.

Grossman RB, Bemis RH, Plesa Skwerer D, Tager-Flusberg H. Lexical and affective prosody in children with high-functioning autism. J Speech Lang Hear Res 2010;53:778–93.

Heikkinen J, Jansson-Verkasalo E, Toivanen J, Suominen K, Väyrynen E, Moilanen I, et al. Perception of basic emotions from speech prosody in adolescents with Asperger's syndrome. Logoped Phoniatr Vocol 2010;35:113–20.

Hyvärinen A, Oja E. Independent component analysis: algorithms and applications. Neural Netw 2000;13:411–30.

Korpilahti P, Jansson-Verkasalo E, Mattila M, Kuusikko S, Suominen K, Rytky S, et al. Processing of affective speech prosody is impaired in Asperger syndrome. J Autism Dev Disord 2007;37:1539–49.

Korpilahti P, Krause C, Holopainen I, Lang H. Early and late mismatch negativity elicited by words and speech-like stimuli in children. Brain Lang 2001;76:332–9.

Korpilahti P, Lang H, Aaltonen O. Is there a late-latency mismatch negativity (MMN) component? Electroencephalogr Clin Neurophysiol 1995;95:P96.

Kujala T, Lepistö T, Näätänen R. The neural basis of aberrant speech and audition in autism spectrum disorders. Neurosci Biobehav Rev 2013;37:697–704.

Kujala T, Näätänen R. The adaptive brain: a neurophysiological perspective. Prog Neurobiol 2010;91:55–67.

Kujala T, Kuuluvainen S, Saalasti S, Jansson-Verkasalo E, Wendt LV, Lepistö T. Speech-feature discrimination in children with Asperger syndrome as determined with the multi-feature mismatch negativity paradigm. Clin Neurophysiol 2010;121:1410–9.

Kujala T, Lepistö T, Nieminen-von Wendt T, Näätänen P, Näätänen R. Neurophysiological evidence for cortical discrimination impairment of prosody in Asperger syndrome. Neurosci Lett 2005;383:260–5.

Le Sourn-Bissaoui S, Aguert M, Girard P, Chevreuil C, Laval V. Emotional speech comprehension in children and adolescents with autism spectrum disorders. J Commun Disord 2013;46:309–20.

Leinonen L, Hiltunen T, Linnankoski I, Laakso M. Expression of emotional-motivational connotations with a one-word utterance. J Acoust Soc Am 1997;102:1853–63.

Lepistö T, Nieminen-von Wendt T, von Wendt L, Näätänen R, Kujala T. Auditory cortical change detection in adults with Asperger syndrome. Neurosci Lett 2007;414:136–40.

Lepistö T, Silokallio S, Nieminen-von Wendt T, Alku P, Näätänen R, Kujala T. Auditory perception and attention as reflected by the brain event-related potentials in children with Asperger syndrome. Clin Neurophysiol 2006;117:2161–71.

Lepistö T, Kujala T, Vanhala R, Alku P, Huotilainen M, Näätänen R. The discrimination of and orienting to speech and non-speech sounds in children with autism. Brain Res 2005;1066:147–57.

Lepistö T, Kajander M, Vanhala R, Alku P, Huotilainen M, Näätänen R, et al. The perception of invariant speech features in children with autism. Biol Psychol 2008;77:25–31.

Lindner JL, Rosén LA. Decoding of emotion through facial expression, prosody and verbal content in children and adolescents with Asperger's syndrome. J Autism Dev Disord 2006;36:769.

Lindström R, Lepistö-Paisley T, Vanhala R, Alén R, Kujala T. Impaired neural discrimination of emotional speech prosody in children with autism spectrum disorder and language impairment. Neurosci Lett 2016;628:47–51.

Lindström R, Lepistö T, Makkonen T, Kujala T. Processing of prosodic changes in natural speech stimuli in school-age children. Int J Psychophysiol 2012;86:229–37.

Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord 1994;24:659–85.

McCann J, Peppé S. Prosody in autism spectrum disorders: a critical review. Int J Lang Commun Disord 2003;38:325–50.

Nakagawa S. A farewell to bonferroni: The problems of low statistical power and publication bias. Behav Ecol 2004;15:1044–5.

O'Connor K. Auditory processing in autism spectrum disorder: a review. Neurosci Biobehav Rev 2012;36:836–54.

Peppé S, McCann J, Gibbon F, O'Hare A, Rutherford M. Receptive and expressive prosodic ability in children with high-functioning autism. J Speech Lang Hear Res 2007;50:1015–28.

Perneger TV. What's wrong with bonferroni adjustments. BMJ 1998;316:1236–8.

Rapin I, Dunn M. Update on the language disorders of individuals on the autistic spectrum. Brain Dev 2003;25:166–72.

Rutherford M, Baron-Cohen S, Wheelwright S. Reading the mind in the voice: A study with normal adults and adults with Asperger syndrome and high functioning autism. J Autism Dev Disord 2002;32:189–94.

Rutter M, LeCouteur A, Lord C. Autism diagnostic interview-revised (ADI-R). Los Angeles: Western Psychological Services; 2003.

Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology 1990;1:43–6.

Schirmer A, Kotz S. Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. Trends Cogn Sci 2006;10:24–30.

Kotz A, Paulmann S. Emotion, language, and the brain. Lang Linguist Compass 2010;5:108–25.

Shriberg LD, Paul R, McSweeny JL, Klin A, Cohen DJ, Volkmar FR. Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. J Speech Lang Hear Res 2001;44:1097–115.

Taylor MJ, Baldeweg T. Application of EEG, ERP and intracranial recordings to the investigation of cognitive functions in children. Dev Sci 2002;5:318–34.

Wagner M, Watson D. Experimental and theoretical advances in prosody: A review. Lang Cogn Process 2010;25:905–45.

Wecshler D. Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III). 3rd ed. San Antonio, TX: Psychological Corporation; 2009.

Wechsler D Wechsler. Intelligence Scale for Children (WISC-III). 3rd. San Antonio, TX: Psychological Corporation; 1991.

Wetzel N, Schröger E. On the development of auditory distraction: A review. Psych J 2014;3:72–91.

Wiethoff S, Wildgruber D, Kreifelts B, Becker H, Herbert C, Grodd W, et al. Cerebral processing of emotional prosody -influence of acoustic parameters and arousal. Neuroimage 2008;39:885–93.

Witteman J, Van Heuven V, Schiller N. Hearing feelings: A quantitative meta-analysis on the neuroimaging literature of emotional prosody perception. Neuropsychologia 2012;50:2752–63.

World Health Organization. International classification of diseases (ICD-10). 10th ed. Geneva: World Health Organization (WHO); 1993.

Zilbovicius M, Garreau B, Samson Y. Delayed maturation of the frontal cortex in childhood autism. Am J Psychiatry 1995;152:248–52.