



**HAL**  
open science

## Atypical sound discrimination in children with ASD as indicated by cortical ERPs

Aurélie Bidet-Caulet, Marianne Latinus, Sylvie Roux, Joëlle Malvy,  
Frédérique Bonnet-Brilhault, Nicole Bruneau

### ► To cite this version:

Aurélie Bidet-Caulet, Marianne Latinus, Sylvie Roux, Joëlle Malvy, Frédérique Bonnet-Brilhault, et al.. Atypical sound discrimination in children with ASD as indicated by cortical ERPs. *Journal of Neurodevelopmental Disorders*, 2017, 9, pp.13. 10.1186/s11689-017-9194-9 . inserm-01502053

**HAL Id: inserm-01502053**

**<https://inserm.hal.science/inserm-01502053>**

Submitted on 5 Apr 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RESEARCH

Open Access



# Atypical sound discrimination in children with ASD as indicated by cortical ERPs

Aurélie Bidet-Caulet<sup>2†</sup>, Marianne Latinus<sup>1\*†</sup>, Sylvie Roux<sup>1</sup>, Joëlle Malvy<sup>1,3</sup>, Frédérique Bonnet-Brilhaut<sup>1,3</sup> and Nicole Bruneau<sup>1</sup>

## Abstract

**Background:** Individuals with autism spectrum disorder (ASD) show a relative indifference to the human voice. Accordingly, and contrarily to their typically developed peers, adults with autism do not show a preferential response to voices in the superior temporal sulcus; this lack of voice-specific response was previously linked to atypical processing of voices. In electroencephalography, a slow event-related potential (ERP) called the fronto-temporal positivity to voice (FTPV) is larger for vocal than for non-vocal sounds, resulting in a voice-sensitive response over right fronto-temporal sites. Here, we investigated the neurophysiological correlates of voice perception in children with and without ASD.

**Methods:** Sixteen children with autism and 16 age-matched typically developing children heard vocal (speech and non-speech) and non-vocal sounds while their electroencephalographic activity was recorded; overall IQ was smaller in the group of children with ASD. ERP amplitudes were compared using non-parametric statistical tests at each electrode and in successive 20-ms time windows. Within each group, differences between conditions were assessed using a non-parametric Quade test between 0 and 400 ms post-stimulus. Inter-group comparisons of ERP amplitudes were performed using non-paired Kruskal-Wallis tests between 140 and 180 ms post-stimulus.

**Results:** Typically developing children showed the classical voice-sensitive response over right fronto-temporal electrodes, for both speech and non-speech vocal sounds. Children with ASD did not show a preferential response to vocal sounds. Inter-group analysis showed no difference in the processing of vocal sounds, both speech and non-speech, but significant differences in the processing of non-vocal sounds over right fronto-temporal sites.

**Conclusions:** Our results demonstrate a lack of voice-preferential response in children with autism spectrum disorders. In contrast to observations in adults with ASD, the lack of voice-preferential response was attributed to an atypical response to non-vocal sounds, which was overall more similar to the event-related potentials evoked by vocal sounds in both groups. This result suggests atypical maturation processes in ASD impeding the specialization of temporal regions in voice processing.

**Keywords:** Voice, FTPV, Speech, Autism, Development, Auditory

## Background

As initially described by Kanner [1], autism is characterized by two core features: (i) impairments in social interactions and abnormal development of verbal and non-verbal communication and (ii) repetitive and ritualized behaviors associated with a restricted range of interests [2]. Recently, and consistent with Kanner's first

observations, abnormal reactivity to sensory stimulations, including sounds, has been introduced in the diagnostic criteria of autism in the DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, 5th edition, [2]), highlighting atypical processing of environmental sounds. Kanner also emphasized the relative indifference of his patients to the human voice. He writes of one of his patients, "He did not register any change of expression when spoken to," and of another, "He did not respond to being called or to any other words addressed to him." Indeed, a striking characteristic of children with

\* Correspondence: mlatinus@gmail.com

†Equal contributors

<sup>1</sup>UMR Inserm U930, Université François Rabelais de Tours, Tours, France  
Full list of author information is available at the end of the article



autism is their poor orienting to the human voice [3–5]. For example, when given a choice between their mothers' speech and a mixture of environmental noises, children with autism either show a lack of orientation for either sound or an active interest in environmental noises only [3, 6]. Experimental investigations using event-related potentials (ERPs) and an oddball paradigm have shown that, in contrast to typically developed (TD) children, children with autism spectrum disorder (ASD) aged 3–4 years [7] or aged 6–12 years [8] do not automatically orient their attention to vocal stimuli. In addition, a brain-imaging study using functional magnetic resonance imaging (fMRI) showed no difference in brain activation to voice (speech and non-speech) and to non-voice stimuli in adults with ASD, mainly attributed to a decreased response to vocal sounds [9]; this result has recently been refuted on a larger sample of ASD participants without accompanying intellectual impairment [10]. Taken together, these results suggest atypical processing of voice stimuli in patients with ASD, which could be central to the deficits in social interaction and communication.

This lack of voice-sensitivity in ASD contrasts with the observation of “voice-sensitive areas”—that is, brain regions that are more activated by vocal than non-vocal stimuli—that have been identified along the upper bank of the superior temporal sulcus (STS) in healthy adults, with greater sensitivity on the right than on the left hemisphere, in several fMRI studies [11–14]. Using electroencephalography (EEG) or magnetoencephalography (MEG), a voice-sensitive response, discriminating between vocal (speech and/or non-speech) and non-vocal stimuli, was identified at fronto-temporal sites, predominantly over the right hemiscalp, in an early latency range between 100 and 300 ms after stimulus onset [15–18]. The voice-sensitive response was mainly driven by a fronto-temporal positivity to voice (FTPV; [19]), a slow event-related potential larger to vocal than non-vocal sounds and thought to reflect the activation of the “voice-selective areas” [17, 18]. Importantly, automatic voice processing has been identified in typically developing children and infants from a very early age. Studies using fMRI and near infrared spectroscopy (NIRS) have suggested that the voice-sensitive brain systems emerged between 4 and 7 months of age [20–22]. Using EEG, the FTPV was found to begin as early as 60 ms after stimulus onset, in 4- to 5-year-old typically developing children, passively hearing human vocal and environmental sounds [19].

The FTPV occurs at the latencies of the successive peaks of the typical auditory response elicited either at temporal or at fronto-central sites. Indeed, child auditory-evoked potential waveforms are characterized by a large response recorded temporally, the T-response,

clearly dissociated from successive positive-negative fronto-central responses peaking, respectively, around 100 (P100) and 200 ms (N250) [23, 24]. The temporal response, prominent in children, consists of three successive deflections: a first negative peak, the Na or N1a; followed by a positive deflection, named Ta; and finally a negative deflection, the Tb or N1c [23–28]. The biphasic Ta-Tb response was initially described in adults and named the T-complex [29–31] while others referred to them as N1a and N1c peaks afterwards [32, 33]; in the current manuscript, we consider all three deflections and refer to them as the “T-response.” The FTPV overlaps the T-response typically recorded over bilateral sites within the first 300 ms following the presentation of a non-vocal sound; therefore, the T-response appears reduced for vocal sounds [19]. On the contrary, the fronto-central response is not voice sensitive [19]. The T-response to pure tones has been shown to be atypical in children with ASD [28, 34] and linked to communicative impairments.

Therefore, the aim of the present study was to investigate voice processing in children with ASD aged 7–12 years by comparing cortical ERPs to vocal sounds (both non-speech and speech sounds) and non-vocal sounds. We hypothesized that, contrary to age-matched typically developing children, children with ASD will not differentially process vocal and non-vocal stimuli. However, whether this lack of sensitivity is related to atypical voice processing or atypical processing of non-vocal sounds remains an open question.

## Methods

### Participants

Sixteen children with ASD (15 boys and 1 girl) aged from 7 years 8 months to 12 years 2 months (mean age  $\pm$  standard deviation 10 years 6 months  $\pm$  1 year 5 months) participated in the study. They were recruited from the Child Psychiatry Centre of the University Hospital of Tours. Children with neurological disorders (including seizures), physical abnormalities, neurologic impairment in motor or sensory function, or genetically defined disorders were excluded. All had normal hearing, verified by subjective or objective (when necessary) audiometric tests performed before ERP recordings.

Diagnosis of ASD was made by experienced clinicians according to DSM-IV-R criteria [35] at the time of electrophysiological recordings and using the Autism Diagnostic Observation Schedule-Generic (ADOS-G; [36]) and/or the Autism Diagnostic Interview-Revised (ADI-R; [37]). Developmental quotients (DQs) were assessed by the Echelles Différentielles d'Efficiences Intellectuelles (EDEI-R; [38]) or the Wechsler Intelligence Scale for Children (WISC III and WISC IV). These two developmental scales provide verbal developmental (vDQ;

mean  $\pm$  SD  $69 \pm 25$ ) and non-verbal developmental (nvDQ; mean  $\pm$  SD  $85 \pm 18$ ) quotients.

Sixteen typically developing (TD) children (15 boys and 1 girl, mean age  $\pm$  standard deviation 10 years 5 months  $\pm$  1 year 5 months) were matched in age and gender with the patients. All typically developed children had normal education level and language development. The Ethics Committee of the University Hospital of Tours approved the protocol (Comité de Protection des Personnes (CPP) Tours Ouest 1; n°2006-R5). Signed informed consent was obtained from parents, and assent was given by the children.

## Paradigm

### Stimuli

The stimuli used were sounds extracted from the vocal and non-vocal sequences used in Belin et al.'s block-design fMRI studies [11, 13] (<http://vnl.psy.gla.ac.uk/resources.php>). The vocal sounds, both non-speech (Voc<sup>NSp</sup>) (e.g., laughing, sighing, and coughing) and speech (Voc<sup>Sp</sup>) (syllables in several languages, e.g., English, Finnish, Arabic), were produced by a large number of speakers of both genders and of different ages. No French words were included in the experiment so as to prevent the influence of linguistic processing on the ERPs because of inherent differences in language development across the two groups. Overall, there were 53 unique vocal non-speech sounds and 67 unique vocal speech sounds. Non-vocal sounds (NVoc) consisted of sounds from a wide variety of sources, including the human environment (such as telephones, alarms, cars), musical instruments (such as bells and orchestral instruments), and nature (such as streams, wind, animal sounds); in total, there were 160 unique non-vocal sounds. Sound duration was adjusted to 500 ms, and an envelope of 50 ms decay time was applied to the end of each sound to minimize clicks at sound offset. All sounds were normalized according to the root mean square of their amplitude.

The stimuli used in this ERP study were previously selected in a pretest session performed with 4- to 5-year-old typically developing children; see [19].

To better characterize the sound categories at the acoustic level, analyses of sound power in the temporal and spectral domains were performed at each time or frequency bin (11.6 ms; 43 Hz) using a statistical test based on randomization [39] and controlling the false discovery rate to correct for multiple comparisons. Randomization consisted of (1) the random constitution of the two samples to compare, (2) the sum of squared sums of values in the two obtained samples, and (3) the computation of the difference between these two statistic values. We performed 50,000 such randomizations to obtain an estimate of the distribution of this difference

under the null hypothesis. From this distribution, we estimated the threshold corresponding to a significant difference between two conditions; this threshold was then compared to the empirical difference between the values in the two conditions. This analysis highlighted significant differences between sound categories both in frequency and time domains (see Fig. 1).

### Design

Stimuli were delivered with Presentation® (Neurobehavioral Systems Inc.), through two loudspeakers placed 1.20 m in front of the subject at approximately 10° on both sides of the interaural axis. Overall stimulus intensity was adjusted to 70 dB SPL at the subject's head. Stimuli were presented with a constant interstimulus interval (offset to onset) of 700 ms.

Two successive sequences were delivered in alternating order between subjects. A sequence comprised one block of stimulation repeated three times. In the vocal sequence, 582 vocal stimuli, 44% non-speech, were presented as standard and non-vocal sounds as deviants (15%). In the non-vocal sequence (NVoc), 582 non-vocal stimuli were presented as standard and vocal sounds as deviants (15%). The present study only reports ERPs recorded to standard stimuli, i.e., Voc<sup>NSp</sup> and Voc<sup>Sp</sup> in the first sequence and NVoc in the second sequence.

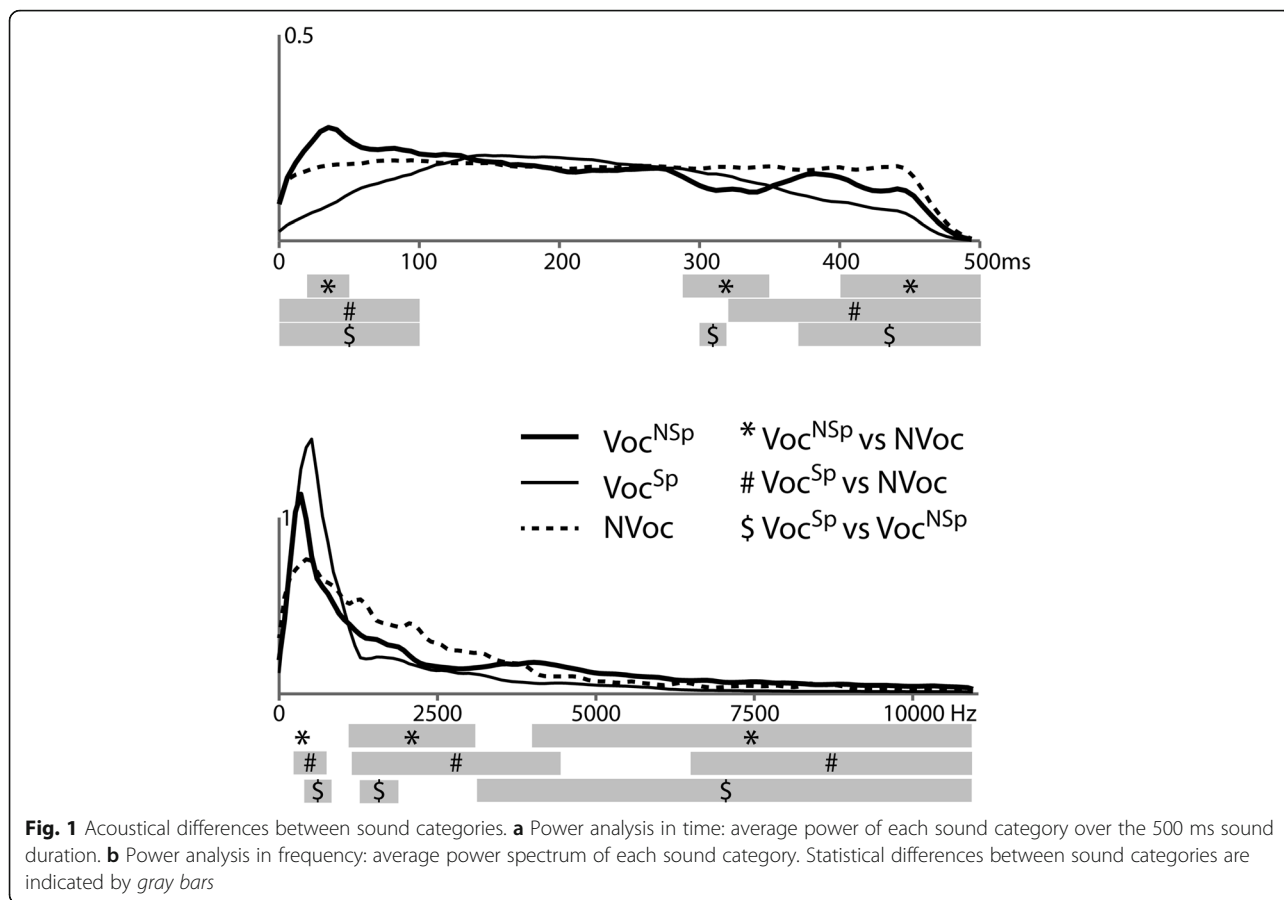
### EEG recording

During the recording session, children sat in a comfortable armchair in a dimly lit, soundproofed room and watched a silent video of their choice.

Auditory event-related potentials were collected using the NeuroScan electrophysiological data acquisition system (SCAN 4.3). In all TD children and in 11 children with ASD, the electroencephalogram was recorded from 28 Ag/AgCl scalp electrodes referenced to the nose. Five children with ASD could only tolerate the placement of 11 scalp electrodes: Fz, Cz, Pz, F7, F8, T7, T8, T5, and T6 placed according to the International 10-20 System and M1 and M2 on the left and right mastoid sites, respectively. Vertical electrooculogram (EOG) activity was recorded from electrodes placed above and below the right eye. All electrode impedance levels were kept below 10 k $\Omega$ . The EEG and EOG were amplified with an analog band-pass filter (0.3–70 Hz; slope 6 dB/octave) and sampled at 500 Hz.

### Data analysis

EEG data were preprocessed within NeuroScan (Compumedics Inc., Neuroscan, 2003). Epochs recorded in response to standard stimuli immediately following a deviant stimulus (i.e., 102 trials) were excluded from the analyses in order to focus on responses to frequent standard stimuli. Epochs corresponding to responses to



specific stimuli, e.g., stimuli deemed ambiguous a posteriori (corresponding to 144 trials in the vocal sequence), and animal vocalization (57 trials in the non-vocal sequence) were excluded from the analysis. Animal vocalizations were excluded as they have the same physiological origin as human vocal sounds. Automatic correction of eye movements was then applied. Eye-movement artifacts were eliminated using a spatial filter developed by NeuroScan (Compumedics Inc., Neuroscan, 2003). The spatial filter is a multi-step procedure that generates an average eye blink, utilizes a spatial singular value decomposition based on principal component analysis (PCA) to extract the first component and covariance values, and then uses those covariance values to develop a filter that retains the EEG activity of interest. EEG periods with movement artifacts were manually rejected. After rejection, the averaged numbers of trials ( $\pm$ SEM) were  $93 \pm 14$ ,  $120 \pm 18$ , and  $274 \pm 37$  in children with ASD and  $108 \pm 17$ ,  $138 \pm 22$ , and  $321 \pm 49$  in TD children, for  $Voc^{NSp}$ ,  $Voc^{Sp}$ , and NVoc, respectively. There were significantly fewer trials in ASD children than in TD children in each condition ( $p < 0.05$ ) as assessed using Kruskal-Wallis tests. However, in all conditions, the average number of trials in the ERPs was always greater than 90, which is sufficient

to accurately measure auditory ERPs in children. For each subject and each stimulus type, ERPs were averaged and baseline corrected according to a 100 ms prestimulus period. A digital zero-phase-shift low-pass filter (30 Hz) was then applied. ERPs were analyzed with the ELAN<sup>®</sup> software [40]. As indicated in previous electrophysiological data [15, 41], responses specifically elicited by voice compared to non-voice occur early; the analysis therefore focused on the first 400 ms after stimulus onset.

Scalp ERP topographies displayed in Fig. 4 were generated using a two-dimensional spherical spline interpolation [42] and a radial projection from Cz (top views), which respects the length of the meridian arcs [42, 43].

**Statistical analysis**

To limit assumptions regarding the data distribution, non-parametric statistical tests were used to compare ERP amplitudes. Differences between conditions ( $Voc^{NSp}$ ,  $Voc^{Sp}$ , and NVoc) within each group were assessed using a non-parametric Quade test at each of the 11 electrodes and in successive 20-ms time windows between 0 and 400 ms post-stimulus. In order to correct for multiple comparisons, results are reported at

threshold corresponding to  $p < 0.00025$  (Bonferroni correction according to the number of tested electrodes and time-windows). Two-by-two post hoc comparisons were assessed using a non-parametric paired Wilcoxon test at the significant electrodes and time windows found with the Quade test.

Inter-groups comparisons of  $Voc^{NSp}$ ,  $Voc^{Sp}$ , and NVoc ERP amplitudes were performed using non-parametric non-paired Kruskal-Wallis tests at each of the 11 electrodes and in successive 20-ms time windows between 140 and 180 ms post-stimulus. The 140–180-ms time window corresponds to the significant voice effect at right fronto-temporal electrode sites in TD children.

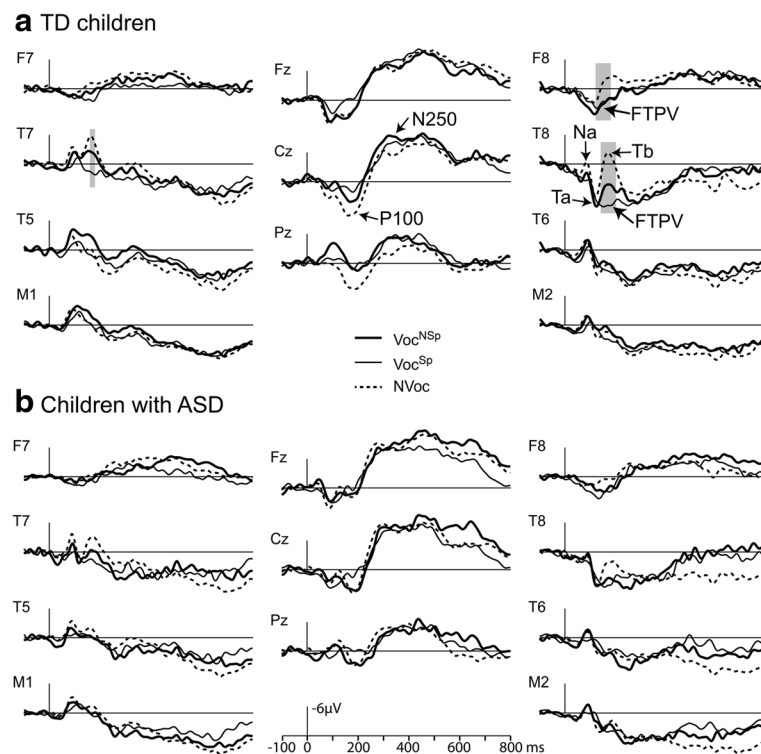
**Results**

Figure 2 presents the grand-averaged ERPs to Voc (either  $Voc^{NSp}$  or  $Voc^{Sp}$ ) and NVoc stimuli recorded from the 11 electrode sites in the two groups of children. Visual inspection of the grand average shows that ERP waveforms varied according to conditions much more in TD children than in children with ASD. This was particularly striking over right fronto-temporal sites (F8 and T8) and was confirmed by results of statistical analyses.

**TD children**

ERPs elicited by the three stimulus categories over fronto-central midline sites (Fig. 2a) displayed similar waveforms with two successive positive-negative deflections peaking at around 200 and 400 ms, respectively, with polarity reversal at mastoid sites, in particular on the left side. ERPs recorded over fronto-temporal sites displayed different waveforms according to stimulus conditions. This was particularly marked at right temporal electrodes (T8): NVoc stimuli evoked the T-response whose peaks culminated at around 80, 120, and 165 ms. This T-response was different in response to Voc stimuli with reduced (for  $Voc^{NSp}$  stimuli) or absent (for  $Voc^{Sp}$  stimuli) Tb peak and more positive preceding peaks (Na and Ta). At the right frontal F8 electrode, ERPs to Voc appeared as a more sustained positive deflection than the response to NVoc stimuli, suggesting that Voc stimuli elicited a positive slow wave, better seen at the right fronto-temporal F8 electrode than at the temporal T8 site where it overlaps the T-response. This positive slow wave previously observed in adults [15–18] and 4- to 5-year-old children [19] corresponds to the FTPV.

Statistical analyses using Quade tests (Bonferroni corrected  $p < 0.00025$ ) indicated significant differences between ERPs evoked by the different stimulus categories over right fronto-temporal sites (F8 120–180 ms, T8



**Fig. 2** ERPs evoked by vocal sounds (either non-speech:  $Voc^{NSp}$ , or speech:  $Voc^{Sp}$ ) and non-vocal sounds (NVoc) at fronto-temporal, temporo-mastoid, and central electrode sites in TD children (a) and children with ASD (b). Significant differences between conditions according to the Quade test (after Bonferroni correction) are indicated by gray bars

140–200 ms) and over a left temporal site (T7 160–180 ms). Post hoc Wilcoxon tests ( $p < 0.05$ ) indicated that over right fronto-temporal sites, Voc<sup>NSp</sup> and Voc<sup>Sp</sup> sounds elicited a greater positivity than NVoc stimuli; this positivity was greater to Voc<sup>Sp</sup> than to Voc<sup>NSp</sup> at T8. Over T7 (160–180 ms), Voc<sup>Sp</sup> sounds elicited a larger positivity than Voc<sup>NSp</sup> and NVoc stimuli.

In summary, a significant FTPV was observed over right fronto-temporal sites in response to vocal speech and non-speech stimuli in comparison to non-vocal sounds in 7–12-aged TD children.

### Children with ASD

As shown on Fig. 2b, ERP waveforms were rather similar for the three conditions in children with ASD at all recorded electrodes. ERPs elicited at fronto-central midline sites displayed the two successive positive-negative deflections also found in TD children. Over temporal sites, irrespective of the stimulus condition, ERPs showed a small Na deflection peaking at around 75 ms, followed by a large slow positive deflection.

Quade tests indicated no significant stimulus-related differences between ERPs to Voc<sup>NSp</sup>, Voc<sup>Sp</sup>, and NVoc sounds. Therefore, in 7–12-aged children with ASD, responses to vocal and non-vocal sounds did not seem to differ as much as in TD children.

### Children with ASD vs. TD children

Group differences were significant for ERPs to NVoc sounds within the significant time window of the FTPV in TD children (140–180 ms). Children with ASD presented a smaller P100 at a central site (Cz 140–160 ms) and a smaller right fronto-temporal negative Tb peak (T8 140–180 ms, F8 160–180 ms) than TD children ( $p < 0.05$ , Figs. 3 and 4). No significant difference was found, between ASD and TD children, in the amplitude of ERPs to Voc<sup>Sp</sup> and Voc<sup>NSp</sup> sounds.

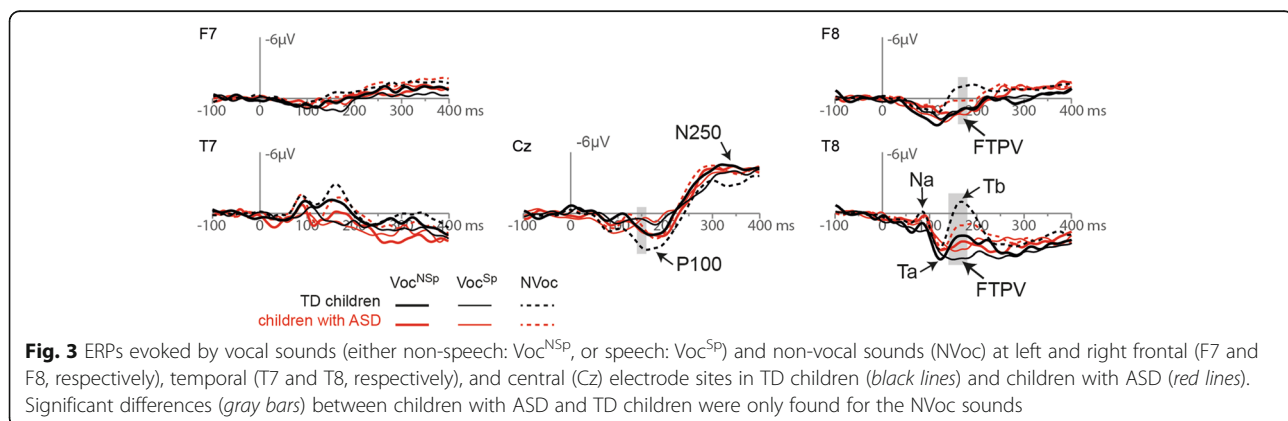
## Discussion

The aim of this study was to investigate voice processing in children with ASD. To that aim, we recorded brain activity evoked by vocal and non-vocal sounds in children with ASD and age-matched typically developed children. The results showed differences in the response to vocal and non-vocal sounds in TD children but not in children with ASD. Comparison across groups revealed significant differences in the response to non-vocal sounds but not to vocal sounds, highlighting that the lack of difference was attributable to atypical processing of non-vocal sounds.

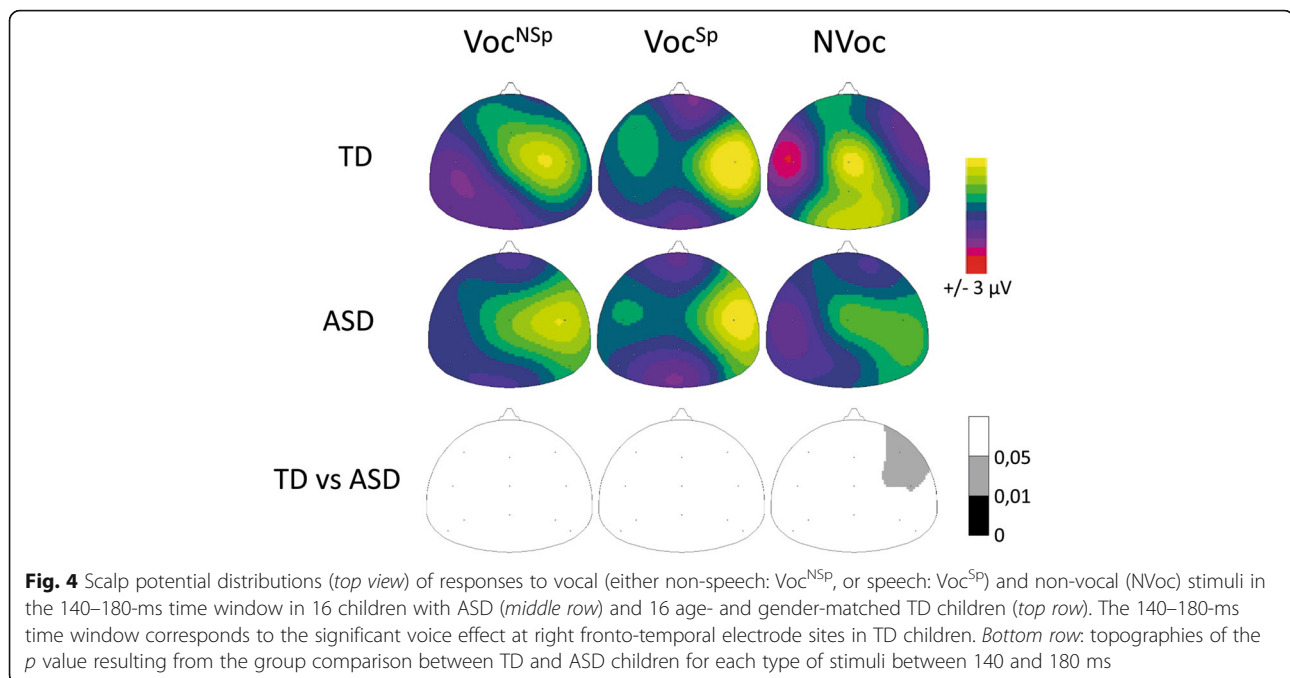
### Processing of vocal and non-vocal stimuli in TD children

In TD children aged 7 to 12 years, non-vocal stimuli elicited the classical T-response, as previously observed in younger TD children aged 4–5 years in response to the same non-vocal sounds [19]. In contrast, vocal stimuli elicited a different pattern over temporal sites with a flattened Tb peak. This result is consistent with a study showing a modulation of Tb amplitude by “speechness” [26]. A thorough observation of the Tb peak reported in studies using speech-like sounds [27] to those using other non-speech sounds [23–25] suggests a similar modulation of the Tb peak by “voiceness.”

We suggest that this flattened T-response to vocal sounds is due to a superimposed positive slow wave spreading over right fronto-temporal sites. This slow wave corresponds to the fronto-temporal positivity to voice (FTPV) previously described in younger children [19] and adults, e.g., [15, 41]; it may also correspond to the early part of the lateral anterior positivity (LAP) that has been associated with the analysis of discourse [44, 45]. The T-response was especially flattened for vocal speech sounds, as illustrated by significant differences within the voice category over bilateral temporal sites in the latency range of the Tb peak. Vocal non-speech sounds elicited a right-lateralized FTPV while speech stimuli elicited a bilateral FTPV. This is consistent with



**Fig. 3** ERPs evoked by vocal sounds (either non-speech: Voc<sup>NSp</sup>, or speech: Voc<sup>Sp</sup>) and non-vocal sounds (NVoc) at left and right frontal (F7 and F8, respectively), temporal (T7 and T8, respectively), and central (Cz) electrode sites in TD children (black lines) and children with ASD (red lines). Significant differences (gray bars) between children with ASD and TD children were only found for the NVoc sounds



previous findings in younger children [19] and in adults since De Lucia et al. [18] and Bruneau et al. [15] described a right-lateralized FTPV to vocal non-speech sounds whereas the FTPV was recorded bilaterally in studies using both vocal speech and non-speech stimuli [16, 17]. ERP differences between both vocal and non-vocal sounds are unlikely to be accounted for by acoustic differences between the stimuli. Before 100 ms, the acoustic power of non-vocal sounds fits in between the acoustic powers of the two vocal sound categories (Fig. 1). Moreover, acoustic differences are also present between the two vocal sound categories, although they evoked similar ERPs.

Previous studies suggested that the FTPV corresponds to the activation of the temporal voice areas described in fMRI studies with adult participants, bilaterally localized in the anterior STS with a right hemisphere predominance [11–13, 46–50]. Consistent with our observations, activity linked to processing the non-verbal aspect of vocal stimuli (e.g., vocal non-speech sounds) is also lateralized to the right hemisphere [13, 47]. Similarly, an fMRI study with 7-month-old infants showed a preferential response to non-speech vocal sounds in the right anterior temporal cortex [20]. Thus, based on the literature [19], we hypothesize that the FTPV observed in children also reflects the activation of the temporal voice areas, although this should be explored with more appropriate methods than the low-density montage used here and non-vocal stimuli matched in spectro-temporal complexity.

The FTPV, as the early LAP, seems to reflect the processing of “voiceness.” This initial analysis of “voiceness”

may feed into the late LAP proposed to reflect discourse processing [44].

#### Processing of vocal and non-vocal stimuli in children with ASD

In contrast to the ERP results found in TD children, children with ASD display few stimulus-related ERP differences. The lack of difference between the ERPs evoked by vocal and non-vocal sounds in ASD children could relate to changes in sensitivity to stimulus-type differences, sustained attention, or developmental quotient. In the current study, most of the children with ASD presented accompanying intellectual disabilities; in order to include these children in the protocol, a passive listening task was used. This ensured that differences in developmental quotient did not influence task comprehension and execution. Participants were not required to perform a task with the auditory stimuli. They were watching a silent video so as to reduce potential confounds in the allocation of attentional resources. It has been shown that TD children automatically attend to speech stimuli, even when watching silent videos, albeit to a smaller extent than when explicitly directing their attention toward the auditory modality [51]. On the contrary, this automatic orientation to speech stimuli appears reduced in children with specific language impairment [51]. Another study has demonstrated that children with ASD showed impaired attention orienting to speech sounds, associated with intact sensory processing [8]. However, an impaired attention orienting is unlikely to explain our results since, in the present



study, group differences were only observed for non-vocal sounds. The most likely explanation of these differences is that participants with ASD are differently sensitive to stimulus-type differences. Children with autism present abnormal responses to all sounds, both at behavioral [1, 52] and neural levels [28, 34]. An abnormal response to sounds could hinder their processing and prevent the discrimination between vocal and non-vocal sounds. This finding is in agreement with a previous fMRI study performed with similar stimulations in adults with ASD, showing no difference in brain activation between voice and non-voice stimuli [9], although it contradicts more recent evidence in adults with ASD with no intellectual disabilities [10]. These fronto-temporal responses are clearly dissociated from the fronto-central response, which appears as a prominent positivity in children around 100–200 ms. The fronto-central response to non-vocal sounds differed between TD children and children with ASD, although it does not discriminate vocal from non-vocal sounds in TD children [19], again highlighting atypical processing of non-vocal sounds.

Interestingly, the present results clearly showed that the ERPs elicited by vocal stimuli are strongly similar between children with ASD and TD children. On the contrary, ERPs to non-vocal stimuli greatly differ in children with ASD and in TD children over central and right fronto-temporal sites. The present results are in contradiction to previous fMRI findings showing, in five adult patients with ASD, similar responses to non-vocal sounds in TD adults and adults with ASD, but different activations to vocal sounds (speech and non-speech), bilaterally along the upper bank of the STS [9]. This discrepancy could arise from the ASD populations tested (e.g., with or without accompanying intellectual impairment), the sample size (only five adults in the fMRI study), and the age of the participants (adults vs. children). This discrepancy echoes data from the face perception literature [53–55]. While children with ASD between 4 and 5 years old showed typical patterns of amplitude in ERPs to facial stimuli and an abnormal response to visual objects [53], in adult participants, group differences between ASD and TD participants were observed mainly in the response to faces [54]. Accordingly, while ERP amplitude to facial emotion differed between ASD and TD adults [55], no group differences were observed in children [55, 56]. Hence, the difference between studies with adults with ASD and children with ASD appear to reflect genuine maturational processes, rather than methodological differences.

Models of the development of face processing argue that brain regions involved in face processing are not initially face-specific and are activated by a broad range of stimuli; with time, their activity tunes to the type of stimuli mostly seen in the environment, e.g., upright

faces, and yield to face-specific activity [57, 58]. Accordingly, we propose that prolonged experience with vocal stimuli during typical development allows the development of “filters” which in time permit the not-initially voice-sensitive regions of the STS to tune their activity toward processing voices. These filters would be derived from the acoustical information present in vocal sounds and may correspond to the voice configuration [59]. Consistent with this hypothesis, it has been shown that the underlying components of the T-response to a vowel sound mature at different rates with the Ta and Tb peaks still not being fully mature at age 8 [60]. This multi-step maturation of the T-response might reflect the development of the “filters” that allows optimizing voice processing. In contrast, the development of these “filters” would be impaired in autism and the STS would not tune to a specific sound category resulting in a lack of specialization in processing voices. Consequently, the activity of the STS would also be strong for non-vocal sounds in children with ASD, resulting in the observation of slow wave components on fronto-temporal electrodes in response to non-vocal sounds; this would in turn yield to a flattened T-response and an absence of stimulus-type-related difference. This observation is consistent with some theories of autism such as the social motivation theory [61] or the weak central coherence (WCC)/enhanced perceptual functioning (EPF) [62, 63]. The social motivation theory stipulates that ASD results from an imbalance in attending to social and non-social stimuli, while the WCC and EPF theories propose that ASD is the consequence of atypical processing of stimuli, resulting from an imbalance between local and global processing. These theories are not mutually exclusive, and whether one or the other yields the current results remains to be tested. Nonetheless, recent evidence points toward impaired global processing of vocal sounds, as people with ASD fail to combine the acoustic features into a coherent percept [52, 64].

## Conclusions

Children with ASD displayed a similar response to vocal and non-vocal sounds over right fronto-temporal sites; therefore, they did not show a voice-preferential response. This lack of difference appears to be driven by an atypical processing of non-vocal sounds.

## Abbreviations

ADI-R: Autism Diagnostic Interview-Revised; ADOS: Autism Diagnostic Observation Schedule; ASDs: Autism spectrum disorders; DQs: Developmental quotients; DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th edition; EDEI-R: Echelles Différentielles d'Efficiences Intellectuelles; EEG: Electroencephalography; ERP: Event-related potential; fMRI: Functional magnetic resonance imaging; FTPV: Fronto-temporal positivity to voice; MEG: Magnetoencephalography; nvDQ: Non-verbal developmental quotient; NVoc: Non-vocal sounds; TD: Typically developed; vDQ: Verbal developmental quotient; Voc<sup>NSP</sup>: Vocal non-speech sounds; Voc<sup>SP</sup>: Vocal speech sounds; WISC: Wechsler Intelligence Scale for Children

**Acknowledgements**

The authors would like to thank Luce Corneau for her help with participant testing. We thank all the volunteers for their time and effort while participating in this study.

**Funding**

This study received financial support from a Regional Clinical Research Program (PHRIO6-VHA) for collection and analysis of the data.

**Availability of data and materials**

Please contact the corresponding author for data requests.

**Authors' contributions**

The study presented here results from teamwork. NB conceptualized the study. SR and NB designed the study. NB, ML, JM, and FBB carried-out the research. NB, SR, ML, and ABC were responsible for data analyses. ABC, ML, SR, and NB were responsible for the statistical analyses. ML, NB, and ABC were responsible for drafting the manuscript. JM and FBB were responsible for clinical evaluation of the patients. All authors read and approved the final manuscript.

**Competing interests**

The authors declare that they have no competing interests.

**Consent for publication**

Not applicable.

**Ethics approval and consent to participate**

The Ethics Committee of the University Hospital of Tours approved the protocol (Comité de Protection des Personnes (CPP) Tours Ouest 1; n°2006-R5). Signed informed consent was obtained from parents, and assent was given by the children.

**Publisher's Note**

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

**Author details**

- <sup>1</sup>UMR Inserm U930, Université François Rabelais de Tours, Tours, France.
- <sup>2</sup>Brain Dynamics and Cognition Team, Lyon Neuroscience Research Center; CRNL, INSERM U1028, CNRS UMR5292, University of Lyon 1, Lyon, France.
- <sup>3</sup>CHRU de Tours, Service de Pédiopsychiatrie, Tours, France.

Received: 18 March 2016 Accepted: 12 March 2017

Published online: 05 April 2017

**References**

1. Kanner L. Autistic disturbances of affective contact: publisher not identified. *Nervous Child*. 1943;2(3):217–50.
2. Association D-AP. *Diagnostic and Statistical Manual of Mental Disorders*. Arlington: American Psychiatric Publishing; 2013.
3. Klin A. Young autistic children's listening preferences in regard to speech: a possible characterization of the symptom of social withdrawal. *J Autism Dev Disord*. 1991;21(1):29–42.
4. Clancy H, McBride G. The autistic process and its treatment. *J Child Psychol Psychiatry*. 1969;10(4):233–44.
5. Ornitz EM, Guthrie D, Farley AH. The early development of autistic children. *J Autism Child Schizophr*. 1977;7(3):207–29.
6. Klin A, Volkmar FR, Sparrow SS. Autistic social dysfunction: some limitations of the theory of mind hypothesis. *J Child Psychol Psychiatry*. 1992;33(5):861–76.
7. Dawson G, Toth K, Abbott R, Osterling J, Munson J, Estes A, Liaw J. Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Dev Psychol*. 2004;40(2):271–83.
8. Ceponiene R, Lepisto T, Shestakova A, Vanhala R, Alku P, Naatanen R, Yaguchi K. Speech-sound-selective auditory impairment in children with autism: they can perceive but do not attend. *Proc Natl Acad Sci U S A*. 2003;100(9):5567–72.
9. Gervais H, Belin P, Boddard N, Leboyer M, Coez A, Sfaello I, Barthelemy C, Brunelle F, Samson Y, Zilbovicius M. Abnormal cortical voice processing in autism. *Nat Neurosci*. 2004;7(8):801–2.
10. Schelinski S, Borowiak K, von Kriegstein K. Temporal voice areas exist in autism spectrum disorder but are dysfunctional for voice identity recognition. *Soc Cogn Affect Neurosci*. 2016;11(11):1812–22.
11. Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex. *Nature*. 2000;403(6767):309–12.
12. Pernet CR, McAleer P, Latinus M, Gorgolewski KJ, Charest I, Bestelmeyer PE, Watson RH, Fleming D, Crabbe F, Valdes-Sosa M, et al. The human voice areas: spatial organization and inter-individual variability in temporal and extra-temporal cortices. *Neuroimage*. 2015;119:164–74.
13. Belin P, Zatorre RJ, Ahad P. Human temporal-lobe response to vocal sounds. *Brain Res Cogn Brain Res*. 2002;13(1):17–26.
14. Belin P, Fecteau S, Bedard C. Thinking the voice: neural correlates of voice perception. *Trends Cogn Sci*. 2004;8(3):129–35.
15. Bruneau N, Roux S, Clery H, Rogier O, Bidet-Caulet A, Barthelemy C. Early neurophysiological correlates of vocal versus non-vocal sound processing in adults. *Brain Res*. 2013;1528:20–7.
16. Charest I, Pernet CR, Rousselet GA, Quinones I, Latinus M, Fillion-Bilodeau S, Chartrand JP, Belin P. Electrophysiological evidence for an early processing of human voices. *BMC Neurosci*. 2009;10:127.
17. Capilla A, Schoffelen JM, Paterson G, Thut G, Gross J. Dissociated alpha-band modulations in the dorsal and ventral visual pathways in visuospatial attention and perception. *Cereb Cortex*. 2013;23(6):1388–95.
18. De Lucia M, Clarke S, Murray MM. A temporal hierarchy for conspecific vocalization discrimination in humans. *J Neurosci*. 2010;30(33):11210–21.
19. Rogier O, Roux S, Belin P, Bonnet-Brilhault F, Bruneau N. An electrophysiological correlate of voice processing in 4- to 5-year-old children. *Int J Psychophysiol*. 2010;75(1):44–7.
20. Blasi A, Mercure E, Lloyd-Fox S, Thomson A, Brammer M, Sauter D, Deeley Q, Barker GJ, Renvall V, Deoni S, et al. Early specialization for voice and emotion processing in the infant brain. *Curr Biol*. 2011; 21(14):1220–4.
21. Grossmann T, Oberecker R, Koch SP, Friederici AD. The developmental origins of voice processing in the human brain. *Neuron*. 2010;65(6):852–8.
22. Belin P, Grosbras MH. Before speech: cerebral voice processing in infants. *Neuron*. 2010;65(6):733–5.
23. Ponton C, Eggermont JJ, Khosla D, Kwong B, Don M. Maturation of human central auditory system activity: separating auditory evoked potentials by dipole source modeling. *Clin Neurophysiol*. 2002;113(3):407–20.
24. Bruneau N, Roux S, Guerin P, Barthelemy C, Lelord G. Temporal prominence of auditory evoked potentials (N1 wave) in 4–8-year-old children. *Psychophysiology*. 1997;34(1):32–8.
25. Bruneau N, Bidet-Caulet A, Roux S, Bonnet-Brilhault F, Gomot M. Asymmetry of temporal auditory T-complex: right ear-left hemisphere advantage in Tb timing in children. *Int J Psychophysiol*. 2015;95(2):94–100.
26. Groen MA, Alku P, Bishop DV. Lateralisation of auditory processing in Down syndrome: a study of T-complex peaks Ta and Tb. *Biol Psychol*. 2008;79(2):148–57.
27. Shafer VL, Schwartz RG, Martin B. Evidence of deficient central speech processing in children with specific language impairment: the T-complex. *Clin Neurophysiol*. 2011;122(6):1137–55.
28. Bruneau N, Bonnet-Brilhault F, Gomot M, Adrien JL, Barthelemy C. Cortical auditory processing and communication in children with autism: electrophysiological/behavioral relations. *Int J Psychophysiol*. 2003;51(1):17–25.
29. Wolpaw JR, Penry JK. A temporal component of the auditory evoked response. *Electroencephalogr Clin Neurophysiol*. 1975;39(6):609–20.
30. Celesia GG. Organization of auditory cortical areas in man. *Brain*. 1976;99(3):403–14.
31. Wood CC, Wolpaw JR. Scalp distribution of human auditory evoked potentials. II. Evidence for overlapping sources and involvement of auditory cortex. *Electroencephalogr Clin Neurophysiol*. 1982;54(1):25–38.
32. McCallum WC, Curry SH. The form and distribution of auditory evoked potentials and CNVs when stimuli and responses are lateralized. *Prog Brain Res*. 1980;54:767–75.
33. Näätänen R, Picton T. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*. 1987;24(4):375–425.
34. Bruneau N, Roux S, Adrien JL, Barthelemy C. Auditory associative cortex dysfunction in children with autism: evidence from late auditory evoked potentials (N1 wave-T complex). *Clin Neurophysiol*. 1999;110(11):1927–34.
35. Association D-I-TAP. *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. Washington: American Psychiatric Association; 2000.

36. Lord C, Risi S, Lambrecht L, Cook Jr EH, Leventhal BL, DiLavore PC, Pickles A, Rutter M. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord*. 2000;30(3):205–23.
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(5):659–85.
38. Perron-Borelli M. Echelles différentielles d'efficacités intellectuelles. Forme révisée (EDEI-R). Issy-les-Moulineaux: Etablissements d'applications psychotechniques EAP; 1996.
39. Edgington ES. Randomization tests. 3rd ed. New York: Marcel Dekker; 1995.
40. Aguera PE, Jerbi K, Caclin A, Bertrand O. ELAN: a software package for analysis and visualization of MEG, EEG, and LFP signals. *Comput Intell Neurosci*. 2011;2011:158970.
41. Bruckert L, Bestelmeyer PEG, Latinus M, Rouger J, Charest I, Rousselet GA, Kawahara H, Belin P. Vocal attractiveness increases by averaging. *Curr Biol*. 2010;20(2):116–20.
42. Perrin F, Pernier J, Bertrand O, Echallier JF. Spherical splines for scalp potential and current density mapping. *Electroencephalogr Clin Neurophysiol*. 1989;72(2):184–7.
43. Perrin F, Pernier J, Bertrand O, Giard MH, Echallier JF. Mapping of scalp potentials by surface spline interpolation. *Electroencephalogr Clin Neurophysiol*. 1987;66(1):75–81.
44. Shafer VL, Kessler KL, Schwartz RG, Morr ML, Kurtzberg D. Electrophysiological indices of brain activity to "the" in discourse. *Brain Lang*. 2005;93(3):277–97.
45. Shafer VL, Schwartz RG, Morr ML, Kessler KL, Kurtzberg D. Deviant neurophysiological asymmetry in children with language impairment. *Neuroreport*. 2000;11(17):3715–8.
46. Belin P, Zatorre RJ. Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport*. 2003;14(16):2105–9.
47. von Kriegstein K, Eger E, Kleinschmidt A, Giraud AL. Modulation of neural responses to speech by directing attention to voices or verbal content. *Brain Res Cogn Brain Res*. 2003;17(1):48–55.
48. Fecteau S, Armony JL, Joanette Y, Belin P. Is voice processing species-specific in human auditory cortex? An fMRI study. *Neuroimage*. 2004;23(3):840–8.
49. Von Kriegstein K, Giraud AL. Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage*. 2004;22:948–55.
50. Koeda M, Takahashi H, Yahata N, Asai K, Okubo Y, Tanaka H. A functional MRI study: cerebral laterality for lexical-semantic processing and human voice perception. *AJNR Am J Neuroradiol*. 2006;27(7):1472–9.
51. Shafer VL, Ponton C, Datta H, Morr ML, Schwartz RG. Neurophysiological indices of attention to speech in children with specific language impairment. *Clin Neurophysiol*. 2007;118(6):1230–43.
52. Lin IF, Agus TR, Sued C, Pressnitzer D, Yamada T, Komine Y, Kato N, Kashino M. Fast response to human voices in autism. *Sci Rep*. 2016;6:26336.
53. Webb SJ, Dawson G, Bernier R, Panagiotides H. ERP evidence of atypical face processing in young children with autism. *J Autism Dev Disord*. 2006;36(7):881–90.
54. Webb SJ, Merkle K, Murias M, Richards T, Aylward E, Dawson G. ERP responses differentiate inverted but not upright face processing in adults with ASD. *Soc Cogn Affect Neurosci*. 2012;7(5):578–87.
55. O'Connor K, Hamm JP, Kirk IJ. The neurophysiological correlates of face processing in adults and children with Asperger's syndrome. *Brain Cogn*. 2005;59(1):82–95.
56. Batty M, Meaux E, Wittemeyer K, Roge B, Taylor MJ. Early processing of emotional faces in children with autism: an event-related potential study. *J Exp Child Psychol*. 2011;109(4):430–44.
57. de Haan M, Johnson MH, Halit H. Development of face-sensitive event-related potentials during infancy: a review. *Int J Psychophysiol*. 2003;51(1):45–58.
58. Nelson CA. The development and neural bases of face recognition. *Infant Child Dev*. 2001;10(1-2):3–18.
59. Latinus M, Belin P. Perceptual auditory aftereffects on voice identity using brief vowel stimuli. *PLoS One*. 2012;7(7):e41384.
60. Shafer VL, Yu YH, Wagner M. Maturation of cortical auditory evoked potentials (CAEPs) to speech recorded from frontocentral and temporal sites: three months to eight years of age. *Int J Psychophysiol*. 2015;95(2):77–93.
61. Chevallier C, Kohls G, Troiani V, Brodwin ES, Schultz RT. The social motivation theory of autism. *Trends Cogn Sci*. 2012;16(4):231–9.
62. Happe F, Briskman J, Frith U. Exploring the cognitive phenotype of autism: weak "central coherence" in parents and siblings of children with autism: I. Experimental tests. *J Child Psychol Psychiatry*. 2001;42(3):299–307.
63. Mottron L, Dawson M, Soulières I, Hubert B, Burack J. Enhanced perceptual functioning in autism: an update, and eight principles of autistic perception. *J Autism Dev Disord*. 2006;36(1):27–43.
64. Schelinski S, Roswandowitz C, von Kriegstein K. Voice identity processing in autism spectrum disorder. *Autism Res*. 2017;10(1):155–68.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

