



Brain mechanisms involved in angry prosody change detection in school-age children and adults, revealed by electrophysiology

Judith Charpentier¹ · Klara Kovarski¹ · Sylvie Roux¹ · Emmanuelle Houy-Durand^{1,2} · Agathe Saby² · Frédérique Bonnet-Brilhaut^{1,2} · Marianne Latinus¹ · Marie Gomot¹

Published online: 7 May 2018
© Psychonomic Society, Inc. 2018

Abstract

Voices transmit social signals through speech and/or prosody. Emotional prosody conveys key information about the emotional state of a speaker and is thus a crucial cue that one has to detect in order to develop efficient social communication. Previous studies in adults reported different brain responses to emotional than to neutral prosodic deviancy. The aim of this study was to characterize such specific emotional deviancy effects in school-age children. The mismatch negativity (MMN) and P3a evoked potentials, reflecting automatic change detection and automatic attention orienting, respectively, were obtained for neutral and emotional angry deviants in both school-age children ($n = 26$) and adults ($n = 14$). Shorter latencies were found for emotional than for neutral preattentive responses in both groups. However, whereas this effect was observed on the MMN in adults, it appeared in an early discriminative negativity preceding the MMN in children. A smaller P3a amplitude was observed for the emotional than for the neutral deviants at all ages. Overall, the brain responses involved in specific emotional change processing are already present during childhood, but responses have not yet reached an adult pattern. We suggest that these processing differences might contribute to the known improvement of emotional prosody perception between childhood and adulthood.

Keywords Mismatch negativity (MMN) · Change detection · P3a · Emotion · Children · Prosody

In human communication, a wealth of information is transmitted by the voice. Besides speech, human voices communicate nonverbal information, in particular through prosody. *Prosody* is defined as the rhythmic and intonational aspect of language, used to change or enhance the meaning of an utterance (Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1991). This goal is reached through the modulation of different acoustic parameters (e.g., fundamental frequency [F0], duration, and intensity; Aziz-Zadeh, Sheng, & Gheytañchi, 2010; Schirmer & Kotz, 2006). In emotional prosody, this modulation conveys the emotional state of a speaker. The interpretation of emotional prosody is a valuable skill that needs to be mastered quickly in order to handle communication in social contexts. Indeed, sensitivity to others' vocal emotional expressions is associated

with social competence in both childhood and adolescence (Goodfellow & Nowicki, 2009).

The ability to recognize emotions from prosody is particularly efficient in adults (Kreiman & Sidtis, 2011). Event-related potential (ERP) studies have shown an early discrimination between neutral and emotional prosody starting around 200 ms (Iredale, Rushby, McDonald, Dimoska-Di Marco, & Swift, 2013; Paulmann & Kotz, 2008), consistent with Schirmer and Kotz's (2006) model, which describes three stages for the processing of emotional prosody: (i) low-level auditory analysis (around 100 ms); (ii) the integration of emotionally significant acoustic cues, around 200 ms; and (iii) higher-level cognition (around 400 ms). More recently, a growing body of research has focused on a particular ERP component—that is, the mismatch negativity (MMN)—to improve understanding of the automatic discrimination of emotional prosody. Elicited by the automatic detection of a change in a regular sequence (Näätänen, Paavilainen, Rinne, & Alho, 2007), the MMN is generally computed by subtracting the brain response to standard stimuli from that to deviant stimuli in an oddball paradigm. In response to tones, the MMN is characterized by a negative deflection that culminates over

✉ Marie Gomot
gomot@univ-tours.fr

¹ UMR1253, iBrain, Université de Tours, Inserm, Tours, France

² Centre Universitaire de Pédopsychiatrie, CHRU de Tours, Tours, France

fronto-central sites around 150–200 ms after stimulus onset (Fishman, 2014). When salient deviants are presented, the MMN is often followed by a P3a, a positive deflection recorded around 300 ms over central sites, which reflects the automatic orientation of attention toward the salient stimuli (Escera, Alho, Schröger, & Winkler, 2000).

MMN studies using vocal stimuli in adults suggest an early integration of emotional prosodic change. Comparing the neutral and emotional MMNs elicited by syllables or pseudowords, studies with various paradigms (oddball, optimum, or control sequences) evidenced a larger MMN amplitude for emotional (happy, angry, sad) than for neutral change (Jiang, Yang, & Yang, 2014; Pakarinen et al., 2014; Pinheiro, Barros, Vasconcelos, Obermeier, & Kotz, 2017; Schirmer, Striano, & Friederici, 2005). Emotional change detection is also characterized by an earlier MMN in studies using syllables (Schirmer et al., 2005), interjections (Schirmer, Escoffier, Cheng, Feng, & Penney, 2016), vocalizations (Pinheiro et al., 2017), or pseudowords (Thönnessen et al., 2010).

In contrast, studies using emotional nonvocal sounds (e.g., tones, instruments) did not replicate any of these differences between neutral and emotional conditions found using vocal sounds (Goydke, Altenmüller, Möller, & Münte, 2004; Leitman, Sehatpour, Garidis, Gomez-Ramirez, & Javitt, 2011). Differences in the automatic detection of emotional and neutral prosody therefore seem to rely on acoustic parameters specifically found in natural voices but not in nonvocal sounds. As mentioned earlier, emotional prosody relies on a number of acoustical factors, such as fundamental frequency (Banse & Scherer, 1996), also known to modulate the MMN (Jacobsen & Schröger, 2001; Sams, Paavilainen, Alho, & Näätänen, 1985). A tight control of acoustical parameters while maintaining the natural voice quality is therefore necessary to avoid any confounding effects of nonspecific acoustic modification (which does not carry emotion) and of emotional deviancy on the MMN.

Given the key role of emotion identification in response to social threat, a great number of behavioral studies have been conducted in infants and children to determine the time course of acquisition of emotional prosody discrimination. Distinction between emotional and neutral voices on the basis of prosodic information arises around the age of five months (Flom & Bahrick, 2007), whereas the identification of exaggerated basic emotions starts around 4–5 years old (Friend, 2000; Quam & Swingley, 2012). These prosodic perception abilities continue to improve with age between 5 and 10 years old (Sauter, Panattoni, & Happé, 2013), which makes childhood a key period for the development of emotional prosodic perception. Despite the behavioral evidence of emotional prosody discrimination in early childhood, ERP studies in young children have provided surprising results. Although the processing of emotional cues affects early obligatory ERP components (e.g., N1, P2) in adults, no such effects have

been observed in children (Chronaki, Benikos, Fairchild, & Sonuga-Barke, 2014; Chronaki et al., 2012). Instead, ERP amplitude difference between angry and neutral prosody have been reported at late latency ranges around 400 ms in infants and children (Chronaki et al., 2012; Grossmann, Striano, & Friederici, 2005). Therefore, behavioral measures in early childhood might not reflect the mature functioning of the brain network underlying vocal emotion processing. For instance, brain regions, such as the amygdala, involved in the processing of emotions such as anger do not reach maturity until adolescence (Schumann et al., 2004). Thus, even if the elementary emotion-processing network emerges in early life, its late anatomic-functional development may delay the fine tuning and automatization of responses specific to emotional (vocal) expressions (Leppänen & Nelson, 2009). In this regard, ERPs allow researchers to obtain fine-tuned information about the functional development of emotional detection in children. Apart from better understanding the development underlying the implicit processing of vocal emotions, better characterizing this somehow late neurodevelopmental course could allow us to track the hallmarks of neuropsychiatric disorders such as autism spectrum disorder, in which emotional prosody is particularly affected.

The MMN is a basic automatic response in which, in response to tones, scalp distributions are similar between children and adults (Gomot, Giard, Roux, Barthélémy, & Bruneau, 2000; Martin, Shafer, Morr, Kreuzer, & Kurtzberg, 2003), although larger amplitude and longer latency have been reported in children (Choudhury, Parascando, & Benasich, 2015; Gomot et al., 2000; Korpilahti & Lang, 1994). MMN findings in response to syllables have been less consistent, with some studies reporting a larger MMN in children than in adults (Kraus et al., 1993; Kraus, McGee, Sharma, Carrell, & Nicol, 1992), and others reporting the opposite (Bishop, Hardiman, & Barry, 2011; Paquette et al., 2013). Several studies have also highlighted significantly longer latencies in children (Ceponiene, Lepistö, Alku, Aro, & Näätänen, 2003; Liu, Chen, & Tsao, 2014; Paquette et al., 2013; Shafer, Yu, & Datta, 2010).

In spite of the great number of studies attempting to understand the maturation of the speech-related MMN, studies using emotional vocal stimuli are few. MMN studies conducted in teenagers, school-age children, or infants with vocal stimuli have reported a fronto-central MMN pattern in response to emotional words or syllables. In infants, Cheng, Lee, Chen, Wang, & Decety (2012) showed a larger MMN for fearful than for happy syllables, with the MMN elicited by happy deviants being more lateralized to the right hemisphere. In school-age children, a biphasic (Korpilahti et al., 2007; Lindström, Lepistö, Makkonen, & Kujala, 2012) and right-lateralized (Korpilahti et al., 2007) MMN was evidenced for a commanding deviant. This MMN was smaller and earlier for a commanding than for a sad deviant (Lindström et al., 2012).

In teenagers, similar MMNs and P3as were found for sadness and fear (Hung, Ahveninen, & Cheng, 2013). Overall, these studies indicated that individuals from infancy to adolescence display brain responses reflecting the automatic detection of emotional vocal changes. Infants and children also showed differentiated responses to different emotions. However, although they compared several emotional vocal MMNs, these developmental studies neither compared the detection of neutral and emotional deviancy nor used an adult comparison group and controlled for acoustic features or neuronal adaptation induced by oddball paradigms. The specific effect of emotional deviancy on prosodic change detection thus remains unknown in children. Since the neural mechanisms involved in auditory perception and discrimination continue to develop throughout childhood (Shafer et al., 2010), we hypothesized that children's brain mechanisms involved in the detection of change in emotional information might differ from those of adults. Hence, the differentiation between neutral and emotional deviancy might be expressed differently in groups of children and adults.

The aim of this study was to address the specific brain response to emotional deviancy with respect to neutral deviancy during automatic change detection in school-age children, and to evaluate whether this response is different from the response in adults. To address these goals, auditory ERPs to neutral and emotional deviants were recorded in children and adults using a tightly controlled paradigm composed of oddball and equiprobable sequences. Using such sequences appeared to be the best option, considering the number of deviants and the choice of natural voices; this also allowed us to control for acoustic differences between the stimuli used in the subtraction process and to reduce the neural adaptation effect. According to the methodology used and the state of the art, we hypothesized that differentiation between neutral and emotional deviancies would not lead to an MMN amplitude difference, but rather to a lateralization of the emotional response in the right hemisphere among children and to a latency difference in adults (shorter emotional MMN than neutral MMN), and possibly in children.

Materials and method

Subjects

Twenty-six school-age children (mean age \pm standard deviation [*SD*]: 9.5 ± 1.5 years, age range: 6.8–12.3; 13 females, 13 males) and 14 adults (24.1 ± 4.4 years, age range: 19–33; seven females, seven males) were included in the study. Recruitment was realized through an e-mail list and flyers at the University of Tours and among Tours Hospital employees. The exclusion criteria comprised evidence of disease of the central nervous system, infectious or metabolic disease,

epilepsy, developmental difficulties in language or walk acquisition, and any psychotropic treatment or medication that would modify the electrogenesis. All subjects had normal audition (subjectively controlled with an audiometer). Intellectual skills in the verbal and nonverbal domains were tested with four subtests (Vocabulary, Similarities, Block Design, and Matrix) of the Wechsler scales WAIS-IV (Wechsler, 2011), for adults, and WISC-IV (Wechsler, 2005), for children, to confirm the absence of intellectual deficiency in the selected groups. Group characteristics are listed in Table 1. Informed written consent was obtained from all adult participants and from the children's parents. The Ethics Committee of the University Hospital of Tours approved the protocol. All procedures were conducted according to the principles of the Declaration of Helsinki.

Stimuli

The vowel /a/ uttered by different female speakers with either neutral or emotional prosody (anger, fear, happiness, surprise, disgust, or sadness) was recorded with Adobe Audio 2.0. The stimuli were edited to have the same duration (400 ms) and intensity (70 dB SPL) through the use of Praat (Boersma, 2002) and Matlab (The MathWorks Inc., Natick, MA, USA). Praat was also used to measure the mean fundamental frequency of each sound.

A set of 37 sounds was presented for validation to 16 adult participants who did not participate in the electroencephalographic (EEG) experiment (24.2 ± 2.2 years, age range: 20–28; 13 females, three males). The subjects were asked to identify the emotion (neutral, anger, fear, happiness, surprise, disgust, or sadness) and to rate the valence and arousal of the sounds on 5-point Likert scales (ranging from *very negative* to *very positive* for valence and from *weak* to *high* for arousal). Two neutral stimuli with closely matched mean fundamental frequencies and good ratings of valence (recognized as neutral by 88% and 94% of the validation group, respectively) and emotion (94% neutral for both) were chosen as the neutral

Table 1 Group characteristics

	School-Age Children	Adults
Age ^a	9.5 ± 1.5 [6–12]	24.1 ± 4.4 [19–33]
Sex Ratio ♂/♀	13/ 13	7/ 7
Wechsler Standard Scores ^b		
Block Design	12 ± 3	12 ± 3
Matrix	13 ± 3	13 ± 2
Similarities	13 ± 3	14 ± 2
Vocabulary	12 ± 3	14 ± 3

^a Mean age (years) \pm *SD* [range]. ^b Standard scores derived from WISC-IV (Wechsler, 2005) for children and from WAIS-IV (Wechsler, 2011) for adults.

standard (neutralStd) and the neutral deviant (neutralDev). An angry stimulus with a mean fundamental frequency similar to those of the two neutral stimuli and with the highest valence rating (100%) and the best level of recognition of angry emotion (69%) was selected as the emotional deviant (angryDev). The arousal values of the selected stimuli were 2.8 ± 0.8 for neutralStd, 3.0 ± 0.7 for neutralDev, and 3.9 ± 0.8 for angryDev, and their mean fundamental frequencies were 226, 228, and 223 Hz, respectively. Different female speakers uttered these stimuli of interest. Five other emotional stimuli (happy, sad, surprised, disgust, and fear) with mean F0s in a 220- to 231 Hz range and with recognition levels above chance were selected as stimuli for the equiprobable sequence. Four different females produced these remaining stimuli in the equiprobable sequence.

The validation procedure for the stimuli was also performed by 18 children (8.8 ± 1.7 years, age range: 6.4–12.2; ten males, eight females) for the valence ratings. High correct ratings were obtained from these children for neutralStd (72%), neutralDev (72%), and angryDev (89%), and also for the stimuli of the equiprobable sequence. The valence ratings of the stimuli of interest (neutralStd, neutralDev, and angryDev) did not differ statistically between groups [two-tailed *t* tests: neutralStd, $t(32) = 1.09$, $p = .285$; neutralDev, $t(32) = 1.66$, $p = .106$; angryDev, $t(32) = 1.37$, $p = .179$].

Acoustic information concerning all stimuli is presented in Fig. 1.

Procedure

During EEG recording, the subjects sat comfortably in a reclining armchair in a sound-attenuated room. Subjects watched a silent movie without subtitles while sounds were delivered through speakers (Logitech Z-2300); they were instructed that they would have to briefly tell the story from the movie at the end of the recording session. This procedure avoided requiring voluntary direct attention toward the auditory stimuli. The speakers and screen were placed at 120 cm from the subjects' heads.

Automatic detection processes were studied using a passive oddball sequence and an equiprobable sequence (see Fig. 1) to control for both stimulus features and neuronal adaptation effects (Jacobsen & Schröger, 2001). The oddball sequence comprised 1,172 neutral standards (neutralStd; Identity 1, probability of occurrence, $p = .830$), 120 neutral deviants (neutralDev; Identity 2, deviant) and 120 angry deviants (angryDev, Identity 3 and emotional “angry” deviant) ($p = .085$ each), with the constraint that any two deviants were separated by a minimum of three standards. The second sequence was composed of eight different stimuli presented with a probability of occurrence close to that of the deviants in the oddball sequence ($p = .125$; 120 stimuli): two neutral stimuli (equiNeutral1, equiNeutral2, which were the neutral standard

[neutralStd] and the neutral deviant [neutralDev] from the oddball sequence) and six emotional stimuli representing the six basic emotions expressed by different speakers (equiHappy, equiSad, equiSurprised, equiDisgust, equiFear, and equiAngry—i.e., the angry stimulus used as angryDev in the oddball sequence). None of the stimuli were repeated more than two times in a row, in order to avoid creation of a regularity pattern. Stimuli were presented with a constant stimulus onset asynchrony (SOA) of 700 ms in both sequences (300 ms interstimulus interval), thus allowing the signal to return to baseline while keeping the duration of the recording session rather short, which is critical in research involving child participants. The total recording session lasted 28 min.

EEG recording and ERP measurements

The EEG was recorded from 64 active electrodes (ActiveTwo Systems, Biosemi, The Netherlands) with a sampling rate of 512 Hz. Horizontal and vertical eye movements were monitored using electrodes placed at the left and right outer canthi and below the left eye. An electrode was placed on a subject's nose for offline re-referencing. The ELAN software package was used for the analysis of EEG–ERP (Aguera, Jerbi, Caclin, & Bertrand, 2011). The EEG signal was amplified and filtered with a 0.3 Hz high-pass filter (Butterworth filter, order 1). Artifacts resulting from eye movements were removed using independent component analysis, and movement artifacts characterized by a high-frequency or high-amplitude signal were discarded manually by the experimenter, who was blind to the trial type. A 30 Hz low-pass filter was applied (Butterworth filter, order 3), and ERPs were averaged in an 800 ms time window including a 100 ms prestimulus baseline. The neutral difference wave was obtained by subtracting the ERP elicited by equiNeutral2 from that elicited by neutralDev (i.e., the same sound in the equiprobable and oddball sequences; Fig. 1). The same subtraction (angryDev–equiAngry) was applied in order to obtain the emotional difference wave (Fig. 1). Since the stimuli in the equiprobable sequence had identical acoustic characteristics and a similar probability of occurrence to the oddball deviants, the resulting difference wave more likely reflects a genuine MMN than does the signal in the oddball paradigm (Kimura, Katayama, Ohira, & Schröger, 2009). Moreover, the application of this subtraction process to the emotional condition allowed us to control for the influence of emotional processing that operates in both sequences, in order to isolate the effect related to emotional deviancy. Finally, the direct comparison between the neutral and emotional difference waves contrasted “identity deviancy” and “identity/emotion deviancy,” leaving only the emotional deviancy as a differential factor between conditions, which allowed for the assessment of a specific emotional deviancy effect. In adults, for each stimulus of interest the

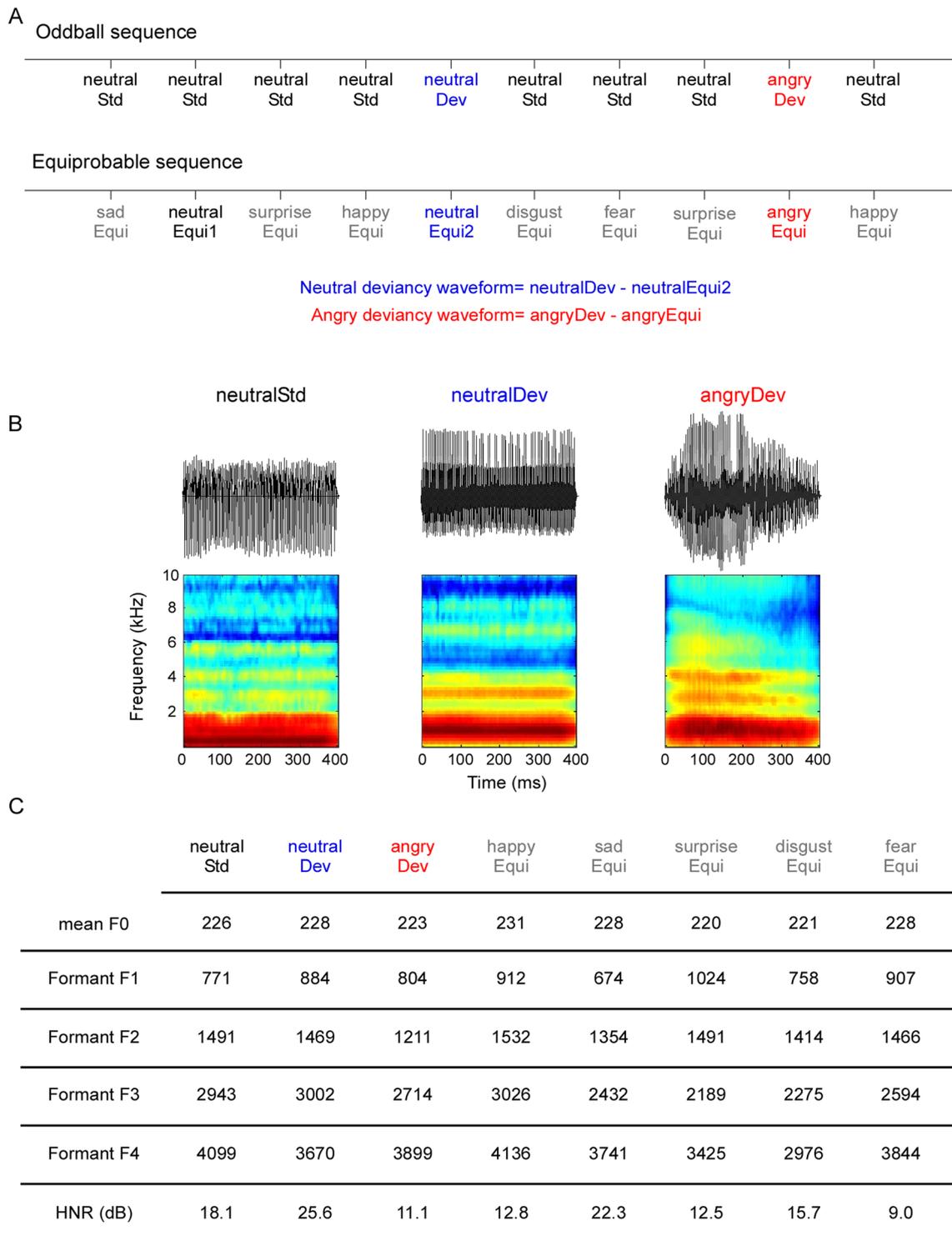


Fig. 1 (A) Illustration of the oddball and equiprobable sequences, composed of neutral standard (neutralStd) and neutral and angry deviants (neutralDev and angryDev) in the oddball sequence, and of neutralEqui1, neutralEqui2, angryEqui, sadEqui, surpriseEqui, happyEqui, disgustEqui, and fearEqui in the equiprobable sequence.

The three stimuli of the oddball sequence were also presented in the equiprobable sequence. (B) Waveforms and spectrograms for the neutralStd, neutralDev, and angryDev stimuli. (C) Acoustic properties of all stimuli. The mean fundamental frequency (F0) and formants are reported in hertz (Hz).

average numbers of artifact-free trials were: 102 ± 17 (angryDev), 103 ± 13 (neutralDev), 102 ± 16 (equiAngry), and 100 ± 16 (equiNeutral2). In children, the average numbers

of artifact-free trials was: 90 ± 13 (angryDev), 90 ± 14 (neutralDev), 88 ± 13 (equiAngry), and 90 ± 13 (equiNeutral2). The numbers of artifact-free trials did not

differ between conditions [neutralDev/equiNeutral2, $t(25) = 0.16$, $p = .873$, for children, and $t(13) = 0.75$, $p = .464$, for adults; angryDev/equiAngry, $t(25) = 0.95$, $p = .353$, for children, and $t(13) = -0.10$, $p = .925$, for adults), neutralDev/angryDev, $t(25) = -0.30$, $p = .767$, for children, and $t(13) = 0.66$, $p = .521$, for adults), equiNeutral2/equiAngry, $t(25) = 1.05$, $p = .302$, for children, and $t(13) = -0.62$, $p = .548$, for adults] but did differ between groups [as revealed by two-tailed t tests: neutralDev, $t(38) = -2.90$, $p = .007$; angryDev, $t(38) = -2.38$, $p = .023$; equiNeutral2, $t(38) = -2.33$, $p = .025$; equiAngry, $t(38) = -3.01$, $p = .005$], which is a classic finding in developmental studies.

For both conditions, the MMN was identified as a negative deflection occurring in the 130- to 230 ms time window, and the P3a as a positive deflection occurring in the 250- to 400 ms time window. The peak amplitudes and latencies of the MMN and P3a were measured in each subject by locating the individual peaks within 80 ms and 100 ms time windows, respectively, centered on the peak of the grand mean average of each group. A negative deflection occurring before the MMN was also measured within a 30- to 90 ms time window and analyzed when it was significantly different from 0.

Statistical analysis

One-tailed t tests performed on the MMN and P3a amplitude values were used to determine the statistical significance of the studied deflections (Table 2).

Given that several adults did not display an early discriminative negativity, especially in the emotional condition, the statistical significance of this response from 0 was evaluated using permutations performed in the 30- to 90 ms time window (Guthrie–Buchwald time correction; Guthrie & Buchwald, 1991). Both the neutral and emotional early discriminative negativities appeared significant in children, whereas no significant response was recorded in the adult group for both conditions.

Condition effects were tested for the early discriminative negativity in children only. The amplitude and latency of responses were subjected to a within-subjects analysis of variance (ANOVA) performed on electrodes Fz, Cz, and Pz, with condition (neutral, emotional) and electrode as within-subjects factors.

After visual inspection of the MMN scalp distributions, and considering previous studies that had reported a right-lateralized MMN to emotion (Cheng et al., 2012; Korpilähti et al., 2007; Schirmer et al., 2005; Thönnessen et al., 2010), MMN amplitudes were analyzed with mixed-design ANOVAs performed at F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4, with condition (neutral, emotional), laterality (left, medial, right), and anterior–posterior (frontal, central, parietal) as within-subjects factors, and age group (children, adults) as a between-subjects factor. The same analysis was performed for MMN latencies, but only at the medial electrodes, therefore removing the laterality factor. P3a amplitudes and latencies were analyzed with mixed-design ANOVAs performed at Fz, Cz, and Pz, with condition and anterior–posterior as within-subjects factors, and age group as a between-subjects factor. Post-hoc analyses (Newmann–Keuls) were performed when needed to determine the origins of interactions. For significant results, the effect sizes are shown as partial eta squared (η_p^2).

Group by condition interactions on MMN and P3 latencies, which revealed different effects of condition on children than on adults, were further investigated with a correlation test between age and MMN or P3a latency in the child group. This analysis aimed at determining the origin of this developmental difference. Correlations between latency and age (in months) were performed for each condition in the child group. To assess differences between conditions, the individual-subject data were permuted across conditions; correlations were computed between age and permuted conditions; and theoretical differences between conditions in terms of slope and intercept were computed. This operation was repeated 15,000 times, and the theoretical differences for slope and

Table 2 Mean MMN, P3a, and early discriminative negativity amplitudes (μV) and latencies (ms)

Response	Condition	School-Age Children		Adults	
		Amplitude (mean \pm SD)	Latency	Amplitude	Latency
MMN	Neutral	$-3.1 \pm 3.2^{**}$	182 ± 23	$-2.3 \pm 2.1^{**}$	181 ± 15
	Emotional	$-3.3 \pm 2.5^{**}$	173 ± 21	$-2.2 \pm 1.6^{**}$	141 ± 19
P3a	Neutral	$5.2 \pm 3.9^{**}$	311 ± 31	$4.0 \pm 1.5^{**}$	265 ± 22
	Emotional	$3.6 \pm 3.0^{**}$	300 ± 27	$3.0 \pm 1.4^{**}$	271 ± 23
Early discriminative negativity	Neutral	$-2.7 \pm 2.7^{**}$	68 ± 21		
	Emotional	$-2.0 \pm 1.7^{**}$	56 ± 21		

Data are reported at electrode Cz. Amplitudes significantly differing from the baseline (one-tailed t tests) are marked with asterisks: $^{**}p < .001$.

intercept were recorded. This gave a distribution of differences under the null hypothesis of no difference, centered around 0, for which the 95% confidence intervals (CIs) around 0 were $[-0.53; 0.53]$ for the MMN and $[-0.69; 0.70]$ for the P3a, for the slope, and $[-65.40; 66.35]$ for the MMN and $[-83.37; 83.53]$ for the P3a, for the intercept. Real differences for slope and intercept were deemed significant if they fell outside the 95% CI.

All ANOVAs were performed using Statistica, whereas the correlations were performed with Matlab.

Results

The event-related responses (to neutralDev, equiNeutral2, angryDev, and equiAngry) used to calculate the difference waveforms of each group are displayed in Fig. 2. The neutral and emotional difference waveforms are displayed in Fig. 3, and the associated group comparisons are shown in Fig. 4. Post-hoc test results for all age and condition effects on latencies can be seen in Fig. 5. The mean peak amplitudes and latencies of the MMN and P3a for each group are reported in Table 2, along with values of the early discriminative negativity in children.

In children, both neutral and emotional changes elicited an early discriminative negativity peaking before 100 ms. In adults, even if visual inspection revealed a small early deflection in the neutral condition, no significant early discriminative response was recorded. In all conditions for both groups,

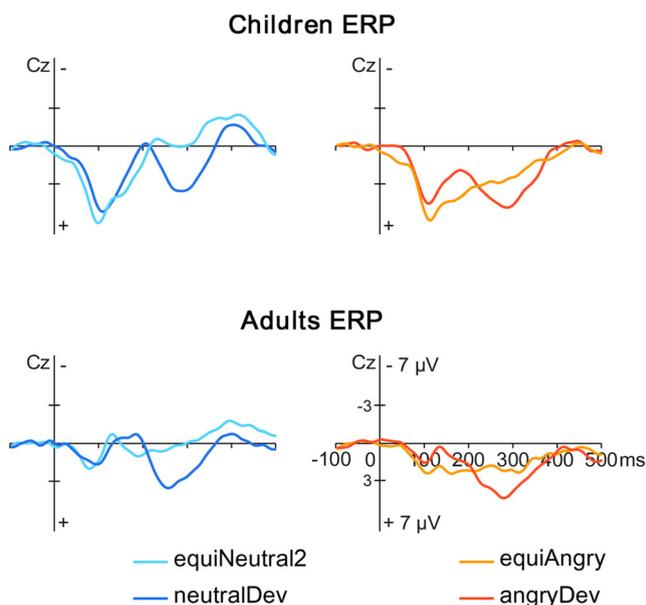


Fig. 2 Event-related potentials (ERPs) to neutral and emotional stimuli in the equiprobable sequence and in the oddball sequence, in which the stimuli were deviants (with the equiprobable equiNeutral2 and equiAngry and deviant stimuli neutralDev and angryDev shown in different colors).

both MMN and P3a deflections were observed. All measured deflections differed significantly from zero, as revealed by one-tailed t tests (Table 2).

All the peak amplitude effects reported below were also observed in the mean amplitude measures.

Early discriminative negativity analysis in children

A significant condition effect was found on the early discriminative negativity latency in children [$F(1, 25) = 9.30, p = .005, \eta_p^2 = .271$] due to a shorter latency for the emotional than for the neutral condition, but no amplitude difference was observed.

MMN analysis

MMN amplitude No age or condition effect and no interaction were observed for MMN amplitudes. A laterality effect was found on MMN amplitudes [$F(2, 76) = 3.45, p = .037, \eta_p^2 = .083$], regardless of age and condition, due to a smaller MMN amplitude over left-hemisphere electrodes than over medial ($p = .013$) electrodes.

MMN latency A group by condition interaction [$F(1, 38) = 18.82, p < .001, \eta_p^2 = .331$] revealed that the emotional MMN latency was significantly shorter than the neutral MMN latency in adults ($p < .001$; emotional condition, $141 \text{ ms} \pm 19$; neutral condition, $181 \text{ ms} \pm 15$), and that the emotional MMN latency was shorter in adults than in children ($173 \text{ ms} \pm 21$; $p < .001$). Otherwise, the MMN latencies did not differ (e.g., the latencies to neutral deviants were similar across groups and similar between the neutral and emotional MMNs in children [neutral condition, $182 \text{ ms} \pm 23$; emotional condition, $173 \text{ ms} \pm 21$]; see Fig. 3). Correlations of MMN latency with age were found to be significant for the neutral condition ($R^2 = .44, F = 19.05, p < .001$), but not for the emotional condition ($R^2 < .01, F < .01, p = .94$) in children (Fig. 5). The correlation slopes differed between conditions (neutral slope [95% CI] = $-0.66 [-0.98; -0.35]$; emotion slope [95% CI] = $0.01 [-0.37; 0.40]$; difference = $0.68, p = .007$), as well as the correlation intercepts (neutral intercept [95% CI] = $256.89 [220.66; 293.12]$; emotion intercept [95% CI] = $170.86 [126.67; 215.05]$; difference = $-86.03, p = .006$).

P3a analysis

P3a amplitude No effect of age was observed on P3a amplitudes. A significant condition effect [$F(1, 38) = 6.52, p = .015, \eta_p^2 = .147$] was observed for P3a amplitudes, with a larger neutral P3a than emotional P3a. An anterior–posterior effect [$F(2, 76) = 10.14, p < .001, \eta_p^2 = .211$] was also observed, due to a larger P3a amplitude over Cz than over either Fz ($p = .027$) or Pz ($p < .001$).

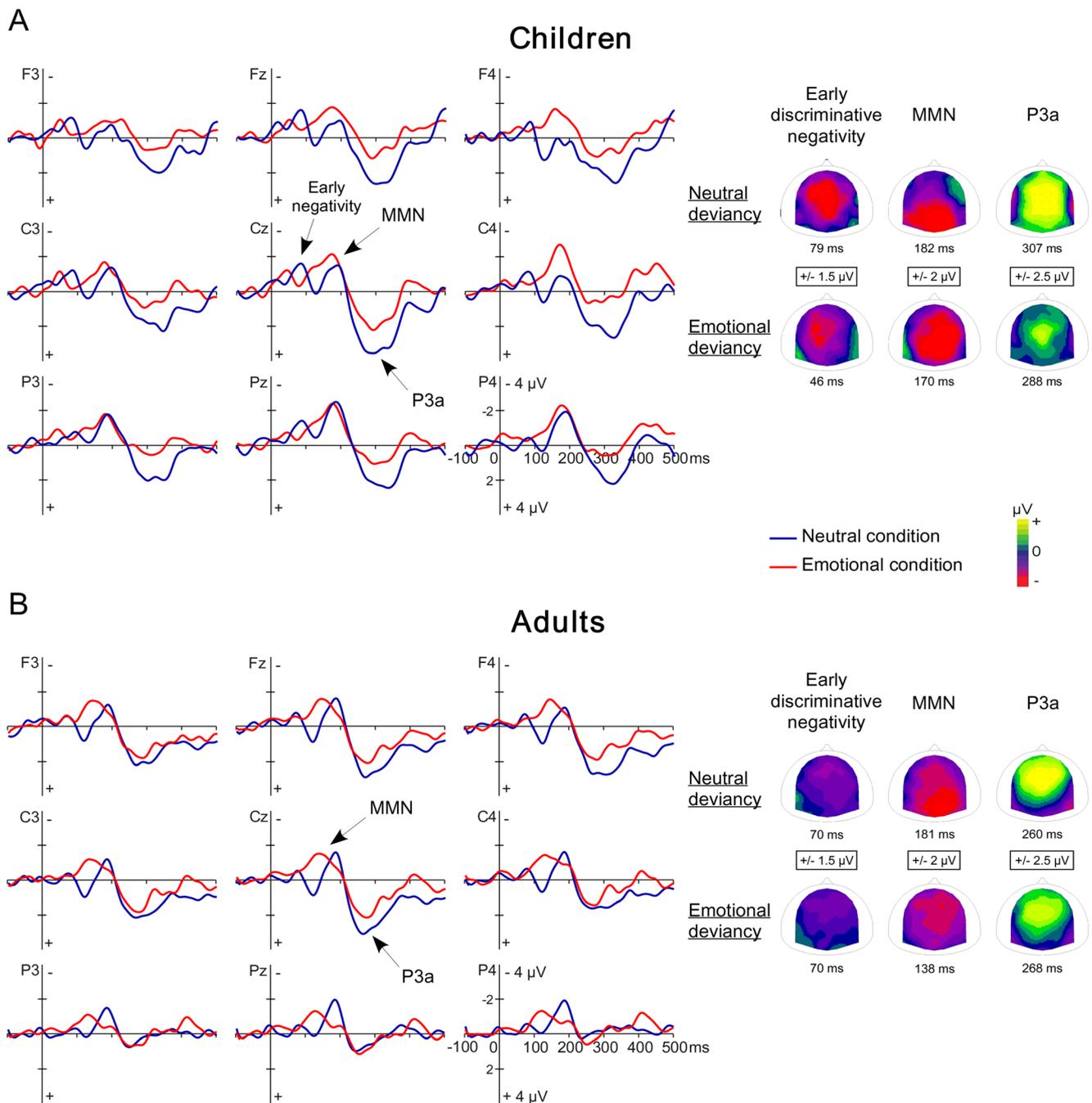


Fig. 3 Brain responses to neutral and emotional prosodic deviancy in (A) children and (B) adults. Graphs show the difference waveforms in response to the neutral (cool color) and emotional (warm color)

prosodic change and scalp potential distributions of the early discriminative negativity, MMN, and P3a.

P3a latency A significant group difference was found for P3a latencies [$F(1, 38) = 33.15, p < .001, \eta_p^2 = .466$], due to shorter P3a latencies in adults (neutral condition, $265 \text{ ms} \pm 22$; emotional condition, $271 \text{ ms} \pm 23$) than in children (neutral condition, $311 \text{ ms} \pm 31$; emotional condition, $300 \text{ ms} \pm 27$). A group by condition interaction [$F(1, 38) = 8.18, p = .007, \eta_p^2 = .177$] further indicated that in addition to the latency differences between groups ($ps \leq .001$), a shorter

P3a latency was found for the emotional than for the neutral condition in children only ($p < .001$). Correlations of P3a latencies with age were significant for both the neutral ($R^2 = .59, F = 34.03, p < .001$) and emotional ($R^2 = .24, F = 7.41, p = .012$) conditions in children (Fig. 5). The correlation slopes did not differ between conditions (neutral slope [95% CI] = $-1.20 [-1.62; -0.77]$; emotion slope [95% CI] = $-0.60 [-1.06; -0.15]$; difference = $0.60, p = .101$), whereas a

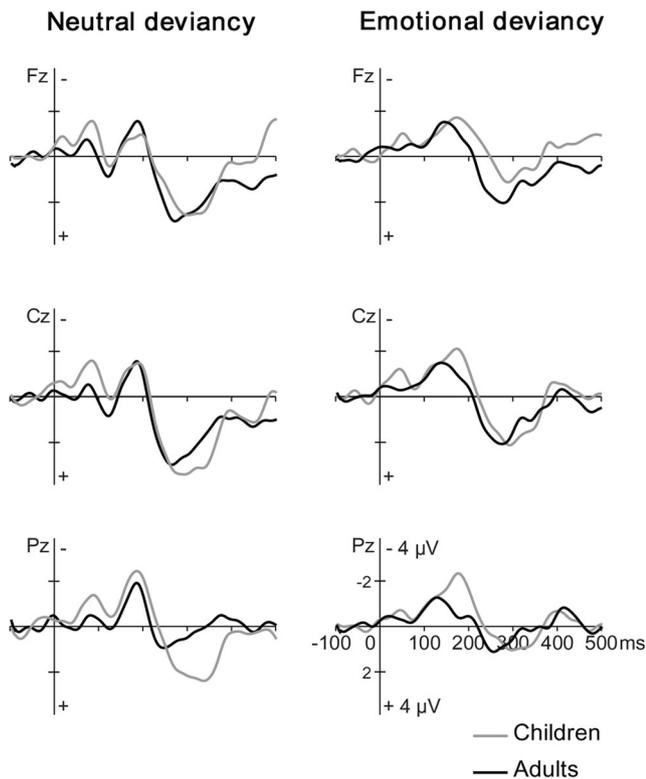


Fig. 4 Age group effects on responses to neutral and emotional prosodic deviancy. Brain responses are represented in gray lines for children and black lines for adults.

significant difference was observed between the correlation intercepts (neutral intercept [95% CI] = 451.46 [402.36; 500.56]; emotion intercept [95% CI] = 365.61 [312.63; 418.59]; difference = -85.85 , $p = .041$).

Given previous findings of a larger P3a amplitude in the emotional than in the neutral condition (Domínguez-Borràs, García-García, & Escera, 2008a, 2008b), we secondarily examined this component in the ERPs in response to deviants with the same mixed-design ANOVA that we had used for the P3a derived from the difference waveforms. Aside from a significant anterior–posterior effect [$F(2, 76) = 21.08$, $p < .001$, $\eta_p^2 = .357$], a significant condition effect was observed [$F(1, 38) = 4.26$, $p = .046$, $\eta_p^2 = .101$], with a larger amplitude for the emotional deviant P3a than for the neutral deviant P3a. A significant anterior–posterior by group by condition interaction [$F(2, 76) = 4.61$, $p = .013$, $\eta_p^2 = .108$] was also reported, due to a larger amplitude to emotional than to neutral deviants at fronto-central sites in children ($p < .001$ at frontal sites, and $p = .016$ at central sites), whereas this effect was only reported at frontal sites in adults ($p = .018$).

Discussion

In the present study, we investigated the early processing of emotional prosodic change in school-age children. The main

aim was to test whether a specific brain response to emotional change was present in children by comparing it to the brain response elicited by neutral change, which had never been done until now. To this end, oddball and equiprobable sequences composed of neutral and emotional stimuli were presented to adults and school-age children. The use of both sequences ensured that the influences of low-level factors (acoustic parameters) and neural adaptation are greatly limited.

In children and adults, both neutral and emotional changes generated MMN and P3a deflections. An additional early negative deflection occurring before 100 ms was also present in children. Both children and adults groups showed differentiated detection of emotional and neutral changes; however, the results clearly highlighted that different brain mechanisms are involved in this process in childhood and adulthood.

Brain responses to prosodic changes in adults

In adults, shorter MMN latencies were found for angry than for neutral deviancy. This finding is partly consistent with previous MMN studies showing a shorter latency for positive emotions than for neutral stimuli (Schirmer et al., 2016; Schirmer et al., 2005). These results suggest that this effect might be valence-independent and present for different basic emotions. This latency difference may reflect an earlier MMN for angry change than for neutral change related to a priority processing of emotional deviancy. From its evolutionary significance to its major role in human social interactions, emotion has long been recognized as high priority information. For instance, stimuli of identical duration are perceived as being longer when they are emotional than when they are neutral (Grommet et al., 2011). The observed latency difference might therefore be related to a faster brain activation for the detection of emotional deviancy through the amygdala “fast route” as evidenced in the visual domain (Taylor & Fragopanagos, 2005; Vuilleumier, 2005). In the auditory modality, fMRI studies that have reported an absence of significant amygdala activation in response to emotional prosody of speech or speech-like sounds (Gebauer, Skewes, Hørlyck, & Vuust, 2014; Wiethoff et al., 2008; Wildgruber et al., 2005) did not support the existence of a quick subcortical route for vocal sounds. However, amygdala lesion has been shown to reduce the activation of auditory cortices that is commonly observed in response to the emotional prosody of speech-like stimuli (Frühholz et al., 2015). Moreover, the amygdala was recently considered to be a central region in a large brain network involved in the decoding of affective meaning from sound (Frühholz, Trost, & Kotz, 2016). Therefore, the subcortical-route theory (Liebenthal, Silbersweig, & Stern, 2016) remains a valuable explanation that needs to be considered. This latency difference might also reflect a delayed

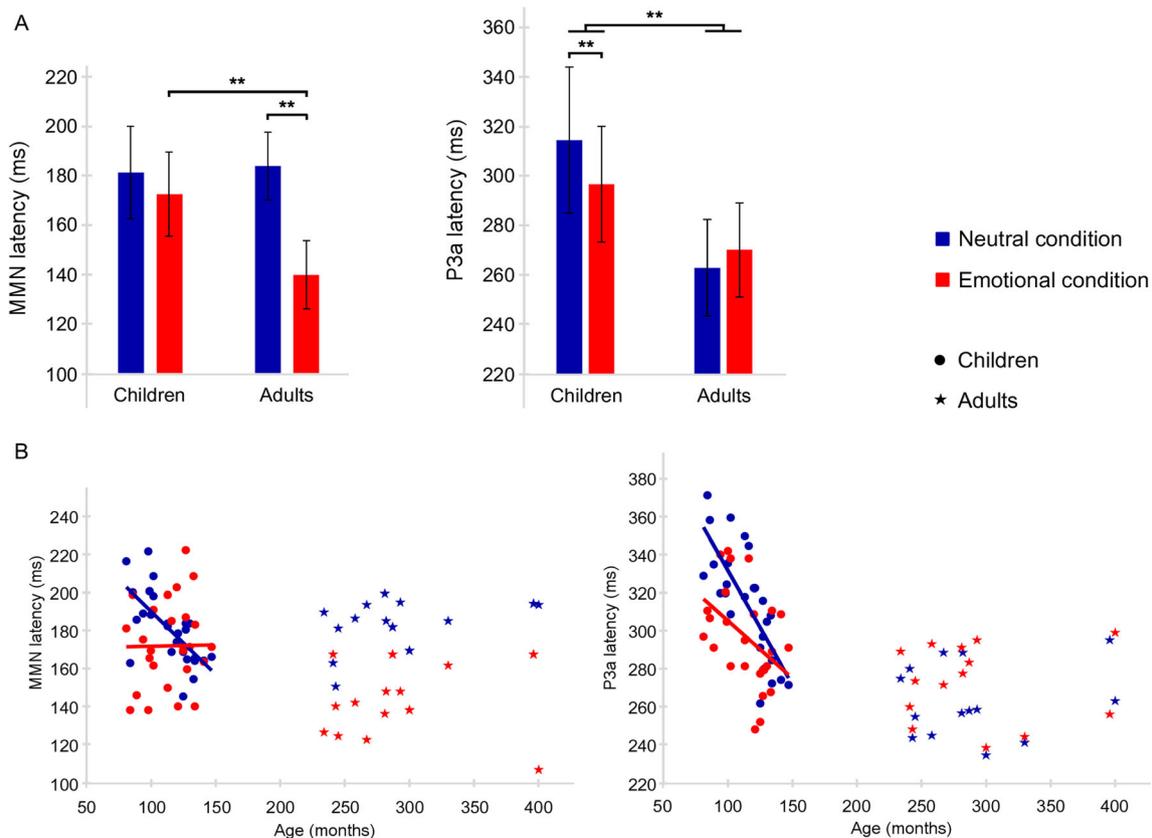


Fig. 5 (A) Post-hoc analyses of age and condition effects on the MMN and P3a latencies (mean amplitudes over Fz, Cz, and Pz). Error bars represent standard deviations. Significant results are marked by ** ($p <$

.001). (B) MMN and P3a latencies as a function of age in children (dots, with correlations between latency and age) and in adults (stars) in both the neutral (cool color) and emotional (warm color) conditions

MMN to neutral change, related to the coexistence of both neutral and angry deviants in the same sequence. Indeed, the presence of emotions can influence behavioral and neuronal responses to neutral stimuli. At the behavioral level, the perception of a neutral stimulus is modified when preceded by an emotional distractor (Lui, Penney, & Schirmer, 2011). At the neuronal level, studies have highlighted that brain responses to change might be delayed when the context contains more salient variations (Campanella et al., 2002). The saliency of emotional angry change could thus have delayed the automatic detection of neutral change in the present study.

Although the first-order parameters of the ecological stimuli displayed in the present study were carefully controlled, one cannot rule out the possibility that the specificity of the overall envelope inherent to the emotional stimulus could have entailed, at least partly, the latency difference between neutral and emotional MMNs. Software such as DAVID (Rachman et al., 2018), which allows researchers to modulate different acoustic parameters of ecological vocal recordings independently, could be useful for determining the dependency of our results on acoustics. Such tools could be valuable to determine the existence of a specific emotion effect, which would be comparable to the specific « vocal effect » recently

evidenced by Agus, Paquette, Suied, Pressnitzer, and Belin (2017; a temporal brain activation was observed specifically to voices but not to acoustic chimeras).

Some previous MMN studies conducted in adults have also reported an effect of emotion on MMN amplitude (Jiang et al., 2014; Pakarinen et al., 2014; Schirmer et al., 2005), whereby the emotional MMN was larger than the neutral MMN. However, potential low-level modulators of the MMN were not always controlled in these studies. Using a tightly controlled paradigm, we did not find any amplitude difference between neutral and emotional deviancy, suggesting that MMN amplitude differences were driven by low-level parameters. Fundamental frequency differences between emotional and neutral stimuli might have influenced MMN amplitude, as evidenced by Jiang et al., (2014) who showed that a combined effect of physical and emotional deviancies resulted in a larger MMN amplitude than did a simple emotional deviancy condition (e.g., no acoustic difference between neutral and emotional stimuli). Even in studies that have controlled fundamental frequency differences with a reverse oddball sequence, acoustical differences between stimuli within a given sequence could have influenced brain responses to these stimuli. Another explanation for the lack of

difference in MMN amplitude between conditions could be linked to the neural adaptation effect observed in oddball paradigms (Jacobsen & Schröger, 2001). In reverse oddball sequences, the MMN is generated from the subtraction of the same stimulus either presented as standard or deviant in two different sequences. This type of paradigm allows to control for acoustic processing but does not control for differences in neural adaptation produced by the different stimuli. The use of an equiprobable sequence that ensures both controlled processing of physical features and close probability of occurrence of stimuli could reduce this neural adaptation effect (Jacobsen & Schröger, 2001), leading to smaller differences between deviancy conditions. Decreased probability of occurrence of emotional as compared to neutral deviants (Pakarinen et al., 2014) or segregation of neutral and emotional deviants in two different sequences (Schirmer et al., 2005) might also have influenced the relative saliencies of deviants. To conclude on the MMN, we did not replicate all the results from the literature, probably because we chose an experimental procedure allowing us to study preattentive emotional categorization while reducing the influence of low-level parameters. This approach aimed at revealing an effect of emotion that was as *genuine* as possible (Grandjean et al., 2005). The control of the saliency of neutral and emotional deviants relative to the neutral standard (i.e., acoustic differences between the two pairs of stimuli were similar) and the use of deviant stimuli behaviorally recognized as neutral and emotional resulted in the absence of an amplitude difference, while a latency effect, which appeared to be specific of the emotional deviancy, was observed.

Finally, the P3a amplitude measured on the difference wave was smaller in the emotional than in the neutral condition. This result is in agreement with a previous study by Schirmer et al. (2005), although their work focused mainly on the MMN. This finding may seem counterintuitive at first glance. Because P3a amplitudes are generally larger for greater deviancy (Alho et al., 1998), a larger P3a amplitude could have been expected for the emotional condition, as previously showed in the visual modality (Keil et al., 2002) and in cross-modal studies (Domínguez-Borràs et al., 2008a, 2008b). A larger P3a was found for the emotional than for the neutral condition, measured directly on deviants ERPs in this study. Modulation of late ERPs by emotion, and more especially by negative emotions, has also been observed in ERP studies (Kanske & Kotz, 2007). The subtraction of this already substantial emotional saliency effect from the regularity violation in the oddball sequence probably caused the smaller P3a amplitude in the difference wave for the emotional than for the neutral condition. In view of these results, previous findings of greater P3a amplitudes in emotional conditions might be considered cautiously.

Brain responses to prosodic changes in children

In children, an early discriminative negativity was elicited regardless of condition, resulting in a clear double-peaked discriminative response (i.e., early negative deflection and the MMN), whereas in adults only the MMN was significant. This early discriminative response is consistent with an early response occurring between 70 and 100 ms after stimulus onset, reported in a previous voice study in children (Rogier, Roux, Belin, Bonnet-Brilhault, & Bruneau, 2010). The present early discriminative negativity is unlikely to be attributed either to first-order acoustic differences, since duration, loudness, and mean fundamental frequency were monitored, or to sensory processing, since each stimulus was used as its own control in the subtraction process. Moreover, several studies in adults have tended to indicate that such early brain responses are elicited by vocal sounds (Bruneau et al., 2013; De Lucia, Clarke, & Murray, 2010; Murray, Camen, Gonzalez Andino, Bovet, & Clarke, 2006; Rigoulot & Armony, 2016) and would be sensitive to irregularities in vocal sequences (Graux et al., 2013).

Pegado et al. (2010) also distinguished two negativities in response to deviancy in adults. The early deflection culminating over frontal sites was sensitive to SOA duration: longer SOAs yielded a decrease in the amplitude of this early deflection. One could assume that this finding might be related to the reduction of neural adaptation to the repeated presentation of standard stimuli for larger SOA. Longer SOAs induce an overall lower stimulation of the neurons responding to standard stimuli, which in turn are able to return to a nonadaptive mode similar to that of neurons responding to deviant stimuli (Lanting, Briley, Sumner, & Krumbholz, 2013; May & Tiitinen, 2010). Pegado et al. interpreted their findings in relation to the echoic memory process: the more the SOA increases, the more the echoic trace fades and the amplitude of the deflection is reduced. In the present study, we controlled for neural adaptation with an equiprobable sequence. Yet an early negative deflection was still present in children. This response could therefore reflect the involvement of echoic memory. Behavioral studies have shown that children rely more on echoic memory than do adults during an auditory working memory task (Engle, Fidler, & Reynolds, 1981). This stronger dependence on echoic memory to process incoming events could explain the existence of a distinct early discriminative negativity for all conditions in children, while this negativity appears to be nonsignificant in adults.

Despite similar response waveforms between neutral and emotional conditions in children (e.g., a double-peaked discriminative response), an emotion effect was observed with a shorter early discriminative negativity in the emotional than in the neutral condition. Aside from average acoustic differences (e.g., frequency and loudness), the modulation of acoustic parameters throughout the sounds allows differentiating

neutral and emotional stimuli. Given that echoic memory is sensitive to physical stimulus characteristics (Engle et al., 1981), differences of some acoustic parameters, such as harmonic-to-noise ratio (which has been proposed to be a major component in hierarchical processing of voices; Lewis et al., 2009), in particular during the first 40 ms of sounds, might engage echoic memory differently and be responsible for the early discrimination of conditions. This discrimination indicates that the high saliency of emotion influences preattentive mechanisms as soon as 6 years of age.

The presence of a double-peaked response for the emotional condition is in accordance with previous MMN studies conducted in children with vocal stimuli, which have reported two negative deflections in response to commanding deviant stimuli, though these occurred at later latencies (Korpilahti et al., 2007; Lindström et al., 2012). This latency difference could have arisen from the use of shorter stimuli in the present study (400 ms long) than in other studies (between 538 and 775 ms). Because the stimulus durations differed greatly, the timing of occurrence of spectro-temporal differences might have influenced the timing of the change detection response. The difference could also be related to the rather short SOA used (700 ms), even though this parameter should have affected MMN amplitudes rather than latencies (Sabri & Campbell, 2001). These findings suggest the existence of sequential processing of prosodic change, requiring two steps in children. The similarity of the neutral and emotional waveforms and scalp distributions suggests that this two-step process relies on the same neuronal networks for both conditions.

In contrast, the MMN amplitudes and latencies did not differ across conditions in children. Previous emotional MMN studies conducted in children have either reported differences between healthy and pathological groups (Korpilahti et al., 2007) or differences between emotions in typically developing subjects (Lindström et al., 2012). Because previous MMN studies conducted in children did not use a neutral MMN comparison, assessment of the emotional deviancy effect per se has not been possible. In our study, the use of both neutral and emotional deviants suggests that in children the MMN reflects the change detection process, but that its amplitude is not modulated by the emotional nature of the change. Lindström et al. (2012) showed significant MMN amplitude and latency differences between commanding, sad, and fearful deviants, suggesting that the MMN might index the discrimination between different emotions during childhood. However, this study did not use a control sequence. Given the absence of such a sequence, MMNs were obtained from the subtraction of standard and deviant stimuli differing in their physical features and probabilities of occurrence, thus preventing them from ruling out an influence of low-level factors in the emotional MMN responses. These findings

emphasize how important it is to use proper control sequences (Ruhnau, Herrmann, & Schröger, 2012), particularly in children, in whom the neural adaptation effect might be specific.

Altogether, our findings indicate that in children the automatic detection of emotional change is indexed by modulation of an early discriminative response rather than by variation in the MMN latency range. Our study also showed early automatic processing of vocal emotional modulations in school-age children. In this way, it provides information beyond previous developmental studies on emotion processing, which showed a late effect of emotion on auditory ERPs. Early effects might not have been observed in these previous studies because they were performed either in 7-month-old infants (Grossmann et al., 2005) or during an explicit detection task (Chronaki et al., 2012).

Comparison of brain responses to prosodic changes in children and in adults

In both children and adults, emotional change was clearly differentiated from neutral change at a preattentive level, in either the early discriminative negativity or the MMN, respectively. These findings are in accordance with behavioral studies that have pointed out an ability for children to detect emotion involuntarily. When testing response times to different prosodic voices through a same–different task, children’s mean response times to commanding stimuli were shorter than their response times to neutral stimuli (Lindström et al., 2012). Moreover, response times were found to be longer in children when they faced contradictory emotional speech and prosody than when the speech and prosody were congruent (Morton & Trehub, 2001).

Although previous studies have reported either a right-lateralized emotional MMN (Korpilahti et al., 2007) or no lateralization of the response in children (Lindström et al., 2012), the present study showed that regardless of condition, the MMN peaked at central sites, revealing no hemispheric dominance. This result was consistent across age groups, even if a nonsignificant tendency toward right-hemisphere lateralization was observed. This absence of significant lateralization corroborates the existence of a bilateral network for the implicit processing of angry prosody (Castelluccio, Myers, Schuh, & Eigsti, 2016; Frühholz, Ceravolo, & Grandjean, 2012), but it might also have been influenced by stimulus type (e.g., stimuli with/without speech content) and attention level (e.g., implicit/explicit; Frühholz & Grandjean, 2013).

Despite these common points, differences between the groups were also reported at the preattentive level. A developmental effect was found on the early negative response. The amplitude of the early deflection was reduced with age and appeared nonsignificant at adulthood. Although no group difference was shown for MMN amplitudes, a significant interaction between age and condition was reported on MMN

latencies. The emotional MMN latency was significantly shorter in adults than in children, whereas no such effect was observed for the neutral MMN. This result highlights larger age-related changes in the emotional condition in order to reach an adult-like preattentive response to emotional change. The correlations between age and MMN latency in the child group revealed that the neutral MMN latency decreased between 6 and 12 years of age, whereas no correlation was found for the emotional MMN latency. However, a smaller intercept was found for the emotional than for the neutral correlation. Because the latency of the emotional MMN is longer in children than in adults, these results would indicate that the shortening of the emotional MMN latency takes place during adolescence.

An emotion-specific response was also evidenced at the P3a level. In both groups, P3a amplitudes were smaller in the emotional than in the neutral condition. However, a shorter P3a latency was shown for the emotional deviant than for the neutral deviant in children, whereas no latency difference was present in adults. Analyses of linear correlations of the slopes and intercepts of P3a latencies with age tend to indicate that this finding reflects an earlier, rather than a faster, maturation of the orientation of attention to emotional relative to neutral change.

Overall, neutral change detection responses appeared similar between adults and children. In contrast, emotional change detection responses changed with age. Since developmental changes in the MMN take place along with brain maturation, the difference between children and adults for the emotional MMN but not for the neutral MMN response highlights the specificity of the automatic detection of emotional deviancy.

The presence of an early discriminative response in both conditions suggests that children more likely use low-level acoustic factors to differentiate neutral and emotional conditions. Yet children present an emotion-specific P3a response and an earlier maturation of attention orientation to emotional than to neutral change. We hypothesize that this early emotional P3a maturation promotes an automatization of preattentive brain responses to emotional change during adolescence. This developmental change would allow the growing brain to process emotional change faster than neutral change and to be more independent of echoic memory processes.

Limitations

Although our study provides new insight into emotion-specific discrimination in school-age children, it has some limitations. First, more subjects would allow researchers to (1) evaluate the developmental trajectory through the use of multiple age groups and (2) estimate potential gender effects, as previously reported in adults (Schirmer et al., 2005).

Moreover, in this study, conclusions drawn about age-related differences were based on group comparisons, and longitudinal studies on the development of prosodic change detection could provide more reliable information about the maturation of neutral and emotional prosodic change detection. Second, behavioral tests of emotional prosody perception and production would have been beneficial for this study. Such tests would allow researchers to establish more clearly the developmental relevance of an emotion-specific response in children. Finally, we used only an angry deviant in this study. In this regard, the observed effects might not be generalizable to all basic emotions, or even to all negative basic emotions. Further work will be necessary to assess whether similar responses are present for other emotions and how they relate to the children's social behavior.

Conclusions

This study has allowed, for the first time, the identification of specific effects of emotional deviancy in school-age children. The paradigm used made it possible to overcome confounding effects related to emotion processing in order to observe brain responses to emotional change per se. The findings underlined that despite developmental differences, both children and adults present an emotion-specific response in the automatic detection of prosodic change. Similar research into neurodevelopmental pathologies such as autism spectrum disorder, characterized by both social and sensory difficulties, would be of great interest for determining why the processing of changing emotional stimuli is so challenging for these patients.

Author note We thank all the volunteers for their time and effort while participating in this study. We are grateful to Luce Corneau for her help during the EEG recording, Remy Magne for technical support, and Joëlle Malvy and Mathieu Lemaire for including subjects. This work was supported by a French National Research Agency grant (ANR-12-JSH2-0001-01-AUTATTEN). J.C. was supported by an INSERM-Région Centre grant. The funding sources had no involvement in the study design, data collection and analysis, or the writing of the report. All authors declare that they have no conflicts of interest.

References

- Aguera, P.-E., Jerbi, K., Caclin, A., & Bertrand, O. (2011). ELAN: A software package for analysis and visualization of MEG, EEG, and LFP signals. *Computational Intelligence and Neuroscience*, 2011, 158970. <https://doi.org/10.1155/2011/158970>

- Agus, T. R., Paquette, S., Suied, C., Pressnitzer, D., & Belin, P. (2017). Voice selectivity in the temporal voice area despite matched low-level acoustic cues. *Scientific Reports*, *7*, 11526. <https://doi.org/10.1038/s41598-017-11684-1>
- Alho, K., Winkler, I., Escera, C., Huotilainen, M., Virtanen, J., Jääskeläinen, I. P., . . . Ilmoniemi, R. J. (1998). Processing of novel sounds and frequency changes in the human auditory cortex: magnetoencephalographic recordings. *Psychophysiology*, *35*, 211–224.
- Aziz-Zadeh, L., Sheng, T., & Gheyntanchi, A. (2010). Common premotor regions for the perception and production of prosody and correlations with empathy and prosodic ability. *PLoS ONE*, *5*, e8759. <https://doi.org/10.1371/journal.pone.0008759>
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*, 614–636.
- Bishop, D. V., Hardiman, M. J., & Barry, J. G. (2011). Is auditory discrimination mature by middle childhood? A study using time–frequency analysis of mismatch responses from 7 years to adulthood. *Developmental Science*, *14*, 402–416. <https://doi.org/10.1111/j.1467-7687.2010.00990.x>
- Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glott International*, *5*, 341–345.
- Bruneau, N., Roux, S., Cléry, H., Rogier, O., Bidet-Caulet, A., & Barthélémy, C. (2013). Early neurophysiological correlates of vocal versus non-vocal sound processing in adults. *Brain Research*, *1528*, 20–27. <https://doi.org/10.1016/j.brainres.2013.06.008>
- Campanella, S., Gaspard, C., Debatisse, D., Bruyer, R., Crommelinck, M., & Guérit, J.-M. (2002). Discrimination of emotional facial expressions in a visual oddball task: An ERP study. *Biological Psychology*, *59*, 171–186.
- Castelluccio, B. C., Myers, E. B., Schuh, J. M., & Eigsti, I.-M. (2016). Neural substrates of processing anger in language: Contributions of prosody and semantics. *Journal of Psycholinguistic Research*, *45*, 1359–1367. <https://doi.org/10.1007/s10936-015-9405-z>
- Ceponiene, R., Lepistö, T., Alku, P., Aro, H., & Näätänen, R. (2003). Event-related potential indices of auditory vowel processing in 3-year-old children. *Clinical Neurophysiology*, *114*, 652–661.
- Cheng, Y., Lee, S.-Y., Chen, H.-Y., Wang, P.-Y., & Decety, J. (2012). Voice and emotion processing in the human neonatal brain. *Journal of Cognitive Neuroscience*, *24*, 1411–1419. https://doi.org/10.1162/jocn_a_00214
- Choudhury, N. A., Parascando, J. A., & Benasich, A. A. (2015). Effects of presentation rate and attention on auditory discrimination: A comparison of long-latency auditory evoked potentials in school-aged children and adults. *PLoS ONE*, *10*, e0138160. <https://doi.org/10.1371/journal.pone.0138160>
- Chronaki, G., Benikos, N., Fairchild, G., & Sonuga-Barke, E. J. S. (2014). A typical neural responses to vocal anger in attention-deficit/hyperactivity disorder. *Journal of Child Psychology and Psychiatry*, *56*, 477–487. <https://doi.org/10.1111/jcpp.12312>
- Chronaki, G., Broyd, S., Garner, M., Hadwin, J. A., Thompson, M. J. J., & Sonuga-Barke, E. J. S. (2012). Isolating N400 as neural marker of vocal anger processing in 6–11-year old children. *Developmental Cognitive Neuroscience*, *2*, 268–276. <https://doi.org/10.1016/j.dcn.2011.11.007>
- De Lucia, M., Clarke, S., & Murray, M. M. (2010). A temporal hierarchy for conspecific vocalization discrimination in humans. *Journal of Neuroscience*, *30*, 11210–11221. <https://doi.org/10.1523/JNEUROSCI.2239-10.2010>
- Domínguez-Borràs, J., García-García, M., & Escera, C. (2008a). Emotional context enhances auditory novelty processing: Behavioural and electrophysiological evidence. *European Journal of Neuroscience*, *28*, 1199–1206. <https://doi.org/10.1111/j.1460-9568.2008.06411.x>
- Domínguez-Borràs, J., García-García, M., & Escera, C. (2008b). Negative emotional context enhances auditory novelty processing. *NeuroReport*, *19*, 503–507. <https://doi.org/10.1097/WNR.0b013e3282f85bec>
- Engle, R. W., Fidler, D. S., & Reynolds, L. H. (1981). Does echoic memory develop? *Journal of Experimental Child Psychology*, *32*, 459–473.
- Escera, C., Alho, K., Schröger, E., & Winkler, I. (2000). Involuntary attention and distractibility as evaluated with event-related brain potentials. *Audiology and Neuro-Otology*, *5*, 151–166. <https://doi.org/10.1159/000013877>
- Fishman, Y. I. (2014). The mechanisms and meaning of the mismatch negativity. *Brain Topography*, *27*, 500–526. <https://doi.org/10.1007/s10548-013-0337-3>
- Flom, R., & Bahrick, L. E. (2007). The development of infant discrimination of affect in multimodal and unimodal stimulation: The role of intersensory redundancy. *Developmental Psychology*, *43*, 238–252. <https://doi.org/10.1037/0012-1649.43.1.238>
- Friend, M. (2000). Developmental changes in sensitivity to vocal paralinguistic. *Developmental Science*, *3*, 148–162. <https://doi.org/10.1111/1467-7687.00108>
- Frühholz, S., Ceravolo, L., & Grandjean, D. (2012). Specific brain networks during explicit and implicit decoding of emotional prosody. *Cerebral Cortex*, *22*, 1107–1117. <https://doi.org/10.1093/cercor/bhr184>
- Frühholz, S., & Grandjean, D. (2013). Multiple subregions in superior temporal cortex are differentially sensitive to vocal expressions: A quantitative meta-analysis. *Neuroscience & Biobehavioral Reviews*, *37*, 24–35. <https://doi.org/10.1016/j.neubiorev.2012.11.002>
- Frühholz, S., Hofstetter, C., Cristinzio, C., Saj, A., Seeck, M., Vuilleumier, P., & Grandjean, D. (2015). Asymmetrical effects of unilateral right or left amygdala damage on auditory cortical processing of vocal emotions. *Proceedings of the National Academy of Sciences*, *112*, 1583–1588. <https://doi.org/10.1073/pnas.1411315112>
- Frühholz, S., Trost, W., & Kotz, S. A. (2016). The sound of emotions—Towards a unifying neural network perspective of affective sound processing. *Neuroscience & Biobehavioral Reviews*, *68*, 96–110. <https://doi.org/10.1016/j.neubiorev.2016.05.002>
- Gebauer, L., Skewes, J., Hørlyck, L., & Vuust, P. (2014). Atypical perception of affective prosody in Autism Spectrum Disorder. *NeuroImage: Clinical*, *6*, 370–378. <https://doi.org/10.1016/j.nicl.2014.08.025>
- Gomot, M., Giard, M. H., Roux, S., Barthélémy, C., & Bruneau, N. (2000). Maturation of frontal and temporal components of mismatch negativity (MMN) in children. *NeuroReport*, *11*, 3109–3112.
- Goodfellow, S., & Nowicki, S. (2009). Social adjustment, academic adjustment, and the ability to identify emotion in facial expressions of 7-year-old children. *Journal of Genetic Psychology*, *170*, 234–243. <https://doi.org/10.1080/00221320903218281>
- Goydke, K. N., Altenmüller, E., Möller, J., & Münte, T. F. (2004). Changes in emotional tone and instrumental timbre are reflected by the mismatch negativity. *Cognitive Brain Research*, *21*, 351–359. <https://doi.org/10.1016/j.cogbrainres.2004.06.009>
- Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). The voices of wrath: Brain responses to angry prosody in meaningless speech. *Nature Neuroscience*, *8*, 145–146. <https://doi.org/10.1038/nm1392>
- Graux, J., Gomot, M., Roux, S., Bonnet-Brilhault, F., Camus, V., & Bruneau, N. (2013). My voice or yours? An electrophysiological study. *Brain Topography*, *26*, 72–82. <https://doi.org/10.1007/s10548-012-0233-2>
- Grommet, E. K., Droit-Volet, S., Gil, S., Hemmes, N. S., Baker, A. H., & Brown, B. L. (2011). Time estimation of fear cues in human observers. *Behavioural Processes*, *86*, 88–93. <https://doi.org/10.1016/j.beproc.2010.10.003>

- Grossmann, T., Striano, T., & Friederici, A. D. (2005). Infants' electric brain responses to emotional prosody. *NeuroReport*, *16*, 1825–1828. <https://doi.org/10.1097/01.wnr.0000185964.34336.b1>
- Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology*, *28*, 240–244.
- Hung, A.-Y., Ahveninen, J., & Cheng, Y. (2013). Atypical mismatch negativity to distressful voices associated with conduct disorder symptoms. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *54*, 1016–1027. <https://doi.org/10.1111/jcpp.12076>
- Iredale, J. M., Rushby, J. A., McDonald, S., Dimoska-Di Marco, A., & Swift, J. (2013). Emotion in voice matters: neural correlates of emotional prosody perception. *International Journal of Psychophysiology*, *89*, 483–490. <https://doi.org/10.1016/j.ijpsycho.2013.06.025>
- Jacobsen, T., & Schröger, E. (2001). Is there pre-attentive memory-based comparison of pitch? *Psychophysiology*, *38*, 723–727. <https://doi.org/10.1111/1469-8986.3840723>
- Jiang, A., Yang, J., & Yang, Y. (2014). MMN responses during implicit processing of changes in emotional prosody: An ERP study using Chinese pseudo-syllables. *Cognitive Neurodynamics*, *8*, 499–508. <https://doi.org/10.1007/s11571-014-9303-3>
- Kanske, P., & Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Research*, *1148*, 138–148. <https://doi.org/10.1016/j.brainres.2007.02.044>
- Keil, A., Bradley, M. M., Hauk, O., Rockstroh, B., Elbert, T., & Lang, P. J. (2002). Large-scale neural correlates of affective picture processing. *Psychophysiology*, *39*, 641–649. <https://doi.org/10.1111/1469-8986.3950641>
- Kimura, M., Katayama, J., Ohira, H., & Schröger, E. (2009). Visual mismatch negativity: New evidence from the equiprobable paradigm. *Psychophysiology*, *46*, 402–409. <https://doi.org/10.1111/j.1469-8986.2008.00767.x>
- Korpilahti, P., & Lang, H. A. (1994). Auditory ERP components and mismatch negativity in dysphasic children. *Electroencephalography and Clinical Neurophysiology*, *91*, 256–264.
- Korpilahti, P., Jansson-Verkasalo, E., Mattila, M.-L., Kuusikko, S., Suominen, K., Rytty, S., . . . Moilanen, I. (2007). Processing of affective speech prosody is impaired in Asperger syndrome. *Journal of Autism and Developmental Disorders*, *37*, 1539–1549. <https://doi.org/10.1007/s10803-006-0271-2>
- Kraus, N., McGee, T., Micco, A., Sharma, A., Carrell, T., & Nicol, T. (1993). Mismatch negativity in school-age children to speech stimuli that are just perceptibly different. *Electroencephalography and Clinical Neurophysiology*, *88*, 123–130.
- Kraus, N., McGee, T., Sharma, A., Carrell, T., & Nicol, T. (1992). Mismatch negativity event-related potential elicited by speech stimuli. *Ear and Hearing*, *13*, 158–164.
- Kreiman, J., & Sidtis, D. (2011). Foundations of voice studies: An interdisciplinary approach to voice production and perception. Hoboken, NJ: Wiley.
- Lanting, C. P., Briley, P. M., Sumner, C. J., & Krumbholz, K. (2013). Mechanisms of adaptation in human auditory cortex. *Journal of Neurophysiology*, *110*, 973–983. <https://doi.org/10.1152/jn.00547.2012>
- Leitman, D. I., Sehatpour, P., Garidis, C., Gomez-Ramirez, M., & Javitt, D. C. (2011). Preliminary evidence of pre-attentive distinctions of frequency-modulated tones that convey affect. *Frontiers in Human Neuroscience*, *5*, 96. <https://doi.org/10.3389/fnhum.2011.00096>
- Leppänen, J. M., & Nelson, C. A. (2009). Tuning the developing brain to social signals of emotions. *Nature Reviews Neuroscience*, *10*, 37–47. <https://doi.org/10.1038/nrn2554>
- Lewis, J. W., Talkington, W. J., Walker, N. A., Spirou, G. A., Jajosky, A., Frum, C., & Brefczynski-Lewis, J. A. (2009). Human cortical organization for processing vocalizations indicates representation of harmonic structure as a signal attribute. *Journal of Neuroscience*, *29*, 2283–2296. <https://doi.org/10.1523/JNEUROSCI.4145-08.2009>
- Liebethal, E., Silbersweig, D. A., & Stern, E. (2016). The language, tone and prosody of emotions: neural substrates and dynamics of spoken-word emotion perception. *Frontiers in Neuroscience*, *10*, 506. <https://doi.org/10.3389/fnins.2016.00506>
- Lindström, R., Lepistö, T., Makkonen, T., & Kujala, T. (2012). Processing of prosodic changes in natural speech stimuli in school-age children. *International Journal of Psychophysiology*, *86*, 229–237. <https://doi.org/10.1016/j.ijpsycho.2012.09.010>
- Liu, H.-M., Chen, Y., & Tsao, F.-M. (2014). Developmental changes in mismatch responses to mandarin consonants and lexical tones from early to middle childhood. *PLoS ONE*, *9*, e95587. <https://doi.org/10.1371/journal.pone.0095587>
- Lui, M. A., Penney, T. B., & Schirmer, A. (2011). Emotion effects on timing: Attention versus pacemaker accounts. *PLoS ONE*, *6*, e21829. <https://doi.org/10.1371/journal.pone.0021829>
- Martin, B. A., Shafer, V. L., Morr, M. L., Kreuzer, J. A., & Kurtzberg, D. (2003). Maturation of mismatch negativity: a scalp present density analysis. *Ear and Hearing*, *24*, 463–471. <https://doi.org/10.1097/01.AUD.0000100306.20188.0E>
- May, P. J. C., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, *47*, 66–122. <https://doi.org/10.1111/j.1469-8986.2009.00856.x>
- Morton, J. B., & Trehub, S. E. (2001). Children's understanding of emotion in speech. *Child Development*, *72*, 834–843. <https://doi.org/10.1111/1467-8624.00318>
- Murray, M. M., Camen, C., Gonzalez Andino, S. L., Bovet, P., & Clarke, S. (2006). Rapid brain discrimination of sounds of objects. *Journal of Neuroscience*, *26*, 1293–1302. <https://doi.org/10.1523/JNEUROSCI.4511-05.2006>
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*, 2544–2590. <https://doi.org/10.1016/j.clinph.2007.04.026>
- Pakarinen, S., Sokka, L., Leinikka, M., Henelius, A., Korpela, J., & Huutilainen, M. (2014). Fast determination of MMN and P3a responses to linguistically and emotionally relevant changes in pseudoword stimuli. *Neuroscience Letters*, *577C*, 28–33. <https://doi.org/10.1016/j.neulet.2014.06.004>
- Paquette, N., Vannasing, P., Lefrançois, M., Lefebvre, F., Roy, M.-S., McKerral, M., . . . Gallagher, A. (2013). Neurophysiological correlates of auditory and language development: A mismatch negativity study. *Developmental Neuropsychology*, *38*, 386–401. <https://doi.org/10.1080/87565641.2013.805218>
- Paulmann, S., & Kotz, S. A. (2008). Early emotional prosody perception on the basis of different speaker voices. *NeuroReport*, *19*, 209–213. <https://doi.org/10.1097/WNR.0b013e3282f454db>
- Pegado, F., Bekinschtein, T., Chausson, N., Dehaene, S., Cohen, L., & Naccache, L. (2010). Probing the lifetimes of auditory novelty detection processes. *Neuropsychologia*, *48*, 3145–3154. <https://doi.org/10.1016/j.neuropsychologia.2010.06.030>
- Pinheiro, A. P., Barros, C., Vasconcelos, M., Obermeier, C., & Kotz, S. A. (2017). Is laughter a better vocal change detector than a growl? *Cortex*, *92*, 233–248. <https://doi.org/10.1016/j.cortex.2017.03.018>
- Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America*, *90*, 2956–2970.
- Quam, C., & Swingle, D. (2012). Development in children's interpretation of pitch cues to emotions. *Child Development*, *83*, 236–250. <https://doi.org/10.1111/j.1467-8624.2011.01700.x>
- Rachman, L., Liuni, M., Arias, P., Lind, A., Johansson, P., Hall, L., . . . Aucouturier, J.-J. (2018). DAVID: An open-source platform for real-time transformation of infra-segmental emotional cues in running

- speech. *Behavior Research Methods*, 50, 323–343. <https://doi.org/10.3758/s13428-017-0873-y>
- Rigoulot, S., & Armony, J. L. (2016). Early selectivity for vocal and musical sounds: Electrophysiological evidence from an adaptation paradigm. *European Journal of Neuroscience*, 44, 2786–2794. <https://doi.org/10.1111/ejn.13391>
- Rogier, O., Roux, S., Belin, P., Bonnet-Brilhault, F., & Bruneau, N. (2010). An electrophysiological correlate of voice processing in 4- to 5-year-old children. *International Journal of Psychophysiology*, 75, 44–47. <https://doi.org/10.1016/j.ijpsycho.2009.10.013>
- Ruhnau, P., Herrmann, B., & Schröger, E. (2012). Finding the right control: The mismatch negativity under investigation. *Clinical Neurophysiology*, 123, 507–512. <https://doi.org/10.1016/j.clinph.2011.07.035>
- Sabri, M., & Campbell, K. B. (2001). Effects of sequential and temporal probability of deviant occurrence on mismatch negativity. *Cognitive Brain Research*, 12, 171–180.
- Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology*, 62, 437–448.
- Sauter, D. A., Panattoni, C., & Happé, F. (2013). Children's recognition of emotions from vocal cues. *British Journal of Developmental Psychology*, 31, 97–113. <https://doi.org/10.1111/j.2044-835X.2012.02081.x>
- Schirmer, A., Escoffier, N., Cheng, X., Feng, Y., & Penney, T. B. (2016). Detecting temporal change in dynamic sounds: On the role of stimulus duration, speed, and emotion. *Frontiers in Psychology*, 6, 2055. <https://doi.org/10.3389/fpsyg.2015.02055>
- Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, 10, 24–30. <https://doi.org/10.1016/j.tics.2005.11.009>
- Schirmer, A., Striano, T., & Friederici, A. D. (2005). Sex differences in the preattentive processing of vocal emotional expressions. *NeuroReport*, 16, 635–639.
- Schumann, C. M., Hamstra, J., Goodlin-Jones, B. L., Lotspeich, L. J., Kwon, H., Buonocore, M. H., . . . Amaral, D. G. (2004). The amygdala is enlarged in children but not adolescents with autism; the hippocampus is enlarged at all ages. *Journal of Neuroscience*, 24, 6392–6401. <https://doi.org/10.1523/JNEUROSCI.1297-04.2004>
- Shafer, V. L., Yu, Y. H., & Datta, H. (2010). Maturation of speech discrimination in 4- to 7-yr-old children as indexed by event-related potential mismatch responses. *Ear and Hearing*, 31, 735–745. <https://doi.org/10.1097/AUD.0b013e3181e5d1a7>
- Taylor, J. G., & Fragopanagos, N. F. (2005). The interaction of attention and emotion. *Neural Networks*, 18, 353–369. <https://doi.org/10.1016/j.neunet.2005.03.005>
- Thönnessen, H., Boers, F., Dammers, J., Chen, Y.-H., Norra, C., & Mathiak, K. (2010). Early sensory encoding of affective prosody: Neuromagnetic tomography of emotional category changes. *NeuroImage*, 50, 250–259. <https://doi.org/10.1016/j.neuroimage.2009.11.082>
- Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, 9, 585–594. <https://doi.org/10.1016/j.tics.2005.10.011>
- Wechsler, D. (2005). WISC-IV—Echelle d'intelligence de Wechsler pour enfants et adolescents—Quatrième édition—ECPA. Retrieved April 28, 2016, from www.ecpa.fr/psychologie-clinique/test.asp?id=1391
- Wechsler, D. (2011). WAIS-IV—Echelle d'intelligence de Wechsler pour adultes—Quatrième Edition—ECPA. Retrieved April 28, 2016, from www.ecpa.fr/psychologie-clinique/test.asp?id=1968
- Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., & Ethofer, T. (2008). Cerebral processing of emotional prosody—Influence of acoustic parameters and arousal. *NeuroImage*, 39, 885–893. <https://doi.org/10.1016/j.neuroimage.2007.09.028>
- Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., & Ackermann, H. (2005). Identification of emotional intonation evaluated by fMRI. *NeuroImage*, 24, 1233–1241. <https://doi.org/10.1016/j.neuroimage.2004.10.034>