

Neural correlates of infants' sensitivity to vocal expressions of peers



Manuela Missana^{a,b}, Nicole Altvater-Mackensen^b, Tobias Grossmann^{b,c,*}

^a Department of Early Child Development and Culture, Leipzig University, Marschnerstrasse 29, 04109 Leipzig, Germany

^b Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04103 Leipzig, Germany

^c Department of Psychology, University of Virginia, Gilmer Hall, 485 McCormick Road, Charlottesville, VA 22903, USA

ARTICLE INFO

Keywords:

Emotion

Infants

Vocal expressions

Event-related brain potentials

ABSTRACT

Responding to others' emotional expressions is an essential and early developing social skill among humans. Much research has focused on how infants process facial expressions, while much less is known about infants' processing of vocal expressions. We examined 8-month-old infants' processing of other infants' vocalizations by measuring event-related brain potentials (ERPs) to positive (infant laughter), negative (infant cries), and neutral (adult hummed speech) vocalizations. Our ERP results revealed that hearing another infant cry elicited an enhanced negativity (N200) at temporal electrodes around 200 ms, whereas listening to another infant laugh resulted in an enhanced positivity (P300) at central electrodes around 300 ms. This indexes that infants' brains rapidly respond to a crying peer during early auditory processing stages, but also selectively respond to a laughing peer during later stages associated with familiarity detection processes. These findings provide evidence for infants' sensitivity to vocal expressions of peers and shed new light on the neural processes underpinning emotion processing in infants.

1. Introduction

The human voice is an important social stimulus in our environment. Our ability to analyze and effectively respond to information carried in the voice plays a pivotal role for social functioning. The voice not only carries speech information but it can also be seen as an “auditory face” that enables us to recognize individuals and their emotional states (Belin et al., 2004). Responding to others' emotional expressions is an essential and early developing social skill among humans (Grossmann, 2010; Grossmann and Friederici, 2012; Leppanen and Nelson, 2009). Behavioral and neuroscientific studies provide evidence that within the first year of life, infants begin to reliably discriminate between a variety of positive and negative emotional facial expressions, such as happy, sad, angry and fearful expressions (Grossmann, 2013; Kotsoni et al., 2001; LaBarbera et al., 1976; Nelson and de Haan, 1996; Serrano et al., 1992). Moreover, there is behavioral evidence showing that, by 5–7 months of age infants represent emotional information at the categorical level (e.g., Bornstein and Arterberry, 2003; Nelson et al., 1979). For example, infants recognize the similarity of happy facial expressions over changing identities (Kotsoni et al., 2001; Ludemann and Nelson, 1988). In addition, 6- and 7-month-old infants have been shown to distinguish between varying intensities of facial expressions, such as happy, fearful, and sad faces (Ludemann and Nelson, 1988; Striano et al., 2002). These

studies indicate that in the first year of life, infants develop the ability to discriminate and categorize emotional information from facial expressions. However, most research with infants has focused on visual (facial) emotion processing while much less is known about auditory (vocal) emotion processing during early development and its neural correlates.

From birth infants preferentially orient to voices and their prenatal hearing experiences in utero shape listening preferences exhibited as newborns (DeCasper and Fifer, 1980; DeCasper and Sigafos, 1983; DeCasper and Spence, 1986; Ecklund-Flores and Turkewitz, 1996). Another line of research shows that newborn infants sensitively respond to other infants' emotional vocalizations. Specifically, newborns show behavioral signs of self-distress when listening to a peer's crying but not when they listened to their own cries or non-emotional sounds (Dondi et al., 1999; Simner, 1971). This sensitivity to hearing another infant cry continues beyond the newborn phase as it can be seen in infants up to the age of 9 months (Geangu et al., 2010). The existence of a sensitivity to other infants' emotional vocalizations from birth is particularly striking considering behavioral work showing that it is not until 5 months of age that infants distinguish between vocal emotions presented by adults (Flom and Bahrick, 2007). Moreover, these findings of infants' distress reactions to other infants' cries have been taken as evidence for a rudimentary form of empathic concern, which is known to be an important motivator of prosocial behavior in

* Corresponding author at: Department of Psychology, University of Virginia, Gilmer Hall, 485 McCormick Road, Charlottesville, VA 22903, USA.
E-mail address: grossmann@virginia.edu (T. Grossmann).

<http://dx.doi.org/10.1016/j.dcn.2017.04.003>

Received 26 August 2016; Received in revised form 10 April 2017; Accepted 11 April 2017

Available online 15 April 2017

1878-9293/ © 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

children (Davidov et al., 2013; Geangu et al., 2010; Roth-Hanania et al., 2011). This further supports the notion that sensitive responding to emotional expressions serves as a foundational ability guiding social behavior.

Infants' sensitivity to emotional voices has also been investigated using neuroscientific methods such as event-related brain potentials (ERPs), functional near-infrared spectroscopy (fNIRS), and functional magnetic resonance imaging (fMRI) (Blasi et al., 2011; Graham et al., 2013; Grossmann et al., 2013; Grossmann et al., 2010a,b; Grossmann et al., 2005; Lloyd-Fox et al., 2012). These studies show that cortical regions within the auditory cortex (superior temporal cortex) become specialized in processing human voices between 4 and 7 months of age (see Grossmann and Friederici, 2012 for a review and discussion; Grossmann et al., 2010a,b; Lloyd-Fox et al., 2012). With respect to processing emotional information from voices, negative (angry) emotional voices have been shown to evoke greater brain responses than positive and neutral voices in 7-month-old-infants' superior temporal cortex, whereas positive (happy) emotional voices evoked greater brain responses in inferior frontal cortex (Grossmann et al., 2010a,b). Corresponding evidence for infants' ability to distinguish between positive and negative emotional voices comes from studies using ERPs showing that infants' brains at the age of 7 months discriminate between angry and happy voices (Grossmann et al., 2013, 2005). Specifically, happy voices have been found to elicit a greater positivity compared to a negative shift seen in response to angry voices at frontal and central electrodes peaking around 400 ms (Grossmann et al., 2013, 2005). Critically, the studies reviewed above used emotional speech (tone of voice) to examine emotion processing from voices and measured brain responses from superficially located brain regions because the fNIRS method is limited with respect to picking up responses from deeper or ventrally located brain structures (see Lloyd-Fox et al., 2010). However, there is a study that used non-speech emotional vocalizations and measured infants' brain responses to crying (sad), laughing (happy), and neutral adult vocalizations with fMRI in infants ranging from 3 to 7 months of age (Blasi et al., 2011). The fMRI data from this study show that listening to crying vocalizations when compared to neutral vocalizations resulted in greater activation in infants' insula and orbitofrontal cortex, whereas infant brain responses to laughing did not differ from their responses to neutral vocalizations. This suggests that, from early in ontogeny, listening to crying sounds engages brain regions implicated in emotional processing in adults (see Blasi et al., 2011 for a discussion).

In summary, the findings from the behavioral and neuroscientific studies reviewed here provide compelling evidence for infants' ability to detect and distinguish emotional cues from voices. However, the neural correlates of infants' processing of peer emotional vocalizations have not been investigated. To examine the neural correlates of processing emotional vocalization in peers is particularly important considering the host of prior behavioral work showing that from very early in ontogeny infants respond sensitively to emotional vocalizations of peers, especially the crying of other infants (Dondi et al., 1999; Geangu et al., 2010; Roth-Hanania et al., 2011; Simner, 1971). Furthermore, while the majority of behavioral work in this area has only used negative expressions (distressed crying peer), responses to positive expressions have rarely been considered and compared to negative expressions (for an exception, see Light et al., 2009). Therefore, in the current study, we investigated 8-month-old infants' processing of negative and positive emotional vocal expressions of other infants using ERP measures. Infants were presented with negative vocalizations (distressed crying peer), positive vocalizations (happy laughing peer) and neutral vocalizations (adult hummed speech) in order to investigate whether and how ERPs differ between negative and positive vocalizations as well as between emotionally loaded and neutral vocalizations. On the basis of prior behavioral work on infants' responses to crying peers (Dondi et al., 1999; Geangu et al., 2010; Simner, 1971) and ERP evidence for the rapid neural detection of

distress (fear) vocalizations in adults (Sauter and Eimer, 2010), we hypothesized that detecting crying in other infants would take place during early processing stages in the infant brain. In addition, based on prior ERP work using emotional speech with infants (Grossmann et al., 2005, 2013), we hypothesized that positive and negative emotional vocalizations would elicit differential responses during later stages of voice processing associated with familiarity (mid-latency positivity to familiar vocal sounds, see Grossmann et al., 2013) and affect memory (positive slow wave responses to emotionally charged sounds, see Grossmann et al., 2005). While we tested these specific hypotheses based on previous work, the exact timing and topography of the effects was examined in an exploratory manner since this is the first infant ERP study to use emotional vocalizations rather than emotional speech as stimuli.

2. Material and methods

2.1. Participants

The final sample consisted of 55 8-month-old infants aged between 238 and 261 days (33 females, Median age = 246 days). All infants came from a middle-class background in a medium-sized German city. The infants were born full term (between 37 and 41 weeks) and had a normal birth weight (> 2500 g). 41 additional infants were tested but were excluded from the final sample due to too many artifacts (following recommendations for ERP research with infants, the inclusion criterion applied to each infant was that at least 10 trials per condition were required, see de Haan, 2007). Note that an exclusion rate of 42% as observed in the current study is within the typical range for an ERP study with infants of this age. Ethical approval was obtained from the Ethics Committee of the University of Leipzig. All parents provided written informed consent prior to the study and were paid for their children's participation. The children were given a toy after the session.

2.2. Stimuli

The stimuli consisted of 3 different infant crying vocalizations, 3 different infant laughing vocalizations and 3 non-emotional vocalizations (neutral, hummed speech) from a female adult. The stimuli were 4 s in duration. To create a more natural appearance of the stimuli, they always started with a short 20 ms fade-in and ended with 20 ms fade-out. The infant vocalizations were purchased from the internet sound database (Sound Jay, Soundboard[®]) and were modified with respect to the stimulus length. The neutral vocalizations were taken from a previous study (Merrill et al., 2012). Analyses of the speech stimuli were performed with the "Praat Speech Processing Software" (Boersma & Weenik, Institute of Phonetics Sciences of the University of Amsterdam, Amsterdam, Netherlands). The mean fundamental frequency in Hertz for crying stimuli was 446.60, SD = 46.55 and for laughing stimuli 508.16, SD = 30.30. Since the fundamental frequency for the original neutral stimuli was low (M = 161.21 Hz, SD = 1.52) the fundamental frequency was modified and set to 207.60, SD = 0.32. This fundamental frequency adjustment was performed in order to bring the adult hummed speech stimuli closer to the frequency range previously reported for adult-directed speech (see Fernald et al., 1989). The mean fundamental frequency between crying and laughing sounds were not significantly different from each other, $t(4) = -1.92$, $p = 0.27$. The mean fundamental frequency was significantly higher for crying sounds than for neutral sounds, $t(4) = 8.89$, $p = 0.001$, and was also higher for laughing sounds than for neutral sounds, $t(4) = 17.18$, $p < 0.001$. The mean intensity for all stimuli in was 75 dB and did thus not differ across conditions.

In order to verify that the sounds were recognized as the intended emotional sounds (crying, laughing, neutral), the stimuli were validated in a pilot study with adult participants (N = 12; 6 female; Median

Table 1

This table shows the mean arousal and mean valence ratings obtained from a sample of adult raters.

	Mean arousal (SD)	Mean valence (SD)
crying	6.75 (1.48)	−2.03 (.77)
laughing	6.22 (1.38)	2.64 (1.01)
neutral	2.3 (1.26)	−0.14 (1.07)

age = 24.5 years). The participants were presented with the sound stimuli and were asked to classify the vocalizations in a 3-alternative forced-choice task (crying, laughing, neutral). Crying stimuli were recognized with an accuracy of 100%. Laughter and neutral stimuli were both recognized with an accuracy of 91.7%. Participants were also asked to rate the arousal on a 9-point self-assessment mannequin (SAM; see Bradley and Lang, 1994) scale ranging from “not aroused at all” to “extremely aroused” and valence on a 9-point SAM-scale ranging from “very unpleasant” to “very pleasant”. Arousal ratings did not differ between crying and laughing stimuli. Arousal ratings were higher for crying and laughing stimuli than for neutral stimuli. Crying vocalizations were rated as being more unpleasant than laughing vocalizations and neutral stimuli. Laughing vocalizations were rated as being more pleasant than neutral vocalizations (see Table 1).

2.3. Procedure

The infants were seated on their parent’s lap in a dimly lit, sound-attenuated and electrically shielded room during testing. In order to rule out the possibility that the parents influence the infants’ responses to the stimuli, we asked the parents not to talk to or interact with their infant during the course of the experiment and parents listened to classical music presented via headphones. Furthermore, the sessions were video-recorded so that trials during which the parent interacted with the infant could be excluded from the analysis. The stimuli were presented in a pseudo-randomized order with the exception that no two stimuli from the same emotional content were presented consecutively. The inter-trial interval lasted at least 1000 ms and varied randomly between the trials (see Missana et al., 2015). After a block of 9 stimuli infants were presented with a video clip of an infant moving slowly and sitting quietly in an infant car seat for 6 s. This was done in order to present the stimuli in a more ecologically valid context so that participating infants occasionally saw another infant who might be the likely source of the vocalizations. In addition, infants were presented with a non-social screensaver displaying animated bubbles

in order to keep infants’ attention while the auditory stimuli were presented. The EEG session ended when the infant became fussy or inattentive.

2.4. EEG measurement

The EEG was recorded from 27 Ag/AgCl electrodes attached to an elastic cap (EasyCap GmbH, Germany) using the 10–20 system of electrode placement. The data were online referenced to the CZ electrode and offline re-referenced to the algebraic mean of the left and right mastoid electrode. The horizontal electrooculogram (EOG) was recorded from two electrodes (F9, F10), which are part of the cap located at the outer canthi of both eyes. The vertical EOG was recorded from an electrode on the supraorbital ridge (Fp2), which is part of the cap and an additional single electrode on the infra-orbital ridge of the right eye. The EEG was amplified using a Porti-32/M-REFA amplifier (Twente Medical Systems International) and digitized at a rate of 500 Hz. Electrode impedances were kept between 5–20 kΩ. Data processing for ERP analysis was performed using an in-house software package EEP, commercially available under the name EEProbe™ (Advanced Neuro Technology, Enschede). The raw EEG data were band-pass filtered between 0.3 and 20 Hz (1501 points, 12 db/octave slope; see Grossmann et al., 2005, 2013) and the recordings were segmented into epochs time-locked to the stimulus onset, lasting from 100 ms before stimulus onset until 1000 ms post stimulus onset (total duration 1100 ms). The ERP analysis was time-locked to the stimulus onset, which always started with the 20 ms fade-in period. The epochs were baseline corrected by subtracting the average voltage in the 100-ms baseline period (prior to stimulus onset) from each post-stimulus data point. Data epochs were rejected off-line whenever the standard deviation within a gliding window of 200 ms exceeded 100 μV in any of the two bipolar EOG channels and 80 μV at the remaining EEG electrodes. At each electrode, artifact-free epochs were averaged separately for crying, laughing and neutral vocal expressions to compute the ERPs. The mean number of trials presented per condition was 23.85 (SD = 4.84). The criterion for the minimum number of trials for inclusion in the final ERP average was carried out in accordance with common standards in infant ERP research and required at least 10 artifact-free trials per condition for an infant to be included in the final ERP analysis (see de Haan, 2007). The mean number of trials included in the ERP average was 16.51 (SD = 4.42) for the crying stimuli, 16.29 (SD = 4.57) for laughing stimuli and 17.24 (SD = 4.56) for neutral stimuli. Note that stimulus duration was longer (four s) than in previous ERP studies on vocal emotion processing using speech (e.g., one second,

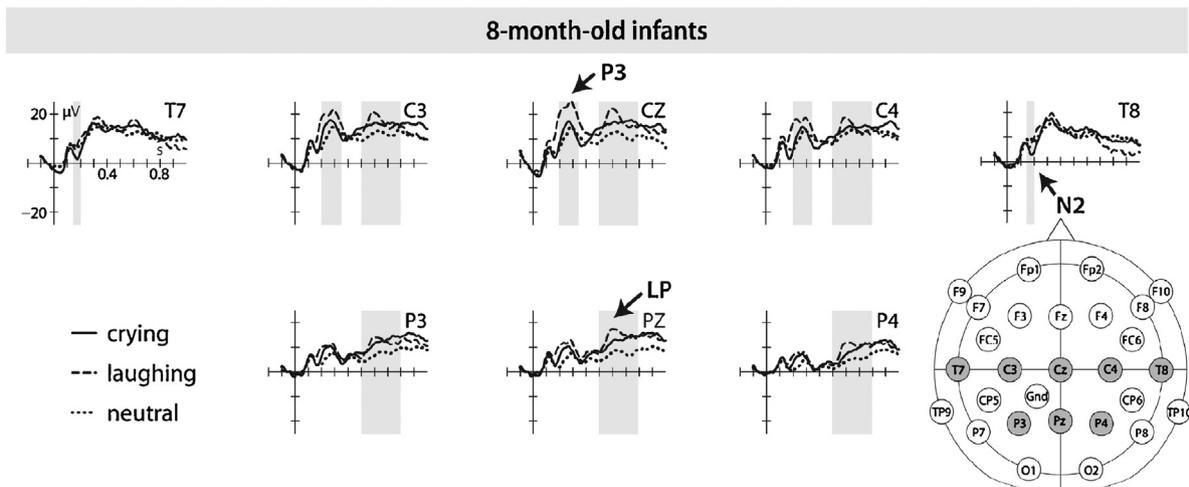


Fig. 1. This figure shows the event-related brain potentials (ERPs) elicited by crying (solid line), laughing (dashed line), and neutral (dotted line) vocal expressions. The time windows during which significant differences between the vocal expressions were observed are marked in grey and the ERP components (N2 [N200], P3 [P300], and LP) included in our analysis are labeled by an arrow.

Grossmann et al., 2005), which may have resulted in the overall lower mean number of trials (between 16 and 17) included per condition in the current study than in previous work (between 20 and 25, see Grossmann et al., 2005). It is also important to mention that our ERP measurement and analysis protocol closely followed the protocol of prior studies that examined emotional speech processing in infants of a similar age (Grossmann et al., 2005; Grossmann et al., 2013) to allow for a better comparison of the current study with these previous ERP studies.

Statistical analyses were based on prior work (Grossmann et al., 2005; Kushnerenko et al., 2002) and on visual inspection of the ERP waveforms and topography. On the basis of this information, mean amplitude effects were assessed during three time-windows for the following regions of interest (ROIs): temporal electrodes (T7, T8) during a time window of 150–200 ms (N200); central electrodes (C3, Cz, C4) for the time window of 200–350 ms (P300); and for central and parietal electrodes (C3, Cz, C4, P3, Pz, P4) for the time window of 500–800 ms (Late Positivity). Repeated measures ANOVAs with vocal expression (crying, laughter, and neutral) as within-subjects' factor were carried out for these regions and time windows to statistically examine effects of emotional expression on infants' ERP response. Note that the labeling of the ERP components (N200 and P300) included in the statistical analysis was carried out in adherence with the common standard of naming ERP components according to their polarity and peak latency (N200: negative deflection around 200 ms and P300: positive deflection around 300 ms). The Late Positivity did not have a clear peak and was therefore labeled according to when it occurred with respect to the other ERP components. It should also be noted that it is unclear how the N200 and P300 observed in the current infant study relate to adult ERP components that might have the same polarity and latency (label). It is unlikely that the ERP components observed in infants are equivalent to similarly labeled adult ERP components. Typically, the latency of infants' early ERP components is delayed by about 100 ms when compared to adults (e.g., see de Haan, 2007). Considering this delay and previous ERP research which directly compared between adults and infants (Grossmann et al., 2013) it is more likely that the N200 and P300 seen in infants are similar to the N100 and P200 in adults. This developmental difference is important to consider also when interpreting the results. Moreover, considering that previous studies did not report any latency differences for ERP components involved in vocal emotion processing and also that the visual inspection of the current ERP data did not suggest any latency differences we refrained from analyzing peak latency.

3. Results

3.1. N200 at temporal electrodes (150–200 ms): cry-sensitive ERP response

Our results revealed a significant main effect of vocal expression at temporal electrodes between 150 and 200 ms, $F(2,108) = 8.759$, $p < 0.001$. A significant difference in mean amplitudes was found between crying and laughing vocal expressions, $t(54) = -3.737$, $p < 0.001$, and between crying and neutral vocal expressions, $t(54) = -3.748$, $p < 0.001$. ERP responses to laughter and neutral vocal expressions did not differ significantly from each other, $t(54) = -0.494$, $p = 0.623$. Specifically, crying vocal expressions elicited a more negative-going ERP response ($M = 3.25 \mu\text{V}$, $SD = 9.49$) than laughing expressions ($M = 7.20 \mu\text{V}$, $SD = 9.24$) and neutral expressions ($M = 7.83 \mu\text{V}$, $SD = 7.20$) (see Fig. 1 and Fig. S1 (in the online version at DOI: <http://dx.doi.org/10.1016/j.dcn.2017.04.003>) for ERP responses on all recorded EEG channels).

3.2. P300 at central electrodes (200–350 ms): laughter-sensitive ERP response

Our analysis further revealed a significant main effect of vocal

expression at central electrodes between 200 and 350 ms, $F(2,108) = 12.68$, $p < 0.001$. A significant difference in mean amplitude was found between laughing and crying vocal expressions, $t(54) = -4.010$, $p < 0.001$, and between laughing and neutral vocal expressions, $t(54) = 4.152$, $p < 0.001$. No difference was found between the processing of crying and neutral vocal expressions, $t(54) = 1.119$, $p = 0.268$. Specifically, laughing vocal expressions elicited greater positivity ($M = 19.57 \mu\text{V}$, $SD = 10.87$) than crying expressions ($M = 13.17 \mu\text{V}$, $SD = 9.16$) and neutral expressions ($M = 11.48 \mu\text{V}$, $SD = 11.49$) (see Fig. 1).

3.3. Late Positivity (LP) at central and parietal electrodes (500–800 ms): emotion-sensitive ERP response

The results also revealed a significant main effect of vocal expression at central and parietal electrodes between 500 and 800 ms, $F(2,108) = 7.573$, $p = 0.001$. Mean amplitudes differed significantly between crying and neutral vocal expressions, $t(54) = 3.199$, $p = 0.002$, and between laughing and neutral vocal expressions, $t(54) = 3.713$, $p < 0.001$. No difference was found between the processing of crying and laughing vocal expressions $t(54) = -0.620$, $p = 0.538$. Specifically, both laughing vocal expressions ($M = 13.88 \mu\text{V}$, $SD = 7.86$) and crying expressions ($M = 12.98 \mu\text{V}$, $SD = 8.41$) elicited a greater positivity than neutral expressions ($M = 8.96 \mu\text{V}$, $SD = 8.96$) (see Fig. 1).

Note that all three main effects (N200, P300, and LP) reported above survive Bonferroni correction for multiple comparisons (Bonferroni corrected p -threshold is 0.017)

4. Discussion

The present study examined 8-month-old infants' neural processing of other infants' vocalizations using ERPs. Our results revealed that hearing another infant cry elicited an enhanced negativity at temporal electrodes around 200 ms, whereas listening to another infant laugh resulted in an enhanced positivity at central electrodes around 300 ms. This indexes that infants' brains rapidly respond to a crying peer during early auditory processing stages, but also selectively respond to a laughing peer during later processing stages associated with attention allocation. Moreover, our data revealed a late positivity (around 600 ms) that was greater in response to both emotional vocalizations (laughing and crying) than for the neutral control condition (adult hummed speech), suggesting a general enhancement of processing affectively loaded sounds in memory. These findings provide critical evidence for infants' sensitivity to vocal expressions of peers and shed new light on the neural processes involved in processing cries and laughter.

The current ERP data show an enhancement of the N200 component over temporal cortex in response to crying voices. This provides evidence for infants' rapid detection of crying sounds during early stages of auditory processing (see Csibra et al., 2008, for a review of infant auditory ERP components). In prior ERP studies using emotional speech no such modulation of the N200 was present (Grossmann et al., 2005, 2013), suggesting that this rapid effect is specific to emotional vocalizations, particularly crying. The current ERP data also add to prior fMRI research showing that, from early in ontogeny, infants' brains sensitively respond to crying sounds by activating insular and orbitofrontal brain regions (see Blasi et al., 2011). The selective enhancement of the early sensory processing when listening to others' crying voices seen in the current ERP study and the involvement of emotional brain regions observed in prior fMRI research (Blasi et al., 2011) might at least partly be linked to infants' behaviorally shown sensitivity to hearing other infants cry (Dondi et al., 1999; Simner, 1971). Moreover, this finding is in line with previous ERP research with adults, which also reported a rapid detection of vocal distress (fear) (Sauter and Eimer, 2010). In this context it is important to note that the

current ERP study critically differed from previous behavioral studies on contagious crying in infants by presenting the crying sounds for a much shorter duration for the purpose of ERP measurement. A shorter presentation time likely prevented (or at least attenuated) any overt expression of distress in the participating infants since signs of distress are typically only observed when crying is presented for prolonged durations (see Dondi et al., 1999; Sagi and Hoffman, 1976; Simner, 1971). Relatedly, in contrast to prior behavioral research with infants, we also presented positive (laughing) voices and neutral voices, which may have further reduced the likelihood of seeing overt signs of distress in the participating infants. Taken together, the result of an enhanced N200 in response to crying peers provides first evidence for the notion that processing vocal distress signals results in rapid brain responses in infants.

Our data also show an enhancement of the P300 component over central electrodes in response to laughing voices. This finding is consistent with previous ERP research using emotional speech showing a greater mid-latency positivity in response to happy speech when compared to angry speech (Grossmann et al., 2005, 2013), which has been taken to reflect a greater familiarity with happy (infant-directed) speech (see also Parise and Csibra, 2013). It is thus possible that the greater P300 to laughing peers is related to processes detecting the familiarity of positive affect expressed in voices. This is in line with the notion that, in typical development, positively-valenced vocal and facial stimuli are more commonly experienced by infants of this age than negatively-valenced ones (see Vaish et al., 2008). Despite the overall similarity, there also exists a difference with respect to prior work using emotional speech. Specifically, the positivity (P300) observed in the current study occurred somewhat earlier (about 50 ms) than in previous studies using emotional speech. This might be explained by the fact that in these prior studies infants needed to parse the speech information while (or in addition to) processing the emotional tone of voice, which may slow down the neural processes involved in emotion detection. The finding of an enhanced P300 in response to laughing peers in the current ERP study adds to the notion that infants' brains are highly sensitive to happy voices and perceive them as familiar sounds.

Furthermore, our ERP data showed an enhanced late positivity (slow wave) at central and parietal electrodes in response to both emotional vocalizations (crying and laughing sounds) when compared to the neutral control vocalizations (hummed speech). Positive slow waves in infant ERP have typically been associated with memory and learning processes, especially with the updating of an existing memory representation (see Grossmann et al., 2009; Nelson et al., 1998). This suggests that emotionally charged sounds might elicit the update of existing memory representations, whereas listening to neutral sounds engenders no such processes (see Grossmann et al., 2005). This underlines infants' sensitivity to emotional information conveyed through the voice and points to a potential mechanism by which emotional vocalizations impact learning and memory. However, this interpretation of the late positivity is limited because differences between the neutral and the emotional sound conditions observed in the current study might alternatively be explained by infants' neural capacity to distinguish between infant (both emotional conditions) and adult (neutral condition) sounds. Research is needed that directly compares between infant and adult vocalizations in order to clarify this issue.

In summary, the current study is the first to investigate the neural correlates of infants' processing of emotional peer vocalizations. These ERP results thus provide novel and valuable insights into the early development of emotion processing and represent a critical link to a host of prior behavioral studies focused on infants' responses to other infants' crying sounds. Our ERP results underscore infants' neural sensitivity to emotional sounds produced by other infants and index the rapid neural detection of other infants' vocal distress signals (crying). The ERP components (N200, P300, and late positivity) found to be sensitive to emotional vocalizations of peers in the current study

may serve as valuable neural markers, which, in future studies, can be used to examine individual differences in emotion processing. Identifying such individual differences in infants' emotion processing might also help predict empathic and helping behaviors in toddlerhood (see Davidov et al., 2013). All in all, the current findings provide further evidence for the notion that responding to others' emotional vocal expressions is an essential and early developing social skill among humans, which includes the category of peer vocalizations.

References

- Belin, P., Fecteau, S., Bédard, C., 2004. Thinking the voice: neural correlates of voice perception. *Trends Cogn. Sci.* 8 (3), 129–135. <http://dx.doi.org/10.1016/j.tics.2004.01.008>.
- Blasi, A., Mercure, E., Lloyd-Fox, S., Thomson, A., Brammer, M., Sauter, D., Deeley, Q., Barker, G.J., Renvall, V., Deoni, S., Gasston, D., Williams, S.C.R., Johnson, M.H., Simmons, A., Murphy, D.G.M., 2011. Early specialization for voice and emotion processing in the infant brain. *Curr. Biol.* 21 (14), 1220–1224. <http://dx.doi.org/10.1016/j.cub.2011.06.009>.
- Bornstein, M.H., Arterberry, M.E., 2003. Recognition, discrimination and categorization of smiling by 5-month-old infants. *Dev. Sci.* 6 (5), 585–599. <http://dx.doi.org/10.1111/1467-7687.00314>.
- Bradley, M.M., Lang, P.J., 1994. Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25 (1), 49–59. [http://dx.doi.org/10.1016/0005-7916\(94\)90063-9](http://dx.doi.org/10.1016/0005-7916(94)90063-9).
- Csibra, G., Kushnerenko, E., Grossmann, T., 2008. Electrophysiological methods in studying infant cognitive development. In: Nelson, C.A., Luciana, M. (Eds.), *Handbook of Developmental Cognitive Neuroscience*, vol. 2nd edition. MIT Press, Cambridge, MA, pp. 247–262.
- Davidov, M., Zahn-Waxler, C., Roth-Hanania, R., Knafo, A., 2013. Concern for others in the first year of life: theory, evidence, and avenues for research. *Child Dev. Perspect.* 7 (2), 126–131. <http://dx.doi.org/10.1111/cdep.12028>.
- DeCasper, A.J., Fifer, W., 1980. Of human bonding: newborns prefer their mothers' voices. *Science* 208 (4448), 1174–1176.
- DeCasper, A.J., Sigafos, A.D., 1983. The intrauterine heartbeat: a potent reinforcer for newborns. *Infant Behav. Dev.* 6 (1), 19–25. [http://dx.doi.org/10.1016/S0163-6383\(83\)80004-6](http://dx.doi.org/10.1016/S0163-6383(83)80004-6).
- DeCasper, A.J., Spence, M.J., 1986. Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behav. Dev.* 9 (2), 133–150. [http://dx.doi.org/10.1016/0163-6383\(86\)90025-1](http://dx.doi.org/10.1016/0163-6383(86)90025-1).
- Dondi, M., Simion, F., Caltran, G., 1999. Can newborns discriminate between their own cry and the cry of another newborn infant? *Dev. Psychol.* 35 (2), 418–426. <http://dx.doi.org/10.1037/0012-1649.35.2.418>.
- Ecklund-Flores, L., Turkewitz, G., 1996. Asymmetric headturning to speech and nonspeech in human newborns. *Dev. Psychobiol.* 29 (3), 205–217. [http://dx.doi.org/10.1002/\(SICI\)1098-2302\(199604\)29:3<205::AID-DEV2>3.0.CO;2-V](http://dx.doi.org/10.1002/(SICI)1098-2302(199604)29:3<205::AID-DEV2>3.0.CO;2-V).
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., Fukui, I., 1989. A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *J. Child Lang.* 16 (03), 477. <http://dx.doi.org/10.1017/S0305000900010679>.
- Flom, R., Bahrick, L.E., 2007. The development of infant discrimination of affect in multimodal and unimodal stimulation: the role of intersensory redundancy. *Dev. Psychol.* 43 (1), 238–252. <http://dx.doi.org/10.1037/0012-1649.43.1.238>.
- Geangu, E., Benga, O., Stahl, D., Striano, T., 2010. Contagious crying beyond the first days of life. *Infant Behav. Dev.* 33 (3), 279–288. <http://dx.doi.org/10.1016/j.infbeh.2010.03.004>.
- Graham, A.M., Fisher, J.H., Pfeifer, J.H., 2013. What sleeping babies hear: a functional MRI study of interparental conflict and infants' emotion processing. *Psychol. Sci.* 24 (5), 782–789. <http://dx.doi.org/10.1177/0956797612458803>.
- Grossmann, T., Friederici, A.D., 2012. When during development do our brains get tuned to the human voice? *Soc. Neurosci.* 7 (4), 369–372. <http://dx.doi.org/10.1080/17470919.2011.628758>.
- Grossmann, T., Striano, T., Friederici, A.D., 2005. Infants electric brain responses to emotional prosody. *Neuroreport* 16 (16), 1825–1828. <http://dx.doi.org/10.1097/01.wnr.0000185964.34336.b1>.
- Grossmann, T., Gliga, T., Johnson, M.H., Mareschal, D., 2009. The neural basis of perceptual category learning in human infants. *J. Cogn. Neurosci.* 21 (12), 2276–2286. <http://dx.doi.org/10.1162/jocn.2009.21188>.
- Grossmann, T., Oberecker, R., Koch, S.P., Friederici, A.D., 2010a. The developmental origins of voice processing in the human brain. *Neuron* 65 (6), 852–858. <http://dx.doi.org/10.1016/j.neuron.2010.03.001>.
- Grossmann, T., Parise, E., Friederici, A., 2010b. The detection of communicative signals directed at the self in infant prefrontal cortex. *Front. Hum. Neurosci.* 4, 201. <http://dx.doi.org/10.3389/fnhum.2010.00201>.
- Grossmann, T., Vaish, A., Franz, J., Schroeder, R., Stoneking, M., Friederici, A.D., 2013. Emotional voice processing: investigating the role of genetic variation in the serotonin transporter across development. *PLoS One* 8 (7), e68377. <http://dx.doi.org/10.1371/journal.pone.0068377>.
- Grossmann, T., 2010. The development of emotion perception in face and voice during infancy. *Restor. Neurol. Neurosci.* 28 (2), 219–236. <http://dx.doi.org/10.3233/RNN-2010-0499>.
- Grossmann, T., 2013. The early development of processing emotions in face and voice. In:

- Belin, P., Campanella, S., Ethofer, T. (Eds.), Integrating Face and Voice in Person Perception. Springer New York, New York, NY, pp. 95–116. http://dx.doi.org/10.1007/978-1-4614-3585-3_5. Retrieved from.
- Kotsoni, E., de Haan, M., Johnson, M.H., 2001. Categorical perception of facial expressions by 7-month-old infants. *Perception* 30 (9), 1115–1125. <http://dx.doi.org/10.1068/p3155>.
- Kushnerenko, E., Ceperoniene, R., Balan, P., Fellman, V., Huotilainen, M., Näätänen, R., 2002. Maturation of the auditory event-related potentials during the first year of life. *Neuroreport* 13 (1), 47–51.
- LaBarbera, J.D., Izard, C.E., Vietze, P., Parisi, S.A., 1976. Four- and six-month-old infants' visual responses to joy, anger, and neutral expressions. *Child Dev.* 47 (2), 535–538. <http://dx.doi.org/10.2307/1128816>.
- Leppänen, J.M., Nelson, C.A., 2009. Tuning the developing brain to social signals of emotions. *Nat. Rev. Neurosci.* 10 (1), 37–47. <http://dx.doi.org/10.1038/nrn2554>.
- Light, S.N., Coan, J.A., Zahn-Waxler, C., Frye, C., Goldsmith, H.H., Davidson, R.J., 2009. Empathy is associated with dynamic change in prefrontal brain electrical activity during positive emotion in children. *Child Dev.* 80 (4), 1210–1231. <http://dx.doi.org/10.1111/j.1467-8624.2009.01326.x>.
- Lloyd-Fox, S., Blasi, A., Elwell, C.E., 2010. Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy. *Neurosci. Biobehav. Rev.* 34 (3), 269–284. <http://dx.doi.org/10.1016/j.neubiorev.2009.07.008>.
- Lloyd-Fox, S., Blasi, A., Mercure, E., Elwell, C.E., Johnson, M.H., 2012. The emergence of cerebral specialization for the human voice over the first months of life. *Soc. Neurosci.* 7 (3), 317–330. <http://dx.doi.org/10.1080/17470919.2011.614696>.
- Ludemann, P.M., Nelson, C.A., 1988. Categorical representation of facial expressions by 7-month-old infants. *Dev. Psychol.* 24 (4), 492–501. <http://dx.doi.org/10.1037/0012-1649.24.4.492>.
- Merrill, J., Sammler, D., Bangert, M., Goldhahn, D., Lohmann, G., Turner, R., Friederici, A.D., 2012. Perception of words and pitch patterns in song and speech. *Front. Psychol.* 3 (76), 1–13. <http://dx.doi.org/10.3389/fpsyg.2012.00076>.
- Missana, M., Atkinson, A.P., Grossmann, T., 2015. Tuning the developing brain to emotional body expressions. *Dev. Sci.* 18 (2), 243–253. <http://dx.doi.org/10.1111/desc.12209>.
- Nelson, C.A., de Haan, M., 1996. Neural correlates of infants' visual responsiveness to facial expressions of emotion. *Dev. Psychobiol.* 29 (7), 577–595. [http://dx.doi.org/10.1002/\(SICI\)1098-2302\(199611\)29:7<577::AID-DEV3>3.0.CO;2-R](http://dx.doi.org/10.1002/(SICI)1098-2302(199611)29:7<577::AID-DEV3>3.0.CO;2-R).
- Nelson, C.A., Morse, P.A., Leavitt, L.A., 1979. Recognition of facial expressions by seven-month-old infants. *Child Dev.* 50 (4), 1239–1242. <http://dx.doi.org/10.2307/1129358>.
- Nelson, C.A., Thomas, K.M., de Haan, M., Wewerka, S.S., 1998. Delayed recognition memory in infants and adults as revealed by event-related potentials. *Int. J. Psychophysiol.* 29 (2), 145–165. [http://dx.doi.org/10.1016/S0167-8760\(98\)00014-2](http://dx.doi.org/10.1016/S0167-8760(98)00014-2).
- Parise, E., Csibra, G., 2013. Neural responses to multimodal ostensive signals in 5-month-old infants. *PLoS One* 8 (8), e72360. <http://dx.doi.org/10.1371/journal.pone.0072360>.
- Roth-Hanania, R., Davidov, M., Zahn-Waxler, C., 2011. Empathy development from 8 to 16 months: early signs of concern for others. *Infant Behav. Dev.* 34 (3), 447–458. <http://dx.doi.org/10.1016/j.infbeh.2011.04.007>.
- Sagi, A., Hoffman, M.L., 1976. Empathic distress in the newborn. *Dev. Psychol.* 12 (2), 175–176. <http://dx.doi.org/10.1037/0012-1649.12.2.175>.
- Sauter, D.A., Eimer, M., 2010. Rapid detection of emotion from human vocalizations. *J. Cogn. Neurosci.* 22 (3), 474–481. <http://dx.doi.org/10.1162/jocn.2009.21215>.
- Serrano, J.M., Iglesias, J., Loeches, A., 1992. Visual discrimination and recognition of facial expressions of anger, fear, and surprise in 4- to 6-month-old infants. *Dev. Psychobiol.* 25 (6), 411–425. <http://dx.doi.org/10.1002/dev.420250603>.
- Simner, M.L., 1971. Newborn's response to the cry of another infant. *Dev. Psychol.* 5 (1), 136–150. <http://dx.doi.org/10.1037/h0031066>.
- Striano, T., Brennan, P.A., Vanman, E.J., 2002. Maternal depressive symptoms and 6-month-old infants' sensitivity to facial expressions. *Infancy* 3 (1), 115–126.
- Vaish, A., Grossmann, T., Woodward, A., 2008. Not all emotions are created equal: the negativity bias in social-emotional development. *Psychol. Bull.* 134 (3), 383–403. <http://dx.doi.org/10.1037/0033-2909.134.3.383>.
- de Haan, M., 2007. Current and future directions in infant electrophysiology. In: de Haan, M. (Ed.), *Infant EEG and Event-Related Potentials*. Psychology Press, New York, N.Y., pp. 305–316.