

Infants' electric brain responses to emotional prosody

Tobias Grossmann^{a,b}, Tricia Striano^{a,b,c} and Angela D. Friederici^a

^aDepartment of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, ^bJunior Scientist Group on Cultural Ontogeny, Max Planck Institute for Evolutionary Anthropology and ^cNeurocognition and Development Group, Center for Advanced Studies, University of Leipzig, Germany.

Correspondence and requests for reprints to Tobias Grossmann, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, D-04103 Leipzig, Germany
E-mail: grossman@cbs.mpg.de

Received 23 August 2005; accepted 12 September 2005

In the current study, we examined 7-month-old infants' processing of emotional prosody using event-related brain potentials. Infants heard semantically neutral words that were spoken with either a happy, angry, or neutral voice. The event-related brain potential data revealed that angry prosody elicited a more negative response in infants' event-related potentials than did happy or neutral prosody, suggesting greater allocation of attention to angry prosody.

A positive slow wave was elicited by angry and happy prosody over temporal electrode sites. This indicates an enhanced sensory processing of the emotionally loaded stimuli (happy and angry). The current findings demonstrate that very early in development, the human brain detects emotionally loaded words and shows differential attentional responses depending on their emotional valence. *NeuroReport* 16:1825–1828 © 2005 Lippincott Williams & Wilkins.

Keywords: development, emotional prosody, emotional speech, event-related brain potentials, infants

Introduction

Humans communicate emotional information through gestures, facial expressions, and language. In spoken language, emotions can be expressed on two separate levels. One level refers to the semantic meaning of the words, such as 'happiness' or 'anger'. On a second level, speech melody, also referred to as emotional prosody, serves to communicate emotions. Emotion produces changes in respiration, phonation, and articulation, which in turn determine the acoustic signal [1]. Emotional prosody depends on different acoustic parameters such as time structure, loudness, roughness, and fundamental frequency. The emotion that is expressed by a speaker is characterized, across cultures, by specific properties of these parameters [2,3]. Adult listeners can reliably and rapidly recognize different emotions on the basis of these vocal cues [4]. Furthermore, evidence shows that emotional prosody is processed non-voluntarily [5], and that the specific acoustic patterns observed in humans in response to certain emotions are very similar to those observed in primates [6].

How adult listeners process emotional prosody has also been studied using event-related brain potentials (ERPs). These studies revealed that ERP components differed between neutral, happy, and angry prosody [5,7,8]. For example, the amplitude of the P200 was larger to happy prosody than to neutral and angry prosody, whereas the N400 differed in its amplitude between neutral and both emotional conditions [9]. These results demonstrate that emotional prosody has differential effects on the ERP at different stages of information processing in adults.

How the processing of emotional prosody develops over the course of ontogeny is only poorly understood. Unlike the extensive behavioral work on processing facial expressions in infancy, studies on infants' perception of vocal expressions of emotion are relatively scarce. Behavioral experiments have been conducted to examine young infants' discrimination of vocal expressions. In these studies, infants were habituated to a visual stimulus, accompanied by the presentation of a vocal expression. The vocal expression was changed after habituation, but the visual stimulus remained the same. An increase in looking time (i.e. dishabituation) to the habituated visual stimulus was interpreted as evidence that the infant had discriminated between the two vocal expressions.

Walker-Andrews and Grolnick [10] habituated 3 and 5-month-old infants to a woman's sad or happy voice, accompanied with a slide of her face expressing the same emotion. The 3-month-old infants showed increased looking time only when the sad voice was presented first, whereas the 5-month-old infants dishabituated for both orders. The authors concluded that infants as young as 5 months can discriminate between happy and sad vocal expressions, when a matching facial expression is presented.

Initially, it was thought that any visual stimulus could accompany the vocal expression, but a study by Walker-Andrews and Lennon [11] revealed that 5-month-old infants discriminated between happy, sad, and angry vocal expressions only when a slide of a face was presented. When the vocal expression was accompanied by a checkerboard pattern, infants' looking time did not reliably increase.

Walker-Andrews and Lennon suggested that the presence of a face might have helped infants to attend to the affective tone of voice.

A developmental sequence has been proposed in which infants learn to discriminate emotional expressions on the basis of multimodal, vocal, and finally, as visual acuity improves, facial cues [12]. This advantage for the auditory over the visual sensory modality might be due to the fact that in all mammals, the auditory system develops much earlier than the visual system [13]. Along these lines, Fernald [14] has shown that infants at 5 months of age respond differentially to positive and negative vocal expressions, whereas a selective responsiveness to facial expressions is not yet exhibited at this age. In a different study, 7-month-old infants dishabituated to a change from happy to angry, and from angry to happy emotions, only when both facial and vocal expressions of a videotaped woman changed [15], but not when the facial expression was presented alone. Together, these behavioral findings suggest that voice provides infants with important affective information that helps them to discriminate between emotions and to selectively respond to them.

The ERP measure has proven to be a valuable tool in assessing infants' processing of emotional information conveyed by the face [16,17]. Previous research [17] has shown that in 7-month-old infants' ERP, a negative component that is thought to reflect the allocation of attention was larger in amplitude when the infants watched a negative facial expression (fearful) than when they watched a positive facial expression (happy). This finding suggests that 7-month-old infants allocated more attentional resources to the negative facial expression. To date, however, there is no ERP study investigating the electrophysiological correlates of emotional prosody in infants.

In the current study, we examined 7-month-old infants' processing of emotional prosody using ERP measures. We had infants listen to words with neutral, happy, and angry prosody in order to investigate whether and how ERP correlates differ between (1) neutral and emotionally charged prosody (happy and angry), and (2) positive emotion (happy) and negative emotion (angry). We hypothesized, on the basis of prior work with facial expressions [17], that if the negative component is also sensitive to emotional information in the voice, it would vary as a function of emotional prosody. Specifically, words with negative (angry) emotional prosody would elicit a larger amplitude in the negative component than words with positive (happy) or neutral prosody.

Materials and methods

Participants

The final sample consisted of 16 7-month-old infants (eight female infants, $M=7$ months 6 days, $SD=6$ days; range=6 months 20 days to 7 months 10 days). An additional eight infants were tested but not included in the final sample because of fussiness ($n=1$) or excessive artifacts ($n=7$). All infants were born full-term (37–42 weeks gestation) and with normal birth weight (>2500 g).

Stimuli

The stimulus material consisted of 74 semantically neutral German verbs previously validated and used by Schirmer and Kotz [7]. A female speaker produced all words with

happy, angry, and neutral prosody. Words were taped with a digital audiotape recorder and digitized at a 16-bit/44.1 kHz sampling rate. Analysis of the speech stimuli was performed with the 'Praat Speech Processing Software' (Boersma & Weenik, Institute of Phonetics Sciences of the University of Amsterdam, Amsterdam, The Netherlands). The following acoustic parameters were evaluated: (1) mean duration in milliseconds (neutral=784.04, $SD=97.64$; happy=862.22, $SD=97.69$; angry=932.11, $SD=118.88$), (2) mean fundamental frequency in Hertz (neutral=250.22, $SD=21.25$; happy=340.96, $SD=52.57$; angry=247.43, $SD=18.89$), and (3) mean intensity in decibels (neutral=66.24, $SD=4.32$; happy=67.84, $SD=4.86$; angry=67.33, $SD=6.51$). These three parameters were then used to compare acoustic differences across the three emotions. The means were compared using *t*-tests, which revealed that angry stimuli were significantly longer in duration than happy stimuli ($t=5.51$, $P<0.001$), and happy stimuli were longer than neutral stimuli ($t=9.93$, $P<0.001$). Furthermore, mean fundamental frequency was significantly higher for happy stimuli than for angry ($t=17.7$, $P<0.001$) and neutral ($t=15.5$, $P<0.001$) stimuli, whereas neutral and angry stimuli did not differ in their fundamental frequency ($t=0.95$, $P>0.34$). The three stimuli did not differ with respect to their mean intensity.

Procedure

Infants were seated on their mother's lap in a dimly lit, sound-attenuated, and electrically shielded room. Mothers were listening to music via headphones during the experimental session so that they could not hear the acoustic stimuli presented to their infant. The session continued until the infant had attended to the maximum number of trials (222) or got tired of the experiment. The experimental session consisted of consecutive presentations of 74 words from each emotional prosody category (happy, angry, and neutral). Stimuli from the different emotional categories were randomly distributed over the session with no more than two stimuli of the same category occurring consecutively. The interstimulus interval varied randomly between 1500 and 2000 ms. During the presentation of the acoustic stimuli, an abstract screensaver without social stimuli was presented to the infants on a computer screen placed at a 60 cm distance in order to reduce eye movement artifacts.

Electroencephalogram measurement and data analysis

The electroencephalogram was recorded with Ag–AgCl electrodes from 19 scalp locations of the 10–20 system, referenced to Cz. Horizontal and vertical electrooculograms were recorded bipolarly. The sampling rate was 250 Hz. Electroencephalogram data were re-referenced to the algebraic mean of the left and the right mastoid electrodes, and band-pass filtered with 0.3–20 Hz (1501 points, 12 dB/octave slope). For eliminating artifacts caused by eye and body movements, electroencephalogram data were rejected off-line whenever the standard deviation within a 200-ms gliding window exceeded $80\mu\text{V}$ for the vertical or horizontal electrooculogram and $50\mu\text{V}$ at any electrode. The mean number of trials was 25.5 ($SD=7.4$) for happy, 24.2 ($SD=5.6$) for angry, and 29.1 ($SD=6.9$) for neutral prosody. Variances were analyzed by repeated measures ANOVAs. The analyzed factors were emotional prosody (happy \times angry \times neutral), scalp distribution [frontal (F3, F4) \times central (C3, C4) \times parietal (P3, P4) \times temporal (T3, T4)], and

lateralization (left × right). Pairwise *t*-tests were corrected for multiple comparisons using a modified Bonferroni procedure.

Results

Negative shift

Seven-month-old infants' ERPs to angry prosody were more negative in their amplitude than ERPs elicited by happy and neutral prosody as early as 300 ms after stimulus onset (see Fig. 1). This difference for emotional prosody reached its peak amplitude around 450 ms, and was maximal over frontal and central electrode sites. A main effect of emotional prosody [$F(1,15)=8.98, P<0.01$] was observed in the 300–600 ms latency interval. Subsequent *t*-tests for the same latency interval revealed that the mean amplitudes differed significantly only between happy and angry ($t=-2.82, P<0.01$) and between neutral and angry ($t=-1.85, P<0.05$) prosody, but not between happy and neutral prosody ($t=0.13, P>0.55$).

Positive slow wave

Seven-month-old infants' ERPs to angry and happy prosody were more positive in their amplitude than ERPs elicited by neutral prosody at temporal sites as early as 500 ms after stimulus onset (see Fig. 2). A significant interaction [$F(1,15)=5.43, P<0.05$] was found in the 500–1000 ms latency interval. Subsequently employed *t*-tests revealed that only at temporal sites the mean amplitudes differed significantly between neutral and angry ($t=-3.41, P<0.01$) and between neutral and happy ($t=-3.52, P<0.01$) prosody, but not between angry and happy ($t=0.54, P>0.7$).

Discussion

In the current ERP study, we examined the processing of words with neutral, happy, and angry prosody in 7-month-old infants. The data revealed that the amplitude of a negative shift and a positive slow wave in infants' ERPs varied as a function of emotional prosody.

In support of our hypothesis, we found that words with an angry prosody elicited a more negative response in infants' ERPs than did words with happy or neutral prosody. The negative shift observed in the current study resembles previous ERP work with 4-month-old infants, in which the mother's voice was compared with unfamiliar

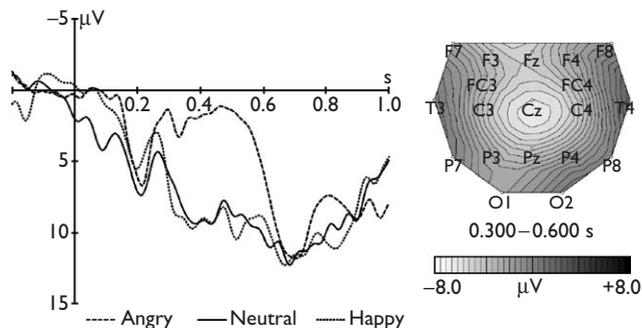


Fig. 1 Event-related brain potentials in response to neutral, happy, and angry emotional prosody in 7-month-old infants at the central midline electrode Cz (left). Potential map depicting the scalp distribution for the observed negative shift (right).

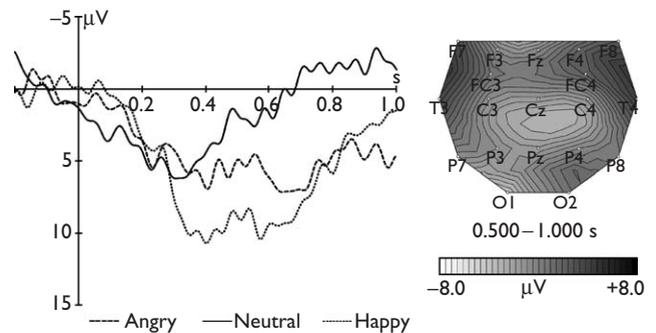


Fig. 2 Event-related brain potentials (ERPs) in response to neutral, happy, and angry emotional prosody in 7-month-old infants at the temporal electrodes (left). The ERPs plotted represent the average of temporal electrode sites (T3 and T4). Potential map depicting the scalp distribution for the observed positive slow wave (right).

voices [18]. In that study, 4-month-old infants' ERPs revealed a negative shift in response to the mother's voice, while in the current study, angry prosody elicited a negative shift in 7-month-old infants' ERPs. In several infant ERP studies on visual processing, it has been suggested that a larger amplitude of a negative component indicates increased allocation of attention [16,17]. On the basis of this view, Purhonen *et al.* [18] argued that the 4-month-old infants in their study allocated more attention to process their own mother's voice than to process unfamiliar voices. Here, we also suggest that the 7-month-old infants in our study allocated more attentional resources to the angry than to the happy or neutral voice.

This interpretation is also in accordance with behavioral findings suggesting an evolutionarily driven propensity to react more strongly to negative (e.g. angry faces or voices) than to positive or neutral stimuli [19]. For example, adults detect angry schematic faces more readily than happy schematic faces [20]. It has been argued that because it is difficult to reverse the consequences of an injurious or fatal assault, the process of natural selection may have resulted in this propensity to react more strongly to negative stimuli [19]. This heightened sensitivity to negative information, termed 'negativity bias', is a reliable psychological phenomenon in adults. Therefore, it is possible that the negative shift revealed in our results indicates that the infant brain allocates more attentional resources to angry voices, because they signify threat and potential negative consequences, which need to be attended to. From the current data, however, it cannot be concluded whether this is an innately programmed capacity or is learned from exposure to vocal emotional information.

Words with angry and happy prosody elicited a positive slow wave in infants' ERPs, whereas ERPs to words with neutral prosody returned to baseline. This effect was observed over temporal electrodes at a latency from 500 to 1000 ms. It has been argued that infants' slow waves reflect more diffuse activation of neural systems [21]. It is thus possible that the observed positive slow wave to happy and angry prosody indexes dispersed activation in auditory (temporal) brain structures to affectively loaded stimuli that are not evoked by neutral voices. This suggests enhanced sensory processing only for the affectively loaded auditory stimuli.

Concordant with this interpretation is evidence from functional magnetic resonance imaging work in adults showing that emotionally charged words undergo more extensive processing than words with neutral prosody [22]. For example, relative to neutral prosody, angry prosody evoked enhanced activity in adults' associative auditory cortex; namely, in the middle portion of the superior temporal sulcus [23]. Similarly, an enhancement in sensory processing driven by emotion was reported in the right mid-fusiform gyrus for fearful relative to neutral faces [24]. Therefore, it has been proposed that enhanced sensory responses to emotional facial and vocal stimuli might be a fundamental neural mechanism. It is possible that this mechanism might also account for the observed positive slow wave to happy and angry prosody over temporal sites in the 7-month-old infants, indicating an enhanced sensory processing of the emotional stimuli. This enhanced processing, which we have demonstrated at an electrophysiological level, could be linked to the behavioral finding that vocal affect facilitates infants' spoken word recognition [25], suggesting a method by which emotional information in the speech signal might help infants develop language comprehension capacities.

The three stimulus categories used in the current experiment significantly differ in duration and fundamental frequency. Effects due to these perceptual differences, however, are likely to be limited to early (exogenous) ERP components. Therefore, at least, they cannot account for the late positive slow wave effect. Clearly, further experiments are needed to tease apart the contribution of perceptual differences and emotional prosody *per se* to the effects observed in the current study.

The present results demonstrate the value of the ERP method to investigate brain processes underlying infants' processing of emotional speech. This might stimulate further ERP research with infants at other ages, and thus shed light on the developmental trajectory of how the human brain learns to process emotional prosody. In order to understand the extent to which the current effects are emotion-specific, this research should be extended to other emotions. Also, it will be important to use high-density ERPs to localize the neural sources and identify the neural circuits that generate the observed ERP effects under the dependency of emotional prosody in infants.

Acknowledgements

We would like to thank the Universitätsfrauenklinik and the Eitingon Krankenhaus for help with participant recruitment, and the parents and infants who participated in this study. We also thank Amrisha Vaish for comments on an earlier version of the manuscript. The research was supported in part by the Sofja Kovalevskaja Award, granted by the Humboldt Foundation, to T. Striano and the Max Planck Society.

References

- Scherer KR. Vocal correlates of emotion. In: Manstead A, Wagner H, editors. *Handbook of psychophysiology: emotion and social behavior*. London: Wiley; 1989. pp. 165–197.
- Scherer KR, Banse R, Wallbott HG. Emotion inferences from vocal expression correlate across languages and cultures. *J Cross Cult Psychol* 2001; **32**:76–92.
- Scherer KR. Vocal affect expression: a review and a model for future research. *Psychol Bull* 1986; **99**:143–165.
- Banse R, Scherer K. Acoustic profiles in vocal emotion expression. *J Pers Soc Psychol* 1996; **70**:614–636.
- Wambacq IJA, Shea-Miller KJ, Abubakr A. Non-voluntary and voluntary processing of emotional prosody: an event-related potentials study. *Neuroreport* 2004; **15**:555–559.
- Scherer KR, Kappas A. Primate vocal expression of affective state. In: Todt D, Goedeke P, Symmes D, editors. *Primate vocal communication*. Berlin: Springer; 1988. pp. 171–194.
- Schirmer A, Kotz S. ERP evidence for a sex-specific Stroop effect in emotional speech. *J Cogn Neurosci* 2003; **15**:1135–1148.
- Wambacq IJA, Jerger JF. Processing of affective prosody and lexical-semantics in spoken utterances as differentiated by event-related potentials. *Cogn Brain Res* 2004; **20**:427–437.
- Alter K, Rank E, Kotz S, Toepel U, Besson M, Schirmer A, et al. Affective encoding in the speech signal and in event-related brain potentials. *Speech Comm* 2003; **40**:61–70.
- Walker-Andrews AS, Grolnick W. Discrimination of vocal expression by young infants. *Infant Behav Dev* 1983; **6**:491–498.
- Walker-Andrews AS, Lennon EM. Infants' discrimination of vocal expressions: contributions of auditory and visual information. *Infant Behav Dev* 1991; **14**:131–142.
- Walker-Andrews AS. Infants' perception of expressive behaviors: differentiation of multimodal information. *Psychol Bull* 1997; **121**:1–20.
- Gottlieb G. Ontogenesis of sensory function in birds and mammals. In: Tobach E, Aronson LR, Shaw E, editors. *The biopsychology of development*. New York: Academic Press; 1971. pp. 67–128.
- Fernald A. Human maternal vocalizations to infants as biologically relevant signals: an evolutionary perspective. In: Barkow JH, Cosmides L, Tooby J, editors. *The adapted mind: evolutionary psychology and the generation of culture*. Oxford: Oxford University Press; 1988. pp. 391–428.
- Caron AJ, Caron RF, MacLean DJ. Infant discrimination of naturalistic emotional expressions: the role of face and voice. *Child Dev* 1988; **59**:604–616.
- de Haan M, Johnson MH, Halit H. Development of face-sensitive event-related potentials during infancy: a review. *Int J Psychophysiol* 2003; **51**:45–58.
- Nelson CA, de Haan M. Neural correlates of infants' visual responsiveness to facial expression of emotion. *Dev Psychobiol* 1996; **29**:577–595.
- Purhonen M, Kilpeläinen-Lees R, Valkonen-Korhonen M, Karhu J, Lehtonen J. Cerebral processing of mother's voice compared to unfamiliar voice in 4-month-old infants. *Int J Psychophysiol* 2004; **52**:257–266.
- Cacioppo JT, Berntson GG. The affect system: architecture and operating characteristics. *Curr Dir Psychol Sci* 1999; **8**:133–137.
- Oehmann A, Lundqvist D, Esteves F. The face in the crowd revisited: a threat advantage with schematic stimuli. *J Pers Soc Psychol* 2001; **80**:381–391.
- de Haan M, Nelson CA. Recognition of the mother's face by six-month-old infants: a neurobehavioral study. *Child Dev* 1997; **68**:187–210.
- Mitchell RLC, Elliott R, Barry M, Cruttenden A, Woodruff PWR. The neural response to emotional prosody as revealed by functional magnetic resonance imaging. *Neuropsychologia* 2003; **41**:1410–1421.
- Grandjean D, Sander D, Pourtois G, Schwartz S, Seghier ML, Scherer KR, et al. The voices of wrath: brain responses to angry prosody in meaningless speech. *Nat Neurosci* 2005; **8**:145–146.
- Vuilleumier P, Armony JL, Driver J, Dolan RJ. Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron* 2001; **30**:829–841.
- Singh L, Morgan JL, White KS. Preference and processing: the role of speech affect in early spoken word recognition. *J Mem Lang* 2004; **51**:173–189.