Developmental Science 9:3 (2006), pp 309–315

REPORT

Crossmodal integration of emotional information from face and voice in the infant brain

Tobias Grossmann,[1,2] Tricia Striano[1,2,3] and Angela D. Friederici[1]

1. *Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany*
2. *Junior Scientist Group on Cultural Ontogeny, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany*
3. *Neurocognition and Development Group, Center for Advanced Studies, University of Leipzig, Germany*

Abstract

We examined 7-month-old infants' processing of emotionally congruent and incongruent face–voice pairs using ERP measures. Infants watched facial expressions (happy or angry) and, after a delay of 400 ms, heard a word spoken with a prosody that was either emotionally congruent or incongruent with the face being presented. The ERP data revealed that the amplitude of a negative component and a subsequent positive component in infants' ERPs varied as a function of crossmodal emotional congruity. An emotionally incongruent prosody elicited a larger negative component in infants' ERPs than did an emotionally congruent prosody. Conversely, the amplitude of infants' positive component was larger to emotionally congruent than to incongruent prosody. Previous work has shown that an attenuation of the negative component and an enhancement of the later positive component in infants' ERPs reflect the recognition of an item. Thus, the current findings suggest that 7-month-olds integrate emotional information across modalities and recognize common affect in the face and voice.

Introduction

The way that emotions are perceived when communicated either by the face or the voice has been extensively studied in each modality separately (for a review, see Murphy, Nimmo-Smith & Lawrence, 2003). However, in most social interactions, emotional information is communicated simultaneously by different modalities such as the face and voice. Thus the question arises how emotional information from the face and voice is integrated.

Recent behavioral findings indicate that adults integrate facial and vocal emotional information. For instance, de Gelder and Vroomen (2000) showed adults face–voice pairs that varied in their levels of emotional congruity, and asked the adults to judge either the facial expression or the emotional tone of voice. Their results indicated that when asked to identify the facial expression, adults' judgment was biased in the direction of the simultaneously presented tone of voice, and, conversely, when asked to judge tone of voice, adults were biased by the accompanying facial expression. This indicates that even when asked to make unimodal judgments, adults nevertheless integrated the available bimodal information.

Electrophysiological experiments by Pourtois and colleagues have also shown the integration of audio-visual emotional information. In a first ERP study using a visual oddball paradigm, participants watched facial expressions (angry or sad) and after a variable delay, heard either an emotionally congruent or incongruent voice while the facial expression was still being presented (Pourtois, de Gelder, Vroomen, Rossion & Crommelick, 2000). When the ERPs were analyzed time-locked to the voice onset, face–voice pairs conveying congruent emotional information elicited a larger N100 than incongruent face–voice pairs. The auditory N100 is an exogenous component (Näätänen, 1992) known to be generated in the auditory cortex. Thus, the finding of an enhanced N100 suggests that facial information that is congruent with auditory information leads to an amplification of auditory processing. This interpretation is in accordance with fMRI data showing a magnetic response amplification in the auditory cortex to audio-visual speech (Calvert, Hansen, Iversen & Brammer, 2001).

In a second ERP study, Pourtois, Debatisse, Despland and de Gelder (2002) used happy and fearful stimuli, and found that emotionally congruent and incongruent

Address for correspondence: Tobias Grossmann, Centre for Brain and Cognitive Development, School of Psychology, Birkbeck College, 32 Torrington Square, London WC1E 7JL, UK; e-mail: t.grossmann@bbk.ac.uk

face–voice pairs elicited a positive ERP component (named the 'P2b component' by Pourtois *et al.*, 2002) which peaked earlier in congruent than in incongruent pairs. This indicates that the processing of emotional speech is delayed in an incongruent facial context. The source generating this effect was localized in the anterior cingulate cortex, which has previously been implicated in error monitoring (Cabeza & Nyberg, 2000; MacLeod & MacDonald, 2000). Together, studies by Pourtois and colleagues suggest that congruent audio-visual emotional information enhances sensory-specific processing, and that incongruent emotional information delays the timing of the ongoing processes.

Unlike the behavioral and neuroscientific work on adults, the mechanisms of audio-visual integration of emotion from the face and voice in infancy are poorly understood. A series of behavioral experiments was conducted to examine infants' recognition of emotional expressions (Soken & Pick, 1992; Walker, 1982; Walker-Andrews, 1986), in which infants had to detect the correspondence between emotional information provided by the face and voice. In these experiments using the intermodal preference technique (Spelke, 1976), 5- and 7-month-old infants were presented simultaneously with two different dynamic facial expressions accompanied by a single vocal expression that affectively matched one of the facial displays.

Using this technique, Walker (1982) showed that infants looked longer at the facial display that affectively matched the voice. In another experiment (Walker-Andrews, 1986), the mouth was occluded so that synchrony between lip movements and vocal expressions could not account for infants' differential looking behavior. The results revealed that 7-month-olds, but not 5-month-olds, looked longer to the facial display that was congruent to the vocal affect. These behavioral findings suggest that 7-month-old infants can detect common affect across audio-visual (bimodal) expressions of emotion, and can do so even in the absence of temporal synchrony between face and voice.

However, to date, the ontogenetic development of the underlying electrophysiological processes for crossmodal integration of emotion have not been examined. The ERP measure has been found to be sensitive to infants' crossmodal (haptic to visual) recognition of objects (Nelson, Henschel & Collins, 1993), and has proven to be a valuable tool in assessing the underlying mechanisms of infants' processing of emotional information conveyed by face (de Haan, Johnson & Halit, 2003; Nelson & de Haan, 1996) and by the voice (Grossmann, Striano, & Friederici, 2005). Furthermore, previous research (Nelson & de Haan, 1996) has shown that two well-documented components in 7-month-old infants' ERPs, i.e. a negative component (Nc), followed by a late positive component (Pc), vary in their amplitudes as a function of emotion. Thus, in the current study, we assessed 7-month-old infants' processing of emotionally congruent and incongruent face–voice pairs using ERPs.

As infants in the present study watched a static facial expression (happy or angry), they heard a word spoken in a tone of voice that was either emotionally congruent or incongruent with the facial expression. Under the assumption that the Nc and Pc are also sensitive to processes related to matching emotional information between the face and the voice, we hypothesized that crossmodal integration of emotion would be reflected in their specific modulations. Specifically, enhanced amplitude of infants' Nc is thought to reflect an attentional orienting response to unexpected/unfamiliar stimuli (Courchesne, Ganz & Norcia 1981; Vaughan & Kurtzberg, 1992). We therefore predicted that incongruent compared to congruent face–voice pairs would elicit an Nc with larger amplitude due to the unexpected/unfamiliar mismatch between facial and vocal emotional information. Previous work on recognition memory indicated that infants' Pc that follows the Nc showed exactly the opposite pattern, namely that it was larger in its amplitude to items recognized as familiar when compared to unfamiliar items (Nelson, Thomas, de Haan & Wewerka, 1998). Thus, we hypothesized that infants in the current study would recognize congruent face–voice pairs as familiar and therefore show a larger Pc to them than to the unfamiliar, incongruent face–voice pairs. Together, we expected that congruent face–voice pairs would elicit more positive-going ERPs than would incongruent face–voice pairs, and that this effect would consist of an attenuation of infants' Nc and an enhancement of infants' Pc to emotionally congruent events.

Method

Participants

The final sample consisted of 18 7-month-old infants (nine females, $M = 6; 26$, SD = 6; Range = 6; 19 to 7; 09). An additional 12 infants were tested but not included in the final sample due to technical problems ($n = 2$) or excessive artifacts ($n = 10$). All infants were born full-term (37–42 weeks gestation) and with normal birth weight (> 2500 g).

Stimuli

The visual stimuli were two color portrait photographs of the same woman posing either a happy or an angry facial expression. The photographs were chosen on the basis of a survey, in which 20 adults (10 females) were asked to rate a set of pictures of seven women posing various facial expressions (happy, angry, sad, surprised, fearful and disgusted). Raters had to identify both the specific emotion displayed in the pictures as well as how arousing each emotion was. The pictures that were finally used as stimuli were correctly identified by 95% of the raters as either happy or angry. Both pictures were rated as equally arousing (happy: $M = 1.6$, $SD = 0.6$; angry: $M = 1.5$, $SD = 0.6$) on a scale from 1 (lightly arousing) to 4 (very strongly arousing).

The auditory stimulus material consisted of 74 neutral German verbs previously validated and used by Schirmer and Kotz (2003). In order to ensure that the words were semantically neutral, semantic word valence was obtained in an earlier study, where 40 participants rated all words on a 5-point scale that ranged from −2 to +2 for their emotional valence. The verbs had a mean valence of 0.05 ($SD = 0.46$), which indicates their neutral semantic meaning. A female speaker produced all words with happy and angry prosody. Words were taped with a DAT recorder and digitized at a 16-bit/44.1 kHz sampling rate.

Procedure

Infants were seated on their mother's lap in a dimly lit, sound-attenuated and electrically shielded room. A trial started with the presentation of a facial expression projected in the center of the screen on a black background, using a 70-Hz, 17-inch computer screen at a distance of 60 cm from the eyes. The image sizes were 27 × 22 cm and the vertical and horizontal visual angles were 12.12° and 10.07°, respectively. Following the onset of the face, there was a 400-ms delay. Then a word was presented via loundspeakers (SPL = 70 dB). The face remained on the screen until the end of the presentation of the word.

Congruent and incongruent face–voice stimulus pairs were presented with equal probability. In a congruent trial, a happy face was followed by a happy voice, or an angry face was followed by an angry voice. In an incongruent trial, a happy face was followed by an angry voice, or an angry face was followed by a happy voice. There were 74 trials per condition. Mothers were instructed to look down at the infant rather than at the computer screen, and they listened to music via headphones during the experimental session so that they could not hear the acoustic stimuli presented to their infant. A camera mounted above the screen recorded a close-up view of the infant's face in order to monitor infants' visual attention to the stimuli. The session continued until the infant had attended the maximum number of trials (296) or became fussy. The inter-trial interval varied randomly between 800 and 1200 ms.

EEG measurement and data analysis

The EEG was recorded with Ag-AgCl electrodes from 19 scalp locations (FZ, F3, F4, F7, F8, FC3, FC4, CZ, C3, C4, T3, T4, T5, T6, PZ, P3, P4, O1, O2) of the 10–20 system, referenced to CZ. Horizontal and Vertical EOGs were recorded bipolarly. Sampling rate was 250 Hz. EEG data was re-referenced to the algebraic mean of the left and the right mastoid electrodes, and band-pass filtered with 0.3 to 20 Hz (1501 points). For elimination of artifacts caused by eye and body movements, EEG data were rejected off-line whenever the standard deviation within a 200-ms gliding window exceeded 80 μV for the vertical or horizontal electro-oculogram and 50 μV at any electrode. ERPs were time-locked to the onset of the word. The mean number of trials was 52.4 for the congruent condition, and 53.6 for the incongruent condition. Amplitude peaks of effects were determined in the difference signals. ERPs were evaluated by computing the following regions of interest (ROIs): left hemisphere (F3, C3, T3, P3), right hemisphere (F4, C4, T4, P4). Variances were analyzed by a repeated measures ANOVA. Analyzed factors in the ANOVA were (1) congruence (congruent × incongruent), (2) lateralization (left × right).

Results

Negative component (Nc)

Seven-month-olds' ERPs to incongruent voices ($M = 4.83$) were more negative in their mean amplitude than ERPs elicited by congruent voices ($M = 12.12$) as early as 350 ms after stimulus onset (see Figure 1). This difference was maximal at frontal and central electrode sites, reaching its peak amplitude around 500 ms. It was statistically significant in the 400–600-ms latency interval, $F(1, 17) = 9.59$, $p = .0065$. There was no significant lateralization of this effect. In addition to the main effect of congruence, subsequent t-tests for the same latency interval (400–600 ms) revealed that the mean amplitudes differed significantly between the two incongruent conditions ($t = -2.84$, $p = .0113$), whereas mean amplitudes did not differ between the two congruent conditions (Figure 2). Mean amplitude to an angry voice presented with a happy face ($M = 2.42$) was more negative than to a happy voice presented with an angry face ($M = 7.87$). Analyses did not reveal any differences in the peak latency for the Nc.

Positive component (Pc)

Seven-month-olds' ERPs to congruent voices ($M = 13.56$) were more positive in their mean amplitude than

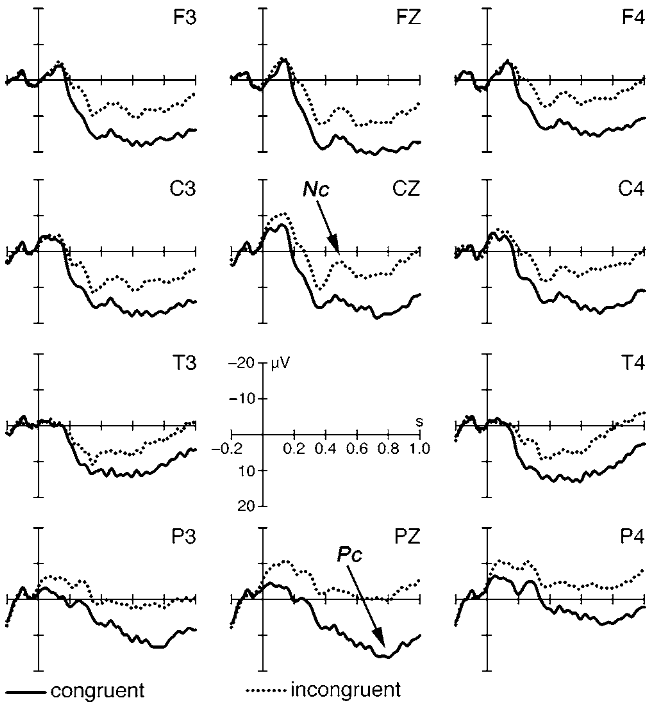

Figure 1 *ERPs in response to congruent (solid) and incongruent (dotted) face–voice pairs in 7-month-old infants.*

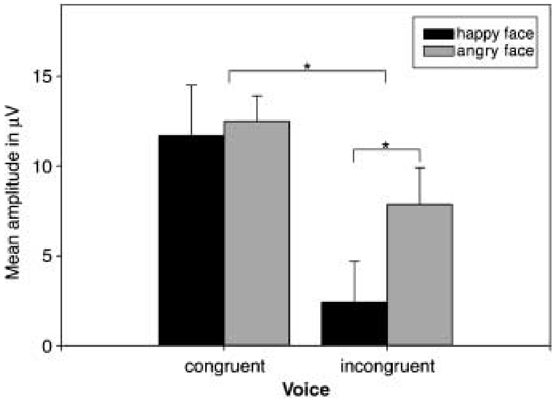

Figure 2 *Mean amplitude scores (SE) for the negative component.*

ERPs elicited by incongruent voices ($M = 3.9$) as early as 600 ms after stimulus onset (see Figure 1). This difference was statistically significant in the 600–1000-ms latency interval, $F(1, 17) = 14.63$, $p = .0014$, and had its maximum at central and parietal electrode sites. In contrast to the Nc, subsequently employed *t*-tests revealed that the mean amplitudes of the Pc did not differ between the two incongruent conditions ($t = -0.52$, $p = .61$) (Figure 3). Thus, mean amplitude to an angry voice presented with a happy face ($M = 3.27$) did not differ from a happy voice presented with an angry face ($M = 4.83$). Similar to the Nc, analyses did not reveal any differences in the peak latency for the Pc.

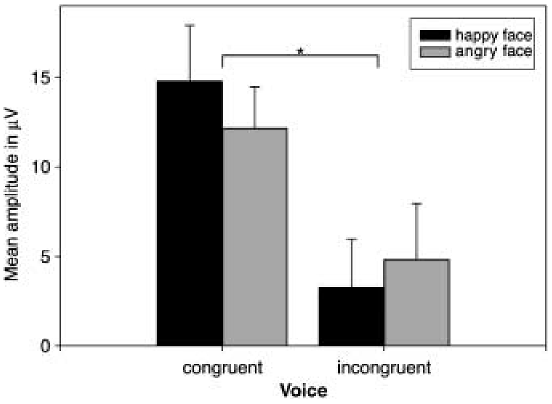

Figure 3 *Mean amplitude scores (SE) for the late positive component.*

Discussion

In the current study, we used ERP measures to examine 7-month-old infants' processing of emotionally congruent and incongruent face–voice pairs. The data revealed that the amplitude of a negative component and a subsequently elicited positive component in infants' ERPs varied as a function of crossmodal emotional congruity. In support of our hypothesis, we found that words spoken with a tone of voice that was emotionally incongruent to the facial expression elicited a larger negative component in infants' ERPs than did emotionally congruent words. Conversely, the amplitude of the positive component was larger to emotionally congruent words than to incongruent words. These findings provide electrophysiological evidence that 7-month-olds recognize common affect across modalities, which is in line with previous behavioral work (Soken & Pick, 1992; Walker, 1982; Walker-Andrews, 1986). However, Kahana-Kalman and Walker-Andrews (2001) found that infants tested with maternal stimuli were able to recognize common affect across modalities already at 3.5 months of age, whereas infants tested with unfamiliar faces and voices did not recognize common affect across modalities until 7 months of age (Walker-Andrews, 1986). Based on these findings effects observed here might well be observable at an earlier age when investigated with maternal expressions.

Extending behavioral findings, the current ERP data reveal insights into the time course and characteristics of the processes underlying the integration of emotional information from face and voice in the infant brain.

Numerous ERP studies in adults have investigated old/new effects in recognition memory tests with different stimuli (for a review, see Rugg & Coles, 1995). The uniform finding across studies is that old (familiar) items evoke more positive-going ERPs than do new (unfamiliar) items. This general old/new effect comprises the modulation of two ERP components: an early negativity (early N400), which consistently shows an attenuated amplitude to old items, and a late positive component or complex (LPC), which shows an enhanced amplitude to old items.

Old (familiar) items have also been found to elicit an attenuated N400 and an enhanced LPC in children's ERPs when compared to new (unfamiliar) items (Friedman, 1991; Friedman, Putnam & Sutton, 1989; Friedman, Putnam, Ritter, Hamberger & Berman, 1992; Coch, Maron, Wolf & Holcomb, 2002). Furthermore, similar effects have been observed in infants' ERPs (Nelson *et al.*, 1998), where old (familiar) items elicited a more positive-going brain response with an attenuated early negative component (Nc) and an enhanced late positive component (Pc). If we assume that the adult and child N400 corresponds with the infant Nc and that the adult and child LPC corresponds with the infant Pc, a coherent picture begins to emerge of the developmental trajectory of recognition memory effects in the ERP.

In the current study, emotionally congruent face–voice pairs elicited similar ERP effects as recognized items in previous memory studies with infants, children and adults. This suggests that 7-month-old infants recognize common affect in face and voice. Since the face–voice pairs presented to the infants were novel to them, the current data not only indicate that these infants recognized common affect, but, moreover, that they applied their knowledge about emotions in face and voice to draw inferences about what might be appropriate emotional face–voice associations when encountering novel bimodal events. Multimodal audio-visual events usually make two kinds of information available: amodal and modality specific information. An example of amodal information is that the movements of the lips and the timing of speech share temporal synchrony, rhythm and tempo, and have common intensity shifts. Since we used static facial expressions, there was no such amodal information available to the infants. Thus, infants could not simply determine that a face and voice belonged together by detecting amodal audio-visual relations; instead, they had to draw inferences based on their prior knowledge.

Another finding was that the amplitude of infants' Nc not only differed between congruent and incongruent face–voice pairs but also between two incongruent conditions. Namely, when a happy face was presented with an angry voice, the Nc was more negative in its amplitude than when an angry face was presented with a happy voice. From behavioral work, we know that prior to the onset of crawling (around 10 months), infants have only little exposure to others' expression of anger, whereas happy emotional expressions are ubiquitous in infants' everyday social interactions (Campos, Anderson, Barbu-Roth, Hubbard, Hertenstein & Witherington, 2000). Based on this observation, it can be assumed that a happy face is more familiar than an angry face (see also Striano, Brennan & Vanman, 2002). It is thus possible that the presentation of the more familiar happy face triggered a stronger expectation about the appropriate emotional prosody, causing an especially strong expectancy violation and a larger Nc when the angry voice was presented. This suggests a sensitivity of infants' Nc to familiarity-based processes, confirming previous research on infants' Nc (Nelson, 1994).

Pourtois *et al.* (2000) found that face–voice pairs conveying congruent emotional information elicited a larger N100 in adults. However, there was no such effect in the current infant ERP data. This might be because Pourtois *et al.* used an oddball paradigm where congruent and incongruent pairs were presented in different blocks, as both standard (85%) and deviant (15%). This interpretation is further supported by the fact that in a different study, when Pourtois *et al.* (2002) presented congruent and incongruent face–voice pairs with equal probability to adults, they did not find a modulation of the N100. Pourtois *et al.* (2002) did, however, find a delay of what they called a P2b to incongruent face–voice pairs in adults, whereas we did not find a delay for any of the infant ERP components to incongruent compared to congruent face–voice pairs despite using a stimulus paradigm very similar to theirs.

This discrepancy between the current findings and Pourtois *et al.*'s (2002) findings could be explained by differences between the specific experimental procedures of the two studies. Pourtois and colleagues instructed their adult participants to pay attention to the voices, which was naturally not the case for the infants in the current experiment. There is a body of evidence suggesting strong attention–emotion interaction effects on the anterior cingulate cortex (for a review, see Ochsner & Gross, 2005). This neural structure was identified as the generator of the delayed P2b in Pourtois *et al.*'s study. It is thus possible that the instruction to pay attention to the voices caused the delayed P2b in adults' ERPs, which was therefore not observed in infants' ERPs.

Future research using dipole modeling with high-density ERPs (Richards, 2005) could be employed to identify the cortical sources generating infants' Nc and Pc in order to shed light on the differentiability and function of the cognitive processes reflected in these ERP components.

Source analysis could not be easily performed with so few electrodes as used in the current study. Furthermore, comparing the cortical sources generating these effects in infants and adults would allow us to make inferences about the development of the neural mechanisms underlying the integration of emotional information from face and voice.

In sum, the current ERP data provide new insights into the cognitive and brain mechanisms underlying infants' integration of emotional information across face and voice. It is our hope that this might stimulate further research on this topic and help us deepen our understanding of the development of crossmodal integration of emotion.

Acknowledgements

We would like to thank the parents and infants who participated in this study. The research was supported in part by the Sofja Kovalevskaja Award to T. Striano provided by the Alexander von Humboldt Foundation, donated by the Federal Ministry of Education and Research. Thanks to Amrisha Vaish for comments on an earlier version of the manuscript.

References

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, **20**, 1–47.

Calvert, G.A., Hansen, P.C., Iversen, S.D., & Brammer, M.J. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *NeuroImage*, **14**, 427–438.

Campos, J.J., Anderson, D.I., Barbu-Roth, M.A., Hubbard, E.M., Hertenstein, M.J., & Witherington, D. (2000). Travel broadens the mind. *Infancy*, **1**, 149–219.

Coch, D., Maron, L., Wolf, M., & Holcomb, P.J. (2002). Word and picture processing in children: an event-related potential study. *Developmental Neuropsychology*, **22**, 373–406.

Courchesne, E., Ganz, L., & Norcia, A.M. (1981). Event-related brain potentials to human faces in infants. *Child Development*, **52**, 804–811.

de Gelder, B., & Vroomen, J. (2000). Perceiving emotion by ear and by eye. *Cognition and Emotion*, **14**, 289–311.

de Haan, M., Johnson, M.H., & Halit, H. (2003). Development of face-sensitive event-related potentials during infancy: a review. *International Journal of Psychophysiology*, **51**, 45–58.

Friedman, D. (1991). The endogenous scalp-recorded brain potentials and their relation to cognitive development. In J.R. Jennings & M.G.H. Coles (Eds.), *Handbook of cognitive psychophysiology: Central and autonomic nervous system approaches* (pp. 621–656). New York: John Wiley.

Friedman, D., Putnam, L., Ritter, W., Hamberger, M., & Berman, S. (1992). A developmental study of picture matching in children, adolescents, and young adults – a replication and extension. *Psychophysiology*, **29**, 593–610.

Friedman, D., Putnam, L., & Sutton, S. (1989). Cognitive brain potential in children, young adults and senior citizens: homologous components and changes in scalp distribution. *Developmental Neuropsychology*, **5**, 33–60.

Grossmann, T., Striano, T., & Friederici, A.D. (2005). Infants' electric brain responses to emotional prosody. *NeuroReport*, **16**, 1825–1828.

Kahana-Kalman, R., & Walker-Andrews, A.S. (2001). The role of person familiarity in young infants' perception of emotional expression. *Child Development*, **72**, 352–369.

MacLeod, C.M., & MacDonald, P.A. (2000). Interdimensional interference in the Stroop effect: uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Science*, **4**, 383–391.

Murphy, F.C., Nimmo-Smith, I., & Lawrence, A.D. (2003). Functional neuroanatomy of emotions: a meta-analysis. *Cognitive, Affective, and Behavioral Neuroscience*, **3**, 207–233.

Näätänen, R. (1992). *Attention and brain function.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Nelson, C.A. (1994). Neurocorrelates of recognition memory in the first postnatal year of life. In G. Dawson & K. Fischer (Eds.), *Human behavior and the developing brain* (pp. 269–313). New York: Guilford Press.

Nelson, C.A., & de Haan, M. (1996). Neural correlates of infants' visual responsiveness to facial expression of emotion. *Developmental Psychobiology*, **29**, 577–595.

Nelson, C.A., Henschel, M., & Collins, P.F. (1993). Neural correlates of cross-modal recognition memory by 8-month-old human infants. *Developmental Psychology*, **29**, 411–420.

Nelson, C.A., Thomas, K.M., de Haan, M., & Wewerka, S. (1998). Delayed recognition memory in infants and adults as revealed by event-related potentials. *International Journal of Psychophysiology*, **29**, 145–165.

Ochsner, K., & Gross, J.J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, **9**, 242–249.

Pourtois, G., Debatisse, D., Despland, P., & de Gelder, B. (2002). Facial expressions modulate the time course of long latency auditory brain potentials. *Cognitive Brain Research*, **14**, 99–105.

Pourtois, G., de Gelder, B., Vroomen, P., Rossion, B., & Crommelick, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *NeuroReport*, **11**, 1329–1333.

Richards, J.E. (2005). Localizing cortical sources of event-related potentials in infants' covert orienting. *Developmental Science*, **8**, 255–278.

Rugg, M.D., & Coles, M.G.H. (1995). *Electrophysiology of mind: Event-related brain potentials and cognition.* Oxford: Oxford University Press.

Schirmer, A., & Kotz, S. (2003). ERP evidence for a sex-specific Stroop effect in emotional speech. *Journal of Cognitive Neuroscience*, **15**, 1135–1148.

Soken, N.H., & Pick, A.D. (1992). Intermodal perception of happy and angry expressive behaviors by seven-month-old infants. *Child Development*, **63**, 787–795.

Spelke, E.S. (1976). Infants' intermodal perception of events. *Cognitive Psychology*, **8**, 553–560.

Striano, T., Brennan, P., & Vanman, E. (2002). Maternal depressive symptoms and 6-month-old infants' sensitivity to facial expressions. *Infancy*, **3**, 115–126.

Vaughan, H.G., & Kurtzberg, D. (1992). Electrophysiologic indices of human brain maturation and cognitive development. In M. Gunnar & C.A. Nelson (Eds.), *Developmental behavioral neuroscience* (pp. 1–36). Hillsdale, NJ: Lawrence Erlbaum Associates.

Walker, A.S. (1982). Intermodal perception of expressive behaviors by human infants. *Journal of Experimental Child Psychology*, **33**, 514–535.

Walker-Andrews, A.S. (1986). Intermodal perception of expressive behaviors: relation of eye and voice? *Developmental Psychology*, **22**, 373–377.

Received: 10 September 2005
Accepted: 7 October 2005