Ψ **Psychology Press**
Taylor & Francis Group

Auditory Recognition Memory in 2-Month-Old Infants as Assessed by Event-Related Potentials

Xiaoqin Mai

Center for Human Growth and Development, University of Michigan, Ann Arbor, Michigan

Lin Xu, Mingyan Li, Jie Shao, and Zhengyan Zhao

Children's Hospital, Zhejiang University School of Medicine, Hangzhou, China

Raye-Ann deRegnier

*Department of Pediatrics, Feinberg School of Medicine, Northwestern University,
Chicago, Illinois*

Charles A. Nelson

Children's Hospital Boston, Harvard Medical School, Boston, Massachusetts

Betsy Lozoff

*Center for Human Growth and Development and Department of Pediatrics and Communicable
Diseases, University of Michigan, Ann Arbor, Michigan*

Previous studies of auditory recognition memory in sleeping newborns reported 2 event-related potential (ERP) components, P2 and negative slow wave (NSW), reflecting voice discrimination and detection of novelty, respectively. In the present study, using high-density recording arrays, ERPs were acquired from 26 2-month-old awake infants as they were presented with a familiar and unfamiliar voice (i.e., mother and stranger) with equal probability. In addition to P2 and NSW, we observed a positive slow wave (PSW) over the right temporo-parietal scalp, indicating memory updating. Our study suggests that infants appear to have the capacity to encode novel stimuli as early as 2 months of age.

Recognition memory, which is a fundamental form of explicit memory, is defined as the ability to remember events that have been encountered before. The medial temporal lobe plays an important role in recognition memory (Clark, Broadbent, Zola, & Squire, 2002; Nelson, 1995; Reed &

Squire, 1997; Richmond & Nelson, 2007; Squire, Schmolck, & Stark, 2001). Most research on infant recognition memory has used visual stimuli and found that older infants recognize stimuli more quickly and prefer more complex stimuli than younger infants do, suggesting that recognition memory improves over the first year of life (cf., Brennan, Ames, & Moore, 1966; Fagan, 1974; Karmel, 1969; Rose, 1983; Rose, Melloycarminar, Gottfried, & Bridger, 1982). More recent studies have used event-related brain potentials (ERPs) to examine the neural correlates of recognition memory in infants. However, it is difficult to examine visual recognition memory with ERPs in newborns and very young infants due to their short attention span and limited visual abilities. Since attention is not necessary when listening to auditory stimuli, examining recognition memory with ERPs in the auditory modality provides the opportunity to understand the neural bases of memory at very young ages.

In the present study, we used the human voice as the auditory stimulus. Voice-specific processing has been found in the human brain. In ERP studies of adult responses to voice and other sound categories (such as bird songs and sounds produced by musical instruments), Charest et al. (2009) reported a fronto-temporal positivity to voices (FTPV) peaking around 200 msec post stimulus, and Levy and colleagues (2001, 2003) observed a voice-specific response (VSR) peaking around 320 msec after stimulus onset. The FTPV has also been observed in 4- to 5-year- old children (Rogier, Roux, Belin, Bonnet-Brilhault, & Bruneau, 2010). Moreover, Belin, Zatorre, Lafaille, Ahad, and Pike (2000), using functional magnetic resonance imaging (fMRI), found that in human adults, voices are processed in specialized brain regions located in the superior temporal cortex. Very recently, however, Grossmann, Oberecker, Koch, and Friederici (2010) using near-infrared spectroscopy found that 7-month-old infants showed increased hemodynamic responses in the superior temporal cortex to the human voice, but this effect was not found in 4-month-old infants, suggesting that voice-sensitive brain systems emerge between 4 and 7 months of age. Therefore, those voice specific ERP components (such as FTPV and VSR) found in adults and children could not be observed in infants younger than 4 months of age. However, studies also showed that infants can discriminate voices and recognize their mothers' voices from the neonatal period (Decasper & Fifer, 1980; Ockleford, Vince, Layton, & Reader, 1988). Thus, the development of voice recognition memory in infant might be different from that of voice-specific processing.

Several studies have examined the neural correlates of voice recognition memory in adults. Von Kriegstein, Kleinschmidt, Sterzer, and Giraud (2005) using fMRI found that recognition of familiar voices compared with unfamiliar involves a distributed brain system, including temporo-occipito-parietal, medial parietal, anterior inferior temporal regions and fusiform cortex. Beauchemin et al. (2006) applied the oddball paradigm in an ERP study in which a familiar voice and an unfamiliar voice were infrequently presented as deviant stimuli and another unfamiliar voice was frequently presented as the standard stimulus and found that familiar-voice deviants elicited a larger mismatch negativity (MMN) and a P3a than unfamiliar-voice deviants did. These results suggested the presence of long-term memory traces for the familiar voice in human adults. The MMN is also proposed to be related to the processing of voice discrimination by Titova and Naatanen (2001). In their study, a female voice was frequently presented as the standard stimulus whereas one male voice and three female voices were infrequently presented as the deviant stimuli. They found positive correlations between the MMN amplitude and the behavioral deviant-standard dissimilarity ratings. Although voice discrimination and recognition are indexed by the same ERP component (MMN) in adults, the study of brain-damaged patients showed that

deficits in the recognition of a familiar voice occurred after damage to the right hemisphere, while deficits in discriminating among unfamiliar voices occurred after damage to either hemisphere, suggesting that these two abilities are independent and engage different cerebral mechanism (Van Lancker & Kreiman, 1987).

Infants seem to be using different (presumably, immature) neural mechanism for voice recognition because they did not show adult-like ERP components (MMN and P3a) in voice recognition. deRegnier, Nelson, Thomas, Wewerka, and Georgieff (2000) investigated the ERP correlates of recognition memory in newborns by presenting the maternal and stranger's voice while infants slept. In this paradigm, the word "baby" spoken by the mother and a stranger served as the familiar stimulus and unfamiliar stimulus, respectively. The two ERP components, a P2 and a negative slow wave (NSW), showed differences to the mother and stranger's voice. The P2, a positive component in the window of 150–400 msec after stimulus onset, was larger in peak amplitude and longer in latency for the maternal voice than for the stranger's voice, indicating rapid discrimination of the voices. In this paradigm, since the strangers "voices" were actually the voices of the mothers (the stranger's voice varied for each infant and was the voice of the previously tested mother), when the data of all infants are averaged together in the grand mean, there should be no systematic acoustic differences between the voices because they are the same voices. Thus, the voices are not differentiated based on acoustic properties in this paradigm.

The NSW, which was more negative for the stranger's than maternal voice, was thought to reflect the detection of novelty. deRegnier, Wewerka, Georgieff, Mattia, and Nelson (2002) further investigated the role of postnatal experience on memory development using the same paradigm. They found that compared to newborns, infants with 2 weeks postnatal experience showed a longer P2 latency to the maternal voice and a more complex waveform. They interpreted these results as greater neural processing of the second syllable of the word ("ba-by") spoken by the mother in the 2-week-old infants, who had more postnatal experience than the newborns. However, there was no evidence of encoding the novel stimulus. The effect of even more experience after birth (such as 2 months) on auditory memory development is unknown.

The late slow wave has also been observed in the studies of visual recognition memory. In studies performed by Nelson and colleagues, infants were randomly presented with three stimuli with different probabilities, frequent-familiar, infrequent familiar, and infrequent-novel faces (that is, 60%, 20%, and 20% probabilities, respectively). In addition to a NSW (800–1,700 msec) to novel faces, they also observed a positive slow wave (PSW, 800–1,700 msec) to the infrequent-familiar faces in infants aged 6 months (Nelson & Collins, 1991). The NSW was interpreted as reflecting novelty detection, and the PSW was thought to represent memory updating for the partially encoded (i.e., infrequent familiar) stimulus. Although no ERP differences in 4-month-olds were found in the study using the three-stimulus oddball paradigm (Nelson & Collins, 1992), a late PSW in 3-month-olds to a novel face was reported in a study using a habituation paradigm (Pascalis, de Haan, Nelson, & de Schonen, 1998). From the previous studies of visual recognition memory, we thus still cannot conclude at what age the abilities of detecting novelty and updating memory emerge.

The purpose of the present study was to examine the electrophysiological correlates of auditory recognition memory in 2-month-old infants using the paradigm developed by deRegnier et al. (2000). The question is how the infant brain responds to familiar and unfamiliar voices (i.e., mother and stranger), after 2 months of postnatal experience. Infants hear their own mother's voice beginning in utero, and by the age of 2 months, they are already familiar with their own

mother's voice. These experiences may have built up a long-term memory template for the maternal voice. After birth infants are exposed to other environmental stimuli, including other people's voices. The input of these stimuli may facilitate brain responses to unfamiliar stimuli. That is, at birth, infants may differentiate a voice as "mother's voice" or "non-mother's voice." After exposure to more voices besides their own mother's, they may be able to process the stimuli as "mother's voice" (familiar) or "the other voice" (novel), which means that they are able to encode novel stimuli. Therefore, in addition to the enhanced P2 amplitude to mother's voice (reflecting voice discrimination) and enhanced NSW amplitude to the stranger's voice (reflecting novelty detection), we might be able to observe the PSW that is related to memory updating for encoded novel stimuli. That is, the process of voice recognition includes not only retrieval of the memory, but also encoding of the unfamiliar voice. However, prior research gives little indication whether this capacity has developed by 2 months. We therefore were not confident in predicting that 2-month-olds would show electrical evidence of encoding a novel stimulus, although we thought that they would have considerable experience with voices other than their own mother's by that age.

Due to the ease of recording and low rate of artifacts, previous ERP studies of auditory recognition memory were performed in newborns during sleep (deRegnier et al., 2000, 2002; Siddappa et al., 2004). However, the ERP was recorded in infants in a state of quiet awake in the present study in order to keep infant state consistent with assessments at older ages in this longitudinal study. State may be an important factor. Different states and attention or arousal can affect cognition in infants (Colombo, Mitchell, & Horowitz, 1988; Gardner & Karmel, 1995; Richards, 1997). Some adult studies using auditory stimuli have found ERP differences during sleeping and waking (Colrain, Di Parsia, & Gora, 2000; Crowley & Colrain, 2004; Nordby, Hugdahl, Stickgold, Bronnick, & Hobson, 1996). For example, the amplitude of P2 evoked by auditory stimuli was smaller when participants were awake than when they were asleep (Colrain et al., 2000).

Another difference from prior research with this paradigm is that we used high-density recording arrays, which allow for greater spatial sampling of brain activity. In previous studies, data were recorded from only five electrodes, and the amplitude and latency of ERP components were the only outcomes possible to analyze. However, scalp topography is an important dimension of ERP characteristics. Using high-density recording arrays allows better characterization of scalp voltage distribution of the P2 and other components elicited in the auditory recognition memory paradigm. Scalp topography provides the pattern of a component's voltage gradient over the scalp and can indicate underlying neuroanatomical activity (Friedman, Cycowicz, & Gaeta, 2001). Thus, this study was designed to extend previous research by assessing young infants during waking and providing information on scalp topography during an auditory recognition task.

METHODS

Participants

ERP recordings were obtained for 49 healthy 2-month-old infants. All infants met the following criteria: singleton full-term birth (37–42 weeks gestation) weighing >2,500 grams; no prenatal complications or congenital malformations; no general undernutrition (<10th percentile

for weight or length); no acute or chronic illness, no multiple or prolonged hospitalizations (>5 days). The sample for analysis consisted of 26 infants, 11 male, 15 female, with mean chronological age of 59.4 days (*SD* 1.8, range 57–62), gestational age of 39.40 weeks (*SD* 1.05, range 37.4–41.1), and birth weight of 3,418 grams (*SD* 507, range 2,700–4,500). Using the Fenton fetal-infant growth curve (Fenton, 2003), no infants were born small for gestational age (<10th percentile birth weight for gestation), 3 (12%) infants were large for gestational age (>90th percentile), and 23 (88%) were appropriate for gestational age (between the 10th and 90th percentile). Twenty-three were excluded from analysis for the following reasons: failed hearing test (n = 1), low iron status at birth (n = 9), or artifact-contaminated data (n = 13). Infants who were anemic at birth (cord blood hemoglobin <130 g/L), presumably due to iron deficiency, were excluded, since previous studies demonstrated that iron deficiency can alter patterns of recognition memory development using ERPs in newborns and 9-month-old infants (Burden et al., 2007; deRegnier, Long, Georgieff, & Nelson, 2007; Siddappa et al., 2004). Mean age of mothers was 28.2 years (*SD* 3.8, range 21–37). Mean duration of mothers' education was 11.1 years (*SD* 2.7, range 6–16). All mothers reported no history of smoking and drug use. The assessments were conducted in the Children's Hospital of Zhejiang University in China. The study was approved by the Institutional Review Board of the University of Michigan and the Children's Hospital of Zhejiang University. Informed written consent for participation in the study was obtained from the parents.

Stimuli and Procedure

The experimental stimuli were natural human voice stimuli, spoken either by the mother or a stranger. The stranger's voice varied for each infant and was the voice of the previously tested mother. The stimulus consisted of the Chinese word "baobao," meaning "baby," digitized and edited to 750 msec. The infant was presented with auditory stimuli of the maternal and stranger's voices with equal probability. A total of 200 trials were presented, randomly ordered with the constraint that the same stimulus was not repeated for more than two consecutive trials. The interstimulus interval was varied randomly from 2,250 to 3,250 msec.

The recording was made in an electrically shielded quiet room. The infant was seated in the lap of his or her mother and tested in a behavioral state of quiet alertness. The sounds were presented through two loudspeakers placed 65 cm away from each side of the infant's head. Stimulus intensity at the infant's head was 65 dB. The parent was instructed to remain as still as possible. We attempted to minimize infant movement by providing a quiet visual stimulus (experimenter gently blowing soap bubbles) (Dawson et al., 1992). The duration of the recording for the present paradigm was around 12 min.

ERP Recording and Analysis

The electroencephalogram (EEG) was recorded with a 64-electrode HydroCel Geodesic Sensor Net (Electrical Geodesics Inc., Eugene, OR). The two EEG electrodes below the infant's eyes (62, 63) and two of the four electrodes on the outside canthii of the eyes (61, 64) were not used,

resulting in 60-electrode recordings. The EEG signal was amplified using a 0.1–100 Hz bandpass and digitized at 500 Hz. Impedances for each electrode were measured prior to recording and kept below 50 kΩ during testing. Recording in every electrode was vertex-referenced.

The EEG data were processed offline using Net Station 4.3 (Eugene, OR). A 30-Hz low-pass filter was applied and trials were constructed that consisted of a 200 msec baseline period and 2,000 msec period following stimulus onset. Data were baseline corrected to the average voltage during the 200 msec prior to stimulus onset. Segmented data were inspected for ocular and motion artifact. The two electrodes above the eyes (5, 10) and the two remaining electrodes on the outside canthii of the eyes (1, 17) were used to identify eye blinks and eye movement. Data from individual electrodes were rejected if there was artifact resulting from poor contact or movement or if signal amplitudes exceeded 150 μV. The entire trial was excluded if more than nine electrodes were rejected or if an eye blink, eye movement, or other significant artifact had occurred. Of the remaining trials, individual electrodes containing artifact were replaced using spherical spline interpolation. Infants who provided at least 15 artifact-free ERP trials per condition were included in the following analysis. Individual participant averages were constructed separately for the maternal (mean = 40.3, range = 18–69 trials) and stranger's voices (mean = 39.9, range = 18–65 trials), and data were re-referenced to the average reference. The number of artifact-free ERP trials did not significantly differ between maternal and stranger conditions ($t(25) = 0.27, p = .79$).

Inspection of the grand-averaged waveforms revealed that two positive deflections at 150–400 (P2) and 500–1,000 msec (here called P750) from stimulus onset, as well as a following negative-going slow wave (NSW, 1,500–2,000 msec) were predominant over frontal and central regions. A positive-going slow wave (here called PSW, 1,500–2,000 msec) was predominant over lateral temporo-parietal scalp sites (TP7/8 electrodes).

In order to keep consistent with previous studies (deRegnier et al., 2000, 2002), we chose to measure peak amplitude of the P2 instead of mean amplitude, even though mean amplitude is thought by some to be superior to peak amplitude (Luck, 2005). For P750, NSW and PSW, mean amplitude was computed because these components, especially NSW and PSW, did not show a prominent peak, and mean amplitude is equivalent to area amplitude used in previous studies (deRegnier et al., 2000; Nelson & Collins, 1991). Nine frontal, fronto-central, and central electrodes were identified for the P2, P750, and NSW analyses (12, 6, 60, 14, 4, 57, 20, VREF, and 50, corresponding to F3, Fz, F4, FC5, FCz, FC6, C3, Cz, and C4). The peak amplitude and latency of the P2 and the mean amplitude of the P750 and NSW were measured in the time window of 150–400, 500–1,000, and 1,500–2,000 msec after stimulus onset, respectively. They were then analyzed using a 2 × 3 ×3 repeated measures analysis of variance (ANOVA) with stimulus type (mother, stranger), anterocentral scalp location (F = frontal, FC = fronto-central, C = central), and lateral scalp location (3 or 5 = left, z = midline, 4 or 6 = right) as within-subjects factors. Two temporo-parietal electrodes were identified for PSW (25 and 48, corresponding to TP7 and TP8). The mean amplitude of the PSW was measured between 1,500 and 2,000 msec, and then analyzed using a 2 × 2 repeated measures ANOVA with stimulus type (mother, stranger) and hemisphere (left, right) as within-subjects factors. Greenhouse-Geisser corrected degrees of freedom were used; post-hoc paired t-tests were conducted when necessary using Bonferroni correction for multiple comparisons. Data are presented as mean ± SEM.

RESULTS

Grand-average ERP waveforms for the maternal and stranger's voice are shown in Figure 1. Consistent with our expectation, we observed P2, NSW, and PSW, which may reflect voice discrimination, novelty detection, and memory updating, respectively. Unexpectedly, we also observed a P750 component that may be related to the processing of the second syllable of the word stimulus.
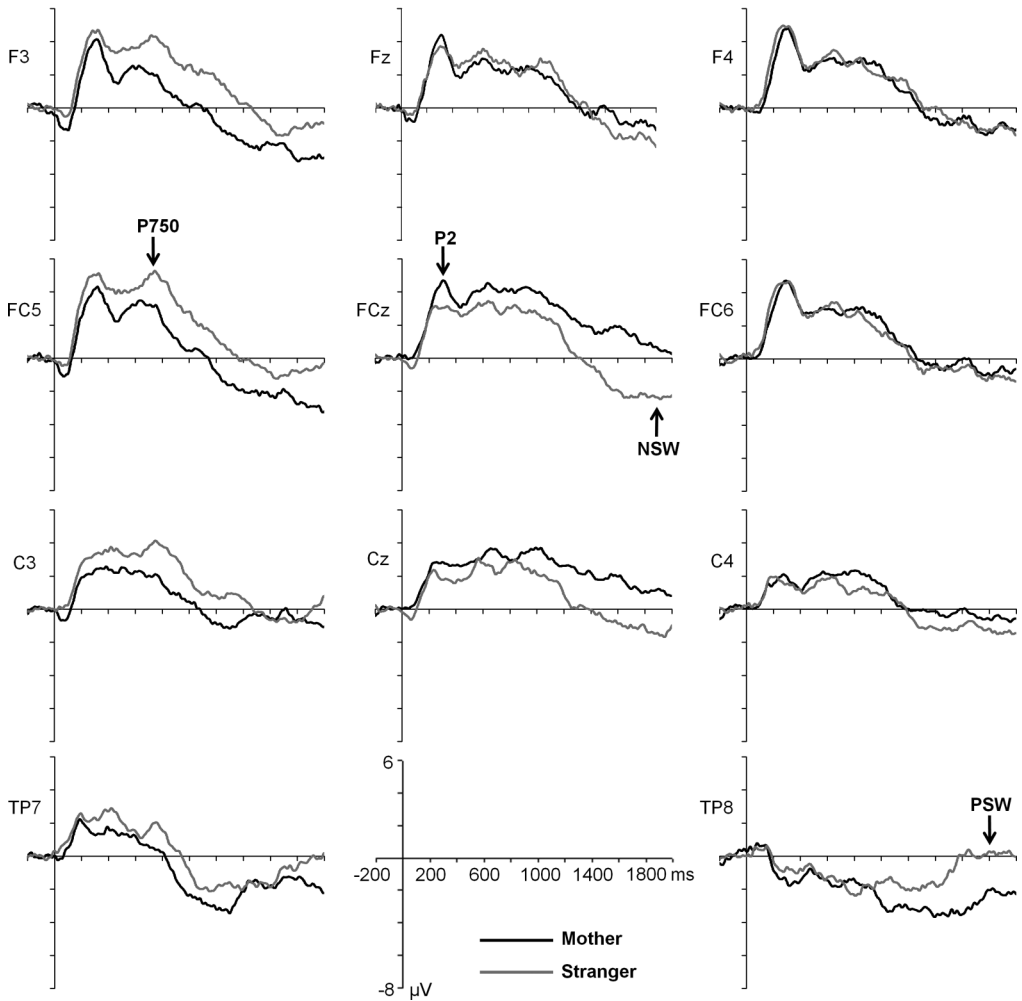
FIGURE 1 Grand average event-related potentials (ERPs) obtained in response to the maternal voice (thick line) and the stranger's voice (thin line). The P2, P750, NSW (negative-going slow wave), and PSW (positive-going slow wave) are indicated.

P2: Voice Discrimination

For P2 amplitude, there was a significant main effect of anterocentral location ($F(2,50) = 20.15$, $p < .001$) and a significant interaction between anterocentral location and laterality ($F(4,100) = 4.34$, $p < .05$), indicating that P2 amplitude was greatest at lateral frontal and fronto-central sites and reduced at more central and midline electrode locations. There was a significant interaction between stimulus type and laterality ($F(2,50) = 4.48$, $p < .05$). Subsequent pairwise comparisons indicated that the P2 elicited by the maternal voice was greater than that by the stranger's voice only at midline electrode locations (mother vs. stranger: 5.04 ± 0.62 μV vs. 3.67 ± 0.36 μV, $p < .05$). The top row of Figure 2 shows the topography of the P2 for the maternal and stranger's voice, respectively. There was a suggestive trend for P2 latency to be longer for the maternal voice than stranger's voice (mother vs. stranger: 294 ± 7 msec vs. 279 ± 8 msec, $F(1,25) = 3.26$, $p = .083$). There were no other significant main effects or interactions for the P2 latency.

P750: Processing of the Second Syllable of the Word Stimuli

There was a significant interaction between stimulus type and laterality ($F(2,50) = 5.76$, $p = .01$). Subsequent pairwise comparisons indicated that average amplitude of the P750 elicited by the stranger's voice was greater than that by the maternal voice only at left electrode locations (stranger vs. mother: 3.73 ± 0.49 μV vs. 2.06 ± 0.58 μV, $p < .05$), also shown in Figure 2 (middle). Within each stimulus type, there was no difference among midline and lateral electrode sites. There was a marginally significant main effect of anteroposterior location ($F(2,50) = 3.64$, $p = .055$), indicating P750 was larger at fronto-central than frontal and central locations.

NSW: Detection of Novelty

The NSW showed a significant interaction between stimulus type and laterality ($F(2,50) = 5.09$, $p < .05$). Subsequent pairwise comparisons indicated that the NSW elicited by the stranger's voice was more negative than that by the maternal voice only at midline electrode locations (stranger vs. mother: -1.58 ± 0.56 μV vs. 0.52 ± 0.66 μV, $p < .05$). Additionally, as shown in Figure 2 (bottom), NSW amplitude to the maternal voice was less negative at midline electrode sites compared to the left electrode sites ($p < .01$).

PSW: Memory Updating

The PSW was more positive for the stranger's voice compared to the maternal voice at temporo-parietal scalp sites ($F(1,25) = 4.45$, $p < .05$). This difference was observed only for the right hemisphere (stranger vs. mother: 0.01 ± 0.90 μV vs. -2.76 ± 0.94 μV, $p < .05$), as shown in Figure 2 (bottom).
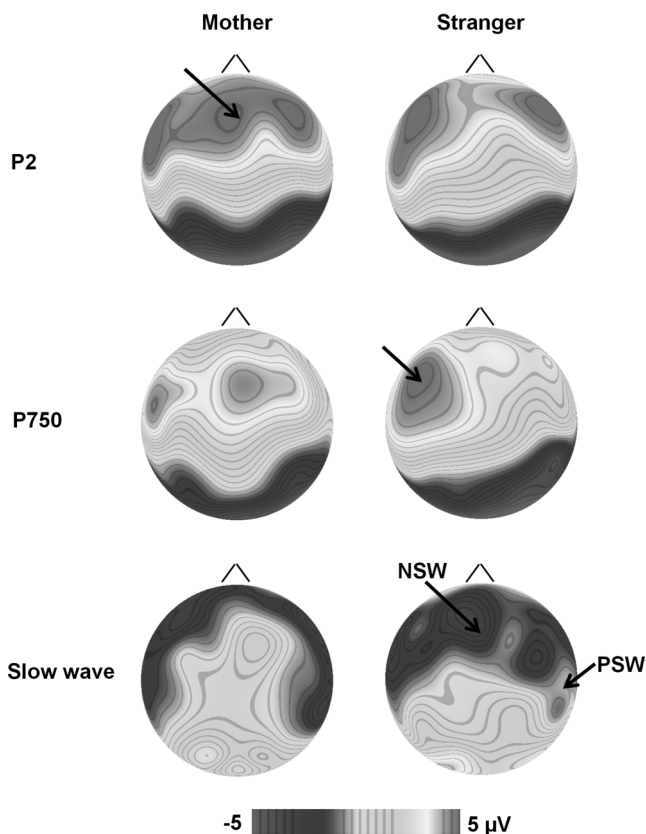
FIGURE 2 Scalp topography of evoked reponses to the maternal and stranger's voice, respectively. Top: Voltage maps of the P2 component at 290 msec following stimulus onset. The arrow indicates that, compared to the evoked responses to the stranger's voice, the P2 is larger in the frontal and central midline sites for the maternal voice. Middle: Voltage maps of the P750 component at 750 msec following stimulus onset. The arrow indicates that, compared to the evoked responses to the maternal voice, the P750 is larger in the left frontal and central sites for the stranger's voice. Bottom: Voltage maps of the slow wave at 1750 msec following stimulus onset. The arrow indicates that, compared to the evoked responses to the maternal voice, the slow wave is more negative in the midline frontal and central sites and more positive in the temporo-parietal sites for the stranger's voice.

DISCUSSION

The present study investigated the auditory ERP to a familiar and unfamiliar stimulus (i.e., maternal and stranger's voice), in healthy awake infants at 2 months of age. In addition to the P2 and NSW reported in studies of sleeping newborns (deRegnier et al., 2000, 2002), the PSW and

P750 showed differences between maternal and stranger's voice in 2-month-olds. These ERP differences between 2-month-old infants in our study and newborns in previous studies suggest that neural correlates underlying auditory recognition memory change over the first 2 months of life. Although previous studies have indicated that cortical auditory ERPs are similar in active sleep and wakefulness in very young infants (Cheour, et al., 2002; Novak, Kurtzberg, Kreuzer, & Vaughan, 1989), it is also possible that these differences might reflect that infants were awake during testing in our study but newborns were in active sleep in previous studies.

Voice Discrimination

The P2 was larger for the maternal voice than stranger's voice, which has also been reported in newborn studies and proposed to reflect rapid discrimination processing (deRegnier et al., 2000, 2002). However, the precise scalp distribution of the P2 was unknown in the previous study, since there were only five electrode sites and the P2 was measured at Fz and Cz. Using high-density electrode arrays in the present study, we found that P2 showed similar amplitude at midline and lateral electrode sites for the maternal voice but smaller amplitude at midline compared to the right electrodes for the stranger's voice. The P2 latency showed a trend to be longer for the maternal voice than the stranger's voice, extending the previous finding of longer P2 latency to the maternal voice after 2 weeks of postnatal experience (deRegnier et al., 2002). Such findings indicate a greater depth of processing of the maternal voice associated with postnatal experience.

Although the time window of the P2 is similar to other voice-specific components found in adults and children, for example, the FTPV that reflects voice-specific processing, the P2 is less like a voice-specific component because it can also be observed when using other auditory stimuli such as pure tones (Novak, Ritter, & Vaughan, 1992). In addition, it is difficult to compare ERP components across adults, children and infant studies because of the variations in task and stimuli.

The P2 has also been considered to reflect an attention-modulated process (Garcia-Larrea, Lukaszewicz, & Mauguiere, 1992; Novak, Ritter, & Vaughan, 1992). In our study, the differences in P2 amplitude between mother's and stranger's voices might indicate that infants attend more to the familiar maternal voice at this young age (2 months). Purhonen, Kilpelainen-Lees, Valkonen-Korhonen, Karhu, and Lehtonen (2004) reported a positivity around 350 msec (P350) in 4-month-old infants that was smaller to the mother's voice as compared to the unfamiliar voice. The P2 in our study may have the same functional significance as the P350 in their study, even though the paradigms differed. In that study, the maternal and unfamiliar voices were presented in intermittent and alternating sequences of four identical stimuli. Although Purhonen et al. (2004) thought that the smaller P350 amplitude to the maternal voice was due to a negative shifting and interpreted this result as infants allocating more attention to process their own mothers' voices, it is also possible that infants at the age of 4 months attend more to unfamiliar voices. This explanation may be supported by findings of behavioral studies on visual memory using the paired-comparison paradigm in which infants are presented with a previous familiarized stimulus paired with a novel one (e.g., Fagan, 1974; Rose et al., 1982). These studies show that infants over the age of 2.5 months spend more time inspecting the novel member of the pair.

Processing of the Second Syllable of the Word Stimuli

A positive deflection (P750) following the P2 was larger for the stranger's voice than for the maternal voice. Although the P750 has not been reported as a distinct component in previous studies, deRegnier et al (2002) described (but did not analyze) the similar positive peak at Cz, which were larger for the maternal voice in full-term infants at 2 weeks (i.e., with 2 weeks of postnatal experience). Because the positive peak seemed to match up with the timing of the second syllables of the word "baby," the authors speculated that it might reflect the greater neural response to the second syllable of the word spoken by the mother. However, we observed that the P750 was larger for the stranger's voice than the maternal voice over the left frontal and central scalp in 2-month-old infants. If deRegnier et al.'s speculation (2002) is correct, our finding would suggest greater neural processing of the second syllable of the word spoken by the stranger. However, it is unclear whether the differing results are due to more electrodes, more postnatal experience, infant states during testing, or different language in our study. In the study performed by deRegnier et al. (2002), the positive peak did not show up in the 1–3-day-olds. Therefore, the appearance of the P750 in older infants suggests that older infants might engage in more processing of the word itself than the newborns, indicating a result of maturation.

In a study of 4-month-old infants, Purhonen, Kilpelainen-Lees, Valkonen-Korhonen, Karhu, and Lehtonen (2005) reported a positive deflection between 500 and 900 msec (P600), which is in the similar time window to P750 in our study. They also observed a smaller P600 for the maternal voice compared to the unfamiliar voice and interpreted it as a memory template for own mother's voice that has been built up by previous experience. However, the stimuli and paradigm in their study differed from ours. In their study, they used a one-syllable word as stimulus and a novelty oddball paradigm in which mother's and stranger's voices as "oddballs" were embedded in a stream of tones. Due to these differences, the functional significance of the P750 in our study may be different from that of the P600 in their study.

The P750 reported in our study was elicited under similar experimental conditions as the late positive peak elicited from newborn infants by deRegnier et al. (2002) and we speculate that our P750 may be a similar peak. deRegnier et al. (2002) speculated that this peak might reflect processing of the second syllable of the word "baby" because the timing of the P2 peak and the late positive peak reflected the bimodal distribution of the sound intensity for the word "baby." However, no further experiments have been performed to confirm or refute this possibility. More evidence is needed to test this theory in the future.

Detection of Novelty and Memory Updating

The fronto-central NSW in our study is similar to the negative slow wave observed in newborn studies. The larger amplitude for the stranger's than maternal voice has been interpreted as detection of novel stimuli against a background of familiar stimuli. The temporo-parietal PSW in our study might be an analogue of the positive slow wave reported in the studies of visual recognition memory, associated with memory updating (Nelson & Collins, 1991, 1992; Nelson & deRegnier, 1992). In our study, infants were presented with a total of 100 trials per stimulus. It is possible that repeated presentations of the unfamiliar stimulus result in encoding of the unfamiliar stimulus so that during the course of the study, the unfamiliar stimulus itself becomes "partially familiar."

Therefore, the PSW might be related to updating of memory for the "partially familiar" stimulus (i.e., the stranger's voice).

This result suggests that 2-month-old infants are able to encode the unfamiliar stimulus but it requires more updating (PSW) than the familiar stimulus. Nelson and Collins (1992) found that there was no difference among the frequent-familiar, infrequent-familiar, and infrequent-novel faces in infants aged 4 months. However, Pascalis et al. (1998) reported a late positive slow wave in 3-month-olds to a novel face in a habituation paradigm in which the novel and familiar faces were presented with equal probability. Other studies using non-face stimuli (visual patterns) in the three-stimulus oddball paradigm observed differences in the late slow wave in infants aged 4.5 months under different attention levels: a frontal NSW following novel-stimulus presentation during attention that was not seen during periods of inattention and a right temporal PSW wave following infrequent-familiar stimulus presentation during attention but not during inattention (Reynolds & Richards, 2005; Richards, 2003). Together, these findings suggest that the updating of memory might emerge at fairly young ages.

According to the present and previous studies, we cannot exclude the possibility that the PSW might be elicited in infants younger than 2 months. In addition, attention has an influence on memory (Reynolds & Richards, 2005; Richards, 2003). Thus, it is also possible that not observing the PSW in newborns might be due to that they were sleeping during the test. To address these issues, future studies are needed to test younger infants or newborns during waking. It should be noted that this may only be tested in the auditory domain but not the visual domain because very young infants have difficulty sustaining visual attention.

The scalp topography of the slow wave in our study is similar to that in Reynold and Richards' study (2005). In that study, the NSW was elicited by the novel stimulus over frontal scalp and the PSW was elicited by the infrequent-familiar stimulus over the right temporal scalp. Using cortical source analysis, the authors further reported that the NSW occurs in the frontal brain regions and the PSW originates from the right temporal regions. The medial temporal lobe structures are required for recognition memory (Clark et al., 2002; Nelson, 1995; Reed & Squire, 1997; Squire et al., 2001). Our findings are consistent with early development of brain areas related to recognition memory. The role of the right hemisphere in voice recognition is also supported by evidence from studies of adults with focal lesion of the brain (Van Lancker & Kreiman, 1987; Van Lancker, Cummings, Kreiman, & Dobkin, 1988). For example, Van Lancker and Kreiman (1987) demonstrated that patients with damage to the right hemisphere were impaired in recognizing familiar voices. The appearance of the right temporo-parietal PSW in 2-month-old infants may suggest the early development of neural systems associated with voice recognition.

One limitation of the present study is that we do not have many artifact-free trials in EEG recording, mainly because infants move too much when they are awake during the test. There were 100 presentations (trials) for each condition but only about 40 artifact-free trials. Therefore, it is difficult to test some of our interpretations with the current dataset. For example, we believe that the PSW might be related to updating of memory for the "partially familiar" stimulus (i.e., the stranger's voice). This explanation could be directly tested by comparing ERPs to the first 25% to the last 25% presentations of the stranger's voice. However, it is unrealistic to do the analysis in the present study because of the few artifact-free trials (about 10 trials) for the first 25 and last 25 presentations. This interpretation should be tested in the future studies by using more trials. Another limitation is that we cannot exclude the possibility that the differences found between our study and previous newborn studies might be due to the fact that infants are in different state

(awake vs. sleeping) during tests. Therefore, longitudinal studies with infants in the same state are needed to verify the differences.

Overall, our study suggests that 2 months of experience with the postnatal environment enables infants to encode an unfamiliar stimulus. The underlying processes of voice recognition at this age might include stimulus discrimination (reflected in the P2), processing the second syllable of the stimulus word (P750), as well as detection of novelty (NSW) and updating of memory for the encoded novel stimulus (PSW). Our findings are important for several reasons. First, our study facilitates longitudinal developmental research on recognition memory by removing infant state as a factor that changes over time. To our knowledge, this is the first ERP study of auditory recognition memory in which infants were awake during the test. Most paradigms for assessing recognition in infants use auditory stimuli in the sleeping infant in the first few months of the postnatal life and visual stimuli in the awake infant after about 3 months. Understanding developmental progressions are likely to be facilitated by having younger and older infants in the same state. Second, our data extend current knowledge of the development of human recognition memory. Infants appear to have the capacity to encode novel stimuli as early as 2 months of age, as reflected in the emergence of a new ERP component (i.e., the PSW). In addition, previous studies have found that the pattern of ERPs evoked by the maternal and stranger's voice was altered in iron-deficient newborns of diabetic mothers and premature newborns compared to typically developing newborns (deRegnier et al., 2002; Siddappa et al., 2004). Future work is needed to examine how auditory recognition memory could be impaired in atypically developing infants after the newborn period.

ACKNOWLEDGMENTS

REFERENCES

Beauchemin, M., De Beaumont, L., Vannasing, P., Turcotte, A., Arcand, C., Belin, P., & Lassonde, M. (2006). Electrophysiological markers of voice familiarity. *European Journal of Neuroscience*, *23*(11), 3081–3086.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*(6767), 309–312.

Brennan, W. M., Ames, E. W., & Moore, R. W. (1966). Age differences in infants' attention to patterns of different complexities. *Science*, *151*(3708), 354–356.

Burden, M. J., Westerlund, A. J., Armony-Sivan, R., Nelson, C. A., Jacobson, S. W., Lozoff, B., . . . Jacobson, J. L. (2007). An event-related potential study of attention and recognition memory in infants with iron deficiency anemia. *Pediatrics*, *120*(2), E336–E345.

Charest, I., Pernet, C. R., Rousselet, G. A., Quinones, I., Latinus, M., Fillion-Bilodeau, S., . . . Belin, P. (2009). Electrophysiological evidence for an early processing of human voices. *BMC Neuroscience*, *10*, 127.

Cheour, M., Martynova, O., Naatanen, R., Erkkola, R., Sillanpaa, M., Kero, P., . . . Hamalainen, H. (2002). Psychobiology—Speech sounds learned by sleeping newborns. *Nature*, *415*(6872), 599–600.

Clark, R. E., Broadbent, N. J., Zola, S. M., & Squire, L. R. (2002). Anterograde amnesia and temporally graded retrograde amnesia for a nonspatial memory task after lesions of hippocampus and subiculum. *Journal of Neuroscience*, *22*(11), 4663–4669.

Colombo, J., Mitchell, D. W., & Horowitz, F. D. (1988). Infant visual attention in the paired-comparison paradigm: Test-retest and attention-performance relations. *Child Development*, *59*(5), 1198–1210.

Colrain, I. M., Di Parsia, P., & Gora, J. (2000). The impact of prestimulus EEG frequency on auditory evoked potentials during sleep onset. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale*, *54*(4), 243–254.

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology*, *115*(4), 732–744.

Decasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, *208*(4448), 1174–1176.

deRegnier, R. A., Long, J. D., Georgieff, M. K., & Nelson, C. A. (2007). Using event-related potentials to study perinatal nutrition and brain development in infants of diabetic mothers. *Developmental Neuropsychology*, *31*(3), 379–396.

deRegnier, R. A., Nelson, C. A., Thomas, K. M., Wewerka, S., & Georgieff, M. K. (2000). Neurophysiologic evaluation of auditory recognition memory in healthy newborn infants and infants of diabetic mothers. *Journal of Pediatrics*, *137*(6), 777–784.

deRegnier, R. A., Wewerka, S., Georgieff, M. K., Mattia, F., & Nelson, C. A. (2002). Influences of postconceptional age and postnatal experience on the development of auditory recognition memory in the newborn infant. *Developmental Psychobiology*, *41*(3), 216–225.

Fagan, J. F. (1974). Infant recognition memory: The effects of length of familiarization and type of discrimination task. *Child Development*, *45*(2), 351–356.

Fenton, T. R. (2003). A new growth chart for preterm babies: Babson and Benda's chart updated with recent data and a new format. *BMC Pediatrics*, *3*, 13.

Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: An event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience and Biobehavioral Reviews*, *25*(4), 355–373.

Garcia-Larrea, L., Lukaszewicz, A. C., & Mauguiere, F. (1992). Revisiting the oddball paradigm—Nontarget vs neurtral stimuli and the evaluation of ERP attentional effects. *Neuropsychologia*, *30*(8), 723–741.

Gardner, J. M., & Karmel, B. Z. (1995). Development of arousal-modulated visual preferences in early infancy. *Developmental Psychology*, *31*(3), 473–482.

Grossmann, T., Oberecker, R., Koch, S. P., & Friederici, A. D. (2010). The developmental origins of voice processing in the human brain. *Neuron*, *65*(6), 852–858.

Karmel, B. Z. (1969). The effect of age, complexity, and amount of contour on pattern preferences in human infants. *Journal of Experimental Child Psychology*, *7*(2), 339–354.

Levy, D. A., Granot, R., & Bentin, S. (2001). Processing specificity for human voice stimuli: electrophysiological evidence. *Neuroreport*, *12*(12), 2653–2657.

Levy, D. A., Granot, R., & Bentin, S. (2003). Neural sensitivity to human voices: ERP evidence of task and attentional influences. *Psychophysiology*, *40*(2), 291–305.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

Nelson, C. A. (1995). The ontogeny of human memory: A cognitive neuroscience perspective. *Developmental Psychology*, *31*(5), 723–738.

Nelson, C. A., & Collins, P. F. (1991). Event-related potential and looking-time analysis of infants' responses to familiar and novel events: Implications for visual recognition memory. *Developmental Psychology*, *27*(1), 50–58.

Nelson, C. A., & Collins, P. F. (1992). Neural and behavioral correlates of visual recognition memory in 4- and 8-month-old infants. *Brain and Cognition*, *19*(1), 105–121.

Nelson, C. A., & deRegnier, R. A. (1992). Neural correlates of attention and memory in the first year of life. *Developmental Neuropsychology*, *8*(2–3), 119–134.

Nordby, H., Hugdahl, K., Stickgold, R., Bronnick, K. S., & Hobson, J. A. (1996). Event-related potentials (ERPs) to deviant auditory stimuli during sleep and waking. *Neuroreport*, *7*(5), 1082–1086.

Novak, G., Ritter, W., & Vaughan, H. G. (1992). Mismatch detection and the latency of temporal judgements. *Psychophysiology*, *29*(4), 398–411.

Novak, G. P., Kurtzberg, D., Kreuzer, J. A., & Vaughan, H. G. (1989). Cortical responses to speech sounds and their formants in normal infants: maturational sequence and spatiotemporal analysis. *Electroencephalography and Clinical Neurophysiology*, *73*(4), 295–305.

Ockleford, E. M., Vince, M. A., Layton, C., & Reader, M. R. (1988). Responses of neonates to parents' and others' voices. *Early Human Development*, *18*(1), 27–36.

Pascalis, O., de Haan, M., Nelson, C. A., & de Schonen, S. (1998). Long-term recognition memory for faces assessed by visual paired comparison in 3- and 6-month-old infants. *Journal of Experimental Psychology-Learning Memory and Cognition*, *24*(1), 249–260.

Purhonen, M., Kilpelainen-Lees, R., Valkonen-Korhonen, M., Karhu, J., & Lehtonen, J. (2004). Cerebral processing of mother's voice compared to unfamiliar voice in 4-month-old infants. *International Journal of Psychophysiology*, *52*(3), 257–266.

Purhonen, M., Kilpelainen-Lees, R., Valkonen-Korhonen, M., Karhu, J., & Lehtonen, J. (2005). Four-month-old infants process own mother's voice faster than unfamiliar voices—Electrical signs of sensitization in infant brain. *Cognitive Brain Research*, *24*(3), 627–633.

Reed, J. M., & Squire, L. R. (1997). Impaired recognition memory in patients with lesions limited to the hippocampal formation. *Behavioral Neuroscience*, *111*(4), 667–675.

Reynolds, G. D., & Richards, J. E. (2005). Familiarization, attention, and recognition memory in infancy: An event-related potential and cortical source localization study. *Developmental Psychology*, *41*(4), 598–615.

Richards, J. E. (1997). Effects of attention on infants' preference for briefly exposed visual stimuli in the paired-comparison recognition-memory paradigm. *Developmental Psychology*, *33*(1), 22–31.

Richards, J. E. (2003). Attention affects the recognition of briefly presented visual stimuli in infants: An ERP study. *Developmental Science*, *6*(3), 312–328.

Richmond, J., & Nelson, C. A. (2007). Accounting for change in declarative memory: A cognitive neuroscience perspective. *Developmental Review*, *27*(3), 349–373.

Rogier, O., Roux, S., Belin, P., Bonnet-Brilhault, F., & Bruneau, N. (2010). An electrophysiological correlate of voice processing in 4- to 5-year-old children. *International Journal of Psychophysiology*, *75*(1), 44–47.

Rose, S. A. (1983). Differential rates of visual information processing in full-term and preterm infants. *Child Development*, *54*(5), 1189–1198.

Rose, S. A., Melloycarminar, P., Gottfried, A. W., & Bridger, W. H. (1982). Familiarity and novelty preferences in infant recognition memory: Implications for information processing. *Developmental Psychology*, *18*(5), 704–713.

Siddappa, A. M., Georgieff, M. K., Wewerka, S., Worwa, C., Nelson, C. A., & DeRegnier, R. A. (2004). Iron deficiency alters auditory recognition memory in newborn infants of diabetic mothers. *Pediatric Research*, *55*(6), 1034–1041.

Squire, L. R., Schmolck, H., & Stark, S. M. (2001). Impaired auditory recognition memory in amnesic patients with medial temporal lobe lesions. *Learning & Memory*, *8*(5), 252–256.

Titova, N., & Naatanen, R. (2001). Preattentive voice discrimination by the human brain as indexed by the mismatch negativity. *Neuroscience Letters*, *308*(1), 63–65.

Van Lancker, D., & Kreiman, J. (1987). Voice discrimination and recognition are separate abilities. *Neuropsychologia*, *25*(5), 829–834.

Van Lancker, D. R., Cummings, J. L., Kreiman, J., & Dobkin, B. H. (1988). Phonagnosia: A dissociation between familiar and unfamiliar voices. *Cortex*, *24*(2), 195–209.

von Kriegstein, K., Kleinschmidt, A., Sterzer, P., & Giraud, A. L. (2005). Interaction of face and voice areas during speaker recognition. *Journal of Cognitive Neuroscience*, *17*(3), 367–376.