Cerebral Hemodynamics in Newborn Infants Exposed to Speech Sounds: A Whole-Head Optical Topography Study

Hiroki Sato,^{1,2*} Yukiko Hirabayashi,^{1,2} Hifumi Tsubokura,^{2,3}
Makoto Kanai,⁴ Takashi Ashida,⁵ Ikuo Konishi,⁶ Mariko Uchida-Ota,^{2,7}
Yukuo Konishi,^{2,8} and Atsushi Maki^{1,2}

¹Hitachi, Ltd., Advanced Research Laboratory, Saitama, Japan

²JST (Japan Science and Technology Agency)/CREST (Core Research for Evolutional Science and Technology), Saitama, Japan

³Department of Pediatrics, Hiroshima City Child Guidance Center, Hiroshima, Japan

⁴Department of Family and Child Nursing and Midwifery, Shinshu University School of Health Sciences, Nagano, Japan

⁵Department of Obstetrics and Gynecology, Shinshu University School of Medicine, Nagano, Japan

⁶Department of Gynecology and Obstetrics, Kyoto University School of Medicine, Kyoto, Japan

⁷The Global Centers of Excellence Program, Keio University, Tokyo, Japan

⁸Center for Baby Science, Doshisha University, Kyoto, Japan

Abstract: Considerable knowledge on neural development related to speech perception has been obtained by functional imaging studies using near-infrared spectroscopy (optical topography). In particular, a pioneering study showed stronger left-dominant activation in the temporal lobe for (normal) forward speech (FW) than for (reversed) backward speech (BW) in neonates. However, it is unclear whether this stronger left-dominant activation for FW is equally observed for any language or is clearer for the mother tongue. We hypothesized that the maternal language elicits clearer activation than a foreign language in newborns because of their prenatal and/or few-day postnatal exposure to the maternal language. To test this hypothesis, we developed a whole-head optode cap for 72-channel optical topography and visualized the spatiotemporal hemodynamics in the brains of 17 Japanese newborns when they were exposed to FW and BW in their maternal language (Japanese) and in a foreign language (English). Statistical analysis showed that all sound stimuli together induced significant activation in the bilateral temporal regions and the frontal region. They also showed that the left temporal-parietal region was significantly more active for Japanese FW than Japanese BW or English FW, while no significant difference between FW and BW was shown for English. This supports our hypothesis and suggests that the few-day-old brain begins to become attuned to the maternal language. Together with a finding of equivalent activation for all sound stimuli in the adjacent measurement positions in the temporal region, these findings further clarify the functional organization of the neonatal brain. *Hum Brain Mapp* 33:2092–2103, 2012. © 2011 Wiley Periodicals, Inc.

Additional Supporting Information may be found in the online version of this article.

Contract grant sponsor: CREST (Japan Science and Technology Agency).

*Correspondence to: Hiroki Sato, Ph.D., Hitachi, Ltd., Advanced Research Laboratory, Hatoyama, Saitama 350-0395, Japan.
E-mail: hiroki.sato.ry@hitachi.com

Received for publication 26 December 2010; Revised 1 April 2011; Accepted 11 April 2011

DOI: 10.1002/hbm.21350

Published online 28 June 2011 in Wiley Online Library (wileyonlinelibrary.com).

Key words: near-infrared spectroscopy (NIRS); neonates; speech perception; language acquisition; auditory cortex

INTRODUCTION

Optical topography (OT), which is a noninvasive functional-imaging modality using near-infrared spectroscopy (NIRS), has been recognized as a powerful tool for investigating the brain functions of infants [Aslin and Mehler, 2005; Koizumi et al., 2003; Minagawa-Kawai et al., 2008; Obrig et al., 2010]. OT measures relative changes in the concentration of oxygenated hemoglobin (Hb) and deoxygenated Hb (oxy-Hb signals and deoxy-Hb signals, respectively) caused by neural activity at multiple locations in the cerebral cortex [Maki et al., 1995]. It can be applied to newborns by virtue of its safety [Ito et al., 2000; Kiguchi et al. 2007] and measurability without body constraint. In addition, OT is free of scanner noise, so it can be used to study sensitive auditory functions such as speech recognition [Sato et al., 1999]. Making use of these advantages of OT, a number of researchers have investigated how the infant brain develops the ability to perceive speech [Bortfeld et al., 2007, 2009; Gervain et al., 2008; Grossmann et al., 2010; Homae et al., 2006, 2007; Kotilahti et al., 2005, 2010; Minagawa-Kawai et al., 2007, 2010; Nakano et al., 2009; Pena et al., 2003; Saito et al., 2007a,b; Telkemeyer et al., 2009].

In particular, researchers have reported remarkable findings for newborn infants [Gervain et al., 2008; Kotilahti et al., 2010; Pena et al., 2003; Saito et al., 2007a,b; Telkemeyer et al., 2009]. The pioneering study of Pena et al. showed a left-dominant activation for speech sounds in newborns (2–5 days after birth); i.e., the total-Hb signal (the summation of the oxy-Hb and deoxy-Hb signals) in the left temporal lobe was significantly greater when they were exposed to recorded maternal-language speech than when they were exposed to the same speech played backward [Pena et al., 2003]. This study provided clear evidence that the functional organization of the newborn brain causes it to process (normal) forward speech (FW) in the left hemisphere but not (reversed) backward speech (BW). With respect to the responses to speech sounds, Kotilahti et al. also observed left-lateralized activation in oxy- and total-Hb signals in response to normal speech sounds and less significant activation in response to music [Kotilahti et al., 2010]. In addition, Telkemeyer et al. reported that the auditory cortex of neonates (2–6 days after birth) exhibits differential sensitivity for processing the varying temporal structure of nonspeech sounds; rapid modulations corresponding to phonemes induced significant bilateral activations whereas slower acoustic modulations induced activations lateralized to the right hemisphere [Telkemeyer et al., 2009]. This finding, based on deoxy-Hb signals, suggests that the development of

speech perception is linked to basic auditory processing capacity. In short, a number of OT studies have provided important information about the neonatal brain mechanism for speech perception.

However, several questions remain unanswered. We focused on two in particular. First, how is cortical activity related to speech sound represented over a wide area of the neonatal brain in OT measurements? Basically, the region measured in previous experiments was limited to parts of the temporal and/or frontal regions [Gervain et al., 2008; Kotilahti et al., 2010; Pena et al., 2003; Saito et al., 2007a,b; Telkemeyer et al., 2009], so the general activation pattern for sound stimuli has not been clarified at the whole-head level. In addition, researchers have not reached a consensus on which Hb signals to use in drawing a conclusion. Pena et al. [2003] used total-Hb signals, and Gervain et al. [2008] primarily used oxy-Hb signals, while Telkemeyer et al. [2009] primarily used deoxy-Hb signals. So which signals should be used to examine functional activation? To address these questions, we developed a whole-head optode cap and visualized the spatiotemporal pattern in oxy-Hb, deoxy-Hb, and total-Hb signals in response to speech sound.

Second, is the left-dominant pattern for FW vs. BW reported in a previous study [Pena et al., 2003] equally observed for any language, or is it clearer for the mother tongue? According to the literature [Binder et al., 2000; Dehaene-Lambertz et al., 2002; Mehler et al., 1988], FW and BW sounds have common overall variations in spectral energy and physical complexity, but BW has spectral changes in different directions and violates several segmental and suprasegmental phonological properties of human speech. Although the amount of phonetic information conveyed by BW sound is not clear [Binder et al., 2000], behavioral studies indicate that BW lacks the phonological information needed for human newborns and tamarin monkeys to discriminate two spoken languages belonging to different rhythmic classes [Mehler et al., 1988; Ramus et al., 2000a]. This means that differences in cortical activity between FW and BW [Pena et al., 2003] could reflect general processes of the primate auditory system for speech sound. As discussed by Minagawa-Kawai et al. [2010], how the early left-dominant activation for FW is related to language processing per se is still unclear. Differences in activation amplitude between FW and BW might reflect low-level acoustic properties: FW contains rapidly changing linguistic segments and/or natural vocal sounds produced by a human vocal tract, which may enhance leftward activation [Minagawa-Kawai et al., 2010]. From this viewpoint, left-dominant activity for FW relative to BW should occur for any language; that is, no activation

differences should be apparent between the maternal language and a foreign language in a FW-BW comparison.

However, we theorize that the maternal language elicits clearer activation than a foreign language between FW and BW in few-day-old newborns for several reasons. To begin with, simple frequency-domain properties do not explain left-lateralized activation in neonates. Telkemeyer et al. [2009] demonstrated right-lateralized activation in response to slow acoustic modulation and bilateral activation in response to rapid modulation. Furthermore, newborns are thought to be sensitive to some auditory properties of their mother's native language due to prenatal exposure to maternal speech; sucking-behavior studies have shown that neonates prefer their mother's voice compared to that of a female stranger [DeCasper and Fifer, 1980], a story read by their mother during late gestation rather than a novel story [DeCasper and Spence, 1986], and their maternal language over a foreign language in a different rhythmic class [Moon et al., 1993]. These findings imply that fetuses can perceive and internalize some acoustic features of their mother's voice, i.e., the primary speech sound present in the womb. More direct evidence supporting this hypothesis was uncovered by Kisilevsky et al. They measured changes in the fetus heart rate and found that near-term fetuses can discriminate their mother's voice from other voices [Kisilevsky et al., 2003] and their mother's native English language from another language (Mandarin) [Kisilevsky et al., 2009]. These findings suggest that fetuses learn some acoustic properties from their mother's voice during their final weeks in the womb. Moreover, Gervain et al. showed that neonates have increasing cortical activation during subsequent brief exposures to a particular speech pattern [Gervain et al., 2008]. They examined the neonatal ability to learn simple speech patterns using syllable sequences containing immediate repetitions (ABB) intermixed with random control sequences (ABC). They found greater responses to the repetition sequences in the temporal and left frontal areas. In addition, subsequent trials showed a further increase in activation for the repetition sequences but not for the random sequences, indicating that recognition of the ABB pattern was enhanced by repeated exposure. On the basis of these findings, we hypothesized that hearing the mother's native language triggers a specific brain activation compared to a foreign language even in few-day-old newborns.

To test this hypothesis, we recorded the cortical activity in Japanese newborns (1–7 days after birth) while they were being exposed to four kinds of sound stimuli produced by the same female voice: their parents' native Japanese language played forward (JFW) and backward (JBW) and the English language played forward (EFW) and backward (EBW). Japanese and English belong to different rhythmic classes, mora-timed and stress-timed, respectively [see Ramus et al., 2000b]. Since newborns discriminate speech sounds in the maternal language from those in a foreign language on the basis of prosodic and rhythmic information [Mehler et al., 1988; Nazzi et al., 1998],

the contrast between Japanese and English spoken by the same person can be sufficient for the present objectives. That is, Japanese and English speech sounds share natural vocal sounds produced by a human vocal tract but differ in linguistic information such as prosody and rhythm.

MATERIALS AND METHODS

Neonates

Seventeen full-term healthy Japanese neonates (11 male and 6 female), ranging in age from 1 to 7 days (mean age: 4.4 days) and having Apgar scores of at least seven at 1 and 5 min after birth, were participated in this study after the parents had given written informed consent. All measurements were conducted with the approval of the Ethics Committee of the Shinshu University School of Medicine, Japan. The gestational ages at their measurement day ranged from 37 weeks and 2 days to 42 weeks and 4 days (mean: 39 weeks and 5.2 days), and the birth weights ranged from 2,554 to 3,482 g (mean: 2,924 g) (see Supporting Information). Prior to the measurements, medical doctors (two of the co-authors) had diagnosed the neonates' health condition and permitted them to be measured.

The experiments were done with the neonates in a state of sleep or quiet rest. The measurements initially failed for nine of them due to their crying or moving but were successful later on the same day or the next day.

Optical Topography (OT) System and Whole-Head Optode Cap

We used an OT system (modified version of model ETG-7000, Hitachi Medical Corporation, Japan) and a whole-head optode cap that covers the whole cortical region with 72 measurement positions (channels: chs) from 20 irradiation positions and 24 detection positions. Figure 1 shows the measurement apparatus and setup. The measurement positions on the brain surface were roughly estimated by using three-dimensional (3D) coordinates determined by scanning a 3D magnetic space digitizer on a head model (Fig. 1E) [Hirabayashi et al., 2008]. However, we could not validate its conformity to the anatomical characteristics of the neonates, so we used the schematic shown in Figure 1D to configure our activation maps.

The OT system simultaneously irradiates light at wavelengths of 695 and 830 nm through an optical fiber to one irradiation point. The average power of each light source was 1.5 mW, and each source was modulated at a distinctive frequency (1–10 kHz) to enable separation by using a lock-in amplifier after detection. The transmitted light was detected every 100 ms with an avalanche photodiode through an optical fiber located 30 mm from the incident position. The optical fibers were supported with fixtures attached to a silicon rubber frame that maintained a distance of 30 mm between the irradiation points and the

Figure 1.

Measurement apparatus and setup: (A) Bed, optode-cap, two stereo speakers, and digital video camera. (B) Inside view of optode cap. Silicon rubber frame with fixtures holding optical fibers; it maintains distance between irradiated points and detection points. Fixtures and frame are arranged on a sponge base, and the frame crosspieces rotate around the fixtures, fitting the cap to the head size of the neonate. (C) Neonate wearing optode cap. (D) Arrangement of measurement positions. (E) Measurement positions estimated using 3D digitizer on model head based on 3D MR images of 2-month-old preterm infant (Hirabayashi et al. 2008).

detection points (Fig. 1B). The measurement positions were defined as the midpoint between the irradiation and detection points. The silicon rubber frame rotated around the fixtures and comfortably secured the optode cap to the neonate's head despite the differences in head size. The fixtures and frame were arranged on a sponge base, so placement of the optode cap on the neonate's head did not awaken him or her.

Stimuli

We used four types of sound stimuli: Japanese played forward (JFW), Japanese played backward (JBW), English played forward (EFW), and English played backward (EBW). These stimuli were created from the recorded voice of a Japanese-English bilingual female (27-year-old Japanese) who was born in England and who had lived there until 2-years old (she had also lived in the United States

from age 5 to 7 and from 12 to 15). She read a portion of a children's story in Japanese and in English to infants (both highly intonated compared to an adult-directed voice). Each recorded speech sound was separated into six 10-s audio clips. The backward speech sounds were made from the audio clips using sound-editing software (Sound Forge® 7.0, Sony Pictures Digital Networks, USA). The fundamental sound properties of the stimuli were analyzed using PRAAT speech signal analysis software (ver. 5.2.17) (www.praat.org). The mean pitch (fundamental frequency: F0) among six audio clips was respectively 240.3 and 245.3 Hz for Japanese and English, with no significant difference ($t(10) = -0.58$, $P = 0.58$). The mean intensity was respectively 74.6 and 74.3 dB (μ E) for Japanese and English, with no significant difference ($t(10) = 1.11$, $P = 0.29$). In total, we prepared 24 audio clips, six for each type of sound stimuli (JFW, JBW, EFW, and EBW).

Although the clips all contained natural pauses, none started with a pause. A 10-s silence audio clip was used as a control stimulus (CTL).

Measurement Procedure

The neonate to be evaluated was laid in the measurement bed and fitted with the optode cap (Fig. 1C) while he or she was asleep or in a quiet rest state. We conducted two sessions for each neonate: a sound session presenting the four types of sound stimuli (JFW, JBW, EFW, and EBW) and a control session presenting the control stimuli (CTL). All stimuli presentations were software controlled (Platform of Stimuli and Tasks, Hitachi, ARL).

The sound session contained 25 random-length silence periods and 24 stimulation periods. The silence periods lasted between 20 and 30 s, and the stimulus periods lasted 10 s. The silence period durations were varied to reduce the effect of synchronization between stimuli exposure and spontaneous oscillations [Pena et al., 2003]. In each stimulus period, the four types of sound stimuli were presented six times each ($4 \times 6 = 24$ audio clips) in a pseudo-randomized order through stereo speakers (MediaMate®, Bose, USA). The volume was controlled so that the sound pressure at the neonate's head was 62–65 dB while the background noise (mainly from an air conditioner and the measurement PC) was about 50–53 dB. At least one sound session was conducted for each neonate; a second one was conducted if the first was not completed because of the neonate moving or crying. When a second sound session was conducted, the data recorded in the first session was used to the extent possible. Therefore, the number of stimuli per neonate ranged from 24 to 34. The data recorded during these sound sessions with silence period (20–30 s) enabled us to analyze the hemodynamic signals in response to sound stimuli, as was done in previous studies [Grossmann et al., 2010; Homae et al., 2006, 2007; Minagawa-Kawai et al., 2010].

The control session contained seven random-length silence periods and six CTL periods, which corresponded

to the repetition time for one of the sound stimuli. The silence periods lasted between 20 and 30 s, and the CTL periods lasted 10 s, the same as in the sound session. Although there were no sound stimuli during either the silence periods or the CTL periods, the cortical activity was recorded exactly as it was during the sound session. The sound session was always conducted before the control session to improve the probability of complete measurement for the sound session, which was more important. The data recorded during these control sessions enabled us to identify the spontaneous variations in hemoglobin signals for time periods corresponding to those of the sound sessions.

Data Analysis

We used plug-in analysis software (Platform of Optical Topography Analysis Tools, Hitachi, ARL) running on MATLAB (The Mathworks, USA) to analyze the OT data. The optical data for two wavelengths measured during each session were transformed into time-series signals (oxy-Hb and deoxy-Hb signals) on the basis of the modified Beer-Lambert law [Delpy et al., 1988; Maki et al., 1995]. The sum of the oxy-Hb and deoxy-Hb signals was defined as the total-Hb signal. These signals were expressed as the products of the changes in hemoglobin concentration (mM) and optical path length (mm) in the activation region (effective optical path length).

The continuous Hb signals for both sessions were divided into 27-s data blocks. Each block consisted of a 1-s silence period before a stimulus period (pre-stimulus period), a 10-s stimulus period (audio clip presentation or CTL period), and a 16-s silence period after the stimulus period (post-stimulus period).

We excluded the data blocks with body-movement artifacts by using an established procedure [Sato et al., 2006], which detects unusual rapid changes due to body movements on the basis of wavelet analysis. The parameters for the total-Hb signals (scale, threshold) were set on the basis of the optimal values determined in a previous study (9, 43) [Sato et al., 2006]. In addition, data blocks with an unusually large variation in any of the Hb signals (> 0.8 mM·mm within 1 s) were considered to be error blocks and discarded. These variations may have been due to insufficient contact between the optodes and scalp. For each neonate, we excluded channels for which we did not obtain at least two good data blocks for each type of sound stimuli. As a result, for each channel, we used data for about 12 neonates (mean $\pm$ SD: 11.8 ± 2.6), comparable to the use of data for 12 neonates by Pena et al. [Pena et al., 2003]. The mean numbers of data blocks per channel per neonate used in the analysis were 5.54 ± 1.26 (JFW), 5.47 ± 1.30 (JBW), 5.55 ± 1.13 (EFW), 5.39 ± 1.18 (EBW), and 4.75 ± 1.07 (CTL).

These valid data blocks were filtered with a 2-s moving average to attenuate the high-frequency noise. In addition,

the baseline was corrected using linear fitting to the mean signal for the first ten time points (1 s) and that for the last ten time points. We then used the filtered data for our statistical analysis. First, the time series of the data blocks was averaged for every 2-s window after stimulus onset until 22 s after the onset in each channel to enable us to observe the primitive activation patterns in each Hb signal at the whole-brain level. A simple *t* test of the mean values was performed for each time window for all sound stimuli together against a zero baseline. In addition, assuming that the average Hb signals were flat in the non-activated regions, we calculated the grand average of the Hb signals for each sound session across all channels and neonates to find a template waveform that represented the activation pattern. We defined the activation periods for the oxy-Hb signal, deoxy-Hb signal, and the total-Hb signals in accordance with the full-width at 75% maximum of the template waveforms. In the channels that showed significant activation for all three hemoglobin signals, we calculated the contrast-to-noise ratio (CNR) for each signal and for each neonate to determine the most sensitive indicator in our data for cortical activation. The contrast was defined as the activation value (mean signal amplitude during activation period) for the average data block for the sound session, and the noise level was defined as the standard deviation for the average data block for the control session. This noise level, which reflected a spontaneous fluctuation in a data block of the control session, was quite fair because the noise itself contained all frequency ranges that were the same as those in the sound session blocks in terms of both system noise and biological fluctuations.

Finally, we performed a two-way ANOVA (sound stimuli $\times$ neonates) for each channel using the activation values for all sound stimuli (JFW, JBW, EFW, EBW, and CTL) in the Hb signal showing the highest CNR, where individual differences (i.e., the "neonate factor") were considered to be a random effect. For a channel with a significant main effect of sound stimuli, a post-hoc test (Tukey-Kramer) was conducted to identify specific differences in the responses to speech sounds. The results of this analysis using data blocks are considered reliable, particularly since they took within-neonate reproducibility of the activation signals into consideration.

RESULTS

General Activation Pattern in Sound Stimuli

First, we determined the spatiotemporal activation pattern for each hemoglobin signal (see Fig. 2). For the oxy-Hb signals, significant increases in response to sound stimuli compared to the zero baseline were found in the frontal and bilateral temporal regions, mainly for the time window 4–14 s after sound onset (*t* test, $P < 0.01$, corrected for 72 channels). In similar regions, the deoxy-Hb signals showed significant decreases, mainly for the 8–18-s time window, following the increases in the oxy-Hb

Figure 2.

Spatiotemporal hemodynamic patterns in response to sound stimuli: Activation maps in 2-s time windows for oxy-Hb, deoxy-Hb, and total-Hb signals. Data from blocks for all sound conditions (JFW, JBW, EFW, and EBW) were used together in the analysis. In each map, channels with significant activation (determined using one-sample t test against zero) are indicated by asterisks (two-tailed, $P < 0.01$, corrected for 72 multiple comparisons).

signals. In contrast, the total-Hb signals resulted in fewer channels and fewer periods that showed statistically significant activation while they showed similar increase patterns as the oxy-Hb signals.

To identify a waveform (temporal feature) representing the activation in each Hb signal, we examined the grand average of the Hb signals for the sound sessions across all channels and neonates. As shown in Figure 3A, the full-widths at 75% maximum of the template waveforms were

6.3–13.6 s for oxy-Hb, 8.9–17.1 s for deoxy-Hb, and 5.2–11.8 s for total-Hb signals after sound onset. The activation channels were determined by conducting t tests for each Hb signal using the mean signals for these time windows (activation values) ($P < 0.01$, corrected for 72 channels) (Fig. 3B). Common activation channels for all Hb signals were found in the posterior part of the left temporal region (ch 46) and anterior parts of the right temporal region (ch 17, 36, and 44), which are marked with red

Figure 3.

Summarized hemodynamic patterns in response to sound stimuli: (A) Grand average of Hb signals for all channels for all neonates. Activation periods were defined as the full-width at 75% maximum for each Hb signal; they are indicated by transparent rectangle for each color. (B) Activation maps derived from activation values. Channels with significant activation (determined using one-sample t-test against zero) are indicated by asterisks (two-tailed, $P < 0.01$, corrected for 72 multiple comparisons). Channels that showed significant changes for all hemoglobin signals are marked with red circles.

circles in Figure 3B. The contrast-to-noise ratios (CNRs) for these activation channels were compared among the Hb signals (Table I). This ratio is the contrast (mean activation value for all sound stimuli) divided by the fluctuation (standard deviation) in the CTL block. The CNR for the oxy-Hb signal (1.942–3.009) was higher than those for the deoxy-Hb (1.214–2.132) and total-Hb signals (1.314–1.836), indicating that the oxy-Hb signal was the most sensitive indicator for cortical activation in our data.

Differences in Activation Signals Among Sound Stimuli

Using the activation values for the oxy-Hb signals for all sound stimuli (JFW, JBW, EFW, EBW, and CTL), we conducted a two-way ANOVA (sound stimuli $\times$ neonates) for each channel, with the neonate factor considered to be a random effect. For a channel with a significant main effect of sound stimuli ($P < 0.05$, uncorrected), a post-hoc test (Tukey-Kramer) was used to identify specific differences in the responses to speech sounds. This analysis revealed that several channels had significant differences between the sound conditions around the temporal areas (see Fig. 4). Ch46 in the left temporal area, showing the main effect of sound stimuli ($F(4,14) = 2.75$, $P < 0.05$), had larger responses for all sound stimuli relative to that for the CTL condition (JFW to CTL and JBW to CTL: $P < 0.01$, EFW to CTL and EBW to CTL: $P < 0.05$). The significantly activated region was expanded around the left temporal area for the contrast between JFW and CTL (ch45, 46, 47, and 48) (Fig. 4A). In particular, a significant difference ($P < 0.05$) was found in a direct comparison between JFW and JBW for ch48 in the temporal-parietal region (main effect of sound stimuli: $F(4,14) = 4.03$, $P < 0.01$). Moreover, a direct comparison between normal speeches (JFW and EFW) showed a significant difference ($P < 0.001$), indicating that responses to the maternal language are greater than to a foreign language. The activation values for ch45 (main effect of sound stimuli: $F(4,9) = 3.35$, $P < 0.02$) in the inferior part of the left temporal cortex were statistically larger for normal speech (JFW and EFW) than for the CTL conditions (JFW to CTL: $P < 0.005$, EFW to CTL: $P < 0.05$), while the BW speech (JBW and EBW) did not elicit statistically significant responses.

In the right temporal area, ch44 showed the main effect of sound stimuli ($F(4,14) = 3.14$, $P < 0.02$) and had larger increases for the English stimuli (EFW and EBW) relative to that for the CTL condition ($P < 0.05$). While activation values for the Japanese stimuli (JFW and JBW) in comparison to that for CTL were not statistically significant, the time courses of the oxy-Hb signals showed parallel responses for all sound stimuli (Fig. 4B). This activation pattern was similar to that shown by ch46 in the left hemisphere.

DISCUSSION

Our study used a whole-head OT system to measure spatiotemporal hemodynamics accompanying speech sound processing in newborn infants (1–7 days after birth). The results demonstrated the basic pattern of hemodynamic responses at the whole-head level and replicated the left-dominant activity for the contrast between FW and BW speech in the maternal language. Overall, our data are consistent with previously reported results [Dehaene-Lambertz et al., 2002; Kotilahti et al., 2005; Pena et al., 2003],

TABLE I. Contrast-to-noise ratio (CNR) for hemoglobin signals

		Oxy-Hb	Deoxy-Hb	Total-Hb
Contrast (activation value) (mM-mm)	(Ch17)	0.063 ± 0.055	-0.025 ± 0.031	0.038 ± 0.043
	(Ch36)	0.098 ± 0.075	-0.048 ± 0.047	0.050 ± 0.073
	(Ch44)	0.100 ± 0.085	-0.035 ± 0.033	0.065 ± 0.072
	(Ch46)	0.105 ± 0.067	-0.033 ± 0.042	0.072 ± 0.055
Noise (mM-mm)	(Ch17)	0.056 ± 0.039	0.032 ± 0.029	0.054 ± 0.044
	(Ch36)	0.048 ± 0.027	0.023 ± 0.011	0.051 ± 0.031
	(Ch44)	0.049 ± 0.027	0.030 ± 0.020	0.050 ± 0.042
	(Ch46)	0.063 ± 0.058	0.029 ± 0.024	0.069 ± 0.083
CNR ($ contrast /noise$)	(Ch17)	1.942 ± 1.993	1.214 ± 1.807	1.314 ± 1.404
	(Ch36)	3.009 ± 2.427	2.132 ± 2.028	1.727 ± 2.186
	(Ch44)	2.498 ± 1.782	1.473 ± 1.683	1.679 ± 1.931
	(Ch46)	2.664 ± 2.261	2.081 ± 3.043	1.836 ± 1.776

The mean $\pm$ SD across neonates is shown for each signal using data from channels that showed significant changes for all hemoglobin signals (Fig. 3b).

showing that our whole-head OT system is useful for obtaining integrated information on neonatal cortical activity in response to speech sounds. Furthermore, we extended the previous findings with our finding that the left temporal-parietal activation (ch48) due to a FW-BW difference is more evident for the maternal language (Japanese) than for a foreign language (English). This finding supports our hypothesis that the left-dominant pattern for FW vs. BW reported in previous studies [Dehaene-Lambertz et al., 2002; Kotilahti et al., 2005; Pena et al., 2003] is clearer for the mother tongue. Although a few concerns about Japanese-English differences in speech properties remain (see *Limitations*), this finding indicates the existence of neural attunement to certain sound features of the maternal language (e.g., intonation and rhythm) within a few days of birth, probably due to prenatal and/or a few days of postnatal exposure to the mother's voice. In addition, we found equivalent activation for all sound stimuli compared to changes for the CTL condition in the middle part of the bilateral temporal regions (left: ch46, right: ch44). This equivalent activation for all sound stimuli might correspond to activation in the primary auditory cortex. Considering these findings together, the present study further clarifies the functional organization of the neonatal brain for speech perception.

Basic Hemodynamic Pattern in Response to Sound Stimuli

We found different patterns in the temporal changes among the Hb signals; the oxy-Hb signal primarily increased before the deoxy-Hb signal decreased in the activation regions. Similar patterns have been commonly observed in a number of NIRS studies on adults [Obrig et al., 1996, 2000; Sato et al., 1999, 2005], but physiological clarification is one of challenges that should be addressed in the future. In this study, we used the temporal difference among hemoglobin signals to estimate cortical activa-

tion. This analysis revealed comparable activation regions for the oxy-Hb and deoxy-Hb signals while the total-Hb signal showed smaller activation region. This indicates that the total-Hb signal, which is the sum of the oxy-Hb and deoxy-Hb signals, is not sensitive to the typical activation pattern in which the oxy-Hb signal increases and the deoxy-Hb signal decreases. Furthermore, comparison of the CNRs for the common activation channels revealed that the oxy-Hb signal is the most sensitive indicator for statistical analysis (Table I). Whereas deoxy-Hb signals have an advantage because they correspond to blood oxygen level dependent (BOLD) signals in functional magnetic resonance imaging (fMRI) [Huppert et al., 2006; Obrig et al., 2000; Telkemeyer et al., 2009], our data suggest that oxy-Hb signals are good for evaluating functional activity.

Our analysis of the oxy-Hb and deoxy-Hb signals enabled us to identify significant activation in the frontal and bilateral temporal regions in response to sound stimuli as a whole (see Fig. 3). Of particular importance is that whole-head imaging revealed significant activation localized in plausible regions. Use of this whole-head OT system will thus contribute to advances in functional network analysis. As Homae et al. demonstrated, OT is a powerful tool for clarifying the organization of cortical networks in infants [Homae et al., 2010].

Activity in the Frontal Area

A basic analysis using all sound stimuli together showed wide activation in the frontal area. While a specific activity depending on sound stimuli could not be identified, this frontal activation is one of the interesting aspects of our data. Although this activation is difficult to interpret, similar responses to speech sounds by newborns [Gervain et al., 2008; Saito et al., 2007a,b] and by 3- and 10-month-old infants [Homae et al., 2006, 2007] have been found in studies using the same modality. Taga et al. found that visual stimuli also activate the frontal cortex in

Figure 4.

Differences among sound stimuli determined using oxy-Hb signals: (A) Difference values for compared pair are shown in color in maps at positions where they were statistically significant (Turkey-Kramer post-hoc test, $P < 0.05$, uncorrected). (B) Average time courses for channels of interest are shown with error bars indicating standard error. Number of neonates on average was 15 for ch 44, ch 46, and ch 48 and 10 for ch 45.

newborns [Taga et al., 2003]. Recent OT studies have illuminated the frontal functions in infants. Minagawa-Kawai et al. found that the anterior part of the orbitofrontal cortex in 9- to 13-month-old infants was activated when the

infants viewed their mothers' smile, probably due to emotional changes [Minagawa-Kawai et al., 2009]. Moreover, the prefrontal cortex was activated in 3-month-old infants in relation to the auditory habituation/dishabituation

process, i.e., activation in relation to novel stimuli [Nakano et al., 2009]. This finding is particularly important for interpreting our finding of frontal activation in response to sound stimuli. Because we used different audio clips for each presentation, even for the same types of stimuli, in each measurement session, one factor in the activation of the prefrontal cortex might have been the perception of a novel sound stimulus. Note that the frontal activation expanded to both hemispheres and that the lateralization was unclear.

Functional Organization for Speech Processing in the Temporal Area

Random-effect analysis to identify the activation differences between sound conditions revealed a functional organization for speech processing in the left temporal region (see Fig. 4). We found significant activation for all sound stimuli relative to the CTL condition for a channel of the middle part of the region (ch 46), while the superior channels (ch 47 and 48) and the inferior channel (ch 45) showed dominant activation for Japanese speech. In relation to the ch 46 activation, a similar hemodynamics in response to every sound stimulus was found for ch 44 in the right temporal region. Although only English sounds (EFW and EBW) showed responses with statistical significance for ch 44 ($P < 0.05$), equivalent temporal changes were observed for Japanese sounds (JFW and JBW). In our data, ch 44 in the right hemisphere was more anterior than ch 46 in the left hemisphere, which agrees with the findings of Kotilahti et al. [2005]. While the cause has not been determined yet, they suggested the possibility of structural and functional asymmetries in the lateral temporal cortices at birth. We speculate that the bilateral activations in ch 46 and ch 44 correspond to activation in the primary auditory cortex, which was found in previous OT studies of basic auditory functions in newborns [Kotilahti et al., 2005] and in 2- to 3-month-old infants [Taga and Asakawa, 2007]. It is reasonable that the primary auditory cortex is equally activated by any sound stimuli that is produced by the same speaker and that is equalized in terms of mean volume and length. In fact, an fMRI study by Dehaene-Lambertz et al. [2002] showed that signals from the temporal lobe do not differ between FW and BW in the maternal language, even in older infants. Unlike in ch 46, which showed activation common to all sounds, a specific activation depending on the speech sound was found in the adjacent channels (ch 45, 47, and 48) in the left hemisphere. In particular, a superior channel (ch 48) showed a significant activation for JFW in comparison not only to CTL but also to JBW and EFW, which suggests that this activation was specific to the maternal language and was different from any other sound condition. Although a few concerns about Japanese-English differences in speech properties remain (see Limitations), this finding indicates the existence of neural attunement to certain

sound features of the maternal language (e.g., intonation and rhythm) within a few days of birth, probably due to prenatal and/or a few days of postnatal exposure to the mother's voice [DeCasper and Fifer, 1980; DeCasper and Spence, 1986; Kisilevsky et al., 2003, 2009; Moon et al., 1993]. We speculate that this selective activity was in the superior part of the primary auditory cortex (temporal-parietal area), around the angular gyrus, although we could not identify the exact region due to technical limitations (see Limitations). Our speculation is supported by the findings of the fMRI study by Dehaene-Lambertz et al. [2002], which revealed that the left angular gyrus is activated more significantly by FW speech than by BW speech in 2- to 3-month-old infants. In addition, a number of studies on adults have shown that the left angular gyrus is involved with speech processing [Binder et al., 2000; Dehaene et al., 1997; Mazoyer et al., 1993; Perani et al., 1996].

The inferior part of the left temporal cortex (ch 45) showed a significant activation not only for Japanese but also for English in comparison to CTL, whereas both BW stimuli did not induce significant differences from the CTL condition. Although we lacked the statistical power needed to identify the difference in a direct comparison between FW and BW, our data supported left-dominant activation for FW sounds in the inferior part of the temporal area [Pena et al., 2003]. Furthermore, our findings extend the findings of Pena et al. [2003]: left-dominant activity for FW speech might be common to two different languages in the inferior region. This finding raises the possibility that the inferior part of the temporal cortex is involved with an innate mechanism that responds to any natural language sounds, whereas the superior part responds to the familiarity of certain sound features of the maternal language due to prenatal and/or postnatal experience. This is simply speculation—further research is needed to fully clarify the functional organization of the neonatal temporal cortex. Together with a finding of equivalent activation for all sound stimuli in the adjacent measurement positions in the temporal region, which might correspond to the primary auditory cortex, our results suggest that the brain of an infant who is only a few days old shows functional differentiation for processing general auditory stimuli and for processing some speech properties.

Limitations

There are several limitations to our study. First, we could not accurately determine the anatomical position for each measurement channel. A previously proposed technique for estimating the channel positions on the brain surface using a real-head model based on MRI data for the infant (Fig. 1E) [Hirabayashi et al., 2008] was not usable partly due to the structure of the optode cap. The middle of the frontal-inferior edge of the cap was positioned

above the nasion, but the channels could not be positioned at regular coordinates due to differences in neonate head size and/or the manner in which the cap was placed on the neonate's head. There were thus differences (a few centimeters or less) in the measurement positions among the neonates. This means that the activation regions might have been underestimated because the activity was evaluated from a somewhat random dispersion of measurement positions among the neonates. We will be able to obtain more reliable results once we have improved the optode cap and have found a way to position the channels at regular coordinates, possibly by using anatomical landmarks.

Second, there are imperfections in our data set. As described in the Materials and methods section, the number of neonates evaluated depended on the channel. Given that the data for different channels was sometimes from different sets of neonates, it is possible that there are systematic differences in the number of neonates for each channel, which could possibly contaminate the results. However, the dispersion of the number of neonates among channels should be random, and we confirmed that the number of neonates did not correlate with the significance level of activation (*p*-values for the oxy-Hb map in Fig. 3B) across channels ($r = -0.03$, $P = 0.81$).

Third, the small number of stimuli (two to six usable data-sets per stimuli condition per neonate) could have resulted in large a standard error among neonates (Fig. 4B). While we could not increase the number of stimulus repetitions (measurement periods) due to practical reasons, the use of more repetitions would produce more reliable results with higher statistical significance. To overcome this limitation, we need to increase the number of successful trials by reducing the effect of body-movement artifacts by possibly using an improved optode cap and by shortening the measurement time by possibly developing a more powerful analysis method for extracting activation signals.

Finally, the differences in speech properties between Japanese and English could have caused activation differences. Perani et al. [1996] found that Japanese generally induces greater brain activation than English in Italian native speaker adults. Moreover, a previous study showing that Japanese is very different from English and Italian in vocalic intervals and consonantal intervals [Ramus et al., 2000b] might provide useful ideas for addressing this issue. Research has shown that newborns can discriminate languages on the basis of their rhythmic properties even if the languages are unfamiliar to them [Nazzi et al., 1998], but differences in neural processing corresponding to different rhythm classes have not been found. Further studies will be needed to investigate the possibility that Japanese, a mora-timed language, has more differences in prosodic properties between FW and BW sounds than English, a stress-timed language.

Nevertheless, our study using whole-head OT has helped clarify the functional organization of the newborn cerebral cortex during speech processing, demonstrating

that the use of whole-head OT will lead to further advances in this field. We believe that international collaboration on a systematic examination of the effects of neonate age, maternal language, measurement modality, and sleeping stage on cortical activation will advance our overall knowledge of the language acquisition process.

ACKNOWLEDGMENTS

The authors thank Dr. J. Mehler for his helpful advice in a previous collaboration, Mr. N. Moriya for developing the measurement devices, Ms. M. Kimura-Shibanuma, Dr. T. Hasegawa, Dr. T. Katura, Mr. H. Atsumori, Ms. K. Yamazaki, and Dr. A. Nakai for their technical assistance, Dr. N. Tanaka, Dr. M. Kiguchi, Ms. Y. Yamamoto, Dr. K. Kubota, Dr. T. Hashizume, Dr. H. Obrig, Dr. F. Homae, and Dr. G. Taga for their helpful comments, and Dr. N. Osakabe and Dr. H. Koizumi for their encouragement. They also thank the newborns and their parents for their participation and the staff at Shinshu University Hospital for their cooperation.

REFERENCES

- Aslin RN, Mehler J (2005): Near-infrared spectroscopy for functional studies of brain activity in human infants: Promise, prospects, and challenges. *J Biomed Opt* 10:11009.
- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET (2000): Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex* 10:512–528.
- Bortfeld H, Wruck E, Boas DA (2007): Assessing infants' cortical response to speech using near-infrared spectroscopy. *Neuroimage* 34:407–415.
- Bortfeld H, Fava E, Boas DA (2009): Identifying cortical lateralization of speech processing in infants using near-infrared spectroscopy. *Dev Neuropsychol* 34:52–65.
- DeCasper AJ, Fifer WP (1980): Of human bonding: Newborns prefer their mothers' voices. *Science* 208:1174–1176.
- DeCasper AJ, Spence MJ (1986): Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behav Dev* 9:133–150.
- Dehaene-Lambertz G, Dehaene S, Hertz-Pannier L (2002): Functional neuroimaging of speech perception in infants. *Science* 298:2013–2015.
- Dehaene S, Dupoux E, Mehler J, Cohen L, Paulesu E, Perani D, van de Moortele PF, Lehericy S, Le Bihan D (1997): Anatomical variability in the cortical representation of first and second language. *Neuroreport* 8:3809–3815.
- Delpy DT, Cope M, van der Zee P, Arridge S, Wray S, Wyatt J (1988): Estimation of optical pathlength through tissue from direct time of flight measurement. *Phys Med Biol* 33:1433–1442.
- Gervain J, Macagno F, Cogoi S, Pena M, Mehler J (2008): The neonate brain detects speech structure. *Proc Natl Acad Sci USA* 105:14222–14227.
- Grossmann T, Oberecker R, Koch SP, Friederici AD (2010): The developmental origins of voice processing in the human brain. *Neuron* 65:852–858.
- Hirabayashi Y, Sato H, Uchida-Ota M, Nakai A, Maki A (2008): Technique for designing and evaluating probe caps used in

- optical topography of infants using a real head model based on three dimensional magnetic resonance images. *Rev Sci Instrum* 79:066106.
- Homae F, Watanabe H, Nakano T, Asakawa K, Taga G (2006): The right hemisphere of sleeping infant perceives sentential prosody. *Neurosci Res* 54:276–280.
- Homae F, Watanabe H, Nakano T, Taga G (2007): Prosodic processing in the developing brain. *Neurosci Res* 59:29–39.
- Huppert TJ, Hoge RD, Diamond SG, Franceschini MA, Boas DA (2006): A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. *NeuroImage* 29:368–382.
- Ito Y, Kennan RP, Watanabe E, Koizumi H (2000): Assessment of heating effects in skin during continuous wave near infrared spectroscopy. *J Biomed Opt* 5:383–390.
- Kiguchi M, Ichikawa N, Atsumori H, Kawaguchi F, Sato H, Maki A, Koizumi H (2007): Comparison of light intensity on the brain surface due to laser exposure during optical topography and solar irradiation. *J Biomed Opt* 12:062108.
- Kisilevsky BS, Hains SM, Lee K, Xie X, Huang H, Ye HH, Zhang K, Wang Z (2003): Effects of experience on fetal voice recognition. *Psychol Sci* 14:220–224.
- Kisilevsky BS, Hains SM, Brown CA, Lee CT, Cowperthwaite B, Stutzman SS, Swansburg ML, Lee K, Xie X, Huang H, et al. (2009): Fetal sensitivity to properties of maternal speech and language. *Infant Behav Dev* 32:59–71.
- Koizumi H, Yamamoto T, Maki A, Yamashita Y, Sato H, Kawaguchi H, Ichikawa N (2003): Optical topography: Practical problems and new applications. *Appl Opt* 42:3054–3062.
- Kotilahti K, Nissila I, Huotilainen M, Makela R, Gavrielides N, Noponen T, Bjorkman P, Fellman V, Katila T (2005): Bilateral hemodynamic responses to auditory stimulation in newborn infants. *Neuroreport* 16:1373–1377.
- Kotilahti K, Nissila I, Nasi T, Lipiainen L, Noponen T, Merilainen P, Huotilainen M, Fellman V (2010): Hemodynamic responses to speech and music in newborn infants. *Hum Brain Mapp* 31:595–603.
- Maki A, Yamashita Y, Ito Y, Watanabe E, Mayanagi Y, Koizumi H (1995): Spatial and temporal analysis of human motor activity using noninvasive NIR topography. *Med Phys* 22:1997–2005.
- Mazoyer BM, Tzourio N, Frak V, Syrota A, Murayama N, Levrier O, Salamon G, Dehaene S, Cohen L, Mehler J (1993): The cortical representation of speech. *J Cogn Neurosci* 5:467–479.
- Mehler J, Jusczyk P, Lambertz G, Halsted N, Bertoncini J, Amiel-Tison C (1988): A precursor of language acquisition in young infants. *Cognition* 29:143–178.
- Minagawa-Kawai Y, Mori K, Naoi N, Kojima S (2007): Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *J Neurosci* 27:315–321.
- Minagawa-Kawai Y, Matsuoka S, Dan I, Naoi N, Nakamura K, Kojima S (2009): Prefrontal activation associated with social attachment: Facial-emotion recognition in mothers and infants. *Cereb Cortex* 19:284–292.
- Minagawa-Kawai Y, Mori K, Hebden JC, Dupoux E (2008): Optical imaging of infants' neurocognitive development: Recent advances and perspectives. *Dev Neurobiol* 68:712–728.
- Minagawa-Kawai Y, van der Lely H, Ramus F, Sato Y, Mazuka R, Dupoux E (2010): Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cereb Cortex* 21:254–261.
- Moon C, Cooper RP, Fifer WP (1993): Two-day-olds prefer their native language. *Infant Behav Dev* 16:495–500.
- Nakano T, Watanabe H, Homae F, Taga G (2009): Prefrontal cortical involvement in young infants' analysis of novelty. *Cereb Cortex* 19:455–463.
- Nazzi T, Bertoncini J, Mehler J (1998): Language discrimination by newborns: Toward an understanding of the role of rhythm. *J Exp Psychol Hum Percept Perform* 24:756–766.
- Obrig H, Hirth C, Junge-Hulsing JG, Doge C, Wolf T, Dirnagl U, Villringer A (1996): Cerebral oxygenation changes in response to motor stimulation. *J Appl Physiol* 81:1174–1183.
- Obrig H, Wenzel R, Kohl M, Horst S, Wobst P, Steinbrink J, Thomas F, Villringer A (2000): Near-infrared spectroscopy: Does it function in functional activation studies of the adult brain? *Int J Psychophysiol* 35:125–142.
- Obrig H, Rossi S, Telkemeyer S, Wartenburger I (2010): From acoustic segmentation to language processing: Evidence from optical imaging. *Front Neuroenergetics* 2:13. doi: 10.3389/fnene.2010.00013
- Pena M, Maki A, Kovacic D, Dehaene-Lambertz G, Koizumi H, Bouquet F, Mehler J (2003): Sounds and silence: An optical topography study of language recognition at birth. *Proc Natl Acad Sci USA* 100:11702–11705.
- Perani D, Dehaene S, Grassi F, Cohen L, Cappa SF, Dupoux E, Fazio F, Mehler J (1996): Brain processing of native and foreign languages. *Neuroreport* 7:2439–2444.
- Ramus F, Hauser MD, Miller C, Morris D, Mehler J (2000a): Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science* 288:349–351.
- Ramus F, Nespor M, Mehler J (2000b): Correlates of linguistic rhythm in the speech signal. *Cognition* 75:AD3–AD30.
- Saito Y, Aoyama S, Kondo T, Fukumoto R, Konishi N, Nakamura K, Kobayashi M, Toshima T (2007a): Frontal cerebral blood flow change associated with infant-directed speech. *Arch Dis Child Fetal Neonatal Ed* 92:F113–F116.
- Saito Y, Kondo T, Aoyama S, Fukumoto R, Konishi N, Nakamura K, Kobayashi M, Toshima T (2007b): The function of the frontal lobe in neonates for response to a prosodic voice. *Early Hum Dev* 83:225–230.
- Sato H, Takeuchi T, Sakai KL (1999): Temporal cortex activation during speech recognition: An optical topography study. *Cognition* 73:B55–B66.
- Sato H, Fuchino Y, Kiguchi M, Katura T, Maki A, Yoro T, Koizumi H (2005): Intersubject variability of near-infrared spectroscopy signals during sensorimotor cortex activation. *J Biomed Opt* 10:044001.
- Sato H, Tanaka N, Uchida M, Hirabayashi Y, Katura T, Kanai M, Ashida T, Konishi I, Maki A (2006): Wavelet analysis for detecting body-movement artifacts in optical topography signals. *NeuroImage* 33:580–587.
- Taga G, Asakawa K (2007): Selectivity and localization of cortical response to auditory and visual stimulation in awake infants aged 2 to 4 months. *Neuroimage* 36:1246–1252.
- Taga G, Asakawa K, Hirasawa K, Konishi Y (2003): Hemodynamic responses to visual stimulation in occipital and frontal cortex of newborn infants: A near-infrared optical topography study. *Early Hum Dev* 75 (Suppl):S203–S210.
- Telkemeyer S, Rossi S, Koch SP, Nierhaus T, Steinbrink J, Poeppel D, Obrig H, Wartenburger I (2009): Sensitivity of newborn auditory cortex to the temporal structure of sounds. *J Neurosci* 29:14726–14733.