Behavioral/Systems/Cognitive

Neural Attunement Processes in Infants during the Acquisition of a Language-Specific Phonemic Contrast

Yasuyo Minagawa-Kawai,[1,2,3,4] **Koichi Mori,**[5] **Nozomi Naoi,**[1] **and Shozo Kojima**[1,2]

[1]Department of Psychology, Keio University, Tokyo 108-8345, Japan, [2]Japan Agency of Science and Technology, Saitama 332-0012, Japan, [3]Laboratoire de Sciences Cognitives et Psycholinguistique, Ecole Normale Supérieure, Ecole des Hautes Etudes en Sciences Sociales, and Centre National de la Recherche Scientifique, Paris 75005, France, [4]Centre for Developmental Language Disorders and Cognitive Neuroscience, University College London, London WC1 NPF, United Kingdom, and [5]Department of Rehabilitation for Sensory Functions, Research Institute, National Rehabilitation Center for Persons with Disabilities, Saitama 359-8555, Japan

To elucidate the developmental neural attunement process in the language-specific phonemic repertoire, cerebral hemodynamic responses to a Japanese durational vowel contrast were measured in Japanese infants using near-infrared spectroscopy. Because only relative durational information distinguishes this particular vowel contrast, both first and second language learners have difficulties in acquiring this phonemically crucial durational difference. Previous cross-linguistic studies conducted on adults showed that phoneme-specific, left-dominant neural responses were observed only for native Japanese listeners. Using the same stimuli, we show that a larger response to the across-category changes than to the within-category changes occurred transiently in the 6- to 7-month-old group before stabilizing in the groups older than 12 months. However, the left dominance of the phoneme-specific response in the auditory area was observed only in the groups of 13 months and above. Thus, the durational phonemic contrast is most likely processed first by a generic auditory circuit at 6–7 months as a result of early auditory experience. The neural processing of the contrast is then switched over to a more linguistic circuit after 12 months, this time with a left dominance similar to native adult listeners.

Key words: speech perception; phonemic acquisition; NIRS; cerebral lateralization; auditory area; development

Introduction

In contrast to their immature visual system, young infants possess a relatively mature auditory system that enables them to detect subtle acoustic–physical differences in sounds (Alho et al., 1990; Huotilainen et al., 2003). In addition to simple sound changes like pitch and intensity, they can discriminate a wide variety of speech sounds (phonemes) used in languages, including those that they have never heard (Kuhl et al., 1992; Jusczyk, 1997). However, this initial ability to discriminate many phonemic units gradually declines as infants are exposed to the phonological system of their ambient language. In other words, an infant's perceptual ability to discriminate native language phonemes becomes so attuned that it becomes language specific. Infants in an English-speaking environment, for instance, can distinguish between Hindi and Salish phonemes at an early stage but lose this ability after a year of exposure to English only (Werker et al., 1981; Werker and Tees, 1984). Since the pioneering work of Eimas et al. (1971), it has been demonstrated that infants become very sensitive to the vowel contrasts of their na-

tive language by ∼6 months of age (Kuhl et al., 1992; Polka and Werker, 1994), and to their native consonant contrasts by ∼10–12 months (Werker and Tees, 1984; Best and McRoberts, 2003).

A question then arises with regard to the neurophysiological mechanism underlying this perceptual change in the first year of life. This issue has been examined by studies on event-related potential (ERP), using mismatch negativities (MMNs) as indicators of language-specific auditory responses (Cheour et al., 1998; Dehaene-Lambertz and Baillet, 1998; Dehaene-Lambertz and Gliga, 2004). Cheour et al. (1998) measured the MMNs of infants at ages of six and 12 months. The results revealed that the 12-month-olds had a weaker response to the non-native vowel contrast than the 6-month-olds. These results provide insights into the different networks involved in various developmental stages. However, there exists a certain issue that neither behavioral nor electrophysiological studies could resolve completely, a developmental process of hemispheric specialization. Many ERP studies as well as dichotic listening studies have attempted to explore this issue; however, their results are diverse (Novak et al., 1989; Simos et al., 1997). The issue remains controversial.

A recent advent, multichannel near-infrared spectroscopy (NIRS), can probably shed light on this issue: NIRS may provide further and additional evidence of the developmental process of left-hemispheric dominance, because it can measure infants non-invasively with a spatial resolution of 2–3 cm (Chance et al., 1993; Villringer et al., 1993). Consequently, this study used NIRS to
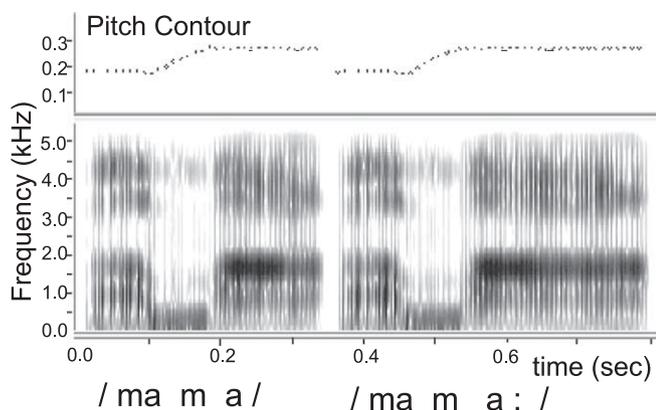
Figure 1. Spectrograms and pitch contours of the stimuli /mama/ and /mama:/ differing in final vowel durations.

assess the neural correlates of the dramatic perceptual change within the first few years of life. Because a previous cross-linguistic adult study (Minagawa-Kawai et al., 2005) revealed that only native Japanese listeners exhibited left-dominant responses to the phonemic-length change, we used the same Japanese vowel stimuli with identical physical differences but with different linguistic information, to examine at what age (1) the specific response to across-category difference is observed and (2) the left-dominant response to phonemic differences appears in the infant's development.

Materials and Methods

Participants. Monolingual infants brought up in Tokyo, Japan, and its suburbs were recruited as paid participants. They were aged from 3 to 4 months ($n = 23$), 6–7 months ($n = 24$), 10–11 months ($n = 28$), 13–14 months ($n = 28$), and 25–28 months ($n = 24$). They were chosen based on the criteria of normal hearing and full-term birth with no history of serious diseases. Age groups were determined from the previous behavioral literature on phonemic acquisition: one group before 6 months, which is a native-vowel acquisition boundary, and two groups after 6 months, when various types of phonological learning occur. We also recruited a 1-year-old group (13–14 months) and 2-year-old group (25–28 months). Handedness of the infants, if applicable, was assessed by the Edinburgh Handedness Inventory; two participants who appeared to be left-handed were excluded. Among the infants considered, 70 participants were excluded from the final data set because of insufficient trials attributable to motion artifacts (49) and excessive fussiness (6), loose probe placement (11), parental interference (2), and experimenter error (2). The final data set included infants aged from 3 to 4 months ($n = 15$; 9 girls and 6 boys), 6–7 months ($n = 14$; 7 girls and 7 boys), 10–11 months ($n = 11$; 5 girls and 6 boys), 13–15 months ($n = 9$; 6 girls and 3 boys), and 25–28 months ($n = 8$; 5 girls and 3 boys). Parents gave informed consent in compliance with a protocol approved by the ethic committee of Keio University (no. 04001) and Research Institute, National Rehabilitation Center for Persons with Disabilities (NRCD).

Stimuli. The stimuli used were four pseudowords from the /mama/-/mama:/ continuum in which the final vowel varied in duration from 151 to 250 ms in steps of 33 ms (stimulus A, 151 ms; B, 184 ms; C, 217 ms; D, 250 ms). The choice of durational spacing was based on our pilot study using 33, 66, and 99 ms steps, among which the latter two evoked relatively large responses even under phonemically noncontrastive conditions and posed a risk of masking responses to phonemic distinctiveness by saturation. Using PARCOR (partial autocorrelation), an analysis–synthesis procedure (Markel and Gray, 1976), these stimuli were resynthesized from the data including pitch and formant information analyzed from a natural spoken word (unaccented pattern) to have a stable pitch contour and formant structure in the second vowel (Fig. 1). The length of the first syllable was 110 ms and the intervocalic /m/ was 90 ms.

The stimulus context in our study is limited to two to three morae. However, the sensitivities to BC, which may have been judged from the relative difference against the preceding syllable duration (Fujisaki et al., 1985), are considered to show the perceptual ability of quantity acquired through linguistic experience of Japanese moraic rhythm.

Three sessions were prepared using two adjacent stimuli pairs from the four. In session AB, stimulus A was repeated every 1.25 s for 20 s as a baseline habituation block; furthermore, stimuli A and B were presented in a pseudorandom order with equal probabilities every 1.25 s for another 20 s period as a target block, which would elicit a graded dishabituation response according to the degree of linguistic distinctiveness between the paired stimuli (Dehaene-Lambertz, 1997; Näätänen et al. 1997). These two blocks were alternated and sequentially repeated for a minimum of five times. The same procedures were performed for sessions BC and CD.

Experimental design. A previous experiment (Minagawa-Kawai et al., 2005) observed that the phonemic boundary of long and short vowels for the native Japanese speakers lay between stimuli B and C. Consequently, compared with those of other sessions, the speakers' cortical responses to stimulus C (dishabituation stimulus) after listening to repetitions of B (habituation) were greater in the left temporal area. In contrast, non-native listeners did not display such phoneme-specific responses, because for them long and short vowel contrasts did not possess phonological (linguistic) value (Minagawa-Kawai et al., 2005). The following experiment was designed for the infants for whom short experimental periods are preferred. Two sessions were conducted for each subject: (1) one was the across-category condition (session BC) in which the stimuli differed linguistically, and (2) and the other within-category condition (either session AB or CD) in which the differences were not linguistic but evoked small comparable responses (Minagawa-Kawai et al., 2002). The number of AB and CD sessions, as well as the order of the two conditions, was counterbalanced within each age group. However, because the brain index is sometimes more sensitive than the behavioral measures (de Haan and Nelson, 1997), we cannot be sure whether significant NIRS responses in each condition mean that the infant is able to discriminate the stimuli behaviorally.

NIRS recording. The changes in hemoglobin (Hb) concentrations and their oxygenation levels in the bilateral temporal areas were recorded using NIRS systems (ETG-100 or ETG-7000; Hitachi, Tokyo, Japan), which emit continuous near-infrared lasers with fixed wavelengths of ~780 and 830 nm. The laser beams were modulated at different frequencies and detected using lock-in amplifiers. These devices can measure localized cortical responses of channels that are present in the optical path of the brain between the nearest pairs of emission and detection probes; these probes were separated by 3 cm on the scalp surface (Fukui et al., 2003).

Recently, some critical issues in practical use of NIRS in baby studies have been raised (Aslin and Mehler, 2005), which include choice of probe spacing and wavelengths, and possible systemic effects on NIRS data. The separation of 30 mm between the source and the detector (SD) was adopted in our infant study as in the previous adult study (Minagawa-Kawai et al., 2005) because of the following considerations. First, judging from the results of the light propagation prediction in neonatal head model using SD of 30 mm (Fukui et al., 2003), the infrared light penetrates approximately down to 24 mm in depth with the sensitivity of 0.1%, which goes deeper than into the adult brain (Villringer et al., 1997), presumably because of less myelinated, less reflective white matter (Fukui et al., 2003). Although this depth might be an overestimation for the age of our infants (3 months and older), the light irradiated from one temporal side would never reach the contralateral hemisphere, because their head diameters were 120 mm or larger. Our results thus reflect the auditory evoked responses strictly in the ipsilateral hemisphere to each recording side, which is essential to the validity of our lateralization assessment. Second, the DPF (differential pathlength factor) in adults is reported to be larger than that of infants (5.73 vs 4.69–4.77 at 832 nm for our subjects) (Duncan et al. 1996), suggesting that measured brain volume in our study was possibly smaller in infants than in adults, which would at least partially counter the effect of the thinner skull and scalp, and smaller brain size of infants than adults. The relative head size dif-
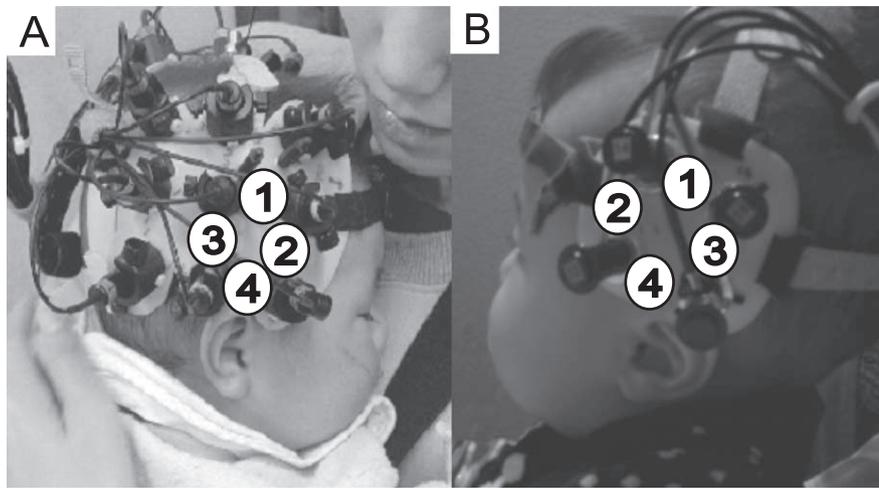
Figure 2. Location of the NIRS probes and channels. Five emission and four detection probes arranged in a 3 × 3 square lattice (**A**) and two emission and two detection probes in a 2 × 2 square lattice (**B**) were fitted on each lateral side of the head.

ference to SD across the age groups still cannot explain the transient appearance and disappearance of the difference between the across- and within-category responses before 1 year of age (see Results). The choice of the laser wavelengths of 780 and 830 nm was based on the similarity in Hb absorbance of the wavelengths across the isoabsorbance ($\sim$800 nm) of oxy- and deoxy-Hb, which should minimize the difference in the regions traveled by the two measuring wavelengths.

Because NIRS only measures Hb in capillaries, venules, and arterioles (Yamamoto and Kato, 2002) <1 mm in diameter (Liu et al. 1995), unlike functional magnetic resonance imaging, NIRS is less affected by systemic circulatory changes. Hemodynamics as measured from the scalp surface, however, does include not only the specific cerebral responses to the stimuli (presumably neural component) but also other unrelated brain activities and systemic vascular effects (Katura et al., 2006) in and outside the brain. The latter components are most likely less localized and less time-locked to the stimuli than the former, as shown in Figure 3. In the current study, global Hb changes across recording channels that were not closely related to the stimulus cycle were labeled as artifact, and the block including the artifact was removed from averaging, although the majority of artifacts were motion related. The arrangement of the bilateral channels in symmetric positions should have also helped reduce any systemic effects entering the functional lateralization analysis.

One of two methods was used for positioning the NIRS probes: (1) five emission and four detection probes arranged in a 3 × 3 square lattice (12 channels on either side) (Fig. 2A) and (2) two emission and two detection probes in a 2 × 2 square lattice (four channels on either side) (Fig. 2B) were fitted on each lateral side of the head. Because we decided to analyze only specific channels in the auditory area (one channel per side), after the completion of more than one-half of the study, we switched to using the 2 × 2 lattice to reduce the number of experimental failures attributable to the resistance of the infant to the fitting of too many probes. The international 10–20 system was referred to when attaching these probes (i.e., the line connecting T3, F7, F8, and T4 was horizontal to the lowest lines of the NIRS probes). Furthermore, in the case of the 3 × 3 configuration, the middle probe (posterior probe for the 2 × 2 configuration) in the lowest line corresponded to T3 on the left and T4 on the right (Fig. 2). In previous NIRS studies conducted on adults, we determined that the channels close to the lateral end of the border between the transverse temporal gyrus and the planum temporale, as projected on a parasagittal magnetic resonance imaging, were the "center" of the auditory area (Furuya and Mori, 2003; Minagawa-Kawai et al., 2005). These channels corresponded to either one of the channels 1, 2, 3, or 4 on the left temporal area and their symmetric positions on the right (Fig. 2A). Based on these previous results, as well as the data from the three-dimensional probabilistic anatomical cranio-cerebral correlation (in accordance with the international 10–20 system) (Okamoto et al., 2004),

we presumed that channels 1, 2, 3, 4, and their symmetric positions on the right approximately covered the auditory area.

Experimental procedures. After positioning the probes, participants passively listened to the stimuli presented via a loudspeaker ($\sim$70 dB sound pressure level) in a sound-attenuated room. The order of the two sessions was counterbalanced for all the infants. The infants were seated on their parents' lap while listening to the stimuli. During the experiment, the experimenter entertained them with silent toys to reduce their body movements.

We used two systems for measuring NIRS: ETG-100 in NRCD and ETG-7000 in Keio University. Although they are almost identical in terms of the ranges of the optical power (1.5–2.0 mW) and wavelengths (ETG-100, 780 $\pm$ 15, 830 $\pm$ 15 nm; ETG-7000 Probe 1 and 2, 781 $\pm$ 1, 827 $\pm$ 1 nm) of the infrared lasers used, the data obtained from the two machines in respective institutions were examined for any biases. Of the final dataset, 51% were from ETG-100, and this ratio was approximately constant in all of the age groups (3–4 months, 47%; 6–7 months, 50%; 10–11 months, 55%; 13–14 months, 56%; 25–28 months, 50%). The data of total Hb changes were analyzed using ANOVA with factors of the machine types and stimulus conditions in each age group and in each of the recording sides (left or right temporal area), but the results showed no significant main effects or interactions in any group [main effect of machine type: $p > 0.05$ for each result; 3–4 months of age, left (L), $F_{(1,26)} = 0.001$; right (R), $F_{(1,26)} = 0.001$; 6–7 months of age, L, $F_{(1,24)} = 0.56$; R, $F_{(1,24)} = 0.08$; 10–11 months of age, L, $F_{(1,18)} = 1.70$; R, $F_{(1,18)} = 2.49$; 13–14 months of age, L, $F_{(1,14)} = 0.93$; R, $F_{(1,14)} = 2.56$; 25–26 months of age, L, $F_{(1,12)} = 1.32$; R, $F_{(1,12)} = 1.79$], indicating that there was no effect on the data attributable to the differences in the experimental environment. A similar analysis on the factor of probe settings revealed no significant effect either [main effect of probe type: $p > 0.05$ for each result; 3–4 months of age, L, $F_{(1,26)} = 0.03$; R, $F_{(1,26)} = 0.18$; 6–7 months of age, L, $F_{(1,24)} = 1.49$; R, $F_{(1,24)} = 0.15$; 10–11 months of age, L, $F_{(1,18)} = 2.60$; R, $F_{(1,18)} = 3.61$; 13–14 months of age, L, $F_{(1,14)} = 0.93$; R, $F_{(1,14)} = 2.56$; 25–28 months of age, L, $F_{(1,12)} = 1.32$; R, $F_{(1,12)} = 1.79$]. Therefore, the rest of the data analysis was conducted on the pooled data from the two machines.

Data analysis. The concentrations of oxygenated (oxy-), deoxygenated (deoxy-) Hb, and the total Hb were calculated from the absorbance changes of 780 and 830 nm laser beams sampled at 10 Hz. After discarding the blocks with artifacts, the Hb concentrations of the remaining blocks were averaged synchronously to the onset of the target blocks and smoothed with a 1 s moving average. The response peaks of the averaged target blocks were evaluated against the 10 s baseline period just before the target block in each channel. NIRS with two infrared wavelengths gives estimation of oxy-, deoxy-, and total Hb concentrations in the tissue, among which total Hb was chosen as the indicator for the local cerebral activation in this study. The reasoning for the choice was that it reflects local cerebral blood volume and is more correlated to a regional cerebral blood flow (CBF) as measured in positron emission tomography than the other parameters (Villringer et al. 1997). Total Hb is also significantly correlated to blood oxygenation level-dependent (BOLD) signals as well as oxy-Hb (Strangman et al., 2002), although BOLD and oxy-Hb exhibit more nonlinearity to sustained stimulation than CBF or total Hb (Sadato and Toyada, 2006). However, suitability of total Hb as a parameter may depend on various factors such as the cortex studied, age, task demand, and individual vascular responsiveness, because these factors influence hemodynamics of deoxy-Hb (Meek et al., 1998; Sakatani et al., 1999; Hoshi et al., 2000; Zaramella et al., 2001). Total Hb as a functional index had given consistent results in previous studies for assessing the lateralization of language function (Watanabe et al., 1998; Furuya and Mori, 2003; Sato et al., 2003) and is regarded to be suitable for our study.

The changes in deoxy-Hb observed in the present study were too small relative to the noise level to be used as a reliable indicator of evoked responses, as shown for representative participants in Figure 3. Total Hb, in contrast, should provide, in many cases, a better signal-to-noise ratio than oxy- or deoxy-Hb because the calculation of total Hb is additive of the absorbance measurements at two wavelengths rather than subtractive as with oxy- and deoxy-Hb. To choose one of the auditory channels for the statistical analysis, we first averaged the absolute total Hb change across the stimulus conditions and then selected one of the channels that exhibited the maximum value on each side. To determine the significance of the response in different conditions, different sides and age groups, the response peaks during the target block were compared with 0 (the average of the 10 s pretarget baseline). A response was deemed significant with t test at $p < 0.05$ (uncorrected for multiple comparisons). To compare the Hb changes between the across- and within-category conditions, a paired t test was conducted for each age group. We applied the Holm corrections to consider the effect of multiple comparisons.

Results

Automatic change-detection responses to the contrast stimuli against the baseline stimuli with or without phonemic differences were recorded in the infants in five different age groups. Although there were variations in their amplitudes, younger infants exhibited hemodynamic responses to the target stimuli in almost every one of the 12 recording sites on either temporal side (Fig. 3A). This is in contrast to the response pattern of the adults (Minagawa-Kawai et al., 2002), in which the evoked responses were found in limited channels in the auditory area (Fig. 3C). Furthermore, infants of 25–28 months exhibited evoked responses in rather limited areas, similar to the adults (Fig. 3B). The following analyses used the results from one channel of the maximum response in the auditory area on each side.

The dishabituation stimuli, regardless of across or within categories, elicited significant responses in all of the age groups [zero test, 3–4 months of age: df = 14; across category (AC), L, $t = 7.13$; R, $t = 7.23$; within category (WC), L, $t = 5.76$; R, $t = 4.79$; 6–7 months of age: df = 13; AC, L, $t = 4.57$; R, $t = 7.65$; WC, L, $t = 5.75$; R, $t = 5.17$; 10–11 months of age: df = 10; AC, L, $t = 3.12$; R, $t = 3.55$; WC, L, $t = 2.19$; R, $t = 2.21$; 13–14 months of age: df = 8; AC, L, $t = 4.36$; R, $t = 5.1$; WC, L, $t = 3.84$; R, $t = 4.18$; 25–28 months of age: df = 7; AC, L, $t = 6.01$; R, $t = 3.45$; WC, L, $t = 6.07$; R, $t = 1.95$]. The 3- to 4-month-old infants exhibited almost the same response amplitudes to the across- and within-category phonemic changes, revealing sensitivity to the physical difference but no specificity to the phonemic durational contrast ($t = 0.66$; $p > 0.5$) (Fig. 4). However, the 6- to 7-month-old infants exhibited significantly larger responses to the across-category change than to the within-category change ($t = 2.93$; $p < 0.05$). In the 10- to 11-month-old infants, no significant differences were observed again in the response amplitudes between the two conditions ($t = 0.05$; $p > 0.10$), whereas both of the older groups aged from 13 to 14 months and 25 to 28 months showed phoneme-specific responses (i.e., significantly larger responses to the across-category stimulus than to the within category stimulus) ($t = 2.9$, $p < 0.05$; $t = 3.44$, $p < 0.05$).
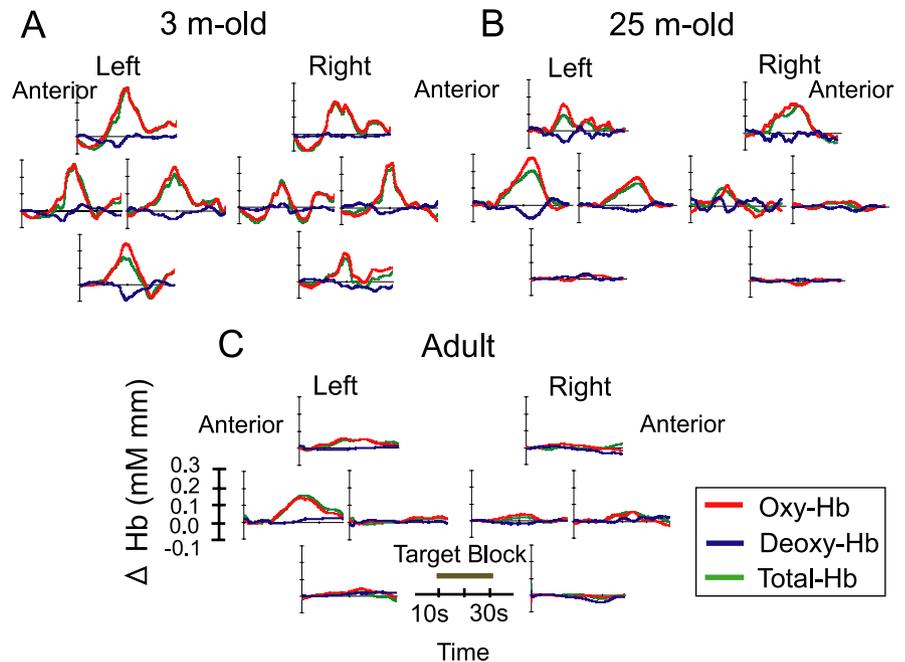

Figure 3. Time course of NIRS responses for eight channels around the temporal areas (left and right) in subjects of different age groups: 3-month-old infant (**A**), 25-month-old infant (**B**), and adult data adapted from Minagawa-Kawai et al. (2002) (**C**). The younger infant showed bilateral and broad activations compared with the adult's pattern with left dominance.
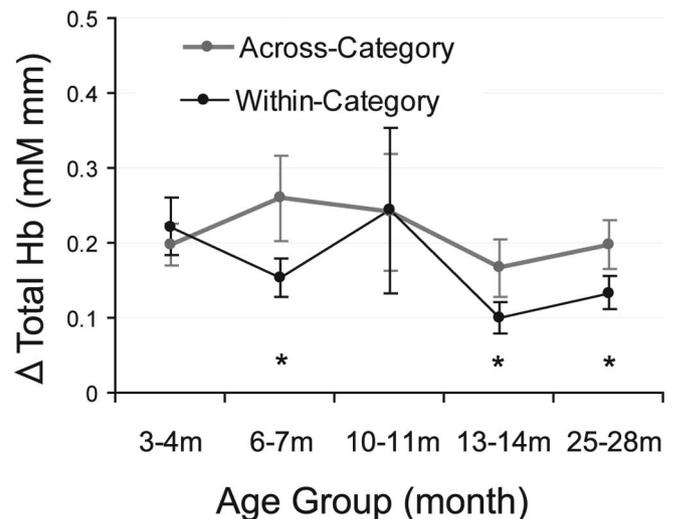

Figure 4. Averaged values of the total Hb changes in the left auditory area (L) in response to the across- and within-phonemic changes of the stimuli. *$p < 0.05$ for the difference between the responses in the across- and within-sessions.

A laterality index was calculated using the formula $(L - R)/(L + R)$, where L and R are the maximal total Hb changes in the left and right auditory channels, respectively. The developmental change in the laterality index (Fig. 5) shows neither leftward nor rightward lateralization in any condition for the three age groups from 3 to 11 months. However, the two groups older than 12 months exhibited a significant left lateralization only for the across-category condition (Mann–Whitney U test: 13–14 months, $p = 0.046$; 25–28 months, $p = 0.036$). Furthermore, the zero test showed that only the across-category conditions in these groups had a laterality index that was significantly greater than zero (13–14 months, $p = 0.003$; 25–28 months, $p = 0.0003$).
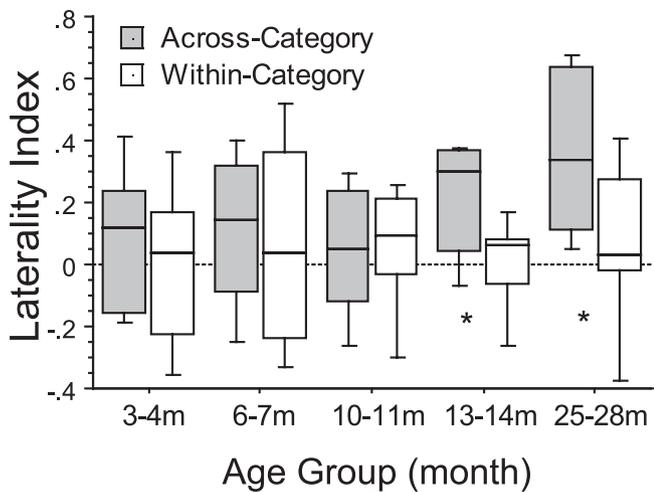
Figure 5. Laterality index in different age groups. Left dominance is observed in the older age groups. *$p < 0.05$. Boxes, The quartiles; bars in the box, the medians; hinges, the ranges.

Discussion

Using phonemic stimulus pairs that differ equally in physical properties but not in linguistic information, this study examined infants' developmental changes and their neural responses to stimulus changes. If the phoneme-specific cerebral response is defined as a stronger response to a phonemic difference than to a nonphonemic difference with the same amount of physical distance, the youngest group that exhibited a phoneme-specific response to the long/short vowel contrast was the 6- to 7-month-old group. However, at this age, the cerebral response was not left dominant, and its phoneme specificity disappeared at 10–11 months. The phonemic response did not become consistently left lateralized until after 12 months, similar to native adults (Minagawa-Kawai et al., 2002; Jacquemot et al. 2003).

As previous NIRS studies have successfully detected cortical activities related to the phonemic or prosodic changes in adults and infants (Sato et al., 1999; Minagawa-Kawai et al., 2002; Furuya and Mori, 2003; Pena et al., 2003; Homae et al., 2006), the present study captured a developmental process in which the neural activations of Japanese infants changed from a nonspecific to a language-specific pattern by assessing the differential cerebral activation and its lateralization in the temporal area.

Phoneme-specific responses

It is shown that the neural responses of the infants to the non-native phonemic contrast that exists at 6–7 months disappear at 11–12 months for vowels (Cheour et al., 1998) and consonants (Rivera-Gaxiola et al., 2005). This agrees well with the behavioral acquisition of native vowels and consonants within the first year of life (Kuhl et al., 1992). Our results reveal a similar tendency: infants of approximately the same age begin to exhibit higher sensitivity toward the native contrast. However, in detail, there are some aspects that differ from those mentioned in previous studies (i.e., we observe larger responses to the across-category contrast even in the 6- to 7-month-olds, whereas our youngest infant's data did not exhibit the phoneme-specific responses). This is consistent with the behavioral data of Japanese quantity acquisition (Bertoncini et al., 1995; Hayashi 2003) but different from the ERP responses to consonantal differences at the same age (Dehaene-Lambertz and Baillet, 1998). The most plausible reason for the difference is the phonemes we targeted; this is because the Japanese long/short vowel contrast (not realized by

spectral variation but solely by durational differences) is the most difficult one for even the Japanese children to grasp (Amano, 1986). Because vowel durations are prolonged or shortened depending on the word context, finding a phonemically distinctive durational difference, which is sometimes only a slight difference, may be difficult.

The temporal disappearance of phoneme-specific responses in the group of 10- to 11-month-olds is slightly counterintuitive. However, speech perception studies have reported similar developmental changes (Pickens et al., 1994; Stager and Werker, 1997; Hayashi et al., 2001), although the stimulus contexts were different. Hayashi et al. (2001) identified a developmental stage at 7–10 months when the preference for infant-directed speech decreases temporarily. The corresponding transient insensitivity to the phonemic length contrast observed in the current study may be a neural correlate of one step in the perceptual reorganization, because the phoneme-specific response reappears, this time, with a leftward dominance at 13–14 months. The infants at 6–7 months might differentiate between the long/short contrast using relatively extensive areas in the bilateral temporal lobes (Fig. 3). However, this may not be through means of a linguistic network, because it had no left dominance, but by means of a network associated with regulations of general temporal information (Kanoh et al., 2004; van Zuijen et al., 2004). It is possible that infants reorganize their brain strategy to process the contrast accurately and efficiently, based on additional language input. After the reorganization process, postulated to occur at 10–11 months, they reestablish their linguistic network with a renewed phoneme-specific subprocessor. The exact age and the progression of the transition, however, may not be uniform across infants, and the perceptual ability or related brain index may even become unstable during the period, which may have contributed to the high SD (Fig. 4) of the total Hb response. As the strategy of infants progresses from general acoustic processing to language specific processing, the neural recruitments shift from bilateral to more unilateral, and are then limited to the Wernicke's area, as observed in the Japanese adults (Minagawa-Kawai et al., 2002; Jacquemot et al., 2003).

Left-dominant responses

Neural responses to the phonemes of a native language have been revealed to be left dominant by electrophysiological measures (Näätänen et al., 1997; Breier et al., 1999) as well as hemodynamic measures (Tervaniemi et al., 2000; Furuya and Mori, 2003; Zevin and McCandliss, 2005). It has been debated whether physical or linguistic properties of the same stimuli contribute more to the functional lateralization of the brain (Gandour et al., 2004; Shtyrov et al., 2005). According to the physical factor hypothesis, slow acoustic transitions, such as pitch change, are preferentially processed in the right hemisphere; rapidly changing sounds, like consonants, in the left (Tallal et al., 1993; Fitch et al., 1997). However, sufficient evidence now exists suggesting that language-specific factors are at least as influential on the hemispheric specialization (Best and Avery, 1999; Gandour et al., 2002, 2004). Tonal languages such as Thai and Chinese employ patterns of pitch changes as one of the lexical determinant factors. The lexical pitch changes predominantly activate the left hemisphere only in native Chinese speakers (Gandour et al., 2004). This finding is also supported by neuronal evidence on second language learning. The left-dominant responses to listening sentences or words are only observed for native listeners in contrast to the bilateral or inconsistent activations for non-native listeners (Bottini et al., 1994; Perani et al., 1996; Dehaene et al., 1997;

Gandour et al., 2004). The leftward lateralization observed in our study also adds another piece of evidence to the linguistic hypothesis. In this study, it was revealed that the left-dominant language-specific neural circuits were developed as a result of the language experience, despite the fact that our stimuli lacked rapidly changing spectral cues, which would presumably have elicited no left-dominant responses in another hypothesis. The stimuli used in the present study are the same ones that did not elicit left-dominant responses in non-native listeners (Minagawa-Kawai et al., 2005).

The age at which leftward lateralization for language processing occurs is a controversial question. Studies by Dehaene-Lambertz and Baillet (1998) and Dehaene-Lambertz (2000) reported a larger response in the left hemisphere to the difference between /ba/ and /da/ in 3- to 4-month-old infants, whereas others obtained different results, either with right dominance (Novak et al., 1989; Simos et al., 1997) or with no asymmetry (Molfese and Hess, 1978; Simos and Molfese, 1997) for infants and children. In contrast, in the present study, a significant leftward lateralization of phoneme-specific responses was observed only in infants >1 year of age. Although the discrepancy between the results of the current study and the ERP studies may be accounted for by the differences in the stimuli, the phonemic type, and their presentation, it is also likely to be attributable to the difference of the stimulus form (syllable vs word). Almost all of the previous ERP studies with infants used isolated syllables differing in consonants or vowels (e.g., /ba/ vs /da/) in contrast to the use of word form stimuli in our study. Because linguistic forms of stimuli crucially influence the lateralization of brain responses (Shtyrov et al., 2005), the neural responses to isolated syllables might be influenced more by the physical features of the stimuli than by their linguistic features.

Recent behavioral findings and theories have shown that various developmental processes are involved in phonemic acquisition before they are used for word recognition (Werker and Curtin, 2005). Maye et al. (2002), for instance, reported that 6- to 8-month-olds are very sensitive to the statistical probabilities of acoustic inputs, whereas infants after 9 months no longer are (Kuhl et al., 2003; Singh et al., 2004). In contrast, infants >9 months of age are influenced by other cues including social cues (Kuhl et al., 2003). These results suggest that qualitatively different phonemic perceptual abilities exist in early and later infancy of the first year. Our findings confirm this view and provide neurophysiological evidence for it: the left-dominant responses observed in 1-year-olds may include an initial process that connects to higher cognitive networks. However, these established frameworks of phonemic acquisition are mostly based on behavioral studies of segmental discrimination, and currently we have only limited information on quantity discrimination by Japanese infants. Thus, diverse developmental models may still be conceivable and the behavioral development of quantity acquisition at different ages is the crucial topic yet to be studied further.

Our current study, together with the existing evidence, suggests that, although the developmental timing when certain acoustic properties of phonemes become discriminable at the lower auditory level may depend on the acoustic salience (e.g., segmental vs quantity), there is a point when those acoustic properties become phonemically represented and integrated into the phonological framework (e.g., syllable structure, phonotactics) of one's language. This point is a necessary and common gateway through which acquisition of every phoneme passes sooner or later, and the transition to the functional leftward asymmetry

observed here should be a plausible correlate of the neural modulation associated with this crucial process.

References

Alho K, Sainio K, Sajaniemi N, Reinikainen K, Näätänen R (1990) Event-related brain potential of human newborns to pitch change of an acoustic stimulus. Electroencephalogr Clin Neurophysiol 77:151–155.

Amano K (1986) Acquisition of phonemic analysis and literacy in children. Ann Rep Ed Psychol Jpn 27:142–164.

Aslin R, Mehler J (2005) Near-infrared spectroscopy for functional studies of brain activity in human infants: promise, prospects, and challenges. J Biomed Opt 10:011009.

Bertoncini J, Floccia C, Nazzi T, Mehler J (1995) Morae and syllables: rhythmical basis of speech representations in neonates. Lang Speech 38:311–329.

Best CC, McRoberts GW (2003) Infant perception of non-native consonant contrasts that adults assimilate in different ways. Lang Speech 46:183–216.

Best CT, Avery RA (1999) Left-hemisphere advantage for click consonants is determined by linguistic significance and experience. Psychol Sci 10:65–70.

Bottini G, Corcoran R, Sterzi R, Paulesu E, Schenone P, Scarpa P, Frackowiak RS, Frith CD (1994) The role of the right hemisphere in the interpretation of figurative aspects of language. A positron emission tomography activation study. Brain 117:1241–1253.

Breier JI, Simos PG, Zouridakis G, Papanicolaou AC (1999) Temporal course of regional brain activation associated with phonological decoding. J Clin Exp Neuropsychol 21:465–476.

Chance B, Zhuang Z, UnAh CC, Alter C, Lipton L (1993) Cognition-activated low-frequency modulation of light absorption in human brain. Proc Natl Acad Sci USA 90:3770–3774.

Cheour M, Ceponiene R, Lehtokovski A, Luuk A, Allik J, Alho K, Näätänen R (1998) Development of language-specific phoneme representations in the infant brain. Nat Neurosci 1:351–353.

de Haan M, Nelson CA (1997) Recognition of the mother's face by six-month-old infants: a neurobehavioral study. Child Dev 68:187–210.

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.

Dehaene-Lambertz G (1997) Electrophysiological correlates of categorical phoneme perception in adults. NeuroReport 8:919–924.

Dehaene-Lambertz G (2000) Cerebral specialization for speech and non-speech stimuli in infants. J Cogn Neurosci 12:449–460.

Dehaene-Lambertz G, Baillet S (1998) A phonological representation in the infant brain. NeuroReport 9:1885–1888.

Dehaene-Lambertz G, Gliga T (2004) Common neural basis for phoneme processing in infants and adults. J Cogn Neurosci 16:1375–1387.

Duncan A, Meek J, Clemence M, Elwell C, Fallon P, Tyszczuk L, Cope M, Delpy DT (1996) Measurement of cranial optical path length as a function of age using phase resolved near infrared spectroscopy. Pediatr Res 39:889–894.

Eimas PD, Siqueland ER, Jusczyk P, Vigorito J (1971) Speech perception in infants. Science 171:303–306.

Fitch RH, Miller S, Tallal P (1997) Neurobiology of speech perception. Annu Rev Neurosci 20:331–353.

Fujisaki H, Nakamura K, Imoto T (1985) Auditory perception of speech and non-speech stimuli. In: Auditory analysis and perception of speech (Fant G, Tatham MAA, eds), pp 197–219. London: Academic.

Fukui Y, Ajichi Y, Okada E (2003) Monte Carlo prediction of near-infrared light propagation in realistic adult and neonatal head models. Appl Opt 42:2881–2887.

Furuya I, Mori K (2003) Cerebral lateralization in spoken language processing measured by multi-channel near-infrared spectroscopy (NIRS). Brain Nerve 55:226–231.

Gandour J, Wong D, Lowe M, Dzemidzic M, Satthamnuwong N, Long Y, Lurito J (2002) Neural circuitry underlying perception of duration depends on language experience. Brain Lang 83:268–290.

Gandour J, Tong Y, Wong D, Talavage T, Dzemidzic M, Xu Y, Li X, Lowe M (2004) Hemispheric roles in the perception of speech prosody. NeuroImage 23:344–357.

Hayashi A (2003) Perception and acquisition of rhythmic units by infants. J Phonetic Soc Japan 7:29–34.

Hayashi A, Tamekawa Y, Kiritani S (2001) Developmental change in auditory preferences for speech stimuli in Japanese infants. J Speech Lang Hear Res 44:1189–1200.

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.

Hoshi Y, Kohri S, Matsumoto Y, Cho K, Matsuda T, Okajima S, Fujimoto S (2000) Hemodynamic responses to photic stimulation in neonates. Pediatr Neurol 23:323–327.

Huotilainen M, Kujala A, Hotakainen M, Shestakova A, Kushnerenko E, Parkkonen L, Fellman V, Näätänen R (2003) Auditory magnetic responses of healthy newborns. NeuroReport 14:1871–1875.

Jacquemot C, Pallier C, LeBihan D, Dehaene S, Dupoux E (2003) Phonological grammar shapes the auditory cortex: a functional magnetic resonance imaging study. J Neurosci 23:9541–9546.

Jusczyk PW (1997) The discovery of spoken language. Cambridge, MA: MIT.

Kanoh S, Futami R, Hoshimiya N (2004) Sequential grouping of tone sequence as reflected by the mismatch negativity. Biol Cybern 91:388–395.

Katura T, Tanaka N, Obata A, Sato H, Maki A (2006) Quantitative evaluation of interrelations between spontaneous low-frequency oscillations in cerebral hemodynamics and systemic cardiovascular dynamics. NeuroImage 31:1592–1600.

Kuhl PK, Williams KA, Lacerda F, Stevens KN, Lindblom B (1992) Linguistic experience alters phonetic perception in infants by 6 months of age. Science 255:606–608.

Kuhl PK, Tsao FM, Liu HM (2003) Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. Proc Natl Acad Sci USA 100:9096–9101.

Liu H, Chance B, Hielscher AH, Jaques SL, Tittel FK (1995) Influence of blood vessels on the measurement of hemoglobin oxygenation as determined by time-resolved reflectance spectroscopy. Med Phys 22:1209–1217.

Markel JD, Gray Jr AH (1976) Linear prediction of speech. Berlin: Springer.

Maye J, Werker JF, Gerken L (2002) Infant sensitivity to distributional information can affect phonetic discrimination. Cognition 82:B101–B111.

Meek JH, Firbank M, Elwell CE, Atkinson J, Braddick O, Wyatt JS (1998) Regional hemodynamic responses to visual stimulation in awake infants. Pediatr Res 43:840–843.

Minagawa-Kawai Y, Mori K, Furuya I, Hayashi R, Sato Y (2002) Assessing cerebral representations of short and long vowel categories by NIRS. NeuroReport 13:581–584.

Minagawa-Kawai Y, Mori K, Sato Y (2005) Different brain strategies underlie the categorical perception of foreign and native phonemes. J Cogn Neurosci 17:1376–1385.

Molfese DL, Hess TM (1978) Hemispheric specialization for VOT perception in the preschool child. J Exp Child Psychol 26:71–84.

Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A, Vainio M, Alku P, Ilmoniemi RJ, Luu A, Allik J, Sinkkonen J, Alho K (1997) Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature 85:432–434.

Novak G, Kurtzberg PD, Kreuzer JA, Vaughan Jr HG (1989) Cortical responses to speech sounds and their formants in normal infants: maturational sequence and spatiotemporal analysis. Clin Electroencephalogr Clin Neurophysiol 73:295–305.

Okamoto M, Dan H, Sakamoto K, Takeo K, Shimizu K, Kohno S, Oda I, Isobe S, Suzuki T, Kohyama K, Dan I (2004) Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10–20 system oriented for transcranial functional brain mapping. NeuroImage 21:99–111.

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.

Pickens J, Field T, Nawrocki T, Martinez A, Soutullo D, Gonzalez J (1994) Full-term and preterm infants' perception of face-voice synchrony. Infant Behav Dev 17:447–455.

Polka L, Werker JF (1994) Developmental changes in perception of nonnative vowel contrasts. J Exp Psychol Hum Percept Perform 20:421–435.

Rivera-Gaxiola M, Silva-Pereyra J, Kuhl PK (2005) Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. Dev Sci 8:162–172.

Sadato N, Toyada H (2006) Simultaneous measurement of NIRS and fMRI. Neurosci Res 55 [Suppl 1]:S31.

Sakatani K, Chen S, Lichty W, Zuo H, Wang YP (1999) Cerebral blood oxygenation changes induced by auditory stimulation in newborn infants measured by near infrared spectroscopy. Early Hum Dev 55:229–236.

Sato H, Takeuchi T, Sakai KL (1999) Temporal cortex activation during speech recognition: an optical topography study. Cognition 73:B55–B66.

Sato Y, Mori K, Furuya I, Hayashi R, Minagawa-Kawai Y, Koizumi T (2003) Developmental changes in cerebral lateralization to spoken language in infants: measured by near-infrared spectroscopy. Jpn J Logoped Phoniatr 44:165–171.

Shtyrov Y, Pihko E, Pulvermuller F (2005) Determinants of dominance: is language laterality explained by physical or linguistic features of speech? NeuroImage 27:37–47.

Simos PG, Molfese DL (1997) Electrophysiological responses from a temporal order continuum in the newborn infant. Neuropsychologia 35:89–98.

Simos PG, Molfese DL, Brenden RA (1997) Behavioral and electrophysiological indices of voicing-cue discrimination: laterality patterns and development. Brain Lang 57:122–150.

Singh L, Morgan JL, White KS (2004) Preference and processing: the role of speech affect in early speech spoken word recognition. J Mem Lang 51:173–189.

Stager CL, Werker JF (1997) Infants listen for more phonetic detail in speech perception than in word-learning tasks. Nature 388:381–382.

Strangman G, Culver JP, Thompson JH, Boas DA (2002) A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation. NeuroImage 17:719–731.

Tallal P, Miller S, Fitch RH (1993) Neurobiological basis of speech: a case for the preeminence of temporal processing. Ann NY Acad Sci 682:27–47.

Tervaniemi M, Medvedev SV, Alho K, Pakhomov SV, Roudas MS, Van Zuijen TL, Näätänen R (2000) Lateralized automatic auditory processing of phonetic versus musical information: a PET study. Hum Brain Mapp 10:74–79.

van Zuijen TL, Sussman E, Winkler I, Näätänen R, Tervaniemi M (2004) Grouping of sequential sounds—an event-related potential study comparing musicians and nonmusicians. J Cogn Neurosc 16:331–338.

Villringer A, Planck J, Hock C, Schleinkofer L, Dirnagl U (1993) Near infrared spectroscopy (NIRS): a new tool to study hemodynamic changes during activation of brain function in human adults. Neurosci Lett 154:101–104.

Villringer A, Minoshima S, Hock C, Obrig H, Ziegler S, Dirnagl U, Schwaiger M, Villringer A (1997) Assessment of local brain activation: a simultaneous PET and near-infrared spectroscopy study. Adv Exp Med Biol 413:149–153.

Watanabe E, Maki A, Kawaguchi F, Takashiro K, Yamashita Y, Koizumi H, Mayanagi Y (1998) Non-invasive assessment of language dominance with near-infrared spectroscopic mapping. Neurosci Lett 256:49–52.

Werker JF, Curtin S (2005) PRIMIR: a developmental framework of infant speech processing. Lang Learn Dev 1:197–234.

Werker JF, Tees RC (1984) Cross-language speech perception: evidence for perceptual reorganization during the first year of life. Infant Behav Dev 7:49–63.

Werker JF, Gilbert JH, Humphrey K, Tees RC (1981) Developmental aspects of cross-language speech perception. Child Dev 52:349–355.

Yamamoto T, Kato T (2002) Paradoxical correlation between signal in functional magnetic resonance imaging and deoxygenated haemoglobin content in capillaries: a new theoretical explanation. Phys Med Biol 47:1121–1141.

Zaramella P, Freato F, Amigoni A, Salvadori S, Marangoni P, Suppjei A, Schiavo B, Chiandetti L (2001) Brain auditory activation measured by near-infrared spectroscopy (NIRS) in neonates. Pediatr Res 49:213–219.

Zevin JD, McCandliss BD (2005) Dishabituation of the BOLD response to speech sounds. Behav Brain Funct 1:4.