Cerebral Cortex Advance Access published May 23, 2010

Cerebral Cortex
doi:10.1093/cercor/bhq082

Optical Brain Imaging Reveals General Auditory and Language-Specific Processing in Early Infant Development

Yasuyo Minagawa-Kawai[1–4], Heather van der Lely[4,5], Franck Ramus[1], Yutaka Sato[2], Reiko Mazuka[2,6] and Emmanuel Dupoux[1]

[1]Laboratoire de Sciences Cognitives et Psycholinguistique, ENS–DEC–EHESS–CNRS, Paris 75005, France, [2]Laboratory for Language Development, Brain Science Institute, RIKEN, Wako, Saitama 351-0198, Japan, [3]Graduate School of Human Relations, Keio University, Minato-ku, Tokyo 108-8345, Japan, [4]Centre for Developmental Language Disorders and Cognitive Neuroscience, University College London, London WC1N 1PF, UK, [5]Department of Psychology, Harvard University, Cambridge, MA 02138, USA and [6]Department of Psychology and Neuroscience, Duke University, Durham, NC 27708, USA

Address correspondence to email: myasuyo@bea.hi-ho.ne.jp.

This study uses near-infrared spectroscopy in young infants in order to elucidate the nature of functional cerebral processing for speech. Previous imaging studies of infants' speech perception revealed left-lateralized responses to native language. However, it is unclear if these activations were due to language per se rather than to some low-level acoustic correlate of spoken language. Here we compare native (L1) and non-native (L2) languages with 3 different nonspeech conditions including emotional voices, monkey calls, and phase scrambled sounds that provide more stringent controls. Hemodynamic responses to these stimuli were measured in the temporal areas of Japanese 4 month-olds. The results show clear left-lateralized responses to speech, prominently to L1, as opposed to various activation patterns in the nonspeech conditions. Furthermore, implementing a new analysis method designed for infants, we discovered a slower hemodynamic time course in awake infants. Our results are largely explained by signal-driven auditory processing. However, stronger activations to L1 than to L2 indicate a language-specific neural factor that modulates these responses. This study is the first to discover a significantly higher sensitivity to L1 in 4 month-olds and reveals a neural precursor of the functional specialization for the higher cognitive network.

Keywords: emotion, infant, laterality, NIRS, speech perception

Introduction

An increasing number of neuroimaging studies have documented that the human language system involves a large network of perisylvian areas predominantly lateralized in the left hemisphere (Branche et al. 1964; Ojeman et al. 1989; Hickok and Poeppel 2000; Friederici 2002). Yet, the developmental process underlying such cerebral specialization for language remains controversial. This is primarily due to the technical difficulties of applying brain imaging techniques to young infants. The introduction of multichannel near-infrared spectroscopy (NIRS) 10 years ago provided a promising technique in this emerging field of developmental neuroscience (for a review, see Minagawa-Kawai et al. 2008). Studies using NIRS have explored the development of the infant brain for several cognitive abilities including social cognition (Grossmann et al. 2008; Lloyd-Fox et al. 2009; Minagawa-Kawai et al. 2009), object recognition (Watanabe et al. 2008; Wilcox et al. 2008, 2009), and speech processing (Peña et al. 2003; Homae et al. 2006, 2007; Bortfeld et al. 2007, 2009; Minagawa-Kawai et al. 2007; Gervain et al. 2008; Nakano et al. 2009; Telkemeyer et al. 2009). NIRS is noninvasive, portable, and scanner noise free and has the advantage of providing reliable brain localization of cerebral responses to sound stimuli, making it ideally suitable for assessing lateralization in infants. In the present study, we utilize NIRS to revisit the issue of brain lateralization for speech and nonspeech sounds in 4-month-old infants.

Neuroimaging studies on speech processing in early infancy are gradually emerging. Using NIRS, Peña et al. (2003) reported that newborn infants' brain responses were stronger in the left temporal areas for sentences in the maternal language (L1) than for the same stimuli played backward (BW). Left dominance for L1 speech processing was also observed in older infants aged 6–9 months by using both forward (FW) speech and silent control together with visual stimuli (Bortfeld et al. 2009). A functional magnetic resonance imaging (fMRI) study of 3-month-old infants (Dehaene-Lambertz et al. 2002) also reported dominance in the left temporal areas for both FW and BW speech, with more activation for FW as compared with BW speech only in the left angular gyrus. Sato, Hirabayashi et al. (2006) extended these findings by using NIRS to compare brain responses to L1 with those to a foreign language (L2) in newborn infants. They found that exclusively in the L1 condition, activations were significantly stronger for FW than for BW speech in the left temporal area. However, for these newborn infants, there were no significant differences in activation amplitudes between L1 and L2. It is likely that in the course of language development, L1-specific cerebral responses will emerge, but as yet, no study has investigated the difference between L1 and L2 activations in older infants. Our study aims to fill this gap in the literature.

Furthermore, although these early brain signals in the left hemisphere for language and BW speech are intriguing, it is still unclear how to relate these activations to language processing per se. One of the plausible explanations for the left-dominant responses to speech is the acoustic properties of the stimuli: speech-containing segments such as consonants and vowels may enhance leftward brain activations due to their rapid acoustic transitions. Several studies (Zatorre and Belin 2001; Boemio et al. 2005; Schönwiesner et al. 2005; Jamison et al. 2006) have revealed that the left hemisphere is preferentially involved in processing rapid spectrotemporal changes, such as the formant transitions in phonemes, whereas the right hemisphere is more engaged in slow changes, such as those in the prosody. Based on this view, a left lateralization in young infants for language could be accounted for by the acoustic properties of speech sounds that contain rapidly changing linguistic segments. Another plausible explanation for stronger activations to FW than to BW speech is the naturalness of the

© The Author 2010. Published by Oxford University Press. All rights reserved.
For permissions, please e-mail: journals.permissions@oxfordjournals.org

Downloaded from http://cercor.oxfordjournals.org at Keio University on May 23, 2010

sounds: BW speech cannot be physically produced by a human vocal tract, and so it is not biologically natural. Specifically, BW speech differs from FW speech in low-level acoustic characteristics: BW speech contains reversed envelope profiles that are impossible to obtain from a vocal tract or any natural physical system. The problem is that the difference between causally possible and impossible physical sounds could be encoded early in the auditory pathway. As highlighted by Galbraith et al. (2004), brainstem-evoked responses differ between FW and BW speech. Furthermore, based on an information theoretical modeling study, Smith and Lewicki (2006) claimed that the spectrotemporal properties of the cochlear code are optimal for natural sounds and FW speech, but not for BW speech. If so, differences in activation amplitude between FW and BW speech, such as those found by Peña et al. (2003) and Sato, Hirabayashi et al. (2006), might reflect low-level acoustic properties, rather than language processing per se. Thus, using BW alone as a control condition for low-level acoustic properties is potentially problematic. Cerebral responses to the FW stimuli also need further investigation, since it is possible that the human brain responds to vocalizations rather than to language per se. Although adult imaging studies have consistently exhibited greater cerebral activations to human vocalizations than to other animal calls such as monkeys and birds (Fecteau et al. 2004; Hashimoto et al. 2006; von Kriegstein et al. 2007), no study has examined whether this species-specific trait exists at the beginning of human development.

Our principal aim is to elucidate the nature of language lateralization and the effects of L1 versus L2 in young infants. Building on previous studies, we focus on infants of 4 months old to test L1 specificity in their brains. Nazzi et al. (2000) reported that infants of 5 months old are able to behaviorally discriminate any language from their L1 in contrast to newborns who confuse L1 and L2 that have the same language rhythms as each other. Because neural signatures tend to appear earlier than behavioral manifestations (Gervain et al. 2008), we predicted that the emergence of L1 specificity in listening to connected speech may exist at 4 months old. We therefore measured left and right temporal activations to L1 (Japanese) and L2 (English) with NIRS. We examined the possible factors that trigger strong activations by employing rigorous control conditions. Instead of using BW speech, we employed 3 different sound stimuli that vary in acoustic properties and biological factors: 1) emotional sounds, which are produced by a human vocal tract but lack the segmental/fast variation structure of speech; 2) monkey vocalizations, which are produced by a nonhuman vocal tract (Ghazanfar and Rendall 2008); and 3) scrambled sounds that are totally artificial nonvocalizations but are matched for energy and long-term spectrum with the other 4 conditions. Using different kinds of controls (see Table 1) elucidates the specificity of language function in relation to acoustic properties of the stimuli, species specificity, and vocalization specificity on left-dominant brain activations.

Table 1
Stimulus type

Type	Native language	Linguistic segment	Same species	Vocally produced	Sound
Language (L1)	+	+	+	+	+
Language (L2)	−	+	+	+	+
Emotional voice	−	−	+	+	+
Monkey call	−	−	−	+	+
Scrambled	−	−	−	−	+

Typically, a block design is used in NIRS for stimulus presentation, whereby relatively large brain response can be obtained (Bortfeld et al. 2007, 2009; Wilcox et al. 2008; Minagawa-Kawai et al. 2009). However, in order to test all these crucial different conditions with the same infant, it is necessary for this study to use a fast event-related design (Zarahn et al. 1997; Friston et al. 1998). This enables us to increase the trial repetition time and allows us to take a new approach to the analysis of these data from the event-related paradigm for infant NIRS. Specifically, we employed a general linear model (GLM) approach (Friston et al. 1994, 1995), which uses a model of the expected hemodynamic response function (HRF). However, it is questionable whether the standard adult HRF model is valid for infants as the brain physiology of 4 month-olds is quite different to that of adults, and indeed infants show a phase delay in their evoked hemodynamic responses (Schroeter et al. 2004; Shimada and Hiraki 2006). We therefore employed a technique based on finite impulse response (FIR) functions (Friston et al. 1995) in order to reconstruct the HRF of the infants before engaging in GLM modeling.

Materials and Methods

Participants

We studied 12 full-term infants (4 girls and 8 boys) with normal birth weights, belonging to monolingual Japanese families. At the time of testing, which was conducted at the RIKEN Brain Science Institute (BSI), their average age was 128 days (standard deviation [SD] = 13.3). An additional 18 infants were tested, but they were excluded from the final sample after consideration of artifacts in the data due to head movements, fussiness, and hair obstruction ($N = 10$); refusal to wear the NIRS holder ($N = 4$); poor positioning of the probe ($N = 1$); and sleeping or drowsiness during testing ($N = 3$). We employed strict criteria to determine our final data set because noise-free hemoglobin (Hb) signals are necessary for a reliable GLM analysis. In accordance with our criteria, first, we excluded all the blocks contaminated with motion artifacts (for details, see Data Analyses); we also discarded whole blocks obtained from sleeping infants. Then, only participants for whom at least 8 blocks survived (of 12 blocks) without motion artifacts for each of the 5 conditions were included in the final data set (see stimuli for details of block and condition). This ensured that the final data set contained more than 40 blocks for each infant. Our experimental design, with its 5 conditions, resulted in a presentation time that was longer time than is typically found in infant studies. As a result, there were many participants who did not complete all the sessions or who were unable to meet our strict criteria for inclusion in the data set as documented above. Parents provided informed consent in compliance with a protocol approved by the ethics committee of Riken, BSI (Wako 3rd-16-12 (10)).

NIRS Recording

This study used NIRS (ETG-4000, Hitachi Medical Co., Tokyo, Japan), which measures the Hb concentration changes of the optical paths in the brain between the nearest pairs of incident and detection probes separated by 3 cm on the scalp surface (Watanabe et al. 1996; Yamashita et al. 1996). This separation enables us to measure hemodynamic changes in the brain 2.5–3 cm deep from the head surface, which corresponds to the gray matter on the outer surface of the brain (Fukui et al. 2003). The instrument emits 2 wavelengths (ca. 695 and 830 nm) of continuous near-infrared lasers, modulated at different frequencies depending on the channels and the wavelengths and detected with the sharp frequency filters of lock-in amplifiers (Watanabe et al. 1996).

Five incident and 4 detection probes arranged in a transformed 3 × 3 grid (12 channels, Fig. 1) were fitted on the temporal and frontal areas

Downloaded from http://cercor.oxfordjournals.org at Keio University on May 23, 2010

of each side of the head using the international 10–20 system. Specifically, we placed the mid-bottom detector to T3 and T4 position for each side and the nasion-to-inion line and the vertex-to-tragus line were used either for the horizontal or the vertical axis (Fig. 1). Brain regions corresponding to NIRS channels were estimated using the virtual registration (Okamoto et al. 2004; Tsuzuki et al. 2007) by taking the small head size of 4 month-olds into account.

Stimuli

The auditory stimuli comprised 5 conditions: native speech (Japanese), non-native speech (British English), emotional voices, monkey calls (Macaque), and scrambled controls. All speech stimuli were made of a concatenation of short sentences recorded from film dialogues or a speech database (Corpus of Spontaneous Japanese) (Maekawa et al. 2004) by either female or male speakers. The emotional voices were human vocalizations with no linguistic content, either with a positive (e.g., admiration and laughing) or with a negative emotional valence (e.g., crying and sigh). Monkey calls were similarly divided into those with positive (e.g., coos, girneys, and harmonic arches) and those with negative emotional valence (e.g., screams and shrill barks). The monkey calls were taken from a sound library compiled by Marc D. Hauser (Harvard University; stimuli recorded from the Island of Cayo Santiago, Puerto Rico, USA). Spectral scrambling was applied to the 4 types of stimuli above to make 4 matched scrambled controls (Japanese, English, emotional voices, and monkey calls). Spectral scrambling sounds that are totally unintelligible sounds consist of the random and independent exchange of bands of spectra within a given stimulus. More precisely, they were synthesized by processing all individual stimuli through a gammatone filter bank with 64 channels. In each channel, the signal was windowed with overlapping Hanning windows of 25-ms duration. The windows were then shuffled randomly within a channel and displaced within a time range of ±500 ms around its original temporal position. The scrambled sounds were finally obtained by putting all frequency channels back together. Thus, the average amplitude and long-term power spectrum of scrambled controls is exactly the same as for their original stimulus. All stimuli ranged from 700 to 1200 ms in length and had equal energy (root mean square). These individual stimuli were used to construct stimulus blocks that lasted 10 s on average for each condition. One block contained a concatenation of 9–12 individual stimuli of the same type (short sentences in Japanese, Monkey vocalizations, etc.) with 200-ms silence between each stimulus. Finally, the blocks were assembled into sessions. One session contained 4 blocks of each condition (20 blocks per session) that were presented to the infants in a random order with a silence period (8–14 s) between the blocks. All infants had at least 3 sessions, each of which lasted for 6.5–6.8 min.

Procedures

NIRS recording was carried out inside a sound-attenuated room. After measuring the infants' head size, experimenters positioned the probe pads onto the infants' head (Fig. 1). Infants, seated on their mothers' lap, listened to the stimuli that were presented via a loudspeaker (ca. 65-dB sound pressure level). To reduce motion artifacts and restlessness, one of the experimenters entertained the infants with silent toys. During the session, both the mother and experimenter listened to other sounds through headphones to prevent any influence from the stimuli on their behavior.

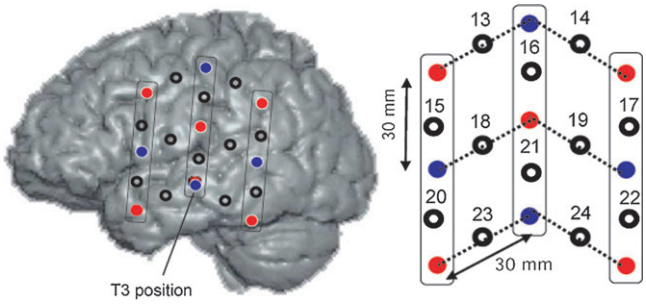

Figure 1. Probe arrangement to attach to the bilateral temporal areas. Black circles on the brain indicate channel positions. Red probes are emitters, and blue ones are detectors. Black circles are NIRS channels between one emitter and detector.

Data Analyses

The concentrations of oxygenated (oxy-) and deoxygenated (deoxy-) Hb were calculated from changes in absorption at 695 and 830 nm. Although some NIRS studies employed objective methods for artifact rejection, including principle component analysis (e.g., Bortfeld et al. 2007) and wavelet analysis (Sato, Tanaka et al. 2006), others have detected movement-induced artifacts by means of visual inspection (e.g., Taga et al. 2003). In this study, we employed the latter method, and the segments of data with motion artifacts were digitally marked by visually assessing the Hb signals and video recordings. Motion artifacts were characterized by sharp and abnormal changes in Hb concentrations. This characterization of motion artifacts was used to define "subruns," that is, continuous stretches of artifact-free data. A subrun that lasted for at least 60 s (i.e., 3 trials on average) was deemed valid, whereas one that lasted less than 60 s was considered to be influenced by artifacts. Because artifact trials could affect the shape and amplitude level of Hb changes in adjacent valid trials, our strict criterion helps us to provide a reliable analysis of the time course of Hb changes using GLM. We employed 2-step processes to analyze the oxy-Hb concentrations using a generalized linear model (GLM) approach. The first analysis collapsed all 5 conditions and was aimed at discovering the time course of the HRF for the infants' brain to auditory stimuli. The second analysis used a time-shifted adult HRF model determined by the first analysis in order to analyze the brain responses to each of the 5 conditions. In both analyses, we introduced sin and cos functions of periods 1, 2, and 4 min as regressors of no interest for each of the runs, in order to model long-term trends in signal strength. Each of the subruns had its own boxcar regressor, in order to model the potential shift in baseline concentration level after each movement artifact. Finally, the artifact samples were silenced during the analysis, that is, they were assigned a weight of zero in the regression. Details of these 2 analyses are further described below.

Results

Estimation of the Infant HRF

We measured oxy- and deoxy-Hb changes from the bilateral temporal areas to the sound stimuli (Fig. 1). Using preprocessed Hb data, we reconstructed the infant's HRF using a FIR function approach (Friston et al. 1995). For each infant, each channel of oxy-Hb data was submitted to a GLM analysis with 20 FIR regressors: each regressor was a 1-s rectangular pulse synchronous with stimulus onset and time shifted, respectively, by 0, 1, 2, etc. We then extracted the 20-β coefficients for each FIR component and averaged these across infants as shown in Figure 2*A*. The 2 channels with the maximum number of FIR components whose βs are significantly different from zero across infants correspond to the left and right auditory areas (CH21 on the left side and CH9 on the right side). The average curve from these 2 channels (Fig. 2*B*) shows an initial dip followed by an increase in Hb, which appears time shifted compared with an adult response. We fitted this curve to the predictions from a time-shifted adult HRF. The best fit was found with a 2.8-s delay and accounted for 97.4% of the variance. The same analysis was performed for deoxy-Hb, where we found a 3.4-s delay accounted for 95.8% of the variance. In the subsequent analyses, we used the adult standard HRF model time shifted by 2.8 s as our infant HRF model.

Downloaded from http://cercor.oxfordjournals.org at Keio University on May 23, 2010

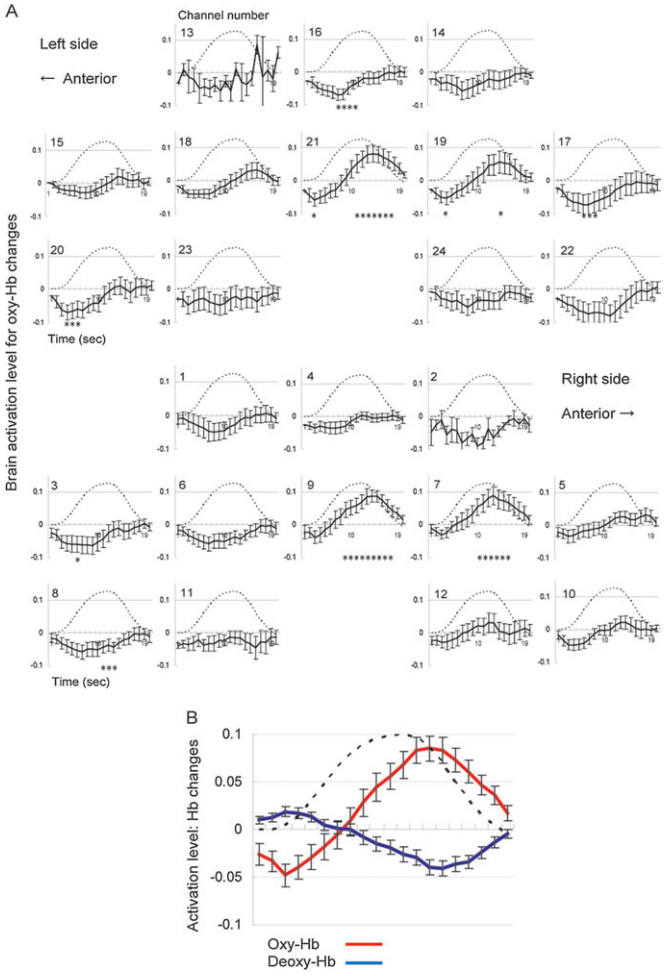

Figure 2. (*A*) FIR results in 24 channels for all the conditions. Dashed line indicates the canonical HRF. (*B*) Time course of Hb changes elicited by the FIR analyses. Dashed line indicates the canonical HRF. Zero point in the time line is a stimulus onset.

Brain Responses to Speech and Nonspeech Stimuli

We performed 2 analyses with a progressively refined level of detail. In the first analysis, all 5 stimulus types were grouped together in a single sound condition. In the second analysis, we declared 5 conditions: Japanese, English, emotion, monkey, and scramble. All these analyses used GLM modeling of the oxy-Hb signal for each condition, channel, and infant separately, using regressors of a boxcar, time locked to each stimulus onset in a given condition, and convolved with the infant HRF model. The regression coefficients (β) for each condition, channel, and each infant were then tested against zero (*t*-test across infants) and corrected for multiple comparisons across channels by using the false discovery rate method. The resulting *t*-tests, uncorrected and corrected, are shown in Table 2. The global sound condition significantly activated the bilateral auditory areas. All of the 5 stimuli activated various regions including the superior temporal gyrus (STG) and inferior frontal gyrus as indicated by a map of across-infant significant βs for each condition (Fig. 3). Although activation patterns around the perisylvian areas were similar for speech and nonspeech vocal sounds, activation levels were higher in the left side for the speech condition in contrast to the larger and broader activation in the right side for the nonspeech vocal condition. Separate analysis for each condition revealed various activations in the two hemispheres (Table 2). While a strong response was only found in the right auditory/STG for the emotional vocalization, significant channels were restricted to the left side for the Japanese, English, and scrambled sounds conditions. The monkey calls broadly activated bilateral temporal areas, including the anterior STG. In order to compare brain activations for native and non-native languages, analysis of variance with factors of laterality and language was performed (Fig. 4). The results revealed significant main effects of laterality ($F_{1,44}$ = 4.41, P = 0.047) and language ($F_{1,44}$ = 5.46, P = 0.029).

Discussion

This study examined the developmental specificity of language processing in infants' brains by using a carefully controlled set of stimuli as well as a new analysis method for infant data. The results revealed localized brain activations with larger left-lateralized responses to speech as opposed to bilateral or rightward activation for various control stimuli in the early infant developing brain. Our data provide new evidence in at least 3 respects: 1) they reveal functional asymmetries for speech and other nonspeech vocalizations such as emotional voices in young infants; 2) they demonstrate neurophysiological evidence for language-specific neural plasticity in 4 month-olds by comparing the brain responses to L1 and L2; and 3) they uncover a slower hemodynamic time course in awake infants' brains compared with that of adults.

Before we focus on the left dominance in speech processing, we first review the results of the nonspeech conditions. For the emotion condition, a significant activation was observed only in the right temporal area. This rightward temporal activation is consistent with many adult imaging studies on vocal emotions (Meyer et al. 2002, 2004; Wildgruber et al. 2002; Wiethoff et al. 2008). So far, only a few studies have examined the neural basis of the perception of emotional prosody in infants. For instance, 7 month-olds showed different electrophysiological responses to different emotional voices, with a greater negative component to angry voices (Grossmann et al. 2005). Although no previous study has revealed the specific brain region associated with emotion processing in young infants, our study demonstrated a trend for the right dominance in processing emotional prosody in 4 month-olds. Our results for the nonspeech conditions for infants generally replicate previous imaging studies with adults. However, we consider it likely that these activations chiefly reflect processing the acoustic properties in the different stimulus conditions. More concretely, sound streams with segmental features or fast spectral changes may enhance activations in the left hemisphere, whereas prosodic pitch contours may enhance activations in the right hemisphere and that this is responsible for the left/right dominance for the different stimuli (Zatorre and Belin 2001; Poeppel 2003; Zatorre and Gandour 2008). We observed that the emotional voices that have slow prosodic changes and fewer linguistic segments evoked rightward activations, while the scrambled sounds with a lot of rapidly changing segments elicited significant leftward activations.

Further to the acoustic factor, our control stimuli enabled us to distinguish 2 additional factors: species specificity and vocalization. In response to the monkey vocalizations, human 4 month-olds showed strong and broad activations in the

Downloaded from http://cercor.oxfordjournals.org at Keio University on May 23, 2010

Table 2
Significant levels in different conditions

Conditions	Brain region	Channel	Side	β	SE	*t*	*P* (unc)	*P* (FDR)
All sounds	Auditory	9	R	0.07	0.08	4.87	0.0003	0.008
All sounds	Auditory	21	L	0.07	0.09	3.53	0.004	0.046
Vocal	pSTG/SMG	6	R	0.05	0.07	3	0.01	0.065
Vocal	Auditory	9	R	0.1	0.11	5.11	0.0002	0.005
Vocal	Auditory	21	L	0.1	0.13	3.28	0.006	0.065
JP	Precentral	16	L	0.06	0.08	3.49	0.004	0.033
JP	pSTG/SMG	19	L	0.09	0.12	3.68	0.003	0.033
JP	Auditory	21	L	0.15	0.16	7.03	0.00001	0.0002
ENG	pSTG/SMG	19	L	0.06	0.09	2.97	0.01	0.131
ENG	Auditory	21	L	0.08	0.11	3.64	0.003	0.075
EM	Auditory	9	R	0.11	0.14	3.61	0.003	0.078
MN	IFG/aSTG	7	R	0.05	0.07	3.07	0.009	0.036
MN	pSTG/SMG	6	R	0.09	0.12	3.29	0.006	0.029
MN	Auditory	9	R	0.09	0.11	4.99	0.0003	0.005
MN	IFG/aSTG	18	L	0.07	0.09	4.5	0.001	0.005
MN	pSTG/SMG	19	L	0.12	0.15	4.07	0.001	0.008
MN	Auditory	21	L	0.11	0.14	4.5	0.001	0.005
SC	Auditory	21	L	0.1	0.14	3.4	0.005	0.116

JP = Japanese, ENG = English, EM = emotional voices, MN = monkey calls, SC = scrambled sounds, SE = standard error, FDR = false discovery rate, IFG = inferior frontal gyrus, SMG = supramarginal gyrus, pSTG = posterior part of superior temporal gyrus, aSTG = anterior part of STG, R = right, L = left, unc = uncorrected.

bilateral temporal areas. Exclusively for this condition, there were significant activations around the bilateral anterior STG. These results are consistent with those of adult fMRI studies examining perception of animal calls (Fecteau et al. 2004; Hashimoto et al. 2006; von Kriegstein et al. 2007). However, human adults typically show weaker brain responses to animal calls than to human vocalizations. Our finding, showing large responses to monkey calls, indicates that 4 month-olds are still sensitive to vocalizations of other species, which is characterized by higher pitch and resonance (Ghazanfar and Rendall 2008). This is congruent with their early visual abilities to discriminate monkey faces in 6 month-olds (Pascalis et al. 2002). Furthermore, recent behavioral results (Vouloumanos et al. 2010) revealed newborns' preferences to both human speech and monkey calls that sounded like human vocalization. These initial abilities may be gradually narrowed during development within the first year of life (Pascalis et al. 2002), as infants tune their sensitivity to conspecific stimuli or their ambient environment such as L1. In contrast to brain responses to vocalizations including monkey calls, artificial scrambled sounds showed only one significant channel (Table 2). Even though vocal versus scrambled conditions have identical long-term spectrum, artificial nonvocal sounds induced fewer activations than the vocal sounds did. This suggests that 4 month-olds' brains are more responsive to biologically natural sounds than to unnatural sounds. This neuronal evidence may explain weaker brain activations for BW speech in the previous imaging studies (Peña et al. 2003; Sato, Hirabayashi et al. 2006). The results are also consistent with infants' behavioral responses showing a preference for speech sounds than for their analogous artificial tones from birth to 6.5 month-old (Vouloumanos and Werker 2004, 2007). Finally, it should be noted that these data from the control stimuli should be taken with careful consideration because although each of these stimuli elicited a different pattern of significant activations, we lacked the power to demonstrate that these patterns of activations were statistically different from one another.

In comparison to the nonspeech stimuli, the speech stimuli evoked greater left-lateralized activation and produced more

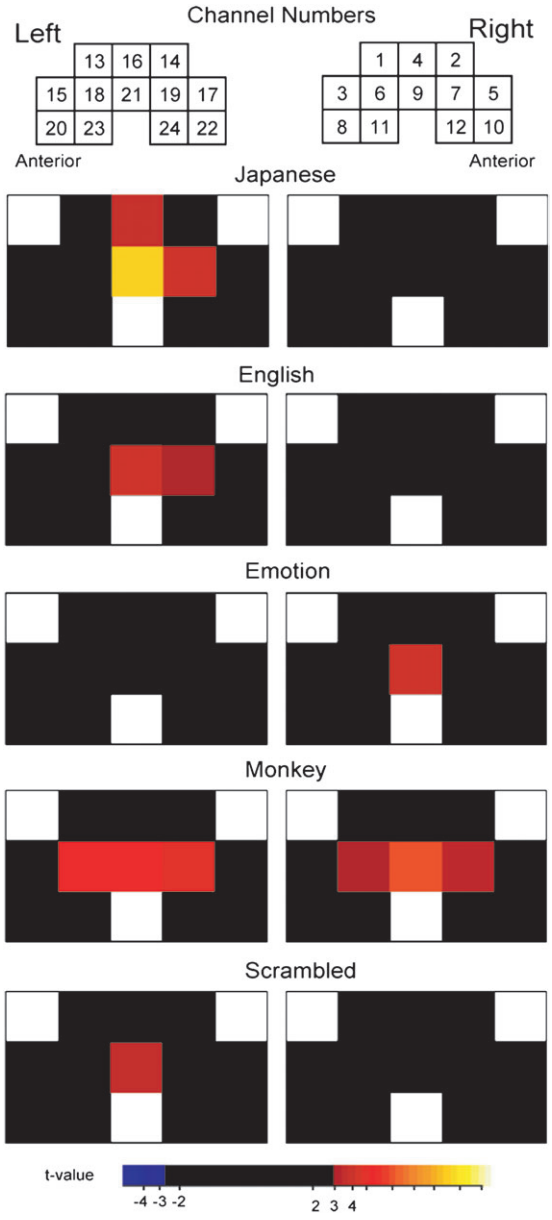

Figure 3. Brain activations in different 5 conditions for 24 channels in the left and right side. *P* values are corrected for multiple comparisons.

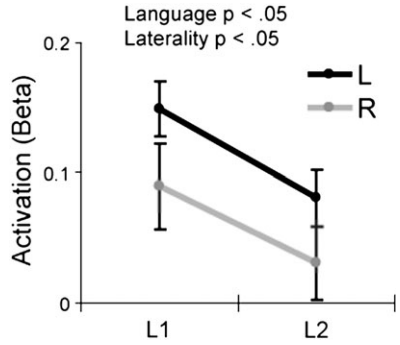

Figure 4. Different brain activation levels for native (L1) and foreign (L2) languages in the left and right auditory areas.

Downloaded from http://cercor.oxfordjournals.org at Keio University on May 23, 2010

significant statistical results. Speech stimuli in native and non-native languages activated the left temporal areas. Like nonspeech stimuli, these activations can also be interpreted to be acoustically driven since speech has rapid temporal features with many consonants and vowels. Four-month-old infants are in their preverbal period during which they cannot process detailed lexical items or syntactic structures. Therefore, in the present study, it is likely that the asymmetries observed in the young infants are chiefly driven by the acoustic properties of the stimuli (Zatorre and Belin 2001; Boemio et al. 2005; Schönwiesner et al. 2005; Jamison et al. 2006), and this acoustical sound processing may be based on the generic auditory system in human infants. Interestingly, a recent NIRS study revealed that hemispheric activations in infants were a function of temporal properties (fast vs. slow) of nonspeech (Telkemeyer et al. 2009), similar to that in adults (Boemio et al. 2005). However, our results demonstrated neural correlates beyond such a stimulus-driven hypothesis by providing evidence of the linguistic factors that affect the development of brain function in infants. Furthermore, both the languages activated the left temporal area, but there was a significant difference in the magnitude of activation between them. The response in the left STG to L1 was significantly greater than that to L2. This means that 4 months of exposure to L1, which is a language-specific factor, enhanced the neural recruitment to process the maternal language. Furthermore, as mentioned earlier, infants begin to discriminate any other language from L1 around this age (Bosch and Sebastian-Galles 1997; Nazzi et al. 2000). Our study has revealed a cerebral basis underlying this behavioral development. This basis is considered to be more intricately structured in the brains of 4-month-old infants than in that of newborn infants, which showed no differences in responses to L1 and L2 (Sato, Hirabayashi et al. 2006). It is assumed that while discriminating between languages, 4-month-old infants cannot process fine linguistic content such as lexical items but are capable of processing general sound structures, such as prosody specific to L1 (Mehler et al. 1988; Nazzi et al. 2000). Accordingly, an increased neural signature for L1 may indicate the facilitation of acoustical sound processing of L1, including its segmental and suprasegmental properties, which is necessary to integrate phonemic, lexical, and grammatical structures in the later stages of language acquisition. In this regard, this enhanced left-dominant activation may be considered a neural precursor for L1-specific language acquisition. In the course of language acquisition, this neural precursor remains in the left temporal area (Bortfeld et al. 2009) and may develop further. However, in older infants, the neural network is expected to be more developed for higher order speech processing beyond phonetic processing.

Because we used an event-related paradigm, our analytical approach was based on the GLM, which is widely used in fMRI as well as in NIRS studies (Schroeter et al. 2004; Shimada and Hiraki 2006; Wartenburger et al. 2007; Watanabe et al. 2010). Although most of the NIRS studies employ an averaging approach by obtaining averaged Hb values for different conditions (Peña et al. 2003; Homae et al. 2006), GLM is also a reliable and sophisticated method that is suitable for data obtained from an event-related paradigm. Indeed, a recent NIRS study (Watanabe et al. 2010) used averaging and GLM approaches for their analysis and reported equivalent results. However, one issue to be further explored for GLM concerns the model to be used for infant data. Specifically, the GLM approach requires a model of the HRF suited to infants. Yet, the standard HRF model used in fMRI analyses with GLM is suited to the adult vascular system. It is unclear that one can apply the same model to infants' data whose vascular system may be less mature and whose peak response has been reported to be slower than that of adults (Schroeter et al. 2004; Shimada and Hiraki 2006). Here, we reconstructed infant's hemodynamic responses with a technique based on FIR (Friston et al. 2005; Schilbach et al. 2008) and estimated the peak latency by pooling all the sound conditions together. We found a response remarkably close to that of the adult, but delayed by approximately 3 s. If we had used an HRF model without a delay, the peak latency of the hemodynamic responses would have unfairly influenced the brain activation levels. Such a delayed time course could be partly derived from immature vascular regulation and myelination in infants (Chugani et al. 1987). However, another possible reason for the delay could be attentional shifts due to the particular paradigm we used: in our study, the infant's attention was captured by visual stimuli (i.e., silent toys), so that it could have taken them a few seconds to notice the sounds played to them, inducing a delay in auditory-related activation. Since no clear details have been provided for the infant's vascular mechanism in relation to neural activities, more physiological studies are required for a conclusive explanation. However, our method provides a principled way to determine the model for GLM in infants instead of using the canonical HRF, or arbitrary windows of analysis for infants' neuroimaging, given that variations in age, stimuli, tasks, and wakefulness may influence the response latency (Meek et al. 1998; Taga et al. 2003; Shimada and Hiraki 2006; Minagawa-Kawai et al. 2008).

By examining hemodynamic responses in 4 month-olds to speech and various nonspeech sounds, the current NIRS study clearly showed a left-lateralized cerebral basis for speech processing. Although these activation patterns could be mostly explained by signal-driven hypotheses of brain lateralization, we revealed a linguistic factor that may contribute to the activation of the language network on the left side and that provides a parsimonious explanation. Moreover, we have provided data relevant to the brain physiology in young infants by showing slow latency of functional hemodynamic responses. Overall, this study illustrated that 4 month-olds' neural development is at a stage where their processing of speech is based on an interaction between generic auditory systems and learning mechanisms that start to extract regularities regarding their native language.

Funding

European Union's sixth-framework program (neuronal origins of language and communication: NEUROCOM (Project no. 012738); Grant-in-Aid for Scientific Research (A) (Project no. 21682002); Wellcome Trust (063713 to H.L.).

Notes

We thank K. Shirasawa, Y. Sogabe, and other staff in the language development laboratory; RIKEN BSI for assistance with NIRS testing; and A. Shestakova, E. Kushnerenko, Y. J. Chang, J. Meek, and other members of the DLDCN Centre for their assistance in conducting the pilot study for this work at the DLDCN Centre, London. We also thank M. D. Hauser for kindly providing the monkey vocalization stimuli. *Conflict of Interest*: None declared.

Downloaded from http://cercor.oxfordjournals.org at Keio University on May 23, 2010

References

Boemio A, Fromm S, Braun A, Poeppel D. 2005. Hierarchical and asymmetric temporal sensitivity in human auditory cortices. Nat Neurosci. 8:389–395.

Bortfeld H, Fava F, Boas DA. 2009. Identifying cortical lateralization of speech processing in infants using near-infrared spectroscopy. Dev Neuropsychol. 34:52–65.

Bortfeld H, Wruck E, Boas DA. 2007. Assessing infants' cortical response to speech using near-infrared spectroscopy. Neuroimage. 34: 407–415.

Bosch L, Sebastian-Galles N. 1997. Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. Cognition. 65:33–69.

Branche C, Milner B, Rasmussen T. 1964. Intracarotid sodium amytal for the lateralization of cerebral speech dominance. J Neurosurg. 21:399–405.

Chugani HT, Phelps ME, Mazziotta JC. 1987. Positron emission tomography study of human brain functional development. Ann Neurol. 22:487–497.

Dehaene-Lambertz G, Dehaene S, Hertz-Pannier L. 2002. Functional neuroimaging of speech perception in infants. Science. 298: 2013–2015.

Fecteau S, Armony JL, Joanette Y, Belin P. 2004. Is voice processing species-specific in human auditory cortex? An fMRI study. Neuroimage. 23:840–848.

Friederici AD. 2002. Towards a neural basis of auditory sentence processing. Trends Cogn Sci. 6:78–84.

Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R. 1998. Event-related fMRI: characterizing differential responses. Neuroimage. 7:30–40.

Friston KJ, Holmes AP, Worsley KJ, Poline JB, Frith CD, Frackowiak RSJ. 1995. Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp. 2:189–210.

Friston KJ, Penny W, David O. 2005. Modeling brain responses. Int Rev Neurobiol. 66:89–124.

Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta JC, Evans AC. 1994. Assessing the significance of focal activations using their spatial extent. Hum Brain Mapp. 1:214–220.

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

Galbraith GC, Amaya EM, Diaz de Rivera JM, Donan NM, Duong MT, Hsu JN, Tran K, Tsang LP. 2004. Brain stem evoked response to forward and reversed speech in humans. Neuroreport. 15:2057–2060.

Gervain J, Macagno F, Cogoi S, Peña M, Mehler J. 2008. The neonate brain detects speech structure. Proc Natl Acad Sci U S A. 105: 14222–14227.

Ghazanfar AA, Rendall D. 2008. Evolution of human vocal production. Curr Biol. 18:R457–R460.

Grossmann T, Striano T, Friederici AD. 2005. Infants' electric brain responses to emotional prosody. Neuroreport. 16:1825–1828.

Grossmann T, Johnson MH, Lloyd-Fox S, Blasi A, Deligianni F, Elwell C, Csibra G. 2008. Early cortical specialization for face-to-face communication in human infants. Proc Biol Sci. 275:2803–2811.

Grossmann T, Striano T, Friederici AD. 2005. Infants' electric brain responses to emotional prosody. Neuroreport. 16:1825–1828.

Hashimoto T, Usui N, Taira M, Nose I, Haji T, Kojima S. 2006. The neural mechanism associated with the processing of onomatopoeic sounds. Neuroimage. 31:1762–1770.

Hickok G, Poeppel D. 2000. Towards a functional neuroanatomy of speech perception. Trends Cogn Sci. 4:131–138.

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.

Jamison HL, Watkins KE, Bishop DV, Matthews PM. 2006. Hemispheric specialization for processing auditory nonspeech stimuli. Cereb Cortex. 16:1266–1275.

Lloyd-Fox S, Blasi A, Volein A, Everdell N, Elwell CE, Johnson MH. 2009. Social perception in infancy: a near infrared spectroscopy study. Child Dev. 80:986–999.

Maekawa K, Kikuchi K, Tsukahara W. 2004. Corpus of Spontaneous Japanese: design, annotation and XML representation. Proceedings of the International Symposium on Large-Scale Knowledge Resources; Tokyo. 2004:19–24.

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.

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.

Meyer M, Alter K, Friederici AD, Lohmann G, von Cramon DY. 2002. Functional MRI reveals brain regions mediating slow prosodic manipulations of spoken sentences. Hum Brain Mapp. 17: 73–88.

Meyer M, Steinhauer K, Alter K, Friederici AD, von Cramon DY. 2004. Brain activity varies with modulation of dynamic pitch variance in sentence melody. Brain Lang. 89:277–289.

Minagawa-Kawai Y, Mori K, Hebden J, Dupoux E. 2008. Optical imaging of infants' neurocognitive development: recent advances and perspectives. Dev Neurobiol. 68:712–728.

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, Naoi N, Kojima S. 2009. New approach to functional neuroimaging: near infrared spectroscopy. Tokyo (Japan): Keio University Press.

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, Juscyzk PW, Johnson EK. 2000. Language discrimination by English-learning 5-month-olds: effects of rhythm and familiality. J Mem Lang. 43:1–19.

Ojeman G, Ojeman J, Lettich B, Berger M. 1989. Cortical language localization in left, dominant hemisphere. J Neurosurg. 71: 316–326.

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

Pascalis O, de Haan M, Nelson CA. 2002. Is face processing species-specific during the first year of life? Science. 296: 1321–1323.

Peña 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 U S A. 100:11702–11705.

Poeppel D. 2003. The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. Speech Commun. 41:245–255.

Sato H, Hirabayashi Y, Tsubokura S, Kanai M, Ashida S, Konishi I, Uchida M, Hasegawa T, Konishi Y, Maki A. 2006. Cortical activation in newborns while listening to sounds of mother tongue and foreign language: an optical topography study. Proceeding of International Conference on Infant Studies, Kyoto, Japan. Presentation number 037-070.

Sato H, Tanaka N, Uchida M, Hirabayashi Y, Kanai M, Ashida T, Konishi I, Maki A. 2006. Wavelet analysis for detecting body-movement artifacts in optical topography signals. Neuroimage. 33:580–587.

Schilbach L, Eickhoff SB, Mojzisch A, Vogeley K. 2008. What's in a smile? Neural correlates of facial embodiment during social interaction. Soc Neurosci. 3:37–50.

Schönwiesner M, Rubsamen R, von Cramon DY. 2005. Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. Eur J Neurosci. 22: 1521–1528.

Schroeter ML, Bucheler MM, Muller K, Uludag K, Obrig H, Lohmann G, Tittgemeyer M, Villringer A, von Cramon DY. 2004. Towards a

Downloaded from http://cercor.oxfordjournals.org at Keio University on May 23, 2010

standard analysis for functional near-infrared imaging. Neuroimage. 21:283–290.
Shimada S, Hiraki K. 2006. Infant's brain responses to live and televised action. Neuroimage. 32:930–939.
Smith EC, Lewicki MS. 2006. Efficient auditory coding. Nature. 439: 978–982.
Taga G, Asakawa K, Maki A, Konishi Y, Koizumi H. 2003. Brain imaging in awake infants by near-infrared optical topography. Proc Natl Acad Sci U S A. 100:10722–10727.
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.
Tsuzuki D, Jurcak V, Singh AK, Okamoto M, Watanabe E, Dan I. 2007. Virtual spatial registration of stand-alone fNIRS data to MNI space. Neuroimage. 34:1506–1518.
von Kriegstein K, Smith DR, Patterson RD, Ives DT, Griffiths TD. 2007. Neural representation of auditory size in the human voice and in sounds from other resonant sources. Curr Biol. 17:1123–1128.
Vouloumanos A, Hauser MD, Werker JF, Martin A. 2010. The tuning of human neonates' preference for speech. Child Dev. 81: 517–527.
Vouloumanos A, Werker JF. 2004. Tuned to the signal: the privileged status of speech for young infants. Dev Sci. 7:270–276.
Vouloumanos A, Werker JF. 2007. Listening to language at birth: evidence for a bias for speech in neonates. Dev Sci. 10:159–164.
Wartenburger I, Steinbrink J, Telkemeyer S, Friedrich M, Friederici AD, Obrig H. 2007. The processing of prosody: evidence of interhemispheric specialization at the age of four. Neuroimage. 34:416–425.
Watanabe E, Yamashita Y, Maki A, Ito Y, Koizumi H. 1996. Non-invasive functional mapping with multi-channel near infra-red spectroscopic topography in humans. Neurosci Lett. 205:41–44.
Watanabe H, Homae F, Nakano T, Taga G. 2008. Functional activation in diverse regions of the developing brain of human infants. Neuroimage. 43:346–357.
Watanabe H, Homae F, Taga G. 2010. General to specific development of functional activation in the cerebral cortexes of 2- to 3-month-old infants. Neuroimage. 50:1536–1344.
Wiethoff S, Wildgruber D, Kreifelts B, Becker H, Herbert C, Grodd W, Ethofer T. 2008. Cerebral processing of emotional prosody—influence of acoustic parameters and arousal. Neuroimage. 39:885–893.
Wilcox T, Bortfeld H, Woods R, Wruck E, Armstrong J, Boas D. 2009. Hemodynamic changes in the infant cortex during the processing of featural and spatiotemporal information. Neuropsychologia. 47: 657–662.
Wilcox T, Bortfeld H, Woods R, Wruck E, Boas DA. 2008. Hemodynamic response to featural changes in the occipital and inferior temporal cortex in infants: a preliminary methodological exploration. Dev Sci. 11:361–370.
Wildgruber D, Pihan H, Ackermann H, Erb M, Grodd W. 2002. Dynamic brain activation during processing of emotional intonation: influence of acoustic parameters, emotional valence, and sex. Neuroimage. 15:856–869.
Yamashita Y, Maki A, Koizumi H. 1996. Near-infrared topographic measurement system: imaging of absorbers localized in a scattering medium. Rev Sci Instrum. 67:730–732.
Zarahn E, Aguirre G, D'Esposito M. 1997. A trial-based experimental design for fMRI. Neuroimage. 6:122–138.
Zatorre RJ, Belin P. 2001. Spectral and temporal processing in human auditory cortex. Cereb Cortex. 11:946–953.
Zatorre RJ, Gandour JT. 2008. Neural specializations for speech and pitch: moving beyond the dichotomies. Philos Trans R Soc Lond B Biol Sci. 363:1087–1104.

Downloaded from http://cercor.oxfordjournals.org at Keio University on May 23, 2010