

# Sensitivity of Newborn Auditory Cortex to the Temporal Structure of Sounds

Silke Telkemeyer,<sup>1,2</sup> Sonja Rossi,<sup>1</sup> Stefan P. Koch,<sup>1</sup> Till Nierhaus,<sup>1</sup> Jens Steinbrink,<sup>1</sup> David Poeppel,<sup>3</sup> Hellmuth Obrig,<sup>1,4,5\*</sup> and Isabell Wartenburger<sup>1,6\*</sup>

<sup>1</sup>Berlin NeuroImaging Center, Charité University Medicine, 10117 Berlin, Germany, <sup>2</sup>Department of Cognitive Psychology, Humboldt-University Berlin, 10099 Berlin, Germany, <sup>3</sup>Department of Psychology, New York University, New York, New York 10003, <sup>4</sup>Department of Cognitive Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, <sup>5</sup>University Hospital, 04103 Leipzig, Germany, and <sup>6</sup>Department of Linguistics, University of Potsdam, 14476 Potsdam, Germany

Understanding the rapidly developing building blocks of speech perception in infancy requires a close look at the auditory prerequisites for speech sound processing. Pioneering studies have demonstrated that hemispheric specializations for language processing are already present in early infancy. However, whether these computational asymmetries can be considered a function of linguistic attributes or a consequence of basic temporal signal properties is under debate. Several studies in adults link hemispheric specialization for certain aspects of speech perception to an asymmetry in cortical tuning and reveal that the auditory cortices are differentially sensitive to spectrotemporal features of speech. Applying concurrent electrophysiological (EEG) and hemodynamic (near-infrared spectroscopy) recording to newborn infants listening to temporally structured nonspeech signals, we provide evidence that newborns process nonlinguistic acoustic stimuli that share critical temporal features with language in a differential manner. The newborn brain preferentially processes temporal modulations especially relevant for phoneme perception. In line with multi-time-resolution conceptions, modulations on the time scale of phonemes elicit strong bilateral cortical responses. Our data furthermore suggest that responses to slow acoustic modulations are lateralized to the right hemisphere. That is, the newborn auditory cortex is sensitive to the temporal structure of the auditory input and shows an emerging tendency for functional asymmetry. Hence, our findings support the hypothesis that development of speech perception is linked to basic capacities in auditory processing. From birth, the brain is tuned to critical temporal properties of linguistic signals to facilitate one of the major needs of humans: to communicate.

## Introduction

To acquire the specifically human faculty of language, infants face the challenging problem of being confronted with a complex auditory signal. Before the infant utters its first word, linguistic input is preferentially processed (Ramus et al., 2000; Vouloumanos and Werker, 2007). At birth newborns are able to distinguish their native language from other rhythmically different languages (Mehler et al., 1988; Kuhl, 2004) and show adult-like, preattentive discrimination of different vowels (Cheour-Luhtanen et al., 1995; Alho et al., 1998). They further differentiate between highly similar consonant–vowel syllables that vary in duration only by milliseconds (Molfese and Molfese, 1979; Bertonicini et al., 1987).

Notably, an early deficit in differentiating such rapid tempo-spectral variations is associated with specific language impairments in childhood (Benasich and Tallal, 2002). Hence, precise identification and analysis of acoustic cues are mandatory for language acquisition. However, which neuronal ensembles support these critical perceptual abilities gating language acquisition?

Whereas many studies have investigated the cortical networks underlying specific linguistic processes in adults (Friederici, 2002; Hickok and Poeppel, 2007), the emergence of this functional organization remains unclear. Pioneering studies showed that forward versus backward speech elicited greater vascular signal changes in left temporal brain regions in newborns and 3 month olds (Dehaene-Lambertz et al., 2002; Pena et al., 2003). Homae et al. (2006) demonstrated right-hemispheric specialization for processing prosodic information in 3 month olds. Evidently, hemispheric specialization already appears in early infancy. However, the origin of the lateralization remains controversial. That is, discrimination abilities between speech versus nonspeech sounds and between different rhythmic or prosodic structures might be driven by differences in basic acoustic properties of the speech signal itself. Several studies link hemispheric specialization for certain aspects of speech perception to an asymmetry in cortical tuning and reveal that the auditory cortices are differentially sensitive to spectrotemporal features (Zatorre and Belin,

Received March 13, 2009; revised July 13, 2009; accepted July 24, 2009.

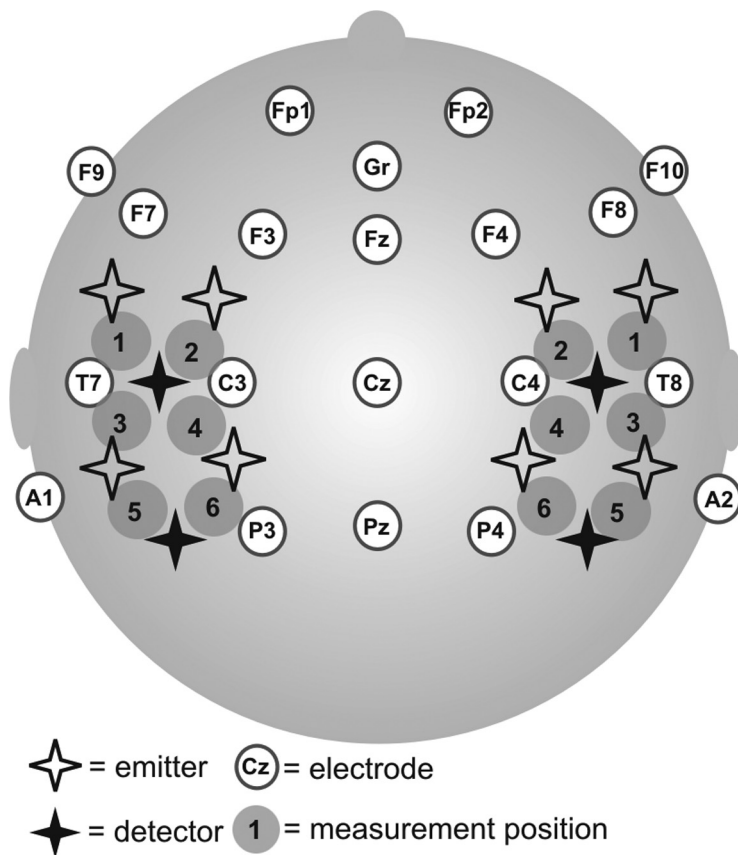
Financial support of the European Union (NEST 012778, EFRE 20002006 2/6, nEUROpt 201076) and Bundesministerium für Bildung und Forschung (Berlin NeuroImaging Center, Bernstein Center for Computational Neuroscience, German–Polish Cooperation FK: 01GZ0710) is gratefully acknowledged. D.P. is supported by National Institutes of Health Grant 2R01DC05660, I.W. by the Stifterverband für die Deutsche Wissenschaft (Claußen-Simon-Stiftung), and S.T. by Charité University Medicine. We thank Prof. Dr. Prof. h.c. mult. J. W. Dudenhausen, Prof. Dr. R. R. Wauer, and their staff from the Department of Obstetrics at Charité Berlin for their support. We would like to express our gratitude to all parents and their children who participated in this study. We thank N. K. Vavatzanidis for helping with data acquisition.

\*H.O. and I.W. contributed equally to this work.

Correspondence should be addressed to Silke Telkemeyer, Berlin NeuroImaging Center, Charité University Medicine, Charitéplatz 1, 10117 Berlin, Germany. E-mail: silke.telkemeyer@charite.de.

DOI:10.1523/JNEUROSCI.1246-09.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/2914726-08\$15.00/0



**Figure 1.** Details of the combined EEG and optical topography setup. EEG was recorded from 17 scalp positions according to the international 10–20 system (Gr = ground, A1/A2 = reference). The following six measurement positions per hemisphere for assessing the vascular response by optical imaging are represented by all available emitter–detector pairs (= measurement position): (1) inferior frontal, (2) superior frontal, (3) inferior temporal, (4) superior temporal, (5) posterior temporal, and (6) temporoparietal.

2001; Schönwiesner et al., 2005). Hickok and Poeppel (2007) postulate that decoding of speech is a multi-time-resolution process, integrating the speech signal analysis on (at least) two different timescales differentially relevant for specific aspects of linguistic analysis. The perception of suprasegmental modulations of language, like prosodic information carried by syllables, requires integration over longer timescales (~150–300 ms) and is expected to predominantly recruit the right-hemispheric auditory cortex. Rapid acoustic modulations conveying segmental information, fundamental to phonetic contrast perception, require the analysis over shorter temporal windows (~20–50 ms). This multi-time-resolution hypothesis has been modified based on growing experimental evidence (Poeppel et al., 2008). For example, Boemio et al. (2005) support these assumptions by demonstrating that right versus left superior temporal cortex shows greater blood oxygenation level-dependent (BOLD) signal changes in response to slow (150–300 ms) acoustic modulations (Overath et al., 2008; Warrier et al., 2009). Rapid temporal transitions, however, are processed bilaterally, leading to symmetrical activation of left and right auditory cortices.

The question of whether temporally modulated acoustic sounds elicit lateralized brain responses at birth has not yet been addressed. Using concurrent electrophysiological and hemodynamic recording, the present study aims at investigating the sensitivity of the auditory cortex in 3-d-old newborns for acoustic stimuli with varying temporal structure. Our results point to a—likely innate—neuronal specialization for basic temporal

characteristics of the auditory input which may be fundamental to language acquisition.

## Materials and Methods

### Subjects

Thirty-four healthy, full-term newborns (age  $3.32 \pm 1.27$  d; range 2–6 d; 16 boys) were tested. The newborns' physical condition was assessed by a local pediatrician with the standard newborn screening. This screening included measuring otoacoustic emissions to assess the newborns' hearing abilities: no hearing deficits, neurological disorders, or prenatal or perinatal complications were reported for any of the subjects. Average birth weight was  $3344 (\pm 591)$  g (range, 2365–4880 g), and the mean gestational age was  $40 (\pm 1.2)$  weeks (range, 38–42 weeks). Information on familial language impairment and handedness was obtained from both parents. In 85% ( $n = 29$ ) of subjects, both parents were right-handed, and in 15% ( $n = 5$ ), one parent was left-handed. In 15% ( $n = 5$ ) of the cases, one parent reported speech or language impairment during childhood [i.e., reading disabilities ( $n = 1$ ) and articulation problems ( $n = 4$ )].

Three subjects had to be excluded from further data analysis as a result of technical problems ( $n = 2$ ) or crying ( $n = 1$ ). Thus, optical topography data from 31 subjects were analyzed. For the electrophysiological recordings, only subjects for whom at least 50% of the analysis epochs (see below) survived artifact correction were included in further EEG analysis. Therefore, we included 22 subjects for further EEG analysis. This procedure corresponds to other previously reported electrophysiological infant studies (Dehaene-Lambertz and Dehaene, 1994).

Informed consent was acquired from both parents. The study protocol was approved of by the local ethics committee at Charité Berlin.

### Stimuli

The stimuli were selected from a set used in a published study investigating lateralization of complex acoustic properties in adults (Boemio et al., 2005). The selected stimuli varied within their temporal structure, whereas the spectral structure remained constant throughout the different conditions. Each tonal stimulus was created by concatenating limited-bandwidth noise segments of varying durations, the specific durations being commensurate with phonetic (shorter modulation frequency) versus syllabic (longer modulation) processing. Whereas each segment has a center frequency in the spectral range relevant for the discrimination of speech formants (1000–1500 Hz), the actual stimuli have no similarity to speech or natural vocalizations. Each 125 Hz bandwidth noise segment was modulated in its center frequency. Throughout the segments, this frequency remained constant. For each stimulus presentation, segment lengths of either 12, 25, 160, or 300 ms were concatenated to form a stimulation period of 9 s (corresponding to segment transition rates of ~83, 40, 6, and 3 segments per second). This does not allow us to delineate preferential sensitivities over the full range of temporal modulation frequencies. However, experimental time is limited in newborns and we focused on two specific “windows” of the continuum. These modulation frequencies were chosen because they encompass fast (phonetic) and slow (syllabic) transition rates of speech (Stevens, 1998). For further descriptions and details on the acoustic properties of the stimuli, see the study by Boemio et al. (2005). An example of each stimulus condition can be heard in supplemental audios 1–4, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material. In a pseudorandomized

order, a total of 23 stimuli per condition were presented with variable interstimulus intervals ranging from 1 to 12 s (mean, 4.1 s).

### Procedure

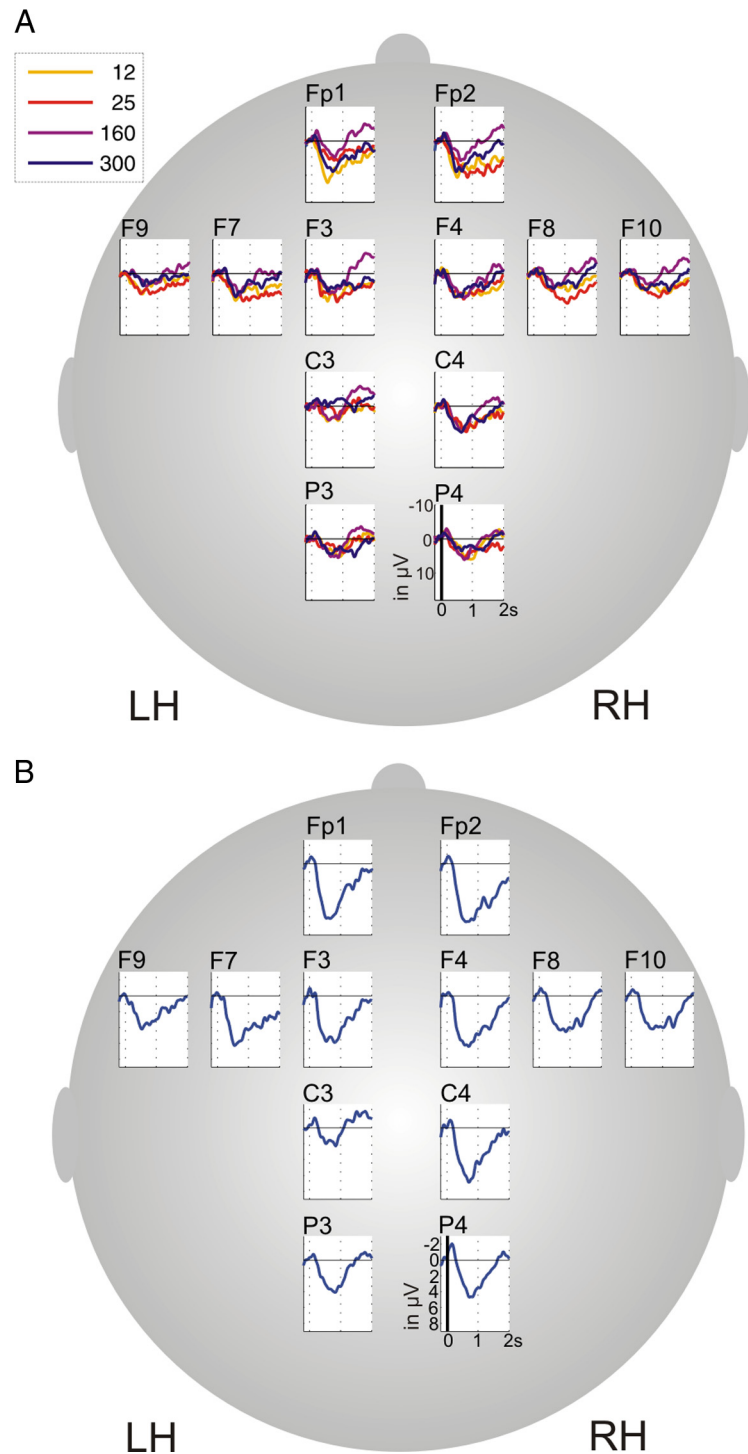
The newborns lay on a parent's lap while the stimuli were presented via two high-quality stereo speakers. The sound level was set to 70 dB<sub>A</sub>. Presentation software (V0.7.1, Neurobehavioral Systems) running on a Microsoft Windows 98 platform was used for controlled stimulus presentation. The total duration of the experiment was 20 min and the whole procedure was tolerated well by the newborns and parents. The experiment was discontinued if the infant showed any sign of discomfort. It should be noted that the very low dropout rate and the high number of acquired data sets indicate that the simultaneous registration of the vascular and the electrophysiological response by means of optical topography and EEG is generally appropriate for newborns.

### Data acquisition

**EEG.** EEG was recorded from 17 sintered Ag/AgCl electrodes (Brainproducts) from the following scalp positions [according to the international 10–20 system (Sharbrough et al., 1991)]: F3, F4, C3, C4, P3, P4, F7, F8, T7, T8, F9, F10, Fp1, Fp2, Fz, Cz, and Pz, online referenced against the left mastoid, with the AFz as the ground electrode (Fig. 1). EEG was recorded at a sampling rate of 1000 Hz by a direct-coupled amplifier (BrainAmp, Brainproducts).

**Optical topography.** To assess the cortical vascular response to the auditory stimuli, we measured cortical oxygenation changes using optical topography based on the principles of near-infrared spectroscopy. Light in the near infrared penetrates biological tissue rather well and thus allows tissue spectroscopy to several centimeters depth, reaching the cerebral cortex. Spectroscopic assessment of cortical concentration changes in oxygenated hemoglobin [oxy-Hb] and deoxygenated hemoglobin [deoxy-Hb] can be derived from changes in attenuation at two wavelengths. Event related focal decreases in [deoxy-Hb] correlate well with activations as detected in BOLD fMRI (Kleinschmidt et al., 1996). Further details of the methodology and the underlying physiology go beyond the scope of this article and were described in detail previously (Obrig and Villringer, 2003).

In the present study we used an optical topography system (Omniat Tissue Oxymeter; ISS) consisting of four light detectors and eight light emitters, separated by an interprobe distance of 2.5 cm. The volumes measured approximately correspond to the cortical area underlying each emitter–detector pair. As illustrated in Figure 1, we recorded six separate positions over each hemisphere covering frontotemporal to temporoparietal regions. Placement of the probe positions partially correspond to the 10–20 system (Sharbrough et al., 1991), as depicted in Figure 1 for the left and right hemisphere: (1) inferior frontal, (2) superior frontal, (3) inferior temporal, (4) superior temporal, (5) posterior temporal, and (6) temporoparietal. Fiber-optic bundles (emitter: 1 mm in diameter; detector: 3 mm) were fixed on the skull by adapting the probe heads to fit into the EEG cap (EASYCAP) used for the concurrently recorded EEG. Op-

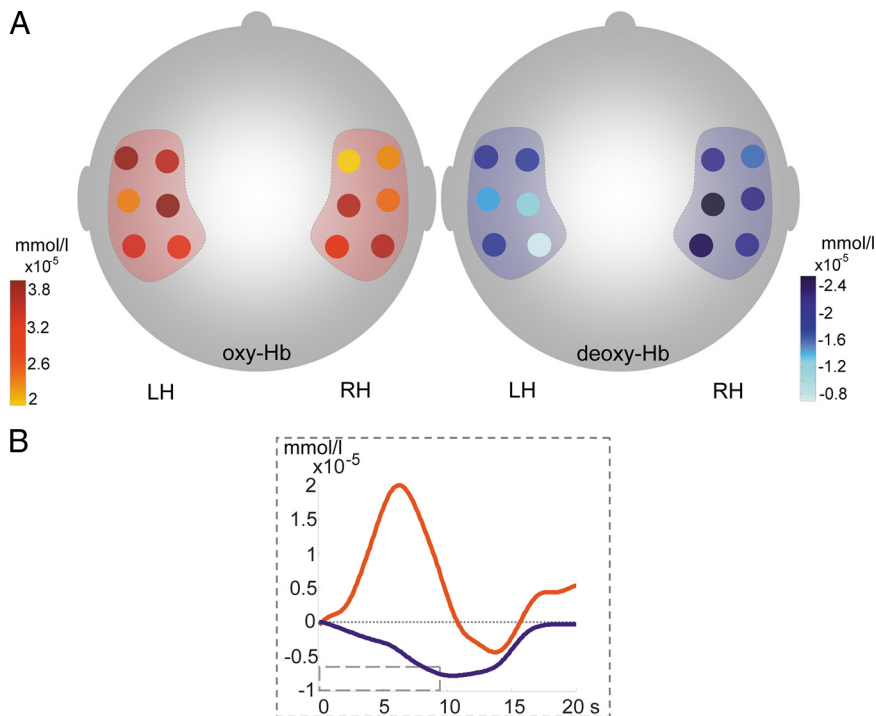


**Figure 2.** Results of the electrophysiological recordings. **A**, Grand average of the AEPs in response to each stimulus condition (12, 25, 160, and 300 ms). **B**, Grand average of the AEPs averaged across all stimulus conditions. LH, Left hemisphere; RH, right hemisphere.

tical topography data were continuously sampled at 10 Hz, and attenuation changes at 690 and 830 nm were converted into changes in [oxy-Hb] and [deoxy-Hb] by the modified Lambert–Beer approach (Cope and Delpy, 1988).

### Data analysis

**EEG.** Off-line analyses of the raw data were performed with BrainVision Analyzer software (Brainproducts). EEG data were downsampled to 500 Hz. The EEG data were filtered off-line using a 0.3 Hz low-cutoff, a 120 Hz high-cutoff, and a 50 Hz notch filter (bandwidth, 5 Hz; 24 dB/octave).



**Figure 3.** Hemodynamic response to sound versus silence periods. **A**, Grand average of [oxy-Hb] increase and [deoxy-Hb] decrease indicating brain activation averaged across all stimulus conditions compared with silence. Each circle represents one probe position. Concentration changes in [oxy-Hb] and [deoxy-Hb] in mmol/L. **B**, Time course of the grand average across all conditions averaged over all measured probe positions. The y-axis represents the concentration changes in millimoles per liter. The x-axis displays the time in seconds. The rectangle along the x-axis illustrates the time of stimulation. The increase in [oxy-Hb] is indicated by the red line, the decrease in [deoxy-Hb] by the blue line. LH, Left hemisphere; RH, right hemisphere.

For each stimulus we extracted an epoch from 200 ms before to 2500 ms after stimulus onset. A manual artifact rejection was performed, and contaminated segments were rejected. Data were re-referenced to the averaged left and right mastoids. Epochs were baseline corrected using a 200 ms prestimulus period (−200 ms to stimulus onset). For the EEG analysis we computed average event-related potentials (ERPs) for each stimulus condition for 2000 ms relative to stimulus onset.

For the statistical analysis of the ERPs, mean amplitudes were calculated within the time windows from 0 to 200, from 200 to 1000, and from 1000 to 1800 ms after stimulus onset and were analyzed using a repeated-measures ANOVA with the factors condition (12, 25, 160, and 300 ms), region (anterior: Fp1, F3; Fp2, F4; posterior: C3, P3; C4, P4), and hemisphere (left, right). A Greenhouse–Geisser correction was applied and is reported here as the corrected significance.

**Optical topography.** After converting attenuation changes into concentration changes of the hemoglobin, high-frequency components—mainly caused by the heartbeat—were attenuated by a low-pass filter at 0.3 Hz (Butterworth, third order). To correct drifts and slow fluctuations, an additional high-pass filter at 0.03 Hz was applied. Because efficient attenuation of movement artifacts is a special challenge in data recorded in infants, an artifact correction procedure was performed. First, a moving variance approach (Busch et al., 2006) was applied separately for each position to detect sudden and abnormal signal changes which mainly originate from movement artifacts. The procedure flagged those time points in which the z-scored data of a probe position within a moving 4 s time interval exceeded 2 SDs. After visual inspection of all positions, the flagged time points were replaced by a linear interpolation of valid data points encompassing the abnormal signal changes. We implemented this interpolation procedure to avoid overly excluding entire trials from data averaging and still attenuate noise stemming from movement of the newborn.

The decrease in [deoxy-Hb] is a reliable parameter for an increase in regional cerebral blood flow, because an increase in regional cerebral blood flow results in a faster washout (reduction) of [deoxy-Hb]. Addi-

tionally, the [deoxy-Hb] decrease is the most relevant physiological parameter for the BOLD signal and it correlates with an increase in the BOLD contrast (Kleinschmidt et al., 1996; Steinbrink et al., 2006). Thus, we focused on signal changes in [deoxy-Hb]. A number of infant studies reported in addition to decreases in [deoxy-Hb] selective increases in [oxy-Hb] or total hemoglobin. Our data, however, showed a reliable decrease in [deoxy-Hb] which is consistent with findings in fMRI studies in infants reporting increases in BOLD contrast in the language network (Dehaene-Lambertz et al., 2002, 2006). The subsequent analysis was based on a general linear model approach in analogy to “Statistical Parametric Mapping,” as is used for other hemodynamic imaging techniques. We computed hemodynamic concentration changes in response to the stimulus conditions. The boxcar functions corresponding to the four stimulus conditions were convolved with a standard hemodynamic response function (Boynton et al., 1996). In line with previous reports from vascular imaging in infants with the BOLD contrast and optical topography (Taga et al., 2003; Homae et al., 2006), we used parameters similar to those used in adult subjects, assuming a peak response at 4 s ( $\tau = 1$ ). Using a general linear model approach, we compared the predictors to the filtered data aiming at computing the  $\beta$  values for each probe position for the different conditions. For the statistical analysis of the  $\beta$  values we chose an ANOVA for repeated measures with the factors condition and hemisphere for each position and computed *post hoc*

paired *t* tests. Greenhouse–Geisser-corrected significances were reported. Because of a strong a priori hypothesis assuming lateralized processing of slow acoustic modulations and bilateral processing of fast acoustic modulations (Boemio et al., 2005; Giraud et al., 2007; Hickok and Poeppel, 2007) we computed one-tailed paired *t* tests for comparing hemispheric effects in those regions showing statistically significant hemispheric effects in the ANOVA. To visualize the time courses of concentration changes, we deconvolved the hemodynamic response of all conditions in those positions in which the ANOVA indicated statistically significant effects (see Fig. 4B).

## Results

We presented four different complex acoustic stimulus conditions with parametrically varying temporal structures to 31 newborns with a mean age of 3 d while assessing electrophysiological and vascular brain responses in a simultaneous EEG and optical topography setup. The stimuli consist of concatenated noise segments of varying length (i.e., 12, 25, 160, and 300 ms), forming four different stimulus conditions with 9-s-long stimuli with varying modulation frequency patterns (Boemio et al., 2005).

### Auditory evoked potentials

To prove that the four stimulus conditions were effective at eliciting robust basic acoustic response, we analyzed the event-related potentials [auditory evoked potentials (AEPs)] in response to the change of acoustic properties at the onset of each stimulus epoch. AEPs were averaged separately for the four stimulus conditions to assess whether the onsets of the different stimuli were processed differentially. We found a prominent positivity without any marked negative deflection in response to stimulus onset for all conditions. Figure 2A shows the grand average for each of the

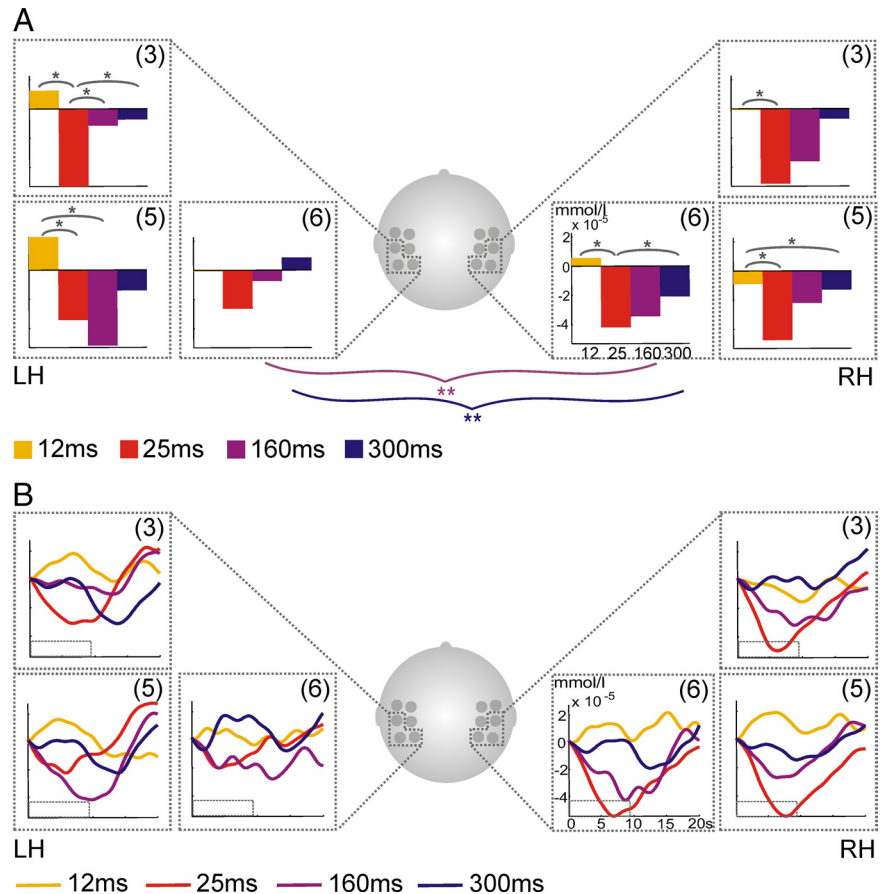
four stimulus conditions. The AEPs in response to each stimulus condition were characterized by a mean maximal amplitude of 5–6  $\mu\text{V}$  and peak latencies at  $\sim 800$  ms before returning to baseline 1800–2000 ms after the onset of the auditory stimulation. The mean AEP is distributed over bilateral frontal and central electrodes (Fig. 2*B*). The ANOVA did not reveal any significant differences between the four conditions in the three time windows from 0 to 200, from 200 to 1000, and from 1000 to 1800 ms in left and right anterior and posterior regions.

### Vascular response

We first averaged across all stimulus conditions to assess the hemodynamic response during sound compared with silence periods. Figure 3*A* shows the typical increase in [oxy-Hb] as well as the decrease in [deoxy-Hb] during the auditory stimulation periods. Figure 3*B* illustrates the mean time course across all conditions averaged over all measured probe positions with a peak at  $\sim 10$  s and a return to baseline some 15 s after stimulus onset. Our results correspond to the well known vascular response dynamics in adults (Cannestra et al., 2003), newborns (Taga et al., 2003), and older infants (Dehaene-Lambertz et al., 2006; Homae et al., 2006; Wartenburger et al., 2007).

Next, we analyzed whether the four stimulus conditions resulted in differential hemodynamic responses. We applied the general linear model and entered the  $\beta$  values for the four different stimuli into a multivariate analysis. The ANOVA yielded a statistically significant main effect of condition in inferior temporal ( $F_{(3,90)} = 3.91$ ,  $p < 0.02$ ) and posterior temporal ( $F_{(3,90)} = 3.71$ ,  $p < 0.02$ ) brain regions. The ANOVA furthermore revealed a significant main effect of hemisphere in the temporoparietal brain region ( $F_{(1,30)} = 4.03$ ,  $p < 0.05$ ). The signal changes in response to the four stimulus conditions in those probe positions with significant ANOVA main effects are illustrated in Figure 4.

The main effect of condition was analyzed *post hoc* using two-tailed paired  $t$  tests. The highest modulation rate stimulus (12 ms segments) elicited the weakest response over all positions investigated (Fig. 4*A*). That is, newborns showed no hemodynamic response modulation to this stimulus (increase in [deoxy-Hb] was not statistically significant), although the stimulus onset was recognized, as indicated by the AEPs. In contrast, the modulation of 25 ms segments—acoustically slightly slower than the 12 ms modulation and associated with a 40 Hz acoustic modulation rate—elicited the greatest bilateral response. When compared with the 12 ms modulation, the 25 ms modulation yielded significantly greater responses in the following brain regions: left inferior temporal ( $t_{(30)} = 3.74$ ,  $p < 0.01$ ) and posterior temporal ( $t_{(30)} = 3.29$ ,  $p < 0.01$ ), right inferior temporal ( $t_{(30)} = 2.64$ ,  $p = 0.01$ ) and posterior temporal ( $t_{(30)} = 2.50$ ,  $p = 0.02$ ), and temporoparietal ( $t_{(30)} = 2.97$ ,  $p = 0.01$ ) regions. Compared with the



**Figure 4.** Hemodynamic response to different temporal acoustic modulations. Grand average of [deoxy-Hb] decrease indicating brain activation for the four stimulus conditions (12, 25, 160, 300 ms) in the following measurement positions with significant ANOVA main effects: (3) inferior temporal; (5) posterior temporal; (6) temporoparietal area. Measurement positions are plotted following the same placement as in Figure 1*A*. The y-axis shows concentration changes of [deoxy-Hb] in millimoles per liter. The x-axis displays the stimulus condition. \* $p < 0.05$ , Statistically significant paired  $t$  tests between conditions; \*\* $p < 0.05$ , statistically significant paired  $t$  tests between hemispheres. **B**, Time courses of the [deoxy-Hb] decrease, indicating brain activation for the four stimulus conditions. The y-axis displays concentration changes of [deoxy-Hb] in millimoles per liter. The x-axis displays time in seconds. The rectangle along the x-axis indicates time of stimulation.

two slow modulation frequencies (segment lengths of 160 and 300 ms, respectively), the 25 ms modulation showed a statistically significant greater response in the left inferior temporal brain region (25 vs 160 ms:  $t_{(30)} = -2.46$ ,  $p = 0.02$ ; 25 vs 300 ms:  $t_{(30)} = -3.00$ ,  $p = 0.01$ ). Notably, there was no statistically significant difference between the two slowly modulated stimuli (160 vs 300 ms and vice versa). Because of the predicted direction of the hemispheric effects (see above), we computed one-tailed paired  $t$  tests investigating potential hemispheric lateralization effects for the superior temporoparietal brain region. The one-tailed  $t$  tests revealed no statistically significant hemispheric differences for the two fast modulated stimuli (12 ms:  $t_{(30)} = 0.08$ ,  $p = 0.47$ ; 25 ms:  $t_{(30)} = 1.41$ ,  $p = 0.08$ ). However, a significantly greater response was found in the right temporoparietal cortex region for the two slowly modulated stimuli (160 ms:  $t_{(30)} = 1.98$ ,  $p = 0.03$ ; 300 ms:  $t_{(30)} = 1.62$ ,  $p = 0.05$ ). Signal time courses for each stimulus condition are presented in Figure 4*B*.

### Discussion

Our study demonstrates that the auditory cortex of newborns exhibits differential sensitivity for processing varying temporal structure of acoustic signals. Newborn infants were presented

with four kinds of temporally varying stimuli while their brain activation was measured using EEG and optical topography.

The electrophysiological recordings show an evoked response (AEP) after stimulus onset which was similar for all stimulus conditions. The AEP is characterized by a large positive component with the maximal amplitude over frontocentral regions. Our findings are in line with previous work on both the auditory (Kushnerenko et al., 2002) and the visual (Lippe et al., 2007) evoked potentials of newborns, demonstrating monophasic and much slower evoked potentials when compared with the AEPs in adults; the results also replicate previous findings on AEPs in newborns (Dehaene-Lambertz and Pena, 2001; Picton and Taylor, 2007). AEPs indicate change detection during the initial analysis of the acoustic signal rather than reflecting the featural analysis of the stimuli (Fellman and Huotilainen, 2006). Thus, according to our assumptions, the analysis of the AEPs across stimulus conditions did not reveal statistically significant differences between conditions or hemispheres and indicate that all stimulus conditions were equally well detected by the newborns.

The vascular response averaged across all stimulus conditions compared with silence indicates an increased cortical activation during stimulus processing. The analysis of the vascular response for each stimulus condition demonstrates that the response in the auditory cortex was modulated by the four different temporal variations, indicating that the auditory cortex of newborns is sensitive to the temporal structure of a sustained acoustic signal. Notably, the differential vascular response to the four temporal variations cannot be explained by simple change perception, because we found similar AEPs in response to each stimulus condition (see above).

Comparing the amplitudes of the responses to the four stimulus conditions, the greatest response was elicited by the modulation of 25 ms, which can be considered the most relevant temporal frequency modulation with respect to speech processing at the phonetic level. The integration of 25–50 ms variations is of exceptional significance for the perception and encoding of speech, because this time window is relevant for the extraction of segmental information from the speech signal, for example the perception of phonemes and phonetic contrast (Rosen, 1992). This effect cannot be explained by a general increased attention to fast modulated stimuli, since the fastest modulated stimulus condition (12 ms) resulted in the weakest vascular response. The observation that temporal modulation of ~12 ms does not result in significant hemodynamic changes is in line with fMRI findings in adults (Boemio et al., 2005).

What underlying mechanisms guide the newborns' preference for the specific temporal modulation of 25 ms? Gerhardt and Abrams (2000) reported that *in utero* the fetus detects only vowels, whereas consonants, which are generally characterized by faster acoustic variations, are unavailable to the fetus. Therefore, the greatest bilateral response to the stimulus condition modulated with 25 ms can be judged as an indication of novelty preference. Notably, the first months of infancy represent a sensitive period to discriminate all phonetic units used in human languages. Our results might reflect this early sensitivity to rapid acoustic changes (Kuhl, 2004).

In the literature, the uniqueness and peculiarity of the 25 ms (= 40 Hz) modulation is also reflected by an electrophysiological response phenomenon, the so-called auditory steady-state response (ASSR). This oscillatory response originates mainly from primary auditory cortex and is driven by a periodical acoustic modulation of 40 Hz (Galambos et al., 1981). The ASSR is phase

locked to the rhythm of the auditory stimulus and follows its temporal fluctuations. Therefore, the ASSR is thought to be related to temporal processing in the auditory system, although the functional relevance of the ASSR phenomenon is still not fully understood (Picton et al., 2003). In our EEG results, we did not find a prominent ASSR for the 25 ms stimulus condition. This might be attributable to the only quasiperiodic modulation of our stimuli (i.e., mean around 25 ms segments but drawn from a distribution of different segment lengths) or to high variability of the ASSR signal and large background noise in the EEG signal of newborns (John et al., 2004). The prominent hemodynamic response to the modulation of the 25 ms stimulus condition in our study could be related to the ASSR phenomenon, thus indicating similar underlying neuronal processes. However, the question of whether there is a direct correspondence between the oscillatory ASSR and the hemodynamic response is still under debate because a direct comparison between the two measured signals is difficult as a result of the different temporal dynamics (Ross et al., 2005).

The 25 ms modulation yielded the largest hemodynamic response in five of the six temporal and temporoparietal positions, suggesting processing predominantly in bilateral auditory cortices. Although in the inferior temporoparietal position (Fig. 4, position 5) the response magnitude for the 25 ms stimulus seems slightly lateralized to the right, whereas the 160 ms stimulus appears to be left lateralized, neither difference proved significant in the statistical analysis. The only hemispheric difference revealed by the ANOVA was over the left versus right superior temporoparietal positions. Interestingly, at this specific location the 25 ms stimulus elicited a strong right-hemispheric response; however, the statistical analysis revealed only a marginal trend for a right-lateralized response. Ross et al. (2005) reported a right-hemispheric dominance of the 40 Hz ASSR obtained by magnetencephalography in response to monaurally and binaurally presented amplitude-modulated sounds in adults. This right-lateralized distribution of the ASSR seems to occur when attention is not specifically directed toward the stimulus (Elhilali et al., 2009). In addition to the disputable correspondence of the oscillatory ASSR and the hemodynamic responses, the observed effect in our results did not reach the significance level. Yet, given the broadly bilateral distributed signature of the 25 ms stimulus, we think that this finding may be interpreted along the assumptions of the multi-time-resolution model (Hickok and Poeppel, 2007). However, future research is needed to address longitudinal changes in the cortical representation of fast acoustic modulations during infancy. In contrast to the 25 ms stimulus, the 160 and 300 ms stimulus conditions elicited a focal right-lateralized response in the superior temporoparietal position only, emphasizing the focal sensitivity of the underlying right-hemispheric brain region for slowly modulated conditions. Although comparisons between adults and newborns and between different methodologies are limited, we consider this pattern found in the present study in line with the findings in adults using the exact same stimuli in adults (Boemio et al., 2005). Hence we interpret our data as an indication that the rightward specialization for processing slow modulations prevails from birth.

In sum, we consider our data to indicate an asymmetric processing predominantly of the slow modulations already present in newborns. Thus, our findings may be interpreted along psychoacoustic models proposing that a more general specialization for different acoustic properties of the incoming acoustic signal can be considered the basis for the observed lateralization of later

emerging speech perceptual functions. For example, Zatorre et al. (2002) argue that whereas left-hemispheric areas are predominantly involved during the processing of rapid temporal modulations, thus forming the basis for processing speech sounds, the superior temporal areas of the right hemisphere are specialized for the analysis of spectral information like the analysis of pitch change. A converging model by Poeppel (2003), with subsequent work by Poeppel et al. (2008), suggests that the temporospectral asymmetries result from differences in the size of the temporal integration windows of the neuronal ensembles in these brain areas. The model assumes that the bilateral auditory cortices are equally sensitive to fast acoustic modulations, whereas slow modulations are preferentially processed by the right-hemispheric auditory cortex. The results of our study can be seen in favor of these assumptions and indicate that the lateralization of speech might be a function of temporal signal properties. Our results shed new light on the conclusions derived from studies demonstrating that interhemispheric specialization for specific language functions are present from early infancy (Dehaene-Lambertz et al., 2002; Pena et al., 2003; Homae et al., 2006) by suggesting that differences in the acoustic and not the linguistic properties of the auditory input (at least at this age) might drive the observed lateralization of function.

Our data suggest that from birth the brain is sensitive to the temporal structure of the auditory input. Newborns show an increased response to fast acoustic modulations, especially in the range that is relevant for phoneme perception. Similar to a study in adults using the same stimulus material, our data suggest that the newborn's brain already shows a functional asymmetry for processing slow acoustic modulations, such as prosodic information, predominantly in the right hemisphere. Our results provide further evidence that from birth the brain seems to exhibit structural and functional properties especially tuned to language, to facilitate one of the major needs of humans: to communicate.

## References

- Alho K, Connolly JF, Cheour M, Lehtokoski A, Huotilainen M, Virtanen J, Aulanko R, Ilmoniemi RJ (1998) Hemispheric lateralization in preattentive processing of speech sounds. *Neurosci Lett* 258:9–12.
- Benasich AA, Tallal P (2002) Infant discrimination of rapid auditory cues predicts later language impairment. *Behav Brain Res* 136:31–49.
- Bertoncini J, Bijeljac-Babic R, Blumstein SE, Mehler J (1987) Discrimination in neonates of very short CVs. *J Acoust Soc Am* 82:31–37.
- Boemio A, Fromm S, Braun A, Poeppel D (2005) Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat Neurosci* 8:389–395.
- Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16:4207–4221.
- Busch NA, Herrmann CS, Müller MM, Lenz D, Gruber T (2006) A cross-laboratory study of event-related gamma activity in a standard object recognition paradigm. *Neuroimage* 33:1169–1177.
- Cannestra AF, Wartenburger I, Obrig H, Villringer A, Toga AW (2003) Functional assessment of Broca's area using near infrared spectroscopy in humans. *Neuroreport* 14:1961–1965.
- Cheour-Luhtanen M, Alho K, Kujala T, Sainio K, Reinikainen K, Renlund M, Aaltonen O, Eerola O, Näätänen R (1995) Mismatch negativity indicates vowel discrimination in newborns. *Hear Res* 82:53–58.
- Cope M, Delpy DT (1988) System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination. *Med Biol Eng Comput* 26:289–294.
- Dehaene-Lambertz G, Dehaene S (1994) Speed and cerebral correlates of syllable discrimination in infants. *Nature* 370:292–295.
- Dehaene-Lambertz G, Pena M (2001) Electrophysiological evidence for automatic phonetic processing in neonates. *Neuroreport* 12:3155–3158.
- Dehaene-Lambertz G, Dehaene S, Hertz-Pannier L (2002) Functional neuroimaging of speech perception in infants. *Science* 298:2013–2015.
- Dehaene-Lambertz G, Hertz-Pannier L, Dubois J, Mériaux S, Roche A, Sigman M, Dehaene S (2006) Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proc Natl Acad Sci U S A* 103:14240–14245.
- Elhilali M, Xiang J, Shamma SA, Simon JZ (2009) Interaction between attention and bottom-up saliency mediates the representation of foreground and background in an auditory scene. *PLoS Biol* 7:e1000129.
- Fellman V, Huotilainen M (2006) Cortical auditory event-related potentials in newborn infants. *Semin Fetal Neonatal Med* 11:452–458.
- Friederici AD (2002) Towards a neural basis of auditory sentence processing. *Trends Cogn Sci* 6:78–84.
- Galambos R, Makeig S, Talmachoff PJ (1981) A 40-Hz auditory potential recorded from the human scalp. *Proc Natl Acad Sci U S A* 78:2643–2647.
- Gerhardt KJ, Abrams RM (2000) Fetal exposures to sound and vibroacoustic stimulation. *J Perinatol* 20:S21–S30.
- Giraud AL, Kleinschmidt A, Poeppel D, Lund TE, Frackowiak RS, Laufs H (2007) Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron* 56:1127–1134.
- Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* 8:393–402.
- 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.
- John MS, Brown DK, Muir PJ, Picton TW (2004) Recording auditory steady-state responses in young infants. *Ear Hear* 25:539–553.
- Kleinschmidt A, Obrig H, Requardt M, Merboldt KD, Dirnagl U, Villringer A, Frahm J (1996) Simultaneous recording of cerebral blood oxygenation changes during human brain activation by magnetic resonance imaging and near-infrared spectroscopy. *J Cereb Blood Flow Metab* 16:817–826.
- Kuhl PK (2004) Early language acquisition: cracking the speech code. *Nat Rev Neurosci* 5:831–843.
- Kushnerenko E, Ceponiene R, Balan P, Fellman V, Huotilainen M, Näätänen R (2002) Maturation of the auditory event-related potentials during the first year of life. *Neuroreport* 13:47–51.
- Lippé S, Roy MS, Perchet C, Lassonde M (2007) Electrophysiological markers of visuocortical development. *Cereb Cortex* 17:100–107.
- 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.
- Molfese DL, Molfese VJ (1979) Hemisphere and stimulus differences as reflected in the cortical responses of newborn infants to speech stimuli. *Dev Psychol* 15:505–511.
- Obrig H, Villringer A (2003) Beyond the visible—imaging the human brain with light. *J Cereb Blood Flow Metab* 23:1–18.
- Overath T, Kumar S, von Kriegstein K, Griffiths TD (2008) Encoding of spectral correlation over time in auditory cortex. *J Neurosci* 28:13268–13273.
- 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.
- Picton TW, Taylor MJ (2007) Electrophysiological evaluation of human brain development. *Dev Neuropsychol* 31:249–278.
- Picton TW, John MS, Dimitrijevic A, Purcell D (2003) Human auditory steady-state responses. *Int J Audiol* 42:177–219.
- Poeppel D (2003) The analysis of speech in different temporal integration windows: cerebral lateralization as “asymmetric sampling in time.” *Speech Commun* 41:245–255.
- Poeppel D, Idsardi WJ, van Wassenhove V (2008) Speech perception at the interface of neurobiology and linguistics. *Philos Trans R Soc Lond B Biol Sci* 363:1071–1086.
- Ramus F, Hauser MD, Miller C, Morris D, Mehler J (2000) Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science* 288:349–351.
- Rosen S (1992) Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond B Biol Sci* 336:367–373.
- Ross B, Herdman AT, Pantev C (2005) Right hemispheric laterality of human 40 Hz auditory steady-state responses. *Cereb Cortex* 15:2029–2039.
- Schönwiesner M, Rübsem 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.

- Sharbrough F, Chatrian GE, Lesser RP, Lüders H, Nuwer M, Picton TW (1991) American Electroencephalographic Society guidelines for standard electrode position nomenclature. *J Clin Neurophysiol* 8:200–202.
- Steinbrink J, Villringer A, Kempf F, Haux D, Boden S, Obrig H (2006) Illuminating the BOLD signal: combined fMRI-fNIRS studies. *Magn Reson Imaging* 24:495–505.
- Stevens KN (1998) Acoustic phonetics. Cambridge, MA: MIT.
- Taga G, Asakawa K, Hirasawa K, Konishi Y (2003) Hemodynamic responses to visual stimulation in occipital and frontal cortex of newborn infants: a near-infrared optical topography study. *Early Hum Dev [Suppl]* 75:S203–S210.
- Vouloumanos A, Werker JF (2007) Listening to language at birth: evidence for a bias for speech in neonates. *Dev Sci* 10:159–164.
- Warrier C, Wong P, Penhune V, Zatorre R, Parrish T, Abrams D, Kraus N (2009) Relating structure to function: Heschl's gyrus and acoustic processing. *J Neurosci* 29:61–69.
- 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.
- Zatorre RJ, Belin P (2001) Spectral and temporal processing in human auditory cortex. *Cereb Cortex* 11:946–953.
- Zatorre RJ, Belin P, Penhune VB (2002) Structure and function of auditory cortex: music and speech. *Trends Cogn Sci* 6:37–46.