



# Acoustic processing of temporally modulated sounds in infants: evidence from a combined near-infrared spectroscopy and EEG study

Silke Telkemeyer<sup>1,2,3,4\*</sup>, Sonja Rossi<sup>3,5</sup>, Till Nierhaus<sup>3,5</sup>, Jens Steinbrink<sup>3</sup>, Hellmuth Obrig<sup>3,5†</sup> and Isabell Wartenburger<sup>1,3,4†</sup>

<sup>1</sup> Languages of Emotion Cluster of Excellence, Freie Universität Berlin, Berlin, Germany

<sup>2</sup> Department of Cognitive Psychology, Humboldt-Universität Berlin, Berlin, Germany

<sup>3</sup> Berlin Neuroimaging Center, Department of Neurology, Charité University Medicine, Berlin, Germany

<sup>4</sup> Department of Linguistics, University of Potsdam, Potsdam, Germany

<sup>5</sup> Department of Cognitive Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, University Hospital, Leipzig, Germany

## Edited by:

Judit Gervain, University Paris Descartes, France

## Reviewed by:

Marcela G. Pena, University of Chile, Chile

Martin Meyer, University of Zurich, Switzerland

## \*Correspondence:

Silke Telkemeyer, Languages of Emotion Cluster of Excellence, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany.  
e-mail: silke.telkemeyer@fu-berlin.de

<sup>†</sup>Hellmuth Obrig and Isabell Wartenburger have contributed equally to this work.

Speech perception requires rapid extraction of the linguistic content from the acoustic signal. The ability to efficiently process rapid changes in auditory information is important for decoding speech and thereby crucial during language acquisition. Investigating functional networks of speech perception in infancy might elucidate neuronal ensembles supporting perceptual abilities that gate language acquisition. Interhemispheric specializations for language have been demonstrated in infants. How these asymmetries are shaped by basic temporal acoustic properties is under debate. We recently provided evidence that newborns process non-linguistic sounds sharing temporal features with language in a differential and lateralized fashion. The present study used the same material while measuring brain responses of 6 and 3 month old infants using simultaneous recordings of electroencephalography (EEG) and near-infrared spectroscopy (NIRS). NIRS reveals that the lateralization observed in newborns remains constant over the first months of life. While fast acoustic modulations elicit bilateral neuronal activations, slow modulations lead to right-lateralized responses. Additionally, auditory-evoked potentials and oscillatory EEG responses show differential responses for fast and slow modulations indicating a sensitivity for temporal acoustic variations. Oscillatory responses reveal an effect of development, that is, 6 but not 3 month old infants show stronger theta-band desynchronization for slowly modulated sounds. Whether this developmental effect is due to increasing fine-grained perception for spectrotemporal sounds in general remains speculative. Our findings support the notion that a more general specialization for acoustic properties can be considered the basis for lateralization of speech perception. The results show that concurrent assessment of vascular based imaging and electrophysiological responses have great potential in the research on language acquisition.

**Keywords: infants, speech perception, language acquisition, auditory processing, near-infrared spectroscopy, event related potentials, brain oscillations**

## INTRODUCTION

The analysis of acoustic features in the continuous auditory speech stream is a prerequisite for language acquisition in infancy. Among other functions it serves the segmentation of the speech stream into smaller units, like words and phrases (Mehler et al., 2004; Gervain and Mehler, 2010). This very early step of speech perception necessitates temporal and spectral differentiation of the acoustic input. In the context of speech the differentiation of the temporal structure of the acoustic input is critical, as illustrated by the clinical finding that infants with a deficit in differentiating rapidly varying auditory stimuli are more likely to develop a specific language impairment (Benasich and Tallal, 2002; Choudhury et al., 2007). The relevance of categorical acoustic feature analysis during early language acquisition is uncontroversial. However, knowledge on the underlying neuronal network and its maturation during early development is sparse. While in adults the brain clearly relies on functionally specialized areas to process speech (and language) little is known on

how this efficient network develops from birth and which “inborn” foundations endow the human brain with the unique ability to reach language competence. In newborns and 3 month old infants seminal work using functional magnetic resonance imaging (fMRI) and near-infrared spectroscopy (NIRS) demonstrated asymmetrical responses to forward compared to backward speech especially in the left angular gyrus (Dehaene-Lambertz et al., 2002, 2006; Pena et al., 2003). Additionally, a larger sensitivity of right hemispheric fronto-temporal regions in response to prosodic features has been demonstrated already in 3 month old infants (Homae et al., 2006). The dominance of the right hemispheric auditory cortex for music processing, which relies on melodic and concise pitch information (Zatorre and Belin, 2001), has been recently shown to be present from birth: A NIRS study in newborns revealed right-lateralized activation during the presentation of music excerpts (Perani et al., 2010). Taken together there is converging evidence that basic aspects of lateralization in the network in response to complex auditory

stimuli, necessary for speech and music comprehension, evolve very early. However, because very young infants clearly lack linguistic and musical knowledge, an intriguing question is how acoustic features may “guide” the lateralized processing. With regard to language, psychoacoustic models propose lateralized auditory processing as a more general basis of the lateralization in the language network. Supported by lesion and functional imaging data in adults, these models highlight that hemispheric specialization for different aspects of language processing may partially be driven by the auditory analysis. In this vein different psychoacoustic models proposed functional asymmetries based on spectral and/or temporal feature analysis (Zatorre and Belin, 2001; Poeppel, 2003; Zaehle et al., 2004; Schönwiesner et al., 2005). Though they stress different aspects of the functional anatomy of processing complex auditory stimuli they partly converge because they posit differential specializations for features of such stimuli. As an example, the multi-time resolution model (Hickok and Poeppel, 2007; Poeppel et al., 2008) postulates at least two temporal integration windows for the processing of the auditory speech input, which operate in parallel. According to the model, the integration of rapidly varying acoustic features (20–50 ms window) fundamental for the perception of phonetic contrasts, recruits areas in the left and right auditory cortices. Conversely, modulations of the acoustic signal at slower rates (150–300 ms) – more relevant for suprasegmental feature analysis (e.g., prosody) – are predominantly processed in the right hemisphere. Using noise stimuli that were modulated at different temporal rates within the predicted windows, Boemio et al. (2005) confirmed these predictions in an fMRI study in adults. In newborns a NIRS study used the same stimuli to explore whether similar lateralization can be found during a very early stage of brain development (Telkemeyer et al., 2009). The results indicate that already the newborn’s auditory cortex is sensitive to the different temporal features of the acoustic input. In particular we could show a right hemispheric lateralization for slow acoustic modulations to be present at birth. These results support the notion that basic acoustic features within the speech signal drive the hemispheric lateralization from the earliest stages of language acquisition.

Lateralization of speech perception based on its auditory features and the acquisition of linguistic competence may interact. During human development the lateralization of linguistic processing increases and consolidates (Holland et al., 2001). Experimental evidence for a successive lateralization of linguistic contrasts during early development has been supplied by longitudinal studies investigating how an initially bilateral processing of a phonemic contrast is progressively lateralized with increasing age (Minagawa-Kawai et al., 2007). Since the infant passes crucial milestones of native language acquisition in the first 6 months (Kuhl et al., 1992; Kuhl, 2004; Friederici, 2005) and changes in the underlying neuronal mechanisms have been demonstrated (Kuhl and Rivera-Gaxiola, 2008), we focused on the development of auditory feature analysis of non-linguistic stimuli in this age group.

The rationale was that processing of non-linguistic contrasts, potentially supporting the evolving lateralization of speech perception can be investigated during the maturation of the network by temporally modulated noise stimuli (Boemio et al., 2005; Telkemeyer et al., 2009). In 6 and 3 month old infants we measured the hemodynamic and electrophysiological brain responses

using simultaneous NIRS and EEG. With regard to lateralization we expected to find a pattern in both, the 3 and 6-months age-group, similar to that reported in our previous study in newborns (Telkemeyer et al., 2009). Oxygenation changes as measured by NIRS should lateralize to the right hemisphere for the slowly modulated stimuli in both age groups.

In our study in newborns we did not find differences in the electrophysiological signals simultaneously assessed (Telkemeyer et al., 2009). The transient event-related potentials (ERP) elicited by the onset of the auditory stimulus reliably showed that infants process the auditory input. It did, however, not show sensitivity for the different modulation frequencies. Also a time-frequency analysis (TFA) of the electrophysiological data did not yield a reliable effect in response to the auditory stimulus. In newborns this may be due to a discontinuous EEG (Lippe et al., 2009) which does not allow for analyses as used in data from adults (e.g., Hoogenboom et al., 2006; Koch et al., 2009). However, a recent study by Pena et al. (2010) showed that gamma-band response and ERPs evidence differential processing of linguistic input in infants. They used stimuli in the infants’ native language and stimuli in languages of rhythmically similar or grossly different classes and compared 3 and 6 month old full-term infants to 6 and 9 month old preterm infants. The results are that increased gamma-band power in response to the native language is present at 6 months in full-term but only at 9 months in preterm infants. These results elegantly corroborate the hypothesis that neuronal maturation plays a pivotal role in the earliest steps of language acquisition, most likely due to acoustic features such as rhythmic class. In the present study we report on the results of ERPs and oscillatory EEG response which nicely extend these findings to the processing of non-linguistic complex auditory material. This may be of great relevance to better understand the interaction between language acquisition and auditory feature analysis in the auditory cortex in the first 6 months of life.

With regard to the EEG data we therefore analyzed two parameters and their dependence on neuronal maturation:

- (i) We analyzed the AEP to find out how the latency and amplitude of two major components (N1, P2) change with age (Kushnerenko et al., 2002; Choudhury and Benasich, 2010). Here we expected differences in latency and amplitude of the early AEP components between 3 and 6 month old infants and were interested in whether AEPs would differ between the different modulation frequencies. Studies in adult subjects demonstrated that auditory non-speech stimuli with different grades of temporal variations modulate the evoked electrophysiological response (Zaehle et al., 2007, 2009). Unless in newborns we did not find effects of temporal modulation frequency on the AEP, the above findings (Pena et al., 2010) suggest that this may evolve between 3 and 6 months.
- (ii) We analyzed the oscillatory activity of neural assemblies, which are more sensitive to the sustained activity in response to the stimuli (Shahin et al., 2010). Brain oscillatory systems act as neural communication mechanisms, dynamically integrating signals from different brain regions and are a potential neuronal correlate of feature binding. They are also known to play a crucial role in attentional processes (Lopes da Silva, 1991; Singer, 1993, 1999). Information on the synchronization

and desynchronization between local and distant neuronal ensembles may hence inform our understanding of whether and how sustained auditory features are processed by the infant's brain. Comparing signal power of cortical oscillations in specific frequency bands after stimulus onset to a pre-stimulus (baseline) interval enables quantification of oscillatory EEG responses to the stimulus. Event-related de-/synchronization of neuronal ensembles results in decreased/increased oscillatory activity, respectively (Pfurtscheller and Lopes da Silva, 1999). This is the more relevant because the electrophysiological de/synchronization can be tonic, that is, it provides us with a measure of neuronal activity over the full duration of a stimulus. In the present study we therefore also ask whether the temporal acoustic modulations are reflected in the oscillatory electrophysiological response. In adults a simultaneous EEG and fMRI study showed that the spontaneous gamma power (28–40 Hz) correlates with activation in the left auditory cortex, while the fluctuations in the theta-range (3–6 Hz) correlate with BOLD-contrast changes in the right hemisphere (Giraud et al., 2007). These data nicely fit into the theory of parallel lateralized processing of slow and fast modulation frequencies (Poeppel et al., 2008). In infants the recent study on rhythmic classes in early language perception suggest that the analysis of oscillatory brain responses to auditory stimulation in young infants may be a promising tool to understand maturational processes (Pena et al., 2010). This approach has only rarely been used (e.g., Stefanics et al., 2007; Lippe et al., 2009) and we are aware of the explorative nature of such an analysis potentially complementing our ERP and NIRS results.

In sum, our study investigates the perception of basic auditory precursors relevant to the decoding of speech in infancy. The novel approach to combine EEG and NIRS measurements allows us to monitor the temporal and topographic aspects of neuronal processing. Using the same set-up and experimental procedure as our previous study in newborns and sharing the stimuli with an fMRI study in adults may be of relevance for a better longitudinal understanding of how auditory feature analysis of linguistic material is shaped by more basic principles of auditory perception.

## MATERIALS AND METHODS

### SUBJECTS

We examined two groups of infants: 6 and 3 month olds. Generally the protocol was similar to a previous study we conducted in newborns (Telkemeyer et al., 2009). We acquired informed consent from both parents. The study protocol was approved by the local ethics committee of the Charité University Medicine Berlin.

#### *Six month old infants*

In this group we measured 44 healthy infants (mean age = 185 days,  $\pm 9.5$  days; 17 boys). Their mean gestational age was 40 weeks ( $\pm 1.4$  weeks), and their average birth weight was 3367 g ( $\pm 487$  g). Information on familial language development and handedness was obtained from the parents. In 82% ( $n = 36$ ) of the subjects both parents were right-handed, in 18% ( $n = 8$ ) one parent was left-handed. In 9% ( $n = 4$ ) one parent reported some kind of language

impairment (e.g., articulation/reading problems) during childhood. In 5% ( $n = 2$ ) both parents reported some kind of language impairment during childhood.

Five subjects were excluded from further analysis because the experiment was ended when the infant showed signs of discomfort. Another three infants were excluded from the NIRS analysis as a result of technical problems during data acquisition. Thus, 36 infants entered the NIRS analysis.

For the EEG analysis we included all subjects in whom at least 50% of the segments survived the artifact correction procedure (see section Data Analysis). Twenty-three subjects fulfilled this criteria and were included in the final EEG analysis.

#### *Three month old infants*

In this group 40 healthy infants (mean age = 94 days,  $\pm 10.3$  days; 22 boys) were measured. They had a mean gestational age of 40 weeks ( $\pm 1.4$  weeks), and their average birth weight was 3564 g ( $\pm 529$  g). Information on familial language development and handedness revealed that in 80% ( $n = 32$ ) of the subjects both parents were right-handed, and in 20% ( $n = 8$ ) one parent was left-handed. In 7.5% ( $n = 3$ ) one parent reported some kind of language impairment during childhood, in one case language impairments were reported by both parents.

Two infants showed signs of discomfort leading to a discontinuation of the experiment. Thus, data of 38 subjects entered the NIRS analysis. After artifact correction of the EEG data, the data of 22 infants survived our criteria (more than 50% segments after artifact correction) and entered the final EEG analysis.

### STIMULI

In analogy to our previous study in newborns (Telkemeyer et al., 2009) we selected four different auditory stimuli from a larger stimulus-set published by Boemio et al. (2005). The tonal stimuli with a total duration of 9 s each are formed by concatenated noise-segments. Each noise-segment has a center frequency in the spectral range relevant for the discrimination of speech formants (1000–1500 Hz). While this spectral information is largely kept constant over the stimulus conditions, the temporal modulation of the segments was manipulated. This was achieved by modulating each noise-segment with a bandwidth of 125 Hz around its central frequency thereby yielding segments of varying length. In the present study and our previous study in newborns, segment lengths of 12, 25, 160, or 300 ms were assembled, thus forming four stimulus conditions differing in their temporal modulation. Because experimental time is limited in infants, we did not use the whole range of stimulus conditions used by Boemio et al. (2005), but selected two specific “time-windows” of temporal modulation frequencies: (1) fast (12 and 25 ms) acoustic modulations that correspond to fast such as phonetic modulations within the speech stream; (2) slow (160 and 300 ms) acoustic modulations which are associated with slow (e.g., syllabic) variations within the speech signal (Stevens, 1998). While the two fast acoustic modulations correspond to modulation frequencies of 83 and 40 Hz, the two slowly modulated stimulation patterns correspond to modulation frequencies of 6 and 3 Hz. We presented 23 stimuli per condition with variable inter-stimulus intervals (ISI) ranging from 1 to 12 s (mean 4.1 s) in a pseudo-randomized order. Please note, that the

two fast (12 and 25 ms) and the two slow (160 and 300 ms) acoustic modulations were pooled together for the data analysis (NIRS as well as EEG) to achieve a larger number of trials per condition (i.e., 46 fast modulated stimuli and 46 slowly modulated stimuli).

For audio-examples of the stimuli, please refer to Boemio et al. (2005) and Telkemeyer et al. (2009).

## PROCEDURE

Throughout the experiment the infants sat on their parent's lap. To keep the experiment as transparent as possible parents were not acoustically masked. We considered undesirable influences by the parents' behavior in response to the stimuli relatively unlikely, because the stimulus material consisted of artificial, noise-like stimuli. To sustain the infants' attention a silent video of moving objects was shown temporally unrelated to the acoustic presentation. The auditory stimuli were presented via two stereo speakers (sound level of 70 dB). Stimulus presentation was controlled by Presentation software (V0.7.1, Neurobehavioral Systems). The experiment consisted of two blocks of 10 min each, separated by a variable break that could be used to interact with the infant or parent if necessary. The total duration of the experiment was approximately 20 min. The experiment was interrupted whenever the infant showed any sign of discomfort, and continued only if infant and parent were willing to further participate.

## DATA ACQUISITION

### *Near-infrared spectroscopy*

Cortical oxygenation changes in response to the auditory stimulation were assessed by NIRS. Near-infrared light ( $\lambda = \sim 600\text{--}900$  nm) penetrates biological tissue up to several centimeters depth, reaching the cerebral cortex, when applied on the head. Models of the neuro-vascular coupling (Fox and Raichle, 1986) predict that increases in neuronal activation lead to an increase in regional cerebral blood flow overcompensating the local demand in oxygen. This results in a focal cortical hyperoxygenation which translates into an increase in oxygenated hemoglobin (oxy-Hb) and a decrease in deoxygenated hemoglobin (deoxy-Hb) concentrations. It should be noted that an increase in regional cerebral blood volume and an increase in blood flow velocity is expected in an activated cortical area (Fox and Raichle, 1986). With regard to the NIRS parameters that translates into an increase in oxy-Hb and a decrease in deoxy-Hb (Obrig and Villringer, 2003). The debate on which parameter is more "powerful" is on sensitivity and specificity. Sensitivity is larger for oxy-Hb, partially due to the larger amplitude. Specificity, however, is larger for deoxy-Hb, because an increased washout of deoxy-Hb in an activated area can be considered a specific feature of the cerebral hemodynamic response as opposed to changes in hemodynamics in the extracerebral tissue (Boden et al., 2007). NIRS unfortunately is extremely sensitive to changes in extracerebral hemodynamic changes. Therefore we advocate to always report deoxy-Hb changes also, because deoxy-Hb decreases are the major source of the BOLD-contrast (Steinbrink et al., 2006). The matter is complicated by the debate on the "typical" response pattern in infants, for a more detailed discussion see Obrig et al. (2010). We here report both increases in oxy-Hb and decreases in deoxy-Hb.

Technically light of two different wavelengths is guided to and from the subject's head by fiber-optic bundles. Detector probes are placed pairwise some 2–3 cm from the emitting probes to collect the reflected light. Each source-detector pair defines a sampling volume. Focal changes in oxy- and deoxy-Hb are derived from the changes in attenuation measured at two wavelengths, based on the modified Beer–Lambert law (Cope and Delpy, 1988). Event-related decreases in deoxy-Hb correlate well with BOLD-contrast increases, termed "activation" in the fMRI literature (Kleinschmidt et al., 1996; Obrig and Villringer, 2003).

We used a NIRS system (Omniat Tissue Oxymeter, ISS, USA) consisting of four light detectors and eight light emitters. The instrument works with modulated light sources at 690 and 830 nm. Raw data were sampled at a rate of 10 Hz. All optical probes and the EEG electrodes were integrated into an EEG cap (EASYCAP, Germany). Emitter and detector probes for the NIRS measurement were separated by an interprobe distance of 2.5 cm. The NIRS array resulted in 6 measurement volumes over each hemisphere (see **Figure 1**): (1) inferior frontal, (2) superior frontal, (3) inferior temporal, (4) superior temporal, (5) posterior temporal, and (6) temporo-parietal. The probe placement paralleled the EEG electrode placement and partially corresponded to positions of the 10–20 system (Sharbrough et al., 1991).

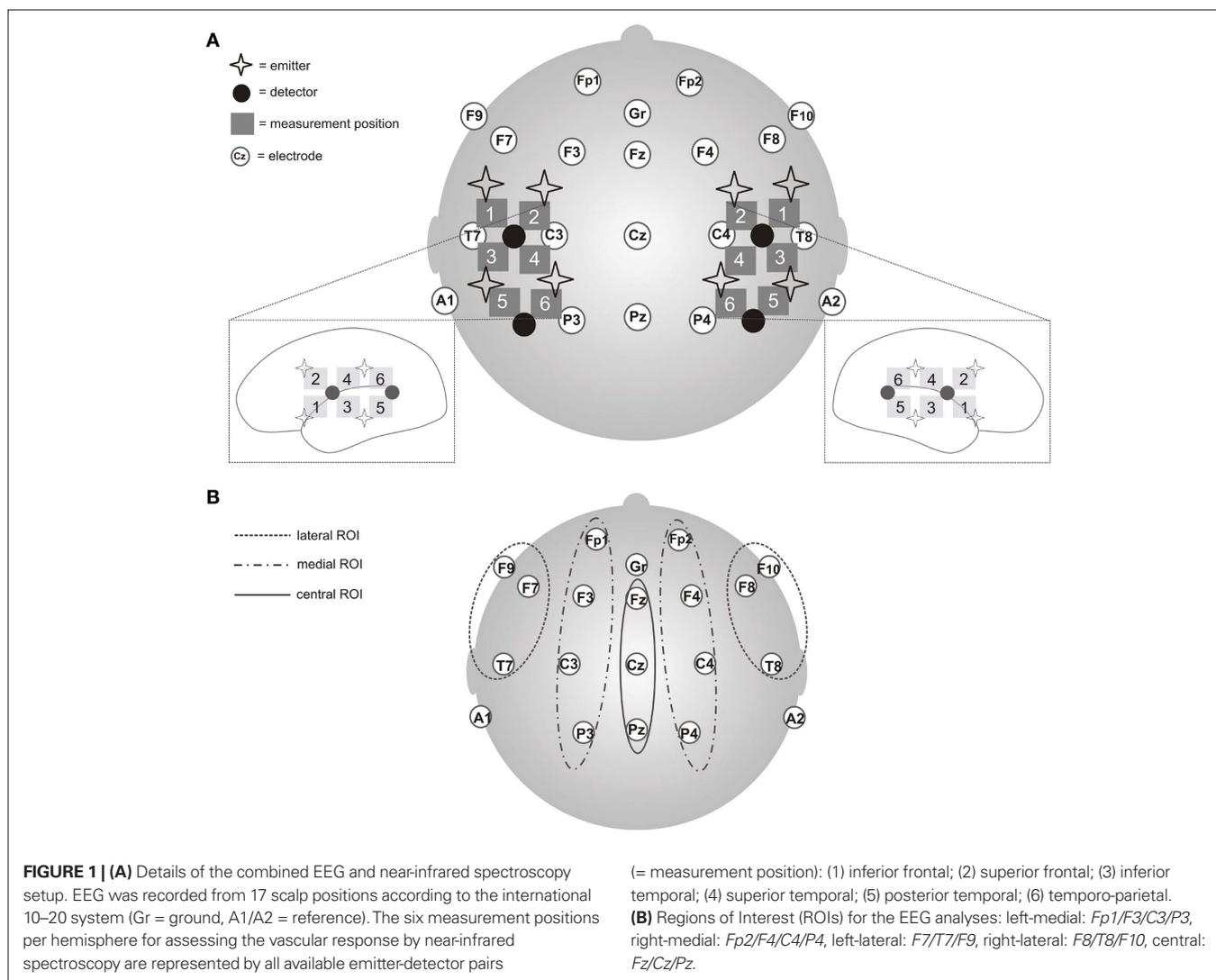
### *Electroencephalography*

Electroencephalography was recorded with 17 Ag/AgCl electrodes (Brainproducts, Germany) also mounted with the elastic EEG cap (EASYCAP, Germany). Electrodes were located according to the 10–20 system (Sharbrough et al., 1991) at the following positions: F3, F4, C3, C4, P3, P4, F7, F8, T7, T8, F9, F10, Fp1, Fp2, Fz, Cz, Pz, online-referenced against the left mastoid, with the AFz as ground electrode (see **Figure 1**). The EEG signal was recorded with a sampling rate of 1000 Hz and digitized online from 0.53 to 120 Hz.

## DATA ANALYSIS

### *Near-infrared spectroscopy*

Attenuation changes at 690 and 830 nm were converted into concentration changes of oxy- and deoxy-Hb using the modified Lambert–Beer law (Cope and Delpy, 1988). Data were low-pass filtered at 0.3 Hz (Butterworth, third order) and additionally high-pass filtered at 0.03 Hz to correct for high-frequency noise and slow drifts and fluctuations. Attenuation of movement artifacts is of special relevance in data recorded in infants. In line with previous infant studies using the same methodology (Taga et al., 2003; Minagawa-Kawai et al., 2011) we detected motion-induced artifacts characterized by sudden and sharp signal changes through visual inspection of the data. Artifacts were digitally marked and replaced by linear interpolation of uncontaminated data-points (10 data-points before and after the artifact), thus avoiding exclusion of whole segments or even whole data-sets. Next, the concentration changes of oxy- and deoxy-Hb were analyzed using a general linear model (GLM) approach. To increase the number of trials per condition the two fast modulated stimulus conditions were pooled (12 and 25 ms) as well as the two slowly modulated stimulus conditions (160 and 300 ms). Thus, the design matrix included two boxcar functions with the stimulus duration of 9 s relative to the onset of each stimulus modeling the pseudo-randomized succession of the fast and slowly modulated stimuli. These predictors were convolved



with the canonical hemodynamic response function (Boynton et al., 1996). The GLM analysis yields beta-values for oxy- and deoxy-Hb for the two stimulus conditions. The contrast between conditions and the *post hoc* statistical analyses (resulting in *t*-values) were performed in analogy to “Statistical Parametric Mapping,” as used for fMRI data. Paired *t*-tests were performed between left and right channels for fast and slow modulations, for each age group and for oxy- and deoxy-Hb separately.

### Electroencephalography

Off-line analyses were performed using Brain Vision Analyzer 2.0. Data were filtered off-line at 0.53 Hz low cutoff, 70 Hz high cutoff, and a 50 Hz notch filter (bandwidth 5 Hz, 24 dB/octave) was applied to attenuate line-voltage artifacts. We re-referenced the data to the averaged left and right mastoids. After the filtering procedure very noisy electrode channels were rejected. These were channels that showed either a flat line or signals stemming from predominantly technical artifacts throughout the whole experiment. Data were segmented into units of 10 s (1 s pre-stimulus onset, 9 s post-stimulus onset). We then applied a semi-automated artifact cor-

rection procedure (Brain Vision Analyzer 2.0). First, each data-set is automatically scanned for segments with maximal voltage step of 50  $\mu\text{V}/\text{ms}$ , and maximal absolute differences of 200  $\mu\text{V}$ . To ensure the quality of this automated procedure, each segment was again checked, and excluded manually if necessary.

Only participants for whom a minimum of 50% of the trials survived the artifact correction were included in the further EEG analyses. In the 23 subjects included in the analysis of the 6 month olds an average of  $30.2 \pm 12.7\%$  of the trials were removed by the artifact correction procedure (fast condition: mean = 32.2%, SD = 14.8%; slow condition: mean = 29.2%, SD = 13.8%). In 3 month olds 22 subjects were included in whom an average of 23.3% (SD = 15.8%) of the trials were removed (fast condition mean = 25.0%, SD = 16.0%; slow condition: mean = 25.3%, SD = 18.6%). A repeated measures analyses of variance (ANOVA) with the factor *condition* (fast versus slow acoustic variations) and *age group* as between-subject factor was performed to assess whether the amount of excluded segments differed across conditions and age groups. The ANOVA did not reveal a significant effect of condition [ $F(1,43) = 1.85$ ;  $p = 0.18$ ], and of condition  $\times$  age group [ $F(1,43) = 2.84$ ;  $p = 0.10$ ].

**Analysis of auditory-evoked potentials.** Auditory-evoked potentials (AEP) upon stimulus onset were computed for each participant and each experimental condition by averaging 1000 ms after stimulus onset referenced to a 100 ms pre-stimulus baseline. We were interested in developmental effects on the general features of the AEPs but also on specific effects of fast and slow acoustic modulations on the AEPs. Therefore we conducted three different AEP analyses:

- (1) Analysis of general features of the AEPs in both age groups  
To investigate how AEPs change with age we computed averaged AEPs across all stimulus conditions for the 6 and 3 month olds separately and performed peak-latency and -amplitude analyses. Previous studies suggest that in infants the most prominent components of the AEP occur within a time window of approximately 500 ms after stimulus onset (e.g., Kushnerenko et al., 2002; Wunderlich et al., 2006). To identify the average peak amplitudes and latencies of the AEP components we therefore ran a peak analysis for the first 500 ms after stimulus onset (Brain Vision Analyzer 2.0). Two peaks were identified within the first 500 ms after stimulus onset, and peak amplitude as well as peak latency for each electrode and each participant of the two age groups were assessed. To evaluate group differences in amplitude and latency we conducted univariate ANOVA separately for latency and amplitude of each peak, using *age group* as between-subject factor. The following electrodes entered statistical analysis subdivided into five regions of interest (ROIs): left-medial: *Fp1/F3/C3/P3*, right-medial: *Fp2/F4/C4/P4*, left-lateral: *F7/T7/F9*, right-lateral: *F8/T8/F10*, central: *Fz/Cz/Pz*.
- (2) Analyses on mean amplitudes of fast and slow acoustic modulations  
Next we analyzed differences of the AEP components with regard to the different stimulus conditions (fast and slow acoustic modulations). To increase the signal to noise ratio the two fast modulations (12 and 25 ms) were pooled and compared to the pooled slowly modulated conditions (160 and 300 ms). To identify peaks of the components the general peak-latency analysis yielded values which are in line with the literature (Kushnerenko et al., 2002; Picton and Taylor, 2007; Lippe et al., 2009). The following time windows were analyzed: for the 6 month olds: 0–100 ms (N1), and 100–225 ms (P2). Due to longer latencies in the younger age group, in the 3 month olds different windows were used: 0–200 ms (N1), and 200–500 ms (P2). Because the time windows differ between 6 and 3 month olds, separate analyses on mean amplitudes were performed for the two age groups. The analysis was performed in the previously specified ROIs (*left-medial*, *right-medial*, *left-lateral*, *right-lateral* and *central*). The repeated measures ANOVAs tested the within-subject factors *condition* (fast versus slow), and *hemisphere* (left versus right) the latter including *left-medial* versus *right-medial*, and *left-lateral* versus *right-lateral* ROIs. For the central ROI we calculated repeated measures ANOVA with the factor *condition*. When an ANOVA revealed a significant ( $p \leq 0.05$ ) main effect or interactions between either *condition* and/or *hemisphere*, *post hoc* paired *t*-tests were calculated between the next levels of the respective factor. Greenhouse and Geisser (1959) corrected significances are reported.

- (3) Analyses on peak amplitudes of fast and slow acoustic modulations

To test for significant differences between fast and slow modulation frequencies we additionally performed *peak* amplitude analyses (Rossi et al., 2010) of the AEP-components. Analysis of *peak* amplitude was performed because: (i) the analysis on *mean* amplitudes did not reveal a significant effect for the P2; (ii) *mean* amplitudes cannot be compared between age groups since the lengths of the time-windows differed between age groups. Statistical analyses of the *peak* amplitudes were performed on the same ROIs following the same schema reported for *mean* amplitude analyses above. However, we now extended the ANOVA with the between-subject factor *age group*.

**Time-frequency analysis.** To investigate tonic differences in the electrophysiological response between conditions and age groups we performed TFA to reveal stimulus-induced changes in oscillatory brain activity in the frequency range from 4 to 70 Hz. Artifact corrected EEG data (see above) were downsampled to 500 Hz. Further analysis was performed using custom-built Matlab scripts (version R2007a, Mathworks, Natick, MA, USA). For calculating the time-frequency representations from 4 to 70 Hz we used segments from –900 to 9000 ms relative to stimulus onset, and performed wavelet analyses (Morlet wavelet) on each trial (Tallon-Baudry and Bertrand, 1999; Jensen et al., 2002). Baseline power was calculated in a 850 ms pre-stimulus interval (–900 to –50 ms prior to stimulus onset to avoid stimulus related contamination of the baseline by smearing effects). One possibility to quantify oscillatory EEG responses is to assess the relative increase or decrease in signal power of cortical oscillations in specific frequency bands in an interval after stimulus onset compared to a pre-stimulus interval. Thereby a resulting event-related synchronization or desynchronization quantifies changes in signal power relative to the event (Pfurtscheller and Lopes da Silva, 1999). Therefore we averaged the time-frequency representations across the trials of the two fast modulated stimuli (12 and 25 ms), and the two slow acoustic modulations (160 and 300 ms), and displayed relative changes to the baseline. Fast and slow conditions comprised up to a maximum of 46 trials for both conditions in each infant. Relative changes were averaged over frequency and time. The frequency windows for the computation of the different frequency bands were chosen according to literature (Pfurtscheller and Lopes da Silva, 1999; Nierhaus et al., 2009). **Figure 5** shows de-/synchronization in the frequency bands from 4–8, and 10–15 Hz from 500 to 9000 ms after stimulus onset. For statistical analysis we computed mean values across the time-frequency windows: 4–8 Hz; 500–8900 ms and 10–15 Hz 500–8900 ms. The mean values entered a repeated measures ANOVA with the within-subject factors *condition* and *hemisphere* and the between-subject factor *age group*, performed in the above defined ROIs.

## RESULTS

### NEAR-INFRARED SPECTROSCOPY

The GLM based on the oxygenation changes yielded  $\beta$ -values of changes in oxy- and deoxy-Hb for fast (12 and 25 ms) and slow (160 and 300 ms) acoustic modulations. To assess lateralization they

were compared by paired *t*-tests between hemispheres. **Figure 2** illustrates in which areas oxy- and/or deoxy-Hb responses showed significant lateralization ( $p \leq 0.05$ ). The upper panel illustrates the results in the 6 month olds the lower those in 3 month olds.

In 6 month olds ( $n = 36$ ) fast acoustic modulations lead to larger hemodynamic responses (oxy-Hb $\uparrow$  and deoxy-Hb $\downarrow$ ) over the left compared to right inferior temporal position [position (3): deoxy-Hb:  $t_{(35)} = -2.37$ ,  $p = 0.012$ ; oxy-Hb:  $t_{(35)} = 2.26$ ,  $p = 0.015$ ]. Fast acoustic modulations additionally elicited a larger hemodynamic response (deoxy-Hb $\downarrow$ ) in the right compared to left temporo-parietal region [position (6); deoxy-Hb:  $t_{(35)} = 1.83$ ,  $p = 0.038$ ].

For the slow acoustic modulations statistics confirmed a larger hemodynamic response (deoxy-Hb $\downarrow$ ) in two right hemispheric positions [position (1): deoxy-Hb:  $t_{(35)} = 1.88$ ,  $p = 0.034$ ; position (6): deoxy-Hb:  $t_{(35)} = 1.89$ ,  $p = 0.033$ ] in inferior frontal and temporo-parietal regions.

The NIRS results for the 3 month olds ( $n = 38$ ) are illustrated in the lower panel of **Figure 2**. In this age group we found a larger increase in oxy-Hb for left compared to right hemispheric brain regions: Left superior frontal and posterior temporal regions

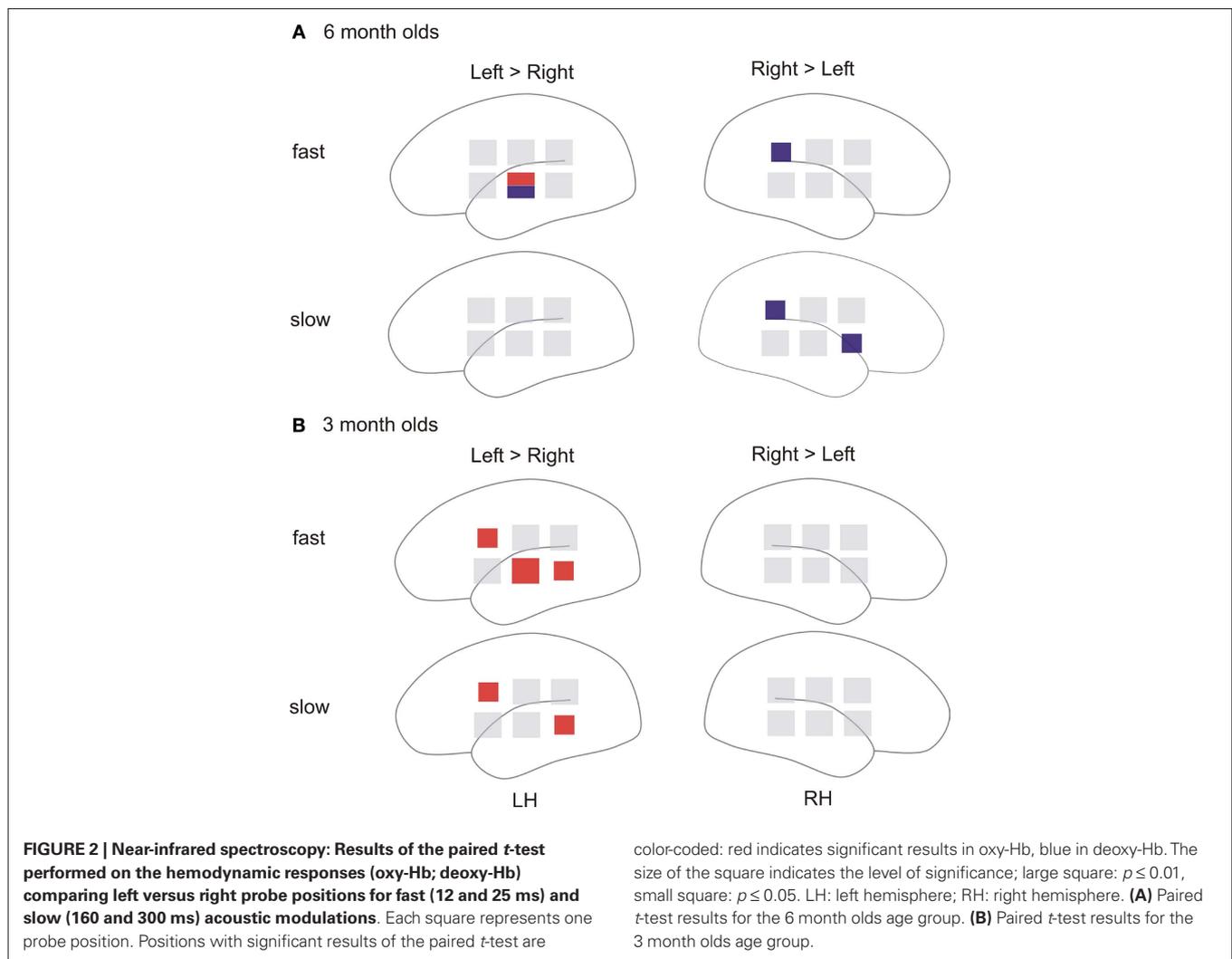
showed increased responses for both, fast [position (2); oxy-Hb:  $t_{(37)} = 1.83$ ,  $p = 0.038$ ; position (5); oxy-Hb:  $t_{(37)} = 1.91$ ,  $p = 0.032$ ], and slow [position (2); oxy-Hb:  $t_{(37)} = 1.72$ ,  $p = 0.047$ ; position (5); oxy-Hb:  $t_{(37)} = 1.96$ ,  $p = 0.029$ ] acoustic modulations. For the fast acoustic modulations we additionally found a stronger response in the left inferior temporal position [position (3); oxy-Hb:  $t_{(37)} = 2.47$ ,  $p = 0.009$ ].

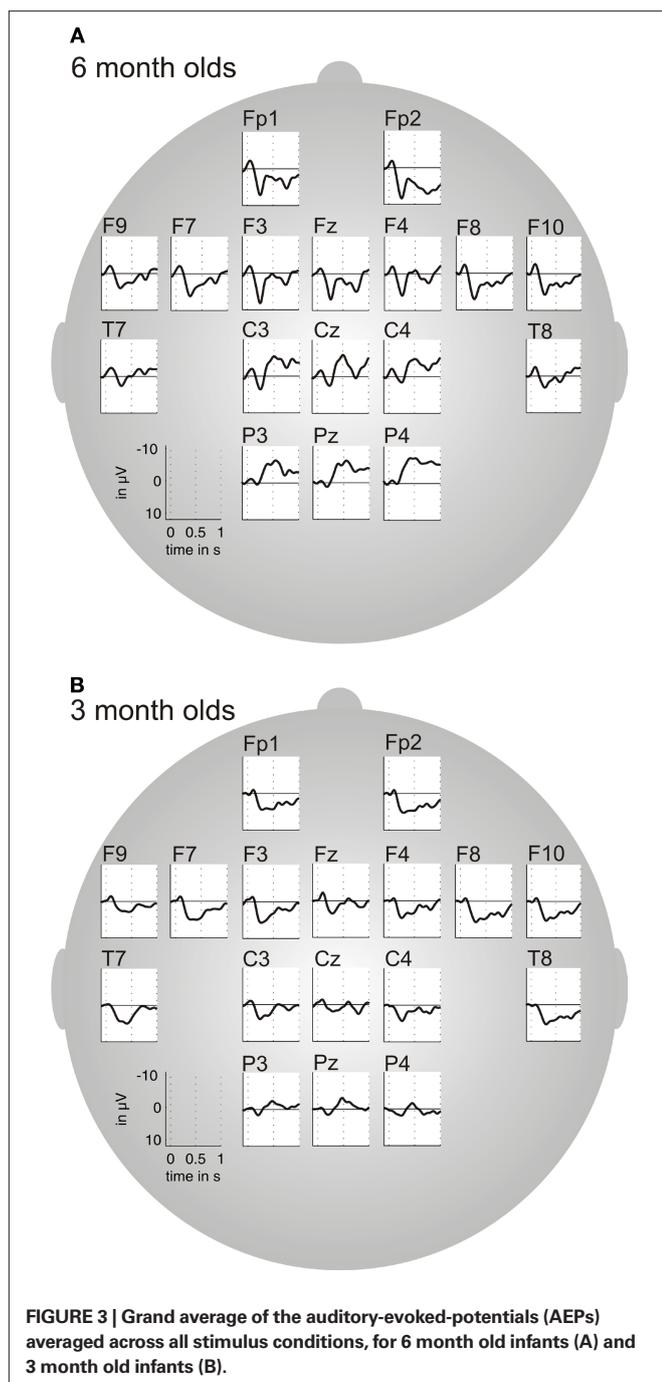
## ELECTROENCEPHALOGRAPHY

The analysis of the EEG data focused on two properties. First we report the results concerning the evoked potentials upon onset of the stimulus periods (AEPs representing the phasic response). To assess the response over the full length of the stimulation period we next report the results of the TFA for two frequency bands at 4–8 and 10–15 Hz (tonic response).

## Auditory-evoked potentials

General features of the AEPs in both age groups. To reveal a general effect of maturation of the AEPs, we first calculated the averaged AEPs across all stimulus conditions. **Figure 3** illustrates the





results for 6 month old infants ( $n = 23$ ) and 3 month old infants ( $n = 22$ ) separately. We performed peak-latency analyses on the AEPs between 0 and 500 ms, which revealed a first peak with a negative polarity (N1) followed by a second component with a positive polarity (P2) in the AEPs of both age groups.

In 6 month olds, the N1 peaked at 58 ms on average (range 25–93 ms,  $SD = 21$  ms). This component showed the same latency in the 3 month olds (mean 59 ms, range 27–98 ms,  $SD = 20$  ms). The univariate ANOVA confirmed that there was no statistically significant difference for *peak latency* in any of the ROIs. On the contrary the *amplitude* of the N1 was larger in 6 compared to 3 month old

infants, which was confirmed by the univariate ANOVA for *peak amplitude* over the *left-lateral ROI*:  $F(1,42) = 7.32$ ,  $p < 0.01$ , and *right-lateral ROI*:  $F(1,43) = 6.95$ ,  $p < 0.01$ . The mean amplitude in the 6 month olds group was  $-4.1 \mu V$  (range  $-13.6$  to  $1.74 \mu V$ ,  $SD = 3.2 \mu V$ ). In 3 month olds the mean amplitude of the N1 was  $-2.4 \mu V$  (range  $-7.3$  to  $2.75 \mu V$ ,  $SD = 2.3 \mu V$ ).

The P2 is clearly visible in the grand averages of both age groups (Figure 3). In 6 month old infants the P2 peaks at 226 ms on average (range 153–277 ms,  $SD = 35$  ms), while in 3 month olds the peak occurs later at around 315 ms (range 154–453 ms,  $SD = 80$  ms). The univariate ANOVA on *peak latency* of the P2 revealed significant differences between age groups for all ROIs (left-medial:  $F(1,43) = 19.97$ ,  $p < 0.001$ ; right-medial:  $F(1,43) = 14.27$ ,  $p < 0.001$ ; left-lateral:  $F(1,42) = 22.61$ ,  $p < 0.001$ ; right-lateral:  $F(1,43) = 18.09$ ,  $p < 0.001$ ; central:  $F(1,43) = 17.02$ ,  $p < 0.001$ ). With regard to the *peak amplitude* of the P2 there was no difference between age groups.

In sum, N1 and P2 components were seen in the AEPs of both age groups. The N1 peaks around 60 ms in both age groups and increases in amplitude with age over bilateral fronto-temporal regions. The P2, on the contrary, decreases in latency with age over all regions but does not change in amplitude.

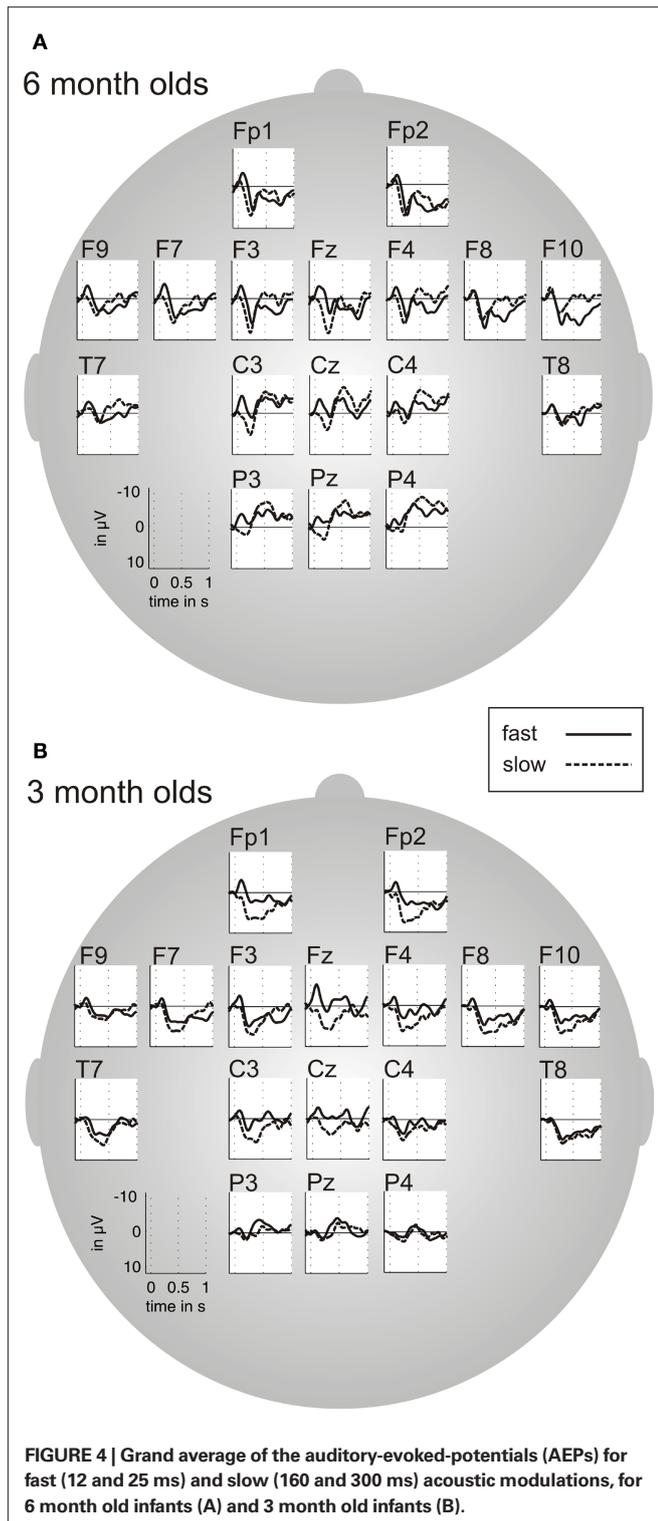
**Analyses on mean amplitudes of fast and slow acoustic modulations.** To test whether fast and slow acoustic modulations elicit differential phasic electrophysiological responses we computed AEPs separately for fast and slow acoustic modulations. Figure 4 shows the results separately for the two different age groups. N1- and P2-component are clearly seen in all conditions.

In 6 month olds the ANOVA for the N1-window (0–100 ms) reveal a significant effect of the factor *condition* only. Therefore we averaged the respective ROI pairs for the paired *t*-tests to compare fast and slow modulations. In the *medial ROI* we found a larger mean amplitude of the N1 for fast compared to slow acoustic modulations ( $F(1,22) = 4.67$ ,  $p < 0.04$ ;  $t_{(22)} = -2.16$ ,  $p = 0.04$ ). In the 3 month olds the ANOVA for the N1-window (0–200 ms) revealed a trend for the main effect *condition* over the *central ROI* ( $F(1,21) = 4.04$ ,  $p < 0.057$ ). Here N1 was larger in amplitude for fast in contrast to slow modulations. The effect was most pronounced over *Fz* (see Figure 4). Separate paired *t*-tests for each of the three midline electrodes confirmed a significantly larger N1 for fast compared to slow acoustic modulations ( $t_{(18)} = -3.65$ ,  $p = 0.002$ ) over *Fz*.

The analysis on the mean amplitude of the P2 (100–225 ms in the 6 months olds and 200–500 ms in the 3 months olds) for fast versus slow stimuli did not yield any statistically significant effects.

In sum, the mean amplitude analysis yielded significant differences between the two conditions only for the N1. In both age groups the N1 was larger for the onset of fast compared to slowly modulated stimuli over bilateral fronto-central ROIs (please also refer to Table 1 for an overview of the results).

**Analyses on peak amplitudes of fast and slow acoustic modulations.** We additionally performed statistical analyses on *peak amplitudes* for N1 and P2. Both peaks (N1 and P2) were identified by an automatic peak detection (see section General Features of the AEPs in Both Age Groups). The within-subject factors *condition*, *hemisphere*, and the between-subject factor *age group* were tested by



repeated measures ANOVAs for the medial and lateral ROIs. For the analysis of the *central* ROI an ANOVA with the within-group factor *condition* and the between-subject factor *age group* was computed. Neither the ANOVA for the peak amplitude of the N1, nor for the P2 did reveal any effect of the between-subject factor *age group*. Therefore we averaged across the two age groups for *post hoc* paired *t*-tests.

For the N1, the ANOVA revealed a significant main effect of *condition* in the *medial*:  $F(1,43) = 6.01, p < 0.02$ , and *central* ROI:  $F(1,43) = 4.95, p < 0.03$ . The *post hoc* paired *t*-test for the averaged left and right *medial* ROI revealed a significantly larger N1 for fast compared to slow acoustic modulations ( $t_{(44)} = -2.48, p = 0.02$ ). The same effect was seen for the *central* ROI ( $t_{(44)} = -2.24, p = 0.03$ ).

With regard to the P2 the ANOVA also revealed significant main effects of *condition* for the *medial* ( $F(1,43) = 4.35, p < 0.04$ ), and *central* ROI:  $F(1,43) = 5.88, p < 0.02$ . The *post hoc* paired *t*-test for the *medial* ROI revealed a significantly larger P2 for slow acoustic modulations ( $t_{(44)} = -2.09, p = 0.04$ ), which also held true for the *central* ROI ( $t_{(44)} = -2.45, p = 0.02$ ).

In summary, the differential peak analyses of the AEPs for fast and slow acoustic modulations showed that the amplitude of the N1 was larger for fast compared to slow modulation frequencies. On the contrary the amplitude of the P2 was larger for slow when compared to fast acoustic modulations. These effects did not differ between age groups (please also see, **Table 1** for an overview).

### Time-frequency analysis

The AEP-analysis reported so far is sensitive only to the onset of the stimuli. To find out whether the differential stimulus features (slow versus fast modulations) elicit a sustained response over the full stimulation period we performed a TFA on the EEG data. Sustained differential synchronizations and desynchronizations have been reported in response to stimulus features in a number of systems in adults (e.g., gamma-synchronization and alpha-desynchronization in the visual system (Koch et al., 2009). Since there are very few reports on TFA in very young infants (e.g., Csibra et al., 2000; Pena et al., 2010) this analysis was explorative in nature and we could not make strong predictions to whether de- or synchronizations were to be expected and in which frequency bands such modulations should be seen. Therefore a TFA over a wide spectral range (4–70 Hz) was performed separately for fast and slow acoustic modulations and in all five ROIs: left-medial: *Fp1/F3/C3/P3*, right-medial: *Fp2/F4/C4/P4*, left-lateral: *F7/T7/F9*, right-lateral: *F8/T8/F10*, central: *Fz/Cz/Pz*.

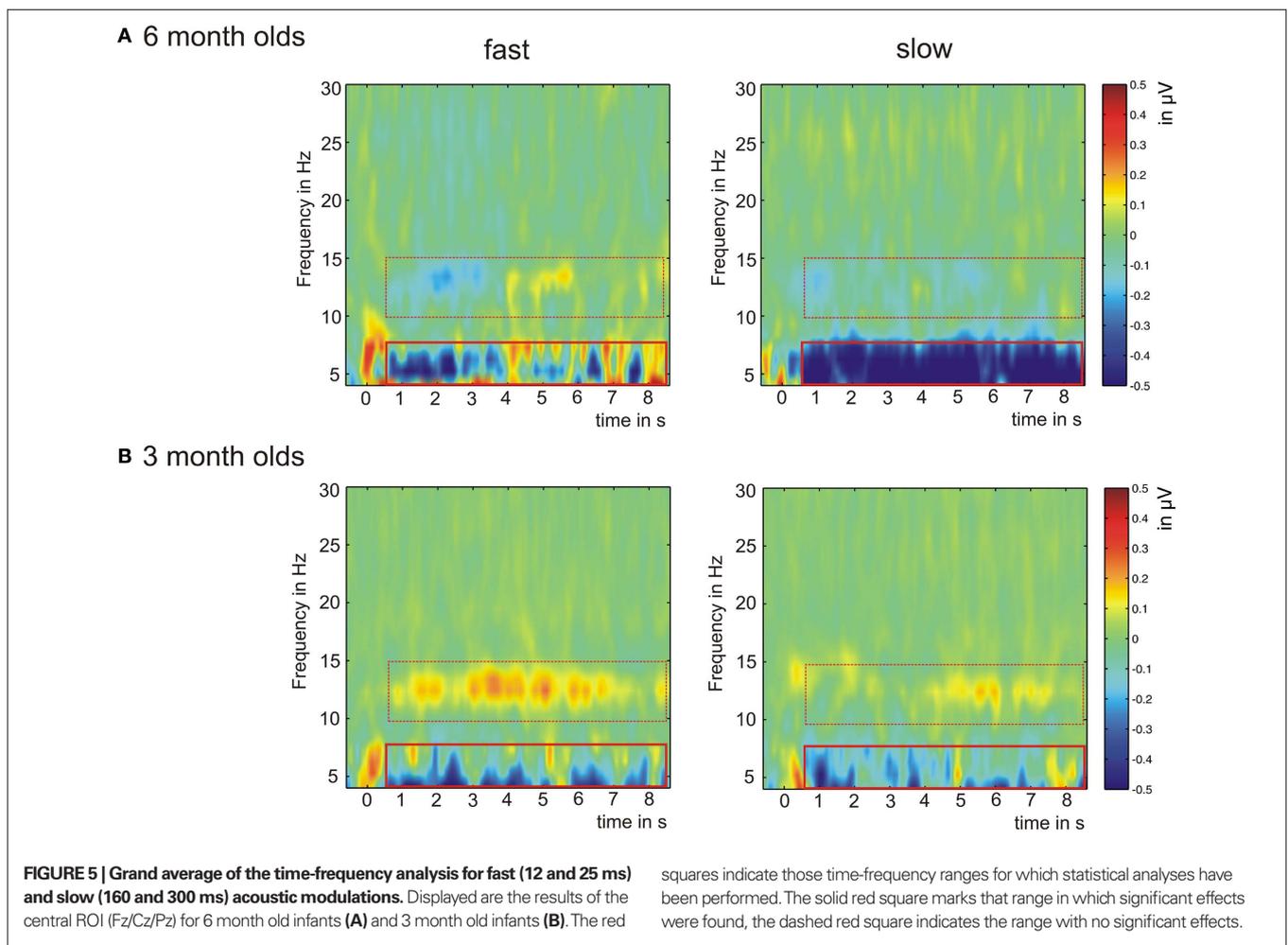
**Figure 5** shows modulations in two frequency bands (averaged across all ROIs), that is, in the theta (4–8 Hz) and alpha-range (10–15 Hz).

Fast and slow acoustic modulations elicited synchronization in the theta-range during the first 300 ms after stimulus onset in both age-groups. Furthermore, ~500 ms after stimulus onset 6 month olds showed a sustained desynchronization in response to the slowly varying stimuli in the theta-range. This desynchronization was less pronounced for the fast modulated stimuli. The 3 month olds did not show this effect. A similar but weaker desynchronization was seen for both conditions in this younger age group. With regard to the higher frequency band (10–15 Hz) only the 3 month olds showed a difference between the conditions. In this frequency band fast modulated stimuli elicited a stronger synchronization when compared to the slowly modulated stimuli. In the 6 month olds a small and unstable desynchronization was seen in this higher frequency range. In all higher frequency bands (including gamma) no de/synchronization was seen in the time-frequency plots, and statistical analysis confirmed this result.

**Table 1 | Overview of the statistically significant EEG and NIRS results.**

		AEP				TFA	NIRS
		Mean amplitude		Peak amplitude			
		N1	P2	N1	P2		
6 mo	Fast	+	=	+	-	-	LH > RH (3) RH > LH (6)
	Slow	-	=	-	+	+	RH > LH (1, 6)
3 mo	Fast	+	=	+	-	-	LH > RH (2, 3, 5)
	Slow	-	=	-	+	-	LH > RH (2, 5)

Displays the comparison of fast and slow acoustic modulations for the two age groups separately. 6 mo, 6 month old infants; 3 mo, 3 month old infants; AEP, auditory-evoked potential; TFA, time-frequency analysis; NIRS, near-infrared spectroscopy; N1, first negativity; P2, second positivity; LH, left hemisphere; RH, right hemisphere; (+), increase; (-), decrease; (=), no difference. Location of the areas measured by means of NIRS are described by the number in brackets (please also refer to Figure 1): (1) inferior frontal, (2) superior frontal, (3) inferior temporal, (4) superior temporal, (5) posterior temporal, and (6) temporo-parietal.



For statistical analysis of the two lower frequency bands we averaged the power changes in the theta-range (4–8 Hz) and the alpha range (10–15 Hz) from 500 to 8900 ms. We chose this time window because early synchronization effects during the first hundreds of milliseconds after stimulus onset are likely due to the evoked response

(see Materials and Methods). The resulting changes in oscillatory amplitude in each ROI entered a repeated measures ANOVA with the within-subject factors *condition*, and *hemisphere*, and *age group* as between-subject factor. The *central ROI* was analyzed using a repeated measures ANOVA with the factors *condition* and *age group*.

The theta desynchronization yielded a significant interaction of *condition*  $\times$  *age group* in every ROI: *left- and right-medial*:  $F(1,43) = 8.09, p < 0.007$ , *left- and right-lateral*:  $F(1,42) = 6.37, p < 0.02$ , and *central*:  $F(1,43) = 7.68, p < 0.008$ . Furthermore, we found a significant main effect of *condition*: *left- and right-medial*:  $F(1,43) = 6.13, p < 0.02$ , *left- and right-lateral*:  $F(1,42) = 5.85, p < 0.02$ , and *central*:  $F(1,43) = 8.76, p < 0.005$ . Based on the results, we computed *post hoc* paired *t*-tests for the two age groups separately, to compare the oscillatory activity for fast and slow acoustic modulations. **Figure 5** shows the results of the TFA analysis for the central ROI exemplarily.

In 6 month olds the paired *t*-tests showed that slow acoustic modulations elicited stronger desynchronization in the theta-range when compared to fast modulations (*lateral*:  $t_{(22)} = 3.11, p = 0.005$ ; *medial*:  $t_{(22)} = 2.91, p = 0.008$ ; *central*:  $t_{(22)} = 3.54, p = 0.002$ ). For the 3 month olds the paired *t*-tests did not reveal significant differences between fast and slow acoustic modulations in the theta-range. In the alpha-range significant effects were found in neither age group and none of the ROIs.

To summarize, slowly modulated stimuli elicit a sustained desynchronization in the theta-range (4–8 Hz) in 6 month old infants. This desynchronization is statistically larger than the response to the fast modulated stimuli. The effect was not seen in the younger infants (please see, **Table 1** for an overview of the results).

## DISCUSSION

### HEMODYNAMIC RESPONSES

Our results show that subtle auditory differences during the processing of complex auditory stimuli elicit a differential pattern of brain activation in infants. NIRS revealed a lateralized brain response for 6 month old infants, similar to the reported findings newborns (Telkemeyer et al., 2009) and in adults (Boemio et al., 2005). Fast acoustic modulations (12 and 25 ms) lead to an activation of bilateral temporal brain regions. On the contrary slow acoustic modulations (160 and 300 ms) resulted in a greater right-lateralized hemodynamic response in the temporal region. These results are in line with the assumptions of the multi-time resolution model linking hemispheric specialization for language features to an asymmetry in cortical tuning (Hickok and Poeppel, 2007; Poeppel et al., 2008). The model proposes that hemispheric lateralization during language perception partially results from the temporal features in the speech signal. Thereby, left and right auditory cortices are differentially specialized for the acoustic analysis in at least two different temporal integration windows (Poeppel, 2003; Poeppel et al., 2008). Bilateral auditory cortex areas are thought to decode fast acoustic modulations, specifically relevant for the decoding of segmental such as phonological information within the speech stream. Slow acoustic modulations, relevant for the perception of suprasegmental language features, like prosodic information, are mainly processed in right hemispheric cortical brain regions. A recent NIRS study in 4 month old infants comparing different speech and non-speech conditions observed right hemispheric activation for slowly modulated emotional voices, whereas speech sounds and scrambled non-speech sounds, both comprising fast acoustic variations, elicited leftward activation (Minagawa-Kawai et al., 2011). Similar to our results in newborns and 6 month

olds, the authors conclude that the observed lateralization might be driven by basic acoustic features. Interestingly, their results also emphasize the influence of linguistic features *per se* (i.e., exposure to the native language) on the modulation of cortical brain responses, because they found stronger left-hemispheric activation during native compared to non-native speech sounds.

It should be noted, that we failed to confirm the right hemispheric specialization for processing slow acoustic modulations in 3 month old infants. The results revealed dominant left-hemispheric responses for fast and slow acoustic modulations. In the light of our previous results in newborns (Telkemeyer et al., 2009) and considering the results in the above mentioned studies in 4 month old infants (Minagawa-Kawai et al., 2011), and adults (Boemio et al., 2005) we do not believe that this result proves a discontinuity of developmental lateralization with regard to complex auditory feature processing. Rather we consider experimental factors constitutive for this negative finding. The different levels of motor activity in infants during development and even more the ability to control and withhold from movement is one factor. Also inter-individual differences in response magnitude and optical parameters contributing to background optical properties vary greatly in adults and infants. Notably the analysis of the NIRS results in 3 month olds in our present study revealed significant results for oxy-Hb only. Both, oxy-Hb increases and deoxy-Hb decreases, are associated with neuronal activation (Obrig and Villringer, 2003). However, oxy-Hb is typically characterized by a larger amplitude compared to deoxy-Hb and thus more likely to yield larger effects. On the other hand, concentration changes in oxy-Hb are more susceptible to extracerebral, systemic changes in the hemodynamics (Boden et al., 2007).

In sum, the NIRS data reported here yield less robust effects compared to our previously reported results in newborns. Beyond differences in movement-artifacts, poorer data quality may also suggest that shorter stimulation periods and more repetitions may be a special requirement in studies of these age groups. During the study design we favored identical stimulation paradigms to allow for a comparison with the data in newborns. However, we recommend the use of shorter stimulation durations for future studies on auditory processing, especially when longitudinal aspects over the first year of life are addressed.

To summarize, despite these limitations we consider the symmetric processing of fast and the asymmetric, right-lateralized processing of slow temporal modulations during the auditory analysis to be rather stable from early development. This lateralization may contribute to the lateralization of differential linguistic feature analyses in the incoming auditory stream evolving in parallel to language competence.

### AUDITORY-EVOKED POTENTIALS

We simultaneously measured EEG response to the stimuli, which provides a superb temporal resolution allowing for an inquiry into temporal aspects of neural activity correlated with auditory processing.

Since temporal features of the stimuli may affect the waveforms of the evoked response we computed AEPs for the time period of 1 s after stimulus onset. We were interested in developmental changes

of the general AEPs in response to auditory stimulation. The averaged AEPs across all stimulus conditions for both age groups are characterized by an early negative component (N1) followed by a large positivity (P2), mainly in fronto-central positions. Our results show, that the latency of the N1 at around 60 ms did not differ between the two age groups, whereas the amplitude increases with age. Previous results also described such a negative component to be present in the AEPs in newborns and young infants (Novak et al., 1989; Wunderlich et al., 2006). Comparable to our data, the N1 increases in amplitude with development until a discrete component is clearly observed in adulthood (Sussman et al., 2008). Kujala and Näätänen (2010) suggest that the increased amplitude of the N1 reflects an increased fine-grained cortical mapping. However, whether the observed component in infants parallels the N1 component in adults remains under debate (Lippe et al., 2009).

In line with previous research our results furthermore indicate that the infant's AEPs are dominated by a large positivity, especially over fronto-central electrodes and show less discrete components compared to adults (Ceponiene et al., 2002; Kushnarenko et al., 2002; Picton and Taylor, 2007). Similar to previous findings (Wunderlich et al., 2006; Pena et al., 2010) the comparison of the two age groups showed, that the P2 decreases in latency, from a mean peak at 315 ms in 3 month olds to 226 ms in 6 month old infants. In adults the peak of the P2 is described at around 150–200 ms (Näätänen and Picton, 1987; Lippe et al., 2009), in infants it varies around 200–250 ms (Picton and Taylor, 2007). Hence, with increasing brain maturation the latency of the P2 decreases.

Besides these developmental effects on the morphology of the averaged AEPs, we investigated whether differences in the temporal features of the stimuli modulate the AEP-components. Thus, we compared the amplitude of the AEPs for fast and slow acoustic modulations in the two age groups. In both age groups we found an increased amplitude of the N1 for fast compared to slowly modulated stimuli, primarily in fronto-central electrodes. In adults, acoustic information is consciously perceived after around 80 ms after stimulus onset (Näätänen and Winkler, 1999). It has been proposed that at least in adults a discriminable change of any feature of a continuous sound would elicit an N1 (Näätänen and Winkler, 1999). Thus, the N1 is associated with sound detection and is sensitive to physical aspects of the auditory stimulus (Näätänen and Picton, 1987) including the temporal modulation of the total acoustic energy (Ceponiene et al., 2005). Hence, the enhanced response to fast acoustic modulations in our results might be associated with the higher number of acoustic changes during the fast condition (acoustic modulations occur every 12, and every 25 ms, respectively), compared to the slow modulation condition (every 160 and 300 ms).

In both age-groups the amplitude of the P2 was larger for slowly modulated stimuli. However, the functional role of this positivity is not fully understood. In contrast to the N1, the P2 is modulated by consciously perceived, stimulus-specific features, such as emotional content or the salience of the stimulus (Ceponiene et al., 2005; Spreckelmeyer et al., 2009). Deregnier et al. (2000) reported an increased amplitude of the P2 elicited by maternal voice compared to a stranger's voice already in newborn infants, suggesting an effect of attention. Therefore, the here observed increased P2 during slow acoustic modulations may indicate an increased atten-

tion or a preference of the infants toward the slowly varying stimuli. Such slow acoustic variations can be found in prosodic features of the speech signal. Studies investigating language acquisition in infancy emphasize the role of suprasegmental, prosodic information (Gleitman and Wanner, 1982; Jusczyk, 1997) during language development as they aid the segmentation of the speech stream into smaller units such as words (Jusczyk et al., 1999). Behavioral studies demonstrated that infants prefer the so called infant-directed speech mode adults use when addressing infants which is characterized by accentuated prosodic features (Werker and McLeod, 1989; Cooper and Aslin, 1990). This finding suggests that infants are more attracted by prosodically modulated features.

### **Oscillatory responses**

In contrast to the AEPs reflecting the effects of temporal variation during the early acoustic analysis, the TFA is a marker of the sustained electrophysiological response. In both age-groups fast and slow acoustic modulations elicited a theta synchronization during the first 300 ms after stimulus onset, which is probably associated with the AEP (Bruneau et al., 1993). This result is in line with Fujioka and Ross (2008) who compared a violin tone to noise-burst stimuli to 4–6 year old children while measuring MEG. The authors report a synchronized theta response during the first ~200 ms after stimulus onset without any difference between the two acoustic stimulations and between hemispheres. Further, the authors reported a desynchronization in the alpha range (8–12 Hz) starting ~400 ms after stimulus onset. These may be similar to the classical Berger-effect in the visual system (Berger, 1929). Our results also revealed a desynchronization beginning ~500 ms after stimulus onset. However, we found this desynchronization in lower frequencies (between 4 and 8 Hz). In 6 but not in 3 month old infants, slow acoustic modulations elicited a significantly stronger desynchronization compared to fast acoustic modulations in that frequency band, hence, suggesting an effect of development. Processing sounds with complex spectrotemporal structure might become more refined with age. A developmental study investigating the phase-locked oscillatory response to musical tones revealed an increase in phase-locking of theta oscillatory activity with age (Shahin et al., 2010). Furthermore, it has been demonstrated that the response to speech sounds in children matures more rapidly than response to non-speech sounds (Pang and Taylor, 2000). Therefore one could speculate, that 6 but not 3 month old infants perceive the slow acoustic modulations at least as more familiar sounds compared to the fast acoustic variations. We did not find oscillatory activity in higher frequency bands, probably due to the fact, that the power of spontaneous oscillations shifts from lower to higher frequencies over early development (Shahin et al., 2010).

### **CONCLUSION**

The present study used simultaneous assessment of hemodynamic and electrophysiological brain responses to investigate the perception of temporal features of non-linguistic complex acoustic stimuli. Subtle auditory differences during the processing of complex auditory stimuli elicit a differential pattern of brain activation in infants. Our NIRS results support the notion that language-specific hemispheric asymmetries are partially driven by acoustic features of

the speech signal. Though the NIRS results in 3 month old infants were inconclusive, we believe that the hemispheric specialization for processing fast and slow temporal modulations during the auditory analysis is rather stable from birth. The AEPs to the onset of the averaged acoustic stimuli indicated an effect of brain maturation on the morphology of the AEPs in general. However, similar to the results of the NIRS no age effect was found in the differential AEP analysis of fast and slow modulations. The larger amplitude of the N1 for fast modulated stimuli may result from higher energy of the acoustic stimulus due to its rapid transitions between different noise bands. On the contrary, the following P2 is affected by more conscious, stimulus-specific features such as attention. Both age groups showed an increased amplitude of the P2 to slow acoustic modulations. Given the importance prosodic features, characterized by slow acoustic modulations, play especially during language acquisition, the increased amplitude might reflect an increased attention of the infants toward the slow modulations. Consistently, the TFA also reveals a stronger theta-band desynchronization for slowly modulated stimuli in the older age group. It is

unclear whether this is due to a more fine-grained processing of complex spectrotemporal sounds in general or whether it is related to effects of attention. To our knowledge, this is the first study investigating slow oscillatory responses to non-linguistic auditory stimulation in early infancy complementing recent results in the language domain (Pena et al., 2010). Though the rather explorative approach precludes a specific interpretation, analyses of the time-frequency representations in infants during language acquisition may shed new light on the way how infants reach instantaneous representations of complex sounds.

## ACKNOWLEDGMENTS

Financial support of the EU (NEST 012778, EFRE 20002006 2/6, nEUROpt 201076), and BMBF (BNIC, Bernstein Center for Computational Neuroscience, German-Polish cooperation FK: 01GZ0710) are gratefully acknowledged. Isabell Wartenburger is supported by the Stifterverband für die Deutsche Wissenschaft (Claussen-Simon-Stiftung). We would like to express our gratitude to all parents and their children who participated in this study.

## REFERENCES

- Benasich, A. A., and Tallal, P. (2002). Infant discrimination of rapid auditory cues predicts later language impairment. *Behav. Brain Res.* 136, 31–49.
- Berger, H. (1929). Über das Elektroenzephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten* 87, 527–550.
- Boden, S., Obrig, H., Kohncke, C., Benav, H., Koch, S. P., and Steinbrink, J. (2007). The oxygenation response to functional stimulation: is there a physiological meaning to the lag between parameters? *Neuroimage* 36, 100–107.
- Boemio, A., Fromm, S., Braun, A., and Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat. Neurosci.* 8, 389–395.
- Boynton, G. M., Engel, S. A., Glover, G. H., and Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *J. Neurosci.* 16, 4207–4221.
- Bruneau, N., Roux, S., Guerin, P., Garreau, B., and Lelord, G. (1993). Auditory stimulus intensity responses and frontal midline theta rhythm. *Electroencephalogr. Clin. Neurophysiol.* 86, 213–216.
- Ceponiene, R., Alku, P., Westerfield, M., Torki, M., and Townsend, J. (2005). ERPs differentiate syllable and non-phonetic sound processing in children and adults. *Psychophysiology* 42, 391–406.
- Ceponiene, R., Rinne, T., and Näätänen, R. (2002). Maturation of cortical sound processing as indexed by event-related potentials. *Clin. Neurophysiol.* 113, 870–882.
- Choudhury, N., and Benasich, A. A. (2010). Maturation of auditory evoked potentials from 6 to 48 months: prediction to 3 and 4 year language and cognitive abilities. *Clin. Neurophysiol.* 122, 320–338.
- Choudhury, N., Leppänen, P. H., Leevers, H. J., and Benasich, A. A. (2007). Infant information processing and family history of specific language impairment: converging evidence for RAP deficits from two paradigms. *Dev. Sci.* 10, 213–236.
- Cooper, R. P., and Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Dev.* 61, 1584–1595.
- Cope, M., and Delpy, D. T. (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.
- Csibra, G., Davis, G., Spratling, M. W., and Johnson, M. H. (2000). Gamma oscillations and object processing in the infant brain. *Science* 290, 1582–1585.
- Dehaene-Lambertz, G., Dehaene, S., and Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science* 298, 2013–2015.
- Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Meriaux, S., Roche, A., Sigman, M., and 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.
- Deregnier, R. A., Nelson, C. A., Thomas, K. M., Wewerka, S., and Georgieff, M. K. (2000). Neurophysiologic evaluation of auditory recognition memory in healthy newborn infants and infants of diabetic mothers. *J. Pediatr.* 137, 777–784.
- Fox, P. T., and Raichle, M. E. (1986). Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. *Proc. Natl. Acad. Sci. U.S.A.* 83, 1140–1144.
- Friederici, A. D. (2005). Neurophysiological markers of early language acquisition: from syllables to sentences. *Trends Cogn. Sci.* 9, 481–488.
- Fujioka, T., and Ross, B. (2008). Auditory processing indexed by stimulus-induced alpha desynchronization in children. *Int. J. Psychophysiol.* 68, 130–140.
- Gervain, J., and Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annu. Rev. Psychol.* 61, 191–218.
- Giraud, A. L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S., and Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron* 56, 1127–1134.
- Gleitman, L. R., and Wanner, E. (1982). “Language acquisition: the state of the state of the art,” in *Language Acquisition: The State of the Art*, eds E. Wanner and L. R. Gleitman (New York: Cambridge University Press), 3–48.
- Greenhouse, S. W., and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika* 24, 95–112.
- Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402.
- Holland, S. K., Plante, E., Weber, B. A., Strawsburg, R. H., Schmithorst, V. J., and Ball, W. S. Jr. (2001). Normal fMRI brain activation patterns in children performing a verb generation task. *Neuroimage* 14, 837–843.
- Homae, F., Watanabe, H., Nakano, T., Asakawa, K., and Taga, G. (2006). The right hemisphere of sleeping infant perceives sentential prosody. *Neurosci Res.* 54, 276–280.
- Hoogenboom, N., Schoffelen, J. M., Oostenveld, R., Parkes, L. M., and Fries, P. (2006). Localizing human visual gamma-band activity in frequency, time and space. *Neuroimage* 29, 764–773.
- Jensen, O., Hari, R., and Kaila, K. (2002). Visually evoked gamma responses in the human brain are enhanced during voluntary hyperventilation. *Neuroimage* 15, 575–586.
- Jusczyk, P. W. (1997). *The Discovery of Spoken Language*. Cambridge, MA: MIT Press.
- Jusczyk, P. W., Houston, D. M., and Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cogn. Psychol.* 39, 159–207.
- Kleinschmidt, A., Obrig, H., Requardt, M., Merboldt, K. D., Dirnagl, U., Villringer, A., and 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.
- Koch, S. P., Werner, P., Steinbrink, J., Fries, P., and Obrig, H. (2009). Stimulus-induced and state-dependent sustained gamma activity is tightly coupled to the hemodynamic response in humans. *J. Neurosci.* 29, 13962–13970.
- Kuhl, P., and Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Annu. Rev. Neurosci.* 31, 511–534.
- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* 5, 831–843.

- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255, 606–608.
- Kujala, T., and Näätänen, R. (2010). The adaptive brain: a neurophysiological perspective. *Prog. Neurobiol.* 91, 55–67.
- Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., Huotilainen, M., and Näätänen, R. (2002). Maturation of the auditory event-related potentials during the first year of life. *Neuroreport* 13, 47–51.
- Lippe, S., Martinez-Montes, E., Arcand, C., and Lassonde, M. (2009). Electrophysiological study of auditory development. *Neuroscience* 164, 1108–1118.
- Lopes da Silva, F. (1991). Neural mechanisms underlying brain waves: from neural membranes to networks. *Electroencephalogr. Clin. Neurophysiol.* 79, 81–93.
- Mehler, J., Sebastian-Galles, N., and Nespore, M. (2004). “Biological foundations of language: language acquisition, cues for parameter setting and the bilingual infant,” in *The New Cognitive Neuroscience*, ed. M. Gazzaniga (Cambridge, MA: MIT Press), 825–836.
- Minagawa-Kawai, Y., Mori, K., Naoi, N., and 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., van der Lely, H., Ramus, E., Sato, Y., Mazuka, R., and Dupoux, E. (2011). Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cereb. Cortex* 21, 254–261.
- Näätänen, R., and Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24, 375–425.
- Näätänen, R., and Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychol. Bull.* 125, 826–859.
- Nierhaus, T., Schön, T., Becker, R., Ritter, P., and Villringer, A. (2009). Background and evoked activity and their interaction in the human brain. *Magn. Reson. Imaging* 27, 1140–1150.
- Novak, G. P., Kurtzberg, D., Kreuzer, J. A., and Vaughan, H. G. Jr. (1989). Cortical responses to speech sounds and their formants in normal infants: maturational sequence and spatiotemporal analysis. *Electroencephalogr. Clin. Neurophysiol.* 73, 295–305.
- Obrig, H., Rossi, S., Telkemeyer, S., and Wartenburger, I. (2010). From acoustic segmentation to language processing: evidence from optical imaging. *Front. Neuroenergetics* 2:13. doi: 10.3389/fnene.2010.00013
- Obrig, H., and Villringer, A. (2003). Beyond the visible-imaging the human brain with light. *J. Cereb. Blood Flow Metab.* 23, 1–18.
- Pang, E. W., and Taylor, M. J. (2000). Tracking the development of the N1 from age 3 to adulthood: an examination of speech and non-speech stimuli. *Clin. Neurophysiol.* 111, 388–397.
- Pena, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., and 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.
- Pena, M., Pittaluga, E., and Mehler, J. (2010). Language acquisition in premature and full-term infants. *Proc. Natl. Acad. Sci. U.S.A.* 107, 3823–3828.
- Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., Baldoli, C., and Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4758–4763.
- Pfurtscheller, G., and Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin. Neurophysiol.* 110, 1842–1857.
- Picton, T. W., and Taylor, M. J. (2007). Electrophysiological evaluation of human brain development. *Dev. Neuropsychol.* 31, 249–278.
- 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, W. J., and van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1071–1086.
- Rossi, S., Jürgenson, I. B., Hanulíková, A., Telkemeyer, S., Wartenburger, I., and Obrig, H. (2010). Implicit processing of phonotactic cues: evidence from electrophysiological and vascular responses. *J. Cogn. Neurosci.* 1, 1–13.
- Schönwiesner, M., Rübsem, R., and von Cramon, D. Y. (2005). Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. *Eur. J. Neurosci.* 22, 1521–1528.
- Shahin, A. J., Trainor, L. J., Roberts, L. E., Backer, K. C., and Miller, L. M. (2010). Development of auditory phase-locked activity for music sounds. *J. Neurophysiol.* 103, 218–229.
- Sharbrough, F., Chatrian, G. E., Lesser, R. P., Lüders, H., Nuwer, M., and Picton, T. W. (1991). American Electroencephalographic Society guidelines for standard electrode position nomenclature. *J. Clin. Neurophysiol.* 8, 200–202.
- Singer, W. (1999). Time as coding space? *Curr. Opin. Neurobiol.* 9, 189–194.
- Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. *Annu. Rev. Physiol.* 55, 349–374.
- Spreckelmeyer, K. N., Kutas, M., Urbach, T., Altenmüller, E., and Münte, T. F. (2009). Neural processing of vocal emotion and identity. *Brain Cogn.* 69, 121–126.
- Stefanics, G., Haden, G., Huotilainen, M., Balazs, L., Sziller, I., Beke, A., Fellman, V., and Winkler, I. (2007). Auditory temporal grouping in newborn infants. *Psychophysiology* 44, 697–702.
- Steinbrink, J., Villringer, A., Kempf, F., Haux, D., Boden, S., and Obrig, H. (2006). Illuminating the BOLD signal: combined fMRI-fNIRS studies. *Magn. Reson. Imaging* 24, 495–505.
- Stevens, K. N. (1998). *Acoustic Phonetics*. Cambridge, MA: MIT Press.
- Sussman, E., Steinschneider, M., Gumenyuk, V., Grushko, J., and Lawson, K. (2008). The maturation of human evoked brain potentials to sounds presented at different stimulus rates. *Hear Res.* 236, 61–79.
- Taga, G., Asakawa, K., Maki, A., Konishi, Y., and Koizumi, H. (2003). Brain imaging in awake infants by near-infrared optical topography. *Proc. Natl. Acad. Sci. U.S.A.* 100, 10722–10727.
- Tallon-Baudry, C., and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn. Sci.* 3, 151–162.
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., Obrig, H., and Wartenburger, I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *J. Neurosci.* 29, 14726–14733.
- Werker, J. F., and McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Can. J. Psychol.* 43, 230–246.
- Wunderlich, J. L., Cone-Wesson, B. K., and Shepherd, R. (2006). Maturation of the cortical auditory evoked potential in infants and young children. *Hear Res.* 212, 185–202.
- Zaehle, T., Jancke, L., Herrmann, C. S., and Meyer, M. (2009). Pre-attentive spectro-temporal feature processing in the human auditory system. *Brain Topogr.* 22, 97–108.
- Zaehle, T., Jancke, L., and Meyer, M. (2007). Electrical brain imaging evidences left auditory cortex involvement in speech and non-speech discrimination based on temporal features. *Behav. Brain Funct.* 3, 63.
- Zaehle, T., Wüstenberg, T., Meyer, M., and Jancke, L. (2004). Evidence for rapid auditory perception as the foundation of speech processing: a sparse temporal sampling fMRI study. *Eur. J. Neurosci.* 20, 2447–2456.
- Zatorre, R. J., and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 12 November 2010; paper pending published: 24 December 2010; accepted: 25 March 2011; published online: 09 April 2011.

Citation: Telkemeyer S, Rossi S, Nierhaus T, Steinbrink J, Obrig H and Wartenburger I (2011) Acoustic processing of temporally modulated sounds in infants: evidence from a combined near-infrared spectroscopy and EEG study. *Front. Psychol.* 2:62. doi: 10.3389/fpsyg.2011.00062

This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*.

Copyright © 2011 Telkemeyer, Rossi, Nierhaus, Steinbrink, Obrig and Wartenburger. This is an open-access article subject to a non-exclusive license between the authors and *Frontiers Media SA*, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other *Frontiers* conditions are complied with.