Physiological detection of interaural phase differences

Bernhard Ross[a)]
Rotman Research Institute, Baycrest Center and University of Toronto, Toronto, M6A 2E1 Canada

Kelly L. Tremblay
Department of Speech and Hearing Sciences, University of Washington, Seattle, Washington 98105-6246 and Rotman Research Institute, Baycrest Center and University of Toronto, Toronto, M6A 2E1 Canada

Terence W. Picton
Rotman Research Institute, Baycrest Center and University of Toronto, Toronto M6A 2E1, Canada

(Received 18 July 2006; revised 23 October 2006; accepted 10 November 2006)

Auditory evoked cortical responses to changes in the interaural phase difference (IPD) were recorded using magnetoencephalography (MEG). Twelve normal-hearing young adults were tested with amplitude-modulated tones with carrier frequencies of 500, 1000, 1250, and 1500 Hz. The onset of the stimuli evoked P1m-N1m-P2m cortical responses, as did the changes in the interaural phase. Significant responses to IPD changes were identified at 500 and 1000 Hz in all subjects and at 1250 Hz in nine subjects, whereas responses were absent in all subjects at 1500 Hz, indicating a group mean threshold for detecting IPDs of 1250 Hz. Behavioral thresholds were found at 1200 Hz using an adaptive two alternative forced choice procedure. Because the physiological responses require phase information, through synchronous bilateral inputs at the level of the auditory brainstem, physiological "change" detection thresholds likely reflect the upper limit of phase synchronous activity in the brainstem. The procedure has potential applications in investigating impaired binaural processing because phase statistic applied to single epoch MEG data allowed individual thresholds to be obtained.
[DOI: 10.1121/1.2404915]

PACS number(s): 43.64.Ri, 43.66.Pn, 43.66.Nm [WPS] Pages: 1017–1027

I. INTRODUCTION

Localization of sound sources in the horizontal plane is mainly based on neural comparison of the sounds arriving at each ear, with the primary cues being interaural time and intensity differences. Other binaural cues used for sound localization include interaural spectral differences associated with head transfer functions and changes in the sound field observed with small head movements. The acoustic cues are likely analyzed in separate neural networks in the brain (Yin, 2002), with some species having particular specializations in these networks.

The ability to localize sound is highly dependent on the ability to detect interaural time differences (ITD) for low frequencies, and interaural intensity differences (IID) for higher frequencies (Stevens, 1936). For steady signals, such as pure tones, the ITD is equivalent to the interaural phase difference (IPD). When presented in free field sound localization becomes ambiguous if the wavelength of the sound is equal to or smaller than the distance between the ears (i.e., $\mathrm{IPD} \geqslant 2\pi$). For instance, the distance between ears in adult humans is less than 20 cm. Given this distance, maximum ITD is 0.7–0.9 ms (depending on frequency and distance of the sound) (Brungart and Rabinowitz, 1999) which corresponds to the period of 1100–1400 Hz tones.

According to behavioral studies, the upper frequency limit for IPD detection is between 1100 and 1300 Hz. For example, Garner and Wertheimer (1951) reported that the upper frequency limit for IPD detection was 1100 Hz for tones presented at 50 dB and 1300 Hz for tones presented at 90 dB. Zwislocki and Feldman (1956) found the upper frequency limit for perceiving IPDs to be 1300 Hz for pure tones presented at 75 dB SL. Schiano *et al.* (1986) reported that the sensitivity for sound lateralization decreases rapidly above 1100 Hz.

IPD processing requires that the phase information of the acoustical signal is preserved in the neural activity at the place of processing. However, phase locked encoding degrades along the ascending auditory pathway. Therefore, neural networks in the early part of the auditory brainstem that receive bilateral input are candidates for processing IPDs. The medial part of the superior olivary complex (MSO) is a likely candidate because it is the first station in the ascending mammalian auditory system that receives bilateral input from both ears. It is believed that MSO neurons likely act as coincidence detectors proposed in Jeffress' model for binaural hearing (Jeffress, 1948). A requirement of this model, that the bipolar MSO cells receive phase-locked input from both ears, seems fulfilled, at least for low frequencies. For example, auditory nerve fibers tuned to low frequencies fire synchronized with particular phases of the sine wave cycle, often at intervals of several cycles (Rose *et al.*, 1967). This phase-locked response to low frequencies is

[a)]Author to whom correspondence should be addressed. Fax: +1(416) 785 2862. Electronic mail: bross@rotman-baycrest.on.ca

enhanced in those cells of the cochlear nucleus that project to the MSO (Joris *et al.*, 1994). When the IPD of pure tones is varied, MSO neurons respond best to a specific IPD. The specialization of MSO neurons to certain IPDs is expressed in highly facilitated responses to favorable IPDs with spike rates significantly exceeding the sum of separate ipsilateral and contralateral stimulation. Additionally, the responses to these IPDs exhibit a far higher degree of phase-locking than to unfavorable IPDs (Batra *et al.*, 1997). As the auditory information is processed beyond the superior olive to the nuclei of the lateral lemniscus and the inferior colliculus, the relative timing information is preserved, but the ability of the neurons to phase lock to the tones decreases (Kuwada *et al.*, 2006).

The aim of our study was to learn more about the encoding of binaural information in the auditory system and how this is reflected in auditory evoked responses. We developed a procedure for the recording of evoked responses to changes in IPD and used these responses to find individual physiological thresholds for the frequency in which binaural information can be processed based on IPD. We used magnetoencephalography (MEG) to record evoked cortical activity in response to binaural stimuli containing phase changes. Our aims were to determine whether IPD transitions evoke physiological "change" responses, and how well physiological IPD thresholds correspond with behavioral thresholds. Discriminating IPD changes would clearly be related to sound localization. Furthermore, our hypothesis was that the responses to IPD changes are limited to the frequency range of phase locked representation of the acoustical input in the auditory brainstem. Thus, the upper frequency limit for processing IPD in dichotic sound could give an indirect measure for phase locking in the auditory pathway. During normal aging or as a result of brain dysfunction phase locked neural encoding may be degraded. In general this threshold may indicate the upper frequency limit for phase locked processing in hearing. The obtained threshold could be an important indicator for the performance in sound localization and for other auditory processes requiring rapid waveform identification for the perception of speech and music.

II. METHODS

A. Subjects

Twelve healthy subjects (seven females), with a mean age of 26.8 years (std dev. 3.1 yrs), provided their informed consent before participating in the study, which was approved by the Research Ethics Board at Baycrest Center. Each participant had normal hearing, defined as thresholds $<$20 dB HL. For frequencies below 2000 Hz absolute threshold differences between left and right ear were less than or equal to 10 dB (mean absolute difference 3.5 dB).

B. Auditory stimuli for the physiological test

A specific stimulus signal was developed for recording auditory evoked responses to transitions in the IPD. Sinusoidal amplitude modulated tones were repeatedly presented for 4.0 s duration with stimulus onset asynchrony uniformly randomized between 7.5 and 8.5 s [Fig. 1(a)]. At 2.0 s after stimulus onset, sudden phase shifts in the carrier signal of +90° in the left and −90° in the right ear, equivalent to 180° IPD, were introduced. Thus, during the first two seconds both ears received the identical tones, but during the last two seconds of stimulation the tones were of opposite polarity in the two ears [Fig. 1(b)]. Participants described the diotic portion of the stimulus during the first two seconds as sound arising from a single source in the center of the head. In contrast, the dichotic portion of the stimulus during the last two seconds was described as sound from separate sources in space without specific localization. A similar method of presenting sounds with different phases from two loudspeakers was used in the early days of sound reproduction to create the illusion of spatially distributed sound sources from a single audio signal and was termed quasi-stereophony (Schroeder, 1961).

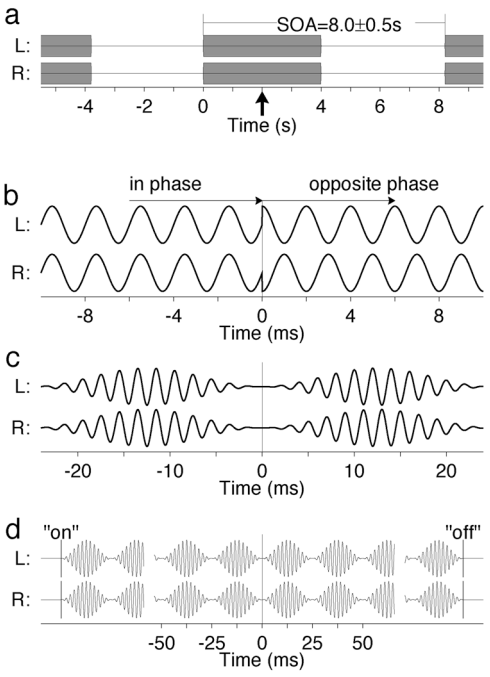

FIG. 1. Description of the auditory stimuli used during MEG testing. (a) The stimulus sequence was comprised of tonebursts (4 s in duration) presented dichotically with a SOA of 7.5–8.5 s. The arrow at 2 s after stimulus onset indicates the time point of the change in the IPD. (b) The 90 degree phase shifts, opposite in direction, result in a polarity reversal between left (L) and right (R) ears. (The time scale is adjusted to the phase shift.) (c) The phase shift occurs at the minimum point of the 40 Hz amplitude modulation (d) The 40 Hz AM envelope defines the shape of stimulus onset, the phase shift, and the stimulus offset.

To prevent the subject from perceiving a discontinuity in the sound at the moment of the phase-shift, the tones were amplitude modulated and the phase-shift was set to occur at the minimum point of modulation [Figs. 1(c) and 1(d)]. In a previous study we demonstrated that 90° diotic phase-shifts (in the same direction in both ears) at the minimum of the AM, did not elicit an evoked response (Ross *et al.*, 2004). Only when the dichotic phase-shift caused a change in the IPD, an evoked response was elicited. In addition, the 12.5 ms onset slope of the modulation envelope eliminates any differences in the onset time as a possible binaural cue. The modulation frequency of 40 Hz was chosen in order to

elicit the strongest auditory steady-state responses. The effect of IPD changes on the steady-state responses will be reported separately.

The experimental parameter in the current study was the carrier frequency of the stimulus, which was set to 500, 1000, 1250, and 1500 Hz. Since the phase change was always half a period, the introduced ITD was 1.0, 0.5, 0.4, and 0.33 ms, respectively. The frequencies were chosen being clearly below (500 and 1000 Hz) and around (1250 and 1500 Hz) the expected frequency threshold. One hundred stimuli with the same carrier frequency were presented in each experimental block of 13 min duration, and each block was repeated once. The eight experimental blocks were randomly presented across two experimental sessions of about one hour duration each on two subsequent days.

Stimuli were presented through Etymotic ER3A transducers connected with 1.5 m of length matched plastic tubing and foam earplugs to the subject's ears. The sound transducer had to be placed in sufficient distance to the MEG sensor in order to avoid any interference between the stimulus signal and the recorded brain activity. Below 2000 Hz the frequency characteristic of the sound transmission system was relatively flat (±6 dB) and the phase characteristic was linear. The stimuli in the current study were defined on a narrow frequency band and, thus, were not distorted by the sound transmission tubes. The phase relation between acoustical signals at both earplugs was tested using a 2 cc coupler.

The stimulus intensity was set to 60 dB above individual sensation thresholds, which were measured immediately before each MEG recording. Insert-earphones provide typically interaural attenuation of more than 75 dB at the frequencies used in this study (Sklare and Denenberg, 1987). Thus, no effects of interaural cross talk were expected.

C. MEG data acquisition and analysis

MEG recordings were performed in a quiet, magnetically shielded room using a 151-channel whole-head neuromagnetometer (VSM-Medtech, Port Coquitlam, BC, Canada). The detection coils of this MEG device are almost equally spaced on the helmet shaped surface and are configured as axial first-order gradiometers (Vrba and Robinson, 2001). After low-pass filtering at 200 Hz, the magnetic field data were sampled at the rate of 625 Hz and stored continuously.

MEG data were collected during passive listening, meaning, that the subjects did not need to attend to the stimuli or execute a task. In order to control for confounding changes in vigilance, the subjects watched a closed captioned movie of their choice, while the auditory stimuli were being presented. Compliance was verified using video monitoring. The subjects were in supine position with the head resting inside the helmet shaped MEG sensor. Head movements were verified to be less than 8 mm during each recording block using three detection coils attached to the subject's nasion and the pre-auricular points. No data had to be rejected because of excessive head movements.

D. MEG data analysis

Each block of continuously recorded MEG data was subdivided into 100 stimulus related epochs of 6000 ms duration including 1000 ms pre- and post-stimulus intervals. For artifact rejection, a principal component analysis was performed on each epoch of magnetic field data. This approach is effective for removing artifacts with amplitudes larger than the brain signals of interest (Lagerlund *et al.*, 1997). Principal components, which exceeded the threshold of 2 pT in at least one channel, were subtracted from the data. This procedure removed artifact primarily related to dental metal and eye-blinks, which are substantially larger than the brain activity (Bardouille *et al.*, 2006). After artifact removal, the magnetic field data were averaged and magnetic source analysis was applied to the ±10 ms time interval around the maximum of the N1m wave, the most prominent response 100 ms after stimulus onset. Source analysis was based on the model of spatio-temporal equivalent current dipoles (ECD) in a spherical volume conductor. A head based Cartesian coordinate system was established by the *x*-axis pointing from the midpoint between the pre-auricular points to the nasion, the *y*-axis running from right to left in the plane formed by the three fiducials and the *z*-axis pointing in superior direction.

Single dipoles in both hemispheres were fit simultaneously to the 151-channel magnetic field distribution. First, the data were modeled with a mirror symmetric pair of dipoles. The resulting source coordinates were then used as starting points to fit the dipole in one hemisphere while the coordinates in the other hemisphere remained fixed. We then switched between hemispheres and repeated the last step until the source coordinates showed no further change. Dipole fits were accepted if their calculated fields explained at least 85% of the variance of the measured magnetic field. Eight estimates (four stimulus frequencies times two repetitions) of the N1m source location were obtained for each subject. The mean spatial coordinates and orientations were used as individual models to measure the source waveforms for the auditory evoked responses. The transient P1m-N1m-P2m components and sustained responses were obtained from source-space projection based on the source coordinates of the N1m "onset" response. To evaluate the validity of this approach, the linear regression of the actual magnetic field at each time point versus the magnetic field observed at the time of the N1m "onset" response was performed on the grand averaged data. This analysis tested how well the magnetic field distribution, at any time point, could be explained with the field distribution of the N1m onset response.

Data were analyzed primarily on dipole moment waveforms representing the source activity in the auditory cortices. MEG responses were modeled by single dipoles in left and right auditory cortices. Previous analyses, using a beamformer approach (Vrba and Robinson, 2001) with no *a priori* assumptions about the source configuration, demonstrated that single dipoles are sufficient to describe the activation of the auditory cortex (Herdman *et al.*, 2003). In this case, magnetic field waveforms of cortical source activity $\mathbf{q}(t)$ are detected by each of the 151 sensors with a distinct sensitivity

depending on the orientations and distances between the source and the sensors. If the sensitivities of all sensors at position $\mathbf{R}$ for the source activity at location $\mathbf{r}$ are given in matrix notation by the lead-field matrix $\mathbf{L}(\mathbf{r},\mathbf{R})$ (Sarvas, 1987) the detected magnetic field $\mathbf{B}(\mathbf{R},t)$ can be described by $\mathbf{B}(\mathbf{R},t)=\mathbf{q}(\mathbf{r},t)\cdot\mathbf{L}(\mathbf{r},\mathbf{R})$. The pseudo-inverse of the lead-field matrix can then be used to obtain the cortical source activity from the measured magnetic field: $\mathbf{q}(\mathbf{r},t)=\mathbf{B}(\mathbf{R},t)\cdot\mathbf{L}^{-1}(\mathbf{r},\mathbf{R})$. This operation, termed source space projection (Ross *et al.*, 2000; Tesche *et al.*, 1995), combines the 151 waveforms of magnetic field strength into a single waveform of a magnetic dipole moment measured in nanoAmpere-meter (nAm). The dipole moment is most sensitive for the localized area in the brain and less sensitive to electro-magnetic sources at other locations. Furthermore, this method of source space projection results in waveforms with higher signal-to-noise ratio than the magnetic field waveforms (Ross *et al.*, 2000).

A further advantage of analysis in source domain is that the dipole moment is independent of the sensor position. The position of the subject's head relative to the MEG sensor can change between sessions and will also be different between subjects. Combination of magnetic field data for group analysis is, therefore, not feasible. In contrast, the waveforms of cortical source activity can be combined across repeated sessions for a subject and across the group of subjects.

Statistical tests for detecting response components were based on the phase distribution of a Morlet wavelet transform with half-maximum-width corresponding to 2.5 periods of the test frequency applied to single trial source waveforms for each subject. The null hypothesis was that the signal phase is uniformly distributed across all trials if the signal contains no stimulus-related components. An evoked response either caused by an additive time-locked signal or by time-locked phase-synchronization is reflected as deviation from uniformly distributed signal phase. The Rayleigh test statistic R_{ij} was calculated for each point in time (i) and frequency (j) as $R_{ij}=\mathrm{abs}(\langle c_{ij}/|c_{ij}|\rangle)$ were c is the complex coefficients obtained from the wavelet transform, c divided by its absolute value gives sine and cosine of the phase, and $\langle\rangle$ denotes calculation of the mean across all trials. Because the distribution of the Rayleigh statistic is known, p-values could be calculated by numerical approximation for all time-frequency coefficients (Fisher, 1993). This method has been successfully used to detect cortical evoked near threshold responses in objective audiometry (Ross *et al.*, 1999). Also, studies have demonstrated that phase tests may be more sensitive than amplitude tests (Sturzebecher and Cebulla, 1997). The minimum p-value for all coefficients in the time interval between 2.0 and 2.4 s and the frequency range from 3 to 10 Hz obtainedfrom left and right hemisphere responses served as the test statistic for detecting a response to the change in the IPD. The selected time-frequency interval contained 1000 coefficients (25 frequencies $\times$40 time points).

To correct the p-values for the multiple tests, we needed to estimate the degrees of freedom (the maximum number of independent time-frequency components). The frequency resolution of the Morlet transform was largest at 3 Hz with $\Delta f=1.25$ Hz and the temporal resolution at 10 Hz with a sample interval of $\Delta t=50.0$ ms. Therefore, the total degrees of freedom were estimated as 44.8 $[(10-3)/1.25^{*}400/50]$. Being conservative we estimated the degree of freedom as 50 and rejected the null hypothesis (absence of a response) on the nominal level of 0.1% corresponding to the corrected $\alpha=5\%$ level.

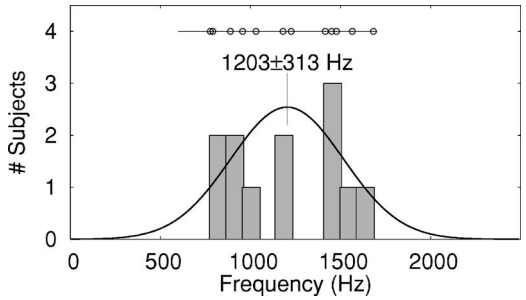

FIG. 2. Distribution of behavioral thresholds ($n=12$). The mean threshold was 1203 Hz, with standard deviation of 313 Hz.

E. Behavioral stimuli and procedure

An adaptive two-alternative forced choice (2AFC) procedure was used for psycho-acoustical testing. Stimuli were sinusoidal amplitude modulated (AM) tones presented through Etymotic ER3A transducers. Two 40 Hz AM tones, 1.0 s in duration, either in phase or out of phase, were presented to both ears with 900 ms inter-stimulus interval. During an initial training phase the subjects learned that the two stimuli could be perceived as sound from a single source (diotic sound with the same interaural phase) or spacious sound in both ears (dichotic sound with different interaural phase). The order of stimuli was randomized during the test and the subjects were asked to identify if the first or the second stimulus sounded like two stimuli in separate ears. Immediately after the subject responded with a button press the next pair of stimuli was presented.

Initially the carrier frequency was 250 Hz; however, the carrier frequency increased by a quarter octave if two responses in a row were correct and decreased if a single response was incorrect (two-down, one-up procedure, Levitt, 1970). Visual feedback was provided, in the form of a green (correct) or red (incorrect) square on the computer screen. One hundred trials were presented in each run, lasting about 10 min. Each run was repeated once.

III. RESULTS

A. Behavioral results

The group mean behavioral threshold was 1203 Hz with a standard deviation of 313 Hz. The distribution of obtained thresholds can be found in Fig. 2. Figure 3(a) shows the time-course of threshold changes, for two individuals, using the adaptive 2AFC procedure. For subject 1, the time-course of frequency threshold changes stabilized immediately after the initial slope, and then oscillated around the assumed threshold. Figure 3(b) illustrates the estimated threshold for this subject (e.g., 1682 Hz). The solid line connecting the black squares denotes the cumulative percentage of downward steps in the graph of Fig. 3(a). The observed curve was

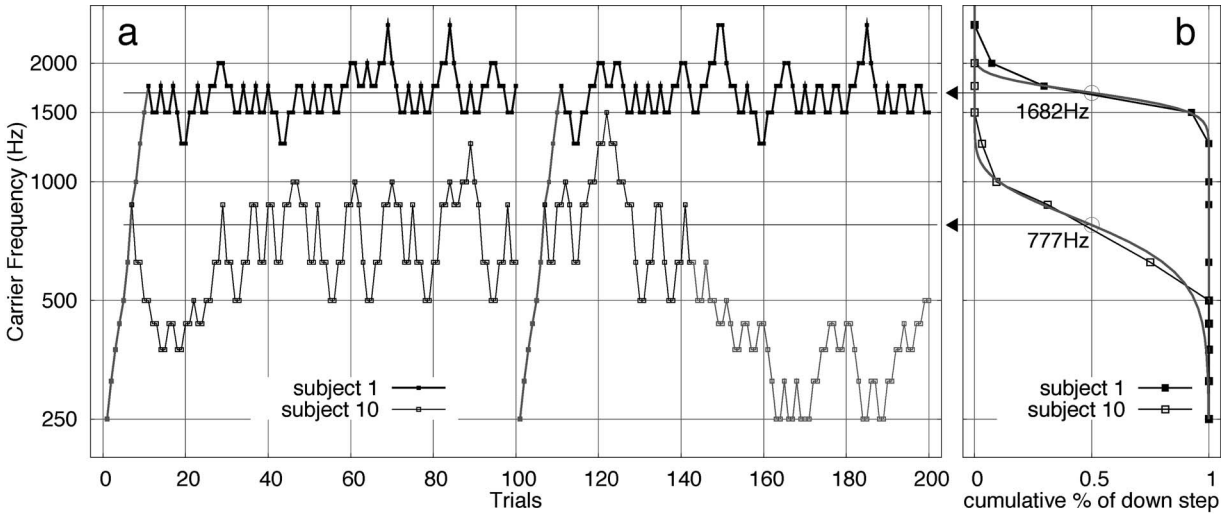

FIG. 3. Behavioral results for two sample subjects. (a) The time course of the test frequency using an adaptive procedure. Both subjects performed two blocks of 100 trials each. The initial phase, before the first reversal (gray line), was excluded from analysis. Whereas subject 1 performed consistently throughout the whole test, subject 10 showed larger fluctuations. Decreases in performance, toward the end of the test, may have resulted from fatigue or reduced alertness. Thus, trials 140–200 (gray line) were not included in the data analysis. (b) Cumulative percentage of downward steps during the adaptive procedure. The observed characteristics were approximated by a normal CDF and the threshold was defined as the 50% value of the CDF.

approximated by a cumulative normal distribution function (CDF) defined by the mean and standard deviation. The mean, equivalent to the 50% of the CDF, was accepted as an estimate for the threshold. When threshold is defined in this manner, it corresponds to the 70% point on the subject's psychometric function (Levitt, 1970).

The second subject in Fig. 3(a) was much more variable than the first subject. Although a relatively consistent threshold was observed in the first block, the performance deteriorated substantially during the second half of the second block. This might have been an effect of fatigue or decreased concentration. Therefore, in this instance, the last 60 trials were excluded from further analysis [shaded in gray in Fig. 3(a)].

B. Auditory evoked fields

Auditory evoked responses, from a sample individual subject, are shown in Fig. 4. Group results are shown in Fig. 5. At all stimulus frequencies, clear P1m-N1m-P2m "onset" and "offset" waveforms are evident. The presence of a P1m-N1m-P2m response signals the physiological detection of a change in the acoustical environment. In this experiment, the response waveforms indicate the onset from silence to sound, the change in IPD, and the offset from sound to silence. Of particular importance is that the presence of the P1m-N1m-P2m "change" response, signaling the physiological detection of phase changes, is absent at 1500 Hz but present at all lower frequencies. For example, on top of the sustained response an additional P1m-N1m-P2m response occurred after the sudden change in the IPD, at 2.0 s after stimulus onset. This change response has about the same amplitude as the onset response at 500 Hz. However, the amplitude of the IPD change response becomes smaller at 1000 Hz and is barely visible at 1250 Hz. Note there is no change response, for this subject, in the 1500 Hz frequency condition.

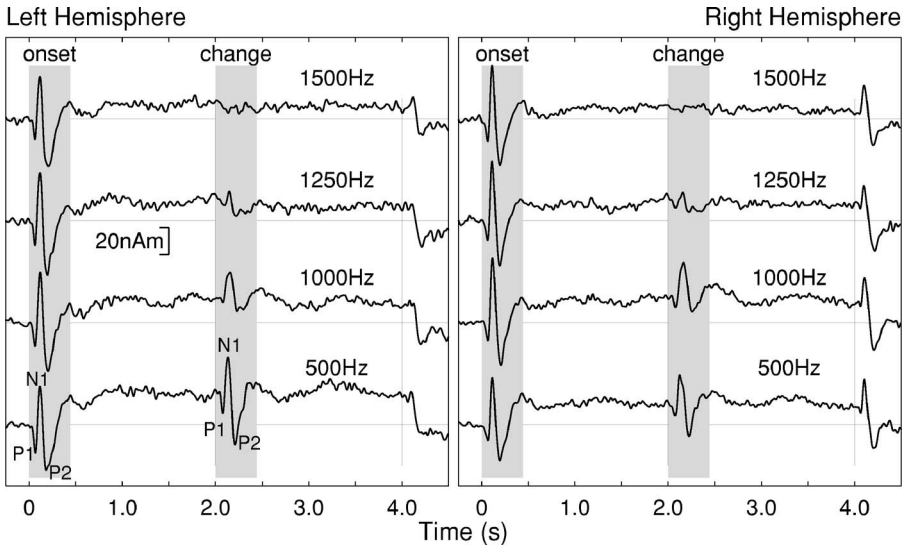

FIG. 4. Individual waveforms of 24 Hz lowpass filtered auditory evoked responses in the left and right hemisphere for each frequency tested. The response waveform at 500 Hz exhibits clear P1m-N1m-P2m responses to the stimulus onset and IPD change (at 2.0 s). Also sustained responses, continuing for the stimulus duration, and N1-P2 offset responses can be seen. Whereas the onset responses are fairly consistent across all stimulus frequencies, the phase change response clearly diminishes with increasing stimulus frequencies and is absent at 1500 Hz.

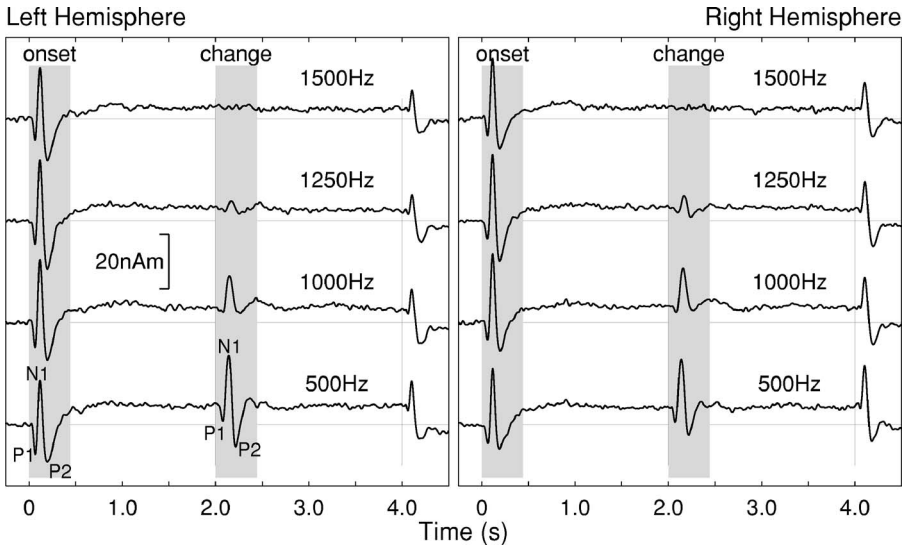

FIG. 5. Grand averaged ($n=12$) waveforms of 24 Hz lowpass filtered auditory evoked responses in the left and right hemisphere for four different stimulus frequencies.

Grand averaged waveforms of the onset and change responses are also shown in Fig. 6(a). This figure shows that the latencies for P1m, N1m, and P2m onset responses did not change across frequency conditions. In contrast, latencies of the change response increased, and amplitudes decreased, when higher frequency stimuli were used to evoked P1m, N1m, and P2m responses. At 500 Hz, the change response was similar in amplitude to the onset response but slightly delayed. For example, P1m latencies of the change response were delayed by about 14 ms when compared to the corresponding onset responses. Detailed information about peak latencies and amplitudes of the P1m-N1m-P2m onset and change responses at 500 Hz stimulus frequency are given in Table I. Differences in the group mean values of the onset and IPD change response peak amplitudes were not significant because of inter subject variability. However, the amplitude ratio of change versus onset response was larger in the left (1.6) than in the right hemisphere (0.6) ($t(11)=2.5$, $p<0.03$) indicating that the IPD "change" response was more left lateralized compared to the onset response.

A summary of N1m amplitude results, as a function of stimulus frequency, is provided in Fig. 6(b). Whereas, amplitudes of the N1m IPD change response decreased monotonically with increasing stimulus frequency; the N1m onset response showed an inverse u-shaped characteristic. From 500 to 1000 Hz the N1m onset response increased significantly [$t(11)=2.87$, $p<0.015$]. The onset N1m amplitude decreased significantly between 1000 and 1500 Hz [$t(11)=2.56$, $p<0.027$] and between 1250 and 1500 Hz [$t(11)=3.52$, $p<0.005$].

C. Source coordinates

Sources of N1m onset responses were localized in all subjects with a goodness of fit larger than 85% (mean square difference between measured and modeled magnetic field). Group mean N1m source locations were converted into Talairach coordinates. N1m sources were found in left and right Planum temporale at ($x=49.2$, $y=-21.5$, $z=5.8$ mm) in the right and at ($x=-47.1$, $y=-26.4$, $z=6.3$ mm) in the left hemisphere with 95% confidence limits for the group means of less than 4 mm in any direction. The source locations represent centers of activity. Because of limited spatial resolution of MEG for simultaneous active sources, the possibility of

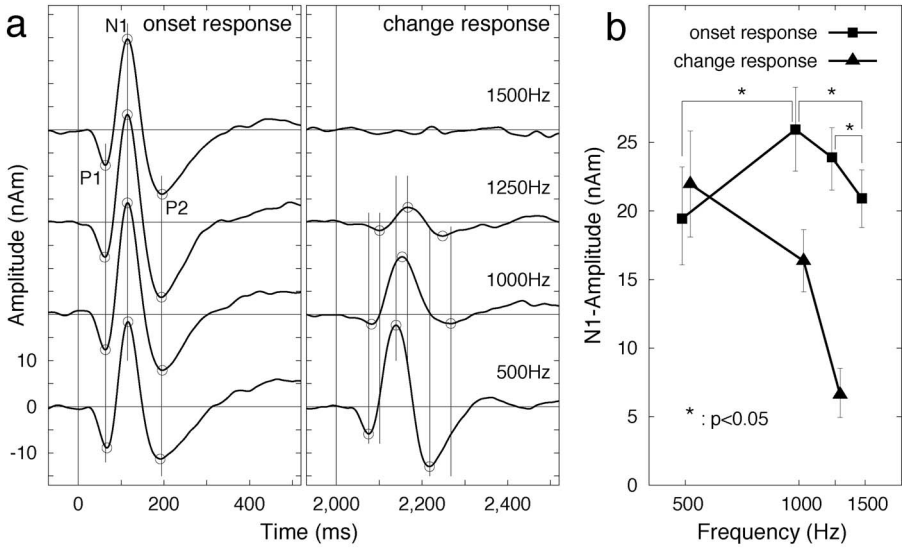

FIG. 6. (a) Averaged waveforms, across left and right hemispheres, for the P1m-N1m-P2m onset (left) and change (right) responses for the four stimulus frequencies. (b) N1m amplitude of the phase change response is affected by stimulus frequency.

TABLE I. Latencies and Amplitudes of onset and phase change responses at 500 Hz.

	Latency (ms)		Amplitude (nAm)	
	onset	phase change	onset	phase-change
P1m right	53±7	65±9*	8.5±4.4	7.0±6.0
P1m left	58±6	74±15*	10.7±3.0	5.9±4.2
N1m right	106±4	129±12*	23.4±9.6	22.4±9.2
N1m left	107±5	133±16*	15.6±6.5	21.4±10.6
P2m right	185±3	219±26*	10.5±9.5	10.4±9.4
P2m left	200±25	214±25	15.2±2.8	15.5±4.8

*Denotes a significant ($p<0.05$) difference between onset and phase change response.

additional contributions from sources in Heschl's gyrus cannot be excluded. The hemispheric asymmetry of a 5 mm more anterior N1m source in the right hemisphere was significant in the group [$t(11)=2.7$, $p<0.02$]. Source locations for the N1m phase change responses were not different from the onset responses suggesting generation of both response types in common or overlapping neural populations. The linear regression analysis applied to the grand averaged magnetic field data revealed that the field variance of the N1m change response could be explained with $r^2=0.956$ by the magnetic field at the maximum of N1m onset response.

D. Individual thresholds

A time-frequency representation of the *p*-values resulting from the Rayleigh test, applied to all time-frequency coefficients, is shown in Fig. 7. Gray shaded areas in the graphics denote time-frequency regions with *p*-values less than 0.1% (corrected <5%).

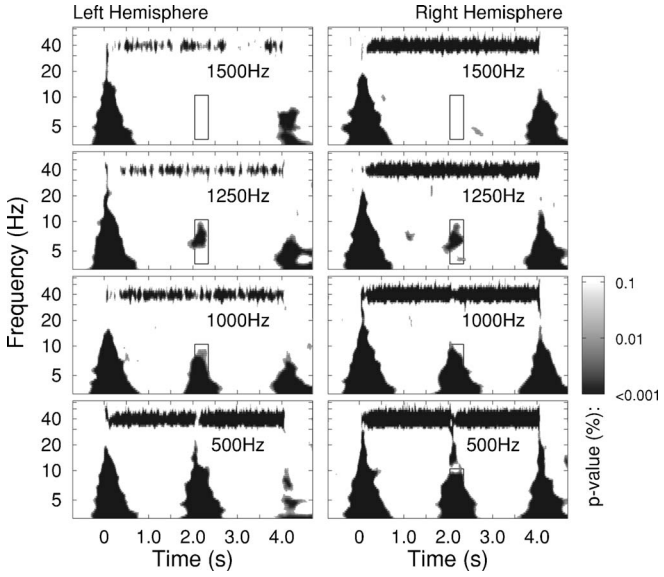

FIG. 7. Individual results of the Rayleigh test applied to the coefficients of a time-frequency decomposition of the response waveforms. The gray shaded areas denote *p*-values less than 0.1%. At 500 Hz, the most prominent response contributions are: the onset, the phase change, and the offset response. Also seen is the 40 Hz steady-state response, seen as a horizontal band of activity at 40 Hz. Rectangular boxes in the 2.0–2.4 s and 3–10 Hz region indicate the range of coefficients used to test for a response to the IPD change.

Characteristic patterns in the time-frequency maps can be related to corresponding time intervals in the response waveforms from the same subject shown in Fig. 4. For instance, the triangular area at the zero time-point (on the *x*-axis) represents the onset response. The triangular shape is partially caused by the temporal resolution of the wavelet transform, which decreases toward low frequencies. A similar pattern can be identified for the offset response at 4.0 s. For this particular subject the offset response is larger in the right than left hemisphere, consistent with the response waveforms shown in Fig. 4. The horizontal bar around 40 Hz reflects the frequency content of the auditory steady-state response evoked by the 40 Hz amplitude modulation. Once again it is more prominent over the right hemisphere, a finding that was consistent across subjects. The 40 Hz steady state response is not seen in Fig. 4, because the waveforms in Fig. 4 were low-pass filtered.

Consistent with the IPD change response shown in Fig. 4, significant time-frequency coefficients are seen at 500 and 1000 Hz stimulus frequency. At 1250 Hz, a smaller but significant set of response coefficients is seen. Upon further analysis, the phase change response was restricted to the 2.0–2.4 s time and 3–10 Hz frequency interval (rectangular boxes in Fig. 7). The boxes in the 1500 Hz region indicate that there was no significant response detected at this stimulus frequency.

The Rayleigh test results in these time-frequency boxes are summarized for all subjects in Fig. 8. Here, the results from left and right auditory cortices have been combined by showing the minimum of the two *p*-values for each time-frequency coefficient. At 500 and 1000 Hz, all twelve subjects showed clearly significant IPD change responses. At 1250 Hz nine subjects showed significant responses and no responses were detected at 1500 Hz. That means that for the subjects 6, 8, and 11, thresholds were above 1000 and below 1250 Hz. For the other nine subjects, thresholds were above 1250 and below 1500 Hz.

A comparison between the physiological and behavioral thresholds is given in Fig. 9. The histogram for the physiological thresholds contains four points on the *x*-axis because four frequencies were tested. The physiologic CDF approximation shows a group mean threshold close to 1250 Hz. The cumulative histogram of behavioral thresholds demonstrates the larger range of individual thresholds. Thresholds fell between 770 and 1682 Hz. The CDF approximation corresponds to the distribution given in Fig. 2, with a group mean of 1203 Hz.

IV. DISCUSSION

We recorded cortical evoked auditory responses to sudden changes in IPD and found the group mean upper frequency limit at 1250 Hz. This physiological threshold corresponded well with the group mean behavioral threshold of 1203 Hz. Both thresholds fall within the 1100–1300 Hz range, commonly reported in the literature as upper frequency limit for ITD/IPD detection.

Cortical responses to changes in ITD have been recorded in EEG studies using noise stimuli (Jones *et al.*,

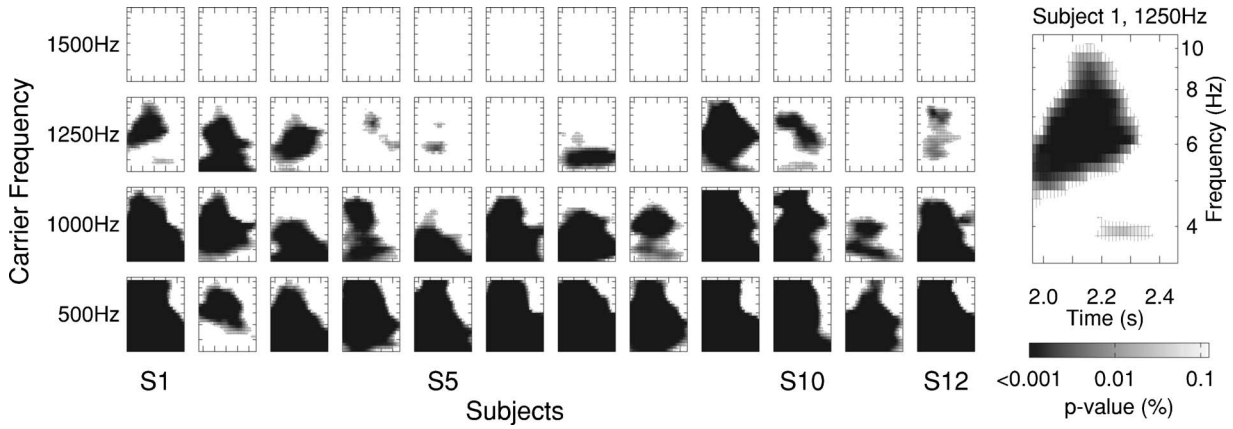

FIG. 8. *P*-values obtained with the Rayleigh test for time-frequency coefficients 3–10 Hz, 0–500 ms after the phase change. Data from subject 1, at 1250 Hz, is shown in the right enlarged panel, as an example. All subjects showed a significant phase change response at 500 and 1000 Hz, nine of twelve subjects showed a response at 1250 Hz and no subject responded to the phase change in the 1500 Hz tone.

1991; McEvoy *et al.*, 1991a, b; McEvoy *et al.*, 1990; Picton *et al.*, 1991). Like in our study, cortical responses were found for interaural changes between noncoherent and fully coherent noise sounds (Chait *et al.*, 2005; Dajani and Picton, 2006; Jones *et al.*, 1991). Several parallels exist between the responses elicited by changes in interaural coherence or ITD and changes in IPD in our study. In all cases the evoked P1-N1-P2 complex was several tens of milliseconds delayed compared to the response to sound onset; suggesting that discriminating changes in interaural disparity requires additional processing time than the mere detection of sound onset. Regardless of stimulus type, strongest effects were found at low frequencies.

A. Possible neural mechanisms underlying the IPD processing

The classical model for processing of ITDs, as proposed by Jeffress (1948), consists of an array of neurons that act as coincidence detectors, responding maximally to synchronous bilateral inputs. Furthermore, the model requires that the neurons receive delayed input from both ears. Such a network creates a topographical representation of the ITD. Recently, an alternative model of sound localization had been proposed, which uses broadly tuned networks of ITD sensitive neurons in the left and right hemisphere and coding of neural firing rate instead of a place code (Harper and McAlpine, 2004; McAlpine, 2005). Common to all models is that they require phase locked bilateral input at the place of binaural computation.

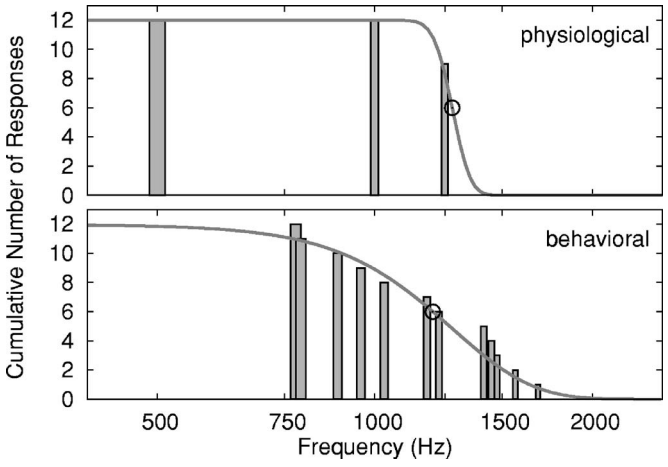

FIG. 9. Cumulative number of responses obtained with the physiological (top) and behavioral (bottom) tests.

Our MEG data do not directly answer the question of underlying structure and cannot directly tell us where or how IPD cues are processed in the brain. Localizing IPD change responses in the auditory cortex could have formally provided evidence for cortical processing of binaural input. However, the upper frequency limit for IPD processing seems too high for a cortical network that directly compares phase information from bilateral auditory inputs. We assume that the IPDs are processed most likely in medial (Yin and Chan, 1990) and lateral (Tollin and Yin, 2005) superior olivary nuclei. Outputs are then projected along the afferent auditory pathway to the cortex. We further speculate that the limitation of useful frequency range for IPD processing occurs at or below the point where bilateral phase locked neural activity converges. An explanation for the upper frequency limit is that phase locked neural activity has a limited temporal acuity. Thus, the thresholds we obtained may reflect the limited phase locking ability in the auditory brainstem or below.

B. Common generators for the change and onset response

The dependency of the N1m onset amplitudes on the stimulus frequency was affected by the presence of a change response. In general we would expect monotonic decrease in N1m amplitude with increasing stimulus frequency (Antinoro *et al.*, 1969; Naatanen and Picton, 1987). In contrast, the N1m amplitude increased between 500 and 1000 Hz (where the change response was largest) and decreased at higher frequencies. A possible explanation is that the onset and change responses share common neural resources. The 500 Hz stimulus elicited the largest IPD change response and could incur increased refractoriness in neural population common for onset and change response. Thus, reduced resources were available for the following onset response. This

situation changed for higher stimulus frequencies where smaller change responses were elicited and the generation of the onset response was unaffected. Thus, we interpret the observed frequency dependency of the onset response as evidence for common neural populations generating both types of responses. The finding that the magnetic field distribution at the peak of the N1m onset response could explain 95% of the variance in the magnetic field at the peak of the change response supports this interpretation.

C. Alternative interpretations

Stimulus frequency was the independent variable in our study. Because the phase change was always a half period, the ITD varied from 1.0 ms at 500 Hz to 0.4 ms at 1250 Hz. It is, therefore, reasonable to question whether the observed threshold reflects the shortest detectable ITD or an upper frequency limit for binaural processing based on ITDs. The ITDs in our study were at least one magnitude larger than those reported in literature for sound localization at corresponding frequencies. According to Zwislocki and Feldman (1956), the threshold for a noticeable change in the ITD of a 1000 Hz pure tone is 20 μs, when tones are presented close to the midline in the horizontal plane. Because of those large differences it seems unreasonable to explain the observed frequency thresholds with the threshold for ITD discrimination. Consequently, we interpret our results as limits for the frequency range that is usable for IPD detection in steady tones.

We should also acknowledge that the current study did not evaluate the effect of stimulus intensity on the detection of IPD changes, because it was not feasible. Therefore, all tests (physiological and behavioral) were conducted using stimulus presentation levels of 60 dB SL. It is reasonable, however, to question if the thresholds obtained in our study would differ with different stimulus intensity levels. Because behavioral performance in sound localization has been shown to be similar, with stimulus intensities between 40 and 90 dB SL, and the upper frequency limit did not change (Zwislocki and Feldman, 1956), we expect that our behavioral thresholds do not depend strongly on intensity. Similarly, McEvoy *et al.* (1991a) showed that EEG responses to changes in the ITD of noise do not change when the stimulus intensity is varied between 55 and 75 dB SL. Thus it is likely that the observed IPD change response is widely independent of the stimulus intensity.

D. Physiological approaches to studying binaural hearing

Two main points should be emphasized. First is the small inter-subject variability in the physiological data, relative to the behavioral data. Second is the good agreement between the group mean results in physiological and behavioral tests even though different stimuli were used. The MEG portion of the study did not involve a behavioral task and instead examined the physiological detection of IPD change contained within the ongoing stimulus. The intention behind using a shorter stimulus, during the behavioral test, was to shorten the test time and prevent subject fatigue. However, a psychoacoustic test that involves discriminating a short tone with or without a change in IPD would be another approach for future studies.

The behavioral threshold data were much more variable than the physiological data. Large inter subject variability in the behavioral data might be explained by the different strategies used by different subjects; individual motivation levels; or varying abilities to attend to the task. Performance variability might have been reduced if we had used highly trained subjects, as is commonly practiced in many psychoacoustical studies. However, we were interested in testing naive listeners under conditions comparable to those that would occur in a clinical test situation. Because physiological results can be less variable, there is interest in developing new techniques for assessing binaural processing in clinical populations with various types of hearing disorders.

Physiological tests, such as the one used in this study, are largely unaffected by the subject's motivation or attention span. For this reason, physiological procedures for assessing binaural interaction components (BIC), in populations with auditory disorders, have been described in the literature. Most BIC tests use auditory evoked brainstem responses to compute the difference between binaurally evoked responses and the sum of left and right monaurally evoked responses (Dobie and Berlin, 1979). The BIC reflects physiological processes underlying ITDs for click stimuli (Furst *et al.*, 1990, 1985) and has been used to assess children with central auditory processing disorders (CAPD) (Delb *et al.*, 2003). However, the brainstem BIC is small and often difficult to record (Stollman *et al.*, 1996). Furthermore, the BIC is derived from responses to click stimuli and likely related to a specific subsystem of binaural processing in the auditory pathway. For these reasons, recording cortical evoked responses to IPDs using low frequency tones, could improve the assessment of central sound processing.

Formally, the BIC could be derived from middle- and long-latency cortical responses and those BIC have larger amplitudes than the brainstem BIC (McPherson and Starr, 1993). However, cortical responses are reasonably well activated by either left or right ear input alone and do not increase to almost twice the size with binaural stimulation (Picton *et al.*, 1985). Thus, it is questionable whether the underlying assumption of linearity for calculating the BIC is valid and cortical BIC could truly represent binaural processing. Because no assumptions about linearity were made, paradigms like the one used in this study might be advantageous for investigating binaural hearing.

The presented method could potentially be used to assess central hearing abilities in individual subjects. One important step toward such application was to demonstrate a statistical test for the significance of individual responses. Another necessary step for clinical application of the paradigm is to demonstrate correlation between individual behavioral and physiological thresholds. Such regression analysis would require larger variation of individual thresholds than those reported in our study. Testing a broad range of individuals with different types of hearing disorders would also permit the opportunity to determine the sensitivity and specificity of this technique. For these reasons, a goal of future

studies is to investigate the physiological detection of IPD in people with various age groups, and with different types of hearing disorders.

V. CONCLUSIONS

Using cortical evoked potentials (P1m-N1m-P2m), we were able to record physiological change responses reflecting IPD processing. The responses could be recorded with stimulus frequencies below an upper limit, which closely approximated the behavioral threshold. Because the physiological responses require phase information through synchronous bilateral inputs at the level of the auditory brainstem, physiological change detection thresholds likely reflect the upper limit of phase synchrony in the brainstem. Therefore, this new approach to studying IPD holds promise for studying normal and disordered physiological processes underlying binaural hearing.

ACKNOWLEDGMENTS

The study was supported by grants from The Hearing Foundation of Canada, the Canadian Institutes for Health Research, and the Canada Foundation for Innovation.

Antinoro, F., Skinner, P. H., and Jones, J. J. (**1969**). "Relation between sound intensity and amplitude of the AER at different stimulus frequencies," J. Acoust. Soc. Am. **46**(6), 1433–6.

Bardouille, T., Picton, T. W., and Ross, B. (**2006**). "Correlates of eye blinking as determined by synthetic aperture magnetometry," Clin. Neurophysiol. **117**(5), 952—8.

Batra, R., Kuwada, S., and Fitzpatrick, D. C. (**1997**). "Sensitivity to interaural temporal disparities of low- and high-frequency neurons in the superior olivary complex. I. Heterogeneity of responses," J. Neurophysiol. **78**(3), 1222–36.

Brungart, D. S., and Rabinowitz, W. M. (**1999**). "Auditory localization of nearby sources. Head-related transfer functions," J. Acoust. Soc. Am. **106**(3 Pt 1), 1465–79.

Chait, M., Poeppel, D., de Cheveigne, A., and Simon, J. Z. (**2005**). "Human auditory cortical processing of changes in interaural correlation," J. Neurosci. **25**(37), 8518–27.

Dajani, H. R., and Picton, T. W. (**2006**). "Human auditory steady-state responses to changes in interaural correlation," Hear. Res. **219**(1-2), 85–100.

Delb, W., Strauss, D. J., Hohenberg, G., and Plinkert, P. K. (**2003**). "The binaural interaction component (BIC) in children with central auditory processing disorders (CAPD)," Int. J. Audiol. **42**(7), 401–12.

Dobie, R. A., and Berlin, C. I. (**1979**). "Binaural interaction in brainstem-evoked responses," Arch. Otolaryngol. **105**(7), 391–8.

Fisher, N. I. (**1993**). *Statistical Analysis of Circular Data* (Cambridge University Press, Cambridge).

Furst, M., Eyal, S., and Korczyn, A. D. (**1990**). "Prediction of binaural click lateralization by brainstem auditory evoked potentials," Hear. Res. **49**(1-3), 347–59.

Furst, M., Levine, R. A., and McGaffigan, P. M. (**1985**). "Click lateralization is related to the beta component of the dichotic brainstem auditory evoked potentials of human subjects," J. Acoust. Soc. Am. **78**(5), 1644–51.

Garner, W. R., and Wertheimer, M. (**1951**). "The effects of interaural phase differences on the perception of pure tones," J. Acoust. Soc. Am. **23**(6), 664–7.

Harper, N. S., and McAlpine, D. (**2004**). "Optimal neural population coding of an auditory spatial cue," Nature (London) **430**(7000), 682–6.

Herdman, A. T., Wollbrink, A., Chau, W., Ishii, R., Ross, B., and Pantev, C. (**2003**). "Determination of activation areas in the human auditory cortex by means of synthetic aperture magnetometry," Neuroimage **20**(2), 995–1005.

Jeffress, L. A. (**1948**). "A place theory of sound localization," J. Comp. Physiol. Psychol. **41**, 35–39.

Jones, S. J., Pitman, J. R., and Halliday, A. M. (**1991**). "Scalp potentials following sudden coherence and discoherence of binaural noise and change in the inter-aural time difference: A specific binaural evoked potential or a "mismatch" response?" Electroencephalogr. Clin. Neurophysiol. **80**(2), 146–54.

Joris, P. X., Carney, L. H., Smith, P. H., and Yin, T. C. (**1994**). "Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency," J. Neurophysiol. **71**(3), 1022–36.

Kuwada, S., Fitzpatrick, D. C., Batra, R., and Ostapoff, E. M. (**2006**). "Sensitivity to interaural time differences in the dorsal nucleus of the lateral lemniscus of the unanesthetized rabbit: Comparison with other structures," J. Neurophysiol. **95**(3), 1309–22.

Lagerlund, T. D., Sharbrough, F. W., and Busacker, N. E. (**1997**). "Spatial filtering of multichannel electroencephalographic recordings through principal component analysis by singular value decomposition," J. Clin. Neurophysiol. **14**(1), 73–82.

Levitt, H. (**1970**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

McAlpine, D. (**2005**). "Creating a sense of auditory space," J. Physiol. (London) **566**(Pt 1), 21–8.

McEvoy, L. K., Picton, T. W., and Champagne, S. C. (**1991a**). "Effects of stimulus parameters on human evoked potentials to shifts in the lateralization of a noise," Audiology **30**(5), 286–302.

McEvoy, L. K., Picton, T. W., and Champagne, S. C. (**1991b**). "The timing of the processes underlying lateralization: Psychophysical and evoked potential measures," Ear Hear. **12**(6), 389–98.

McEvoy, L. K., Picton, T. W., Champagne, S. C., Kellett, A. J., and Kelly, J. B. (**1990**). "Human evoked potentials to shifts in the lateralization of a noise," Audiology **29**(3), 163–80.

McPherson, D. L., and Starr, A. (**1993**). "Binaural interaction in auditory evoked potentials: brainstem middle- and long-latency components," Hear. Res. **66**(1), 91–8.

Naatanen, 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**(4), 375–425.

Picton, T. W., McEvoy, L. K., and Champagne, S. C. (**1991**). "Human evoked potentials and the lateralization of a sound," Acta Oto-Laryngol., Suppl. **491**, 139–43; discussion 144.

Picton, T. W., Rodriguez, R. T., Linden, R. D., and Maiste, A. C. (**1985**). "The neurophysiology of human hearing," Human Communication Canada **9**, 127–136.

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey," J. Neurophysiol. **30**(4), 769–93.

Ross, B., Borgmann, C., Draganova, R., Roberts, L. E., and Pantev, C. (**2000**). "A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones," J. Acoust. Soc. Am. **108**(2), 679–91.

Ross, B., Herdman, A. T., Wollbrink, A., and Pantev, C. (2004). "Auditory cortex responses to the transition from monophonic to pseudo-stereo sound," Neurol. Clin. Neurophysiol. 2004, 18.

Ross, B., Lutkenhoner, B., Pantev, C., and Hoke, M. (**1999**). "Frequency-specific threshold determination with the CERAgram method: Basic principle and retrospective evaluation of data," Audiol. Neuro-Otol. **4**(1), 12–27.

Sarvas, J. (**1987**). "Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem," Phys. Med. Biol. **32**(1), 11–22.

Schiano, J. L., Trahiotis, C., and Bernstein, L. R. (**1986**). "Lateralization of low-frequency tones and narrow bands of noise," J. Acoust. Soc. Am. **79**(5), 1563–70.

Schroeder, M. R. (**1961**). "Improved quasi-stereophony and 'colorless' artificial reverberation," J. Acoust. Soc. Am. **33**, 1061–1064.

Sklare, D. A., and Denenberg, L. J. (**1987**). "Interaural attenuation for tubephone insert earphones," Ear Hear. **8**(5), 298–300.

Stevens, S. S. (**1936**). "The localization of actual sources of sound," Am. J. Psychol. **48**(2), 297–306.

Stollman, M. H., Snik, A. F., Hombergen, G. C., Nieuwenhuys, R., and ten Koppel, P. (**1996**). "Detection of the binaural interaction component in the auditory brainstem response," Br. J. Audiol. **30**(3), 227–32.

Sturzebecher, E., and Cebulla, M. (**1997**). "Objective detection of auditory evoked potentials. Comparison of several statistical tests in the frequency domain on the basis of near-threshold ABR data," Scand. Audiol. **26**(1), 7–14.

Tesche, C. D., Uusitalo, M. A., Ilmoniemi, R. J., Huotilainen, M., Kajola, M., and Salonen, O. (**1995**). "Signal-space projections of MEG data characterize both distributed and well-localized neuronal sources," Electroencephalogr. Clin. Neurophysiol. **95**(3), 189–200.

Tollin, D. J., and Yin, T. C. (**2005**). "Interaural phase and level difference sensitivity in low-frequency neurons in the lateral superior olive," J. Neurosci. **25**(46), 10648–57.

Vrba, J., and Robinson, S. E. (**2001**). "Signal processing in magnetoencephalography," Methods **25**(2), 249–71.

Yin, T. C., and Chan, J. C. (**1990**). "Interaural time sensitivity in medial superior olive of cat," J. Neurophysiol. **64**(2), 465–88.

Yin, T. C. T. (**2002**). *Neural Mechanism of Encoding Localization Cues in the Auditory Brainstem. Integrative Functions in the Mammalian Auditory Pathway*, edited by D. Oertel (Springer, New York), pp. 99–159.

Zwislocki, J., and Feldman, R. S. (**1956**). "Just noticeable differences in dichotic phase," J. Acoust. Soc. Am. **28**(5), 860–864.