

Human Auditory Cortical Dynamics During Perception of Long Acoustic Sequences: Phase Tracking of Carrier Frequency by the Auditory Steady-state Response

Aniruddh D. Patel^{1,*} and Evan Balaban^{1,2,3*}

¹The Neurosciences Institute, 10640 John Jay Hopkins Drive, San Diego, CA 92121 and ²Neurosciences Program, Building 6S Room 320, CUNY College of Staten Island, 2800 Victory Boulevard, Staten Island, NY 10314, USA

³Present address: Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal, Quebec, Canada H3A 1B1

*The authors contributed equally to this work

We recorded human auditory cortical activity during the perception of long, changing acoustic signals and analyzed information provided by dynamic neural population measures over a large range of time intervals (~24 ms–5 s). Participants listened to musical scales that were amplitude modulated at a rate of 41.5 Hz, generating an ongoing, stimulus-related oscillatory brain signal, the auditory steady-state response (aSSR). The aSSR generated energy at the amplitude modulation rate that was recorded using magnetoencephalography. As in previous work, the timing (phase) of this response varied with stimulus carrier frequency over the entire course of minute-long tone sequences ('phase tracking' of carrier frequency). The length of the time interval over which phase was calculated was systematically varied; significant phase tracking was regularly observed at analysis intervals of <50 ms in length. The right auditory cortex exhibited better phase tracking performance than the left at analysis intervals of 24–240 ms, and frequency dependent phase delays were consistently larger than those predicted by cochlear mechanics. Based on these empirical data, a model of the neural populations responsible for phase tracking suggests that it is produced by a subpopulation (~25%) of the cells generating the aSSR.

Keywords: human auditory cortex, magnetoencephalography, music perception, neural dynamics, steady-state responses, tone sequence perception

Introduction

Much research on human auditory cortical physiology has used relatively short and spectrally simple stimuli to study population auditory neural responses called evoked potentials (EPs). In the typical EP paradigm a brief stimulus, such as a 100 ms pure tone, is presented repeatedly, and neural activity time-locked to the stimulus is recorded either electrically (EEG) or magnetically (MEG) and averaged across a large number of trials. The result is a series of negative and positive voltage or magnetic field strength deflections as a function of time. Investigators usually examine the shape and/or overall magnitude of the deflections during a particular time interval. The shapes and magnitudes of these deflections are thought to be the result of the anatomical distribution of active cells, the orientation of their current-producing sources, their degree of synchrony, their size, and their number and overall activity level (Nunez, 1981). EPs are thus well suited to analyzing large-scale brain responses during a predefined short period of time, especially when there is good reason to believe that successive responses to the sounds are identical, so that averaging reveals a robust underlying neural 'signature' normally buried in noise.

For auditory neuroscientists interested in human perception in more naturalistic contexts, the EP paradigm has limitations. The two most important forms of human acoustic communication – speech and music – do not consist of repeated, acoustically similar sounds. They are dynamically changing sound sequences with little exact repetition. Thus it would be highly desirable to find brain measures of *ongoing*, rather than transient, stimulus-related activity associated with the perception of sound sequences. While dynamic physiological imaging methods such as fMRI and PET can provide ongoing measurements with excellent spatial resolution (Cabeza and Kingstone 2001), their temporal resolution of 0.5 s to several seconds at best is inherently limited by slow hemodynamic or biochemical responses that give rise to the measured signals. Auditory processing for both speech and music requires temporal resolution on the order of tens to several hundreds of milliseconds (cf. Patel, 2003). While EEG and MEG both provide such resolution, MEG recording sensors have greater spatial independence (Lewine and Orrison, 1995).

A major challenge for paradigms using ongoing stimulus presentations is to distinguish *stimulus-related* activity from other brain signals recorded during an experimental session. This is possible using a brain response known as the 'auditory steady-state response' (aSSR; Galambos *et al.*, 1981), a cortical signal recorded in response to continuous amplitude modulation (AM) of an acoustic stimulus, which is present as long as the stimulus is on (it is non-refractory). Localization studies suggest that the aSSR arises from sources in each primary auditory cortex (Heschl's gyrus; Gutschalk *et al.*, 1999; Pastor *et al.*, 2002). The aSSR oscillates at the frequency of the acoustic AM rate, and its power is greatest when AM is in the 40-Hz range (Hari *et al.*, 1989; Roß *et al.*, 2000, 2002). Figure 1 illustrates the basic phenomenon of the aSSR. Figure 1a shows a pure tone amplitude modulated at 40 Hz. Figure 1b shows the power spectrum of an MEG signal recorded over auditory cortex when a human listener hears the tone in Figure 1a. The arrow shows a prominent peak at the AM rate; this peak is absent when an unmodulated (non-AM) pure tone of the same carrier frequency is played to the listener.

Most aSSR research has used an event-related approach, focusing on waveform characteristics and source location as determined from average responses to repeated identical stimuli. However, two prior studies have examined how the aSSR changes with time. Galambos and Makeig (1988) examined changes in aSSR magnitude and phase over tens of minutes while participants listened to music or drifted between wakefulness and sleep, and identified ~60 s periodicities which they termed 'minute rhythms'. Roß *et al.* (2002) examined the temporal evolution of the aSSR at the onset of

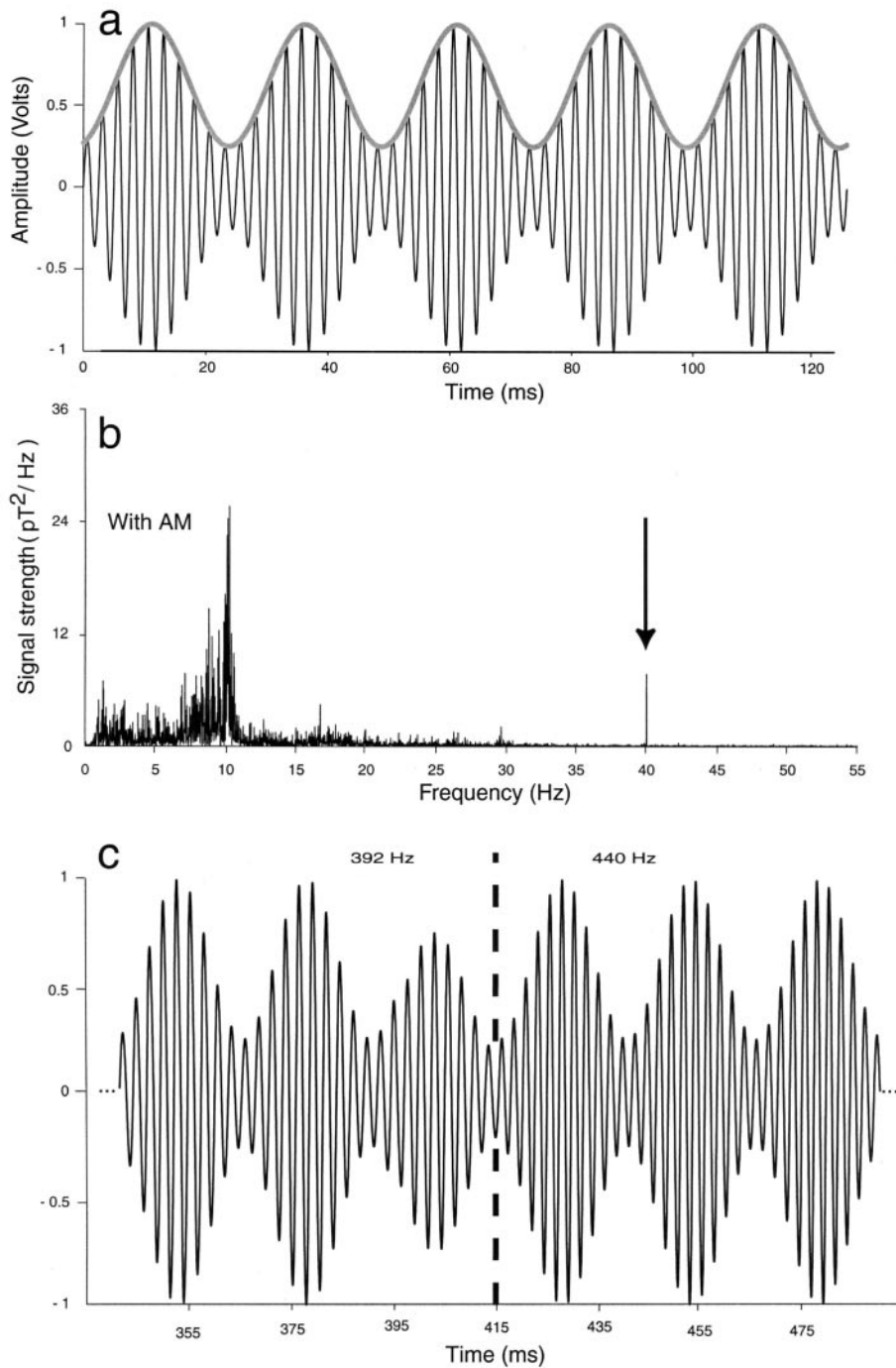


Figure 1. Illustration of the auditory steady-state response (aSSR). (a) A pure tone (carrier frequency = 400 Hz, black line) amplitude modulated at 40 Hz (AM envelope, gray line). (b) The energy spectrum of an MEG signal recorded over auditory cortex in a human listener who heard the tone in (a) played continuously for one minute (pT = picotesla). A peak in the spectrum is clearly visible at 40 Hz, due to the aSSR (arrow). This peak disappears from the MEG power spectrum when the listener hears the same tone without AM. (c) Transition between tones of two different carrier frequencies. The vertical black dotted line denotes that point at which the carrier frequency changes from 392 to 440 Hz. Note the continuity of the AM envelope, in this case at 41.5 Hz. The stimuli in this study changed carrier frequency every 415 ms while the constant AM generated the aSSR.

brief AM tones, showing that the aSSR builds up in magnitude over 200 ms starting ~40 ms after stimulus onset (and is thus distinct from the transient evoked gamma band response to tone onsets). A dynamic approach examining how aSSR characteristics change over time during the presentation of changing acoustic sequences ~1 min in length is presented here.

Previously, dynamic analysis has demonstrated that the relative timing of aSSR oscillations (measured by extracting the signal's phase) significantly co-varies with changing carrier frequencies when subjects are played pure tone sequences modulated with 41.5 Hz AM (Patel and Balaban, 2000). (Throughout this paper 'phase' refers to phase relative to the

acoustic AM, rather than absolute latency between stimulus presentation and cortical response. To study the latter, one needs to take into account phase delays introduced by digital filters, sound conduction devices such as tube-phones or headphones, middle ear transmission time, etc.) As carrier frequency increased, aSSR phase advanced and vice versa, consistent with research based on event-related approaches (Galambos *et al.*, 1981; John and Picton, 2000; Roß *et al.*, 2000). A carrier-frequency-like pattern of phase advances and delays over the 1 min stimulus period could be reliably seen at single sensors within single trials in each participant, a phenomenon termed 'phase tracking'. A question of central interest raised by these findings was whether phase tracking would be observed if shorter or longer analysis epochs had been employed (the original study used 2 s analysis epochs). The present study was designed to examine aSSR dynamics at analysis durations encompassing a few tens of milliseconds to 5 s, during stimuli lasting for 1 min. Changes in the correlation between (brain) phase-time and (stimulus) frequency-time waveforms with varying analysis durations were analyzed to see if they could shed any light on biological mechanisms involved in phase tracking.

If the fundamental neural responses contributing to phase tracking are the result of neural integration operations performed over relatively short time intervals, phase tracking should be observed at relatively short analysis durations. If the integration operations require a minimum time interval, phase tracking should only emerge above a critical analysis length (provided it is longer than the shortest length of ~24 ms used in this study). We find that aSSR phase tracking is consistently evident at analysis durations of <50 ms, and that the temporal characteristics of tracking are consistently different on the right and left sides of the cortex. A specific neural model of phase tracking is also suggested.

Materials and Methods

Participants

The participants were ten right handed individuals (six males) with a mean age of 36.2 years (range 28–47) who gave informed consent and had normal hearing (audiometric testing carried out with a Grayson-Stadler GSI-65 Audiometer). Four had studied music for >5 years, while the others had little or no musical training.

Stimuli

Seven tone sequences were created using SIGNAL (Engineering Design, Belmont, MA). Each sequence was 62.25 s long, and consisted of 150 pure tones of 415 ms each with no pauses. Sequences ascended and descended in frequency in discrete steps according to Western musical scales, with five upward and downward traversals of a scale per sequence. In each sequence, carrier frequencies ranged between 220 and 880 Hz (i.e. musical A3 and A5). The sequences differed slightly from one another in that each was chosen to conform to one of seven Western diatonic musical modes [Ionian ('major scale'), Dorian, Phrygian, Lydian, Mixolydian, Aeolian ('minor scale') and Locrian]. These sequences differ by a semitone (1/12 of an octave) in a few of their constituent notes, thus preserving the same shape of the sequential contour of pitches over time (this study was not designed to examine whether such small differences can be discriminated in the brain response). Table 1 lists the frequencies of the constituent notes for all of the stimuli. The amplitude of each tone was set to ± 1 V, with the last 20 ms of each tone being set to 0.75 V. The entire tone sequence was amplitude-modulated at a rate of 41.5 Hz to a depth of 0.25 of its maximum amplitude using a \cos^2 envelope (a modulation depth of 60%). Thus, while carrier frequency changed every 415 ms, the AM rate stayed constant throughout the sequence. Figure 1c shows the conti-

Table 1

Carrier frequencies of tones used in musical scales

Ionian	Dorian	Phrygian	Lydian	Mixolydian	Aeolian	Locrian
220.00	220.00	220.00	220.00	220.00	220.00	220.00
246.94	246.94	233.08	246.94	246.94	246.94	233.08
277.18	261.63	261.63	277.18	277.18	261.63	261.63
293.66	293.66	293.66	311.13	293.66	293.66	293.66
329.63	329.63	329.63	329.63	329.63	329.63	311.13
369.99	369.99	349.23	369.99	369.99	349.23	349.23
415.30	392.00	392.00	415.30	392.00	392.00	392.00
440.00	440.00	440.00	440.00	440.00	440.00	440.00
493.88	493.88	466.16	493.88	493.88	493.88	466.16
554.37	523.25	523.25	554.37	554.37	523.25	523.25
587.33	587.33	587.33	622.25	587.33	587.33	587.33
659.26	659.26	659.26	659.26	659.26	659.26	622.25
739.99	739.99	698.46	739.99	739.99	698.46	698.46
830.61	783.99	783.99	830.61	783.99	783.99	783.99
880.00	880.00	880.00	880.00	880.00	880.00	880.00

Each column shows the carrier frequency progression (in Hz) for one ascending scale. Tone sequences consisted of five alternating ascending and descending repetitions of one scale with no intervening silences. The music theoretic names of each scale are given at the top of each column of frequencies.

nity of the AM envelope and the stimulus tones at the boundary between tones of two different frequencies. The stimuli are available for downloading or listening at http://www.nsi.edu/users/patel/sound_examples/phase_tracking.

MEG Recordings

Whole-head neuromagnetic signals were collected using a Magnes 2500WH MEG system (4-D Neuroimaging) in a magnetically shielded room, while participants sat in a reclined position. This system provides 148 magnetometer coil sensors (2 cm in diameter) spaced 3 cm apart on an approximately ellipsoidal surface located ~3 cm from the scalp surface. Stimuli were delivered binaurally over non-magnetic ER30 tube-phones (Etymotic Research) at a comfortable level. Participants were instructed to remain awake and attend to the sound sequences. Each participant heard all seven sequences in a different random order, yielding seven runs per individual. Data were sampled at 678.17 Hz and bandpass filtered from 1 – 100 Hz online during data acquisition. Runs with magnetic flux jumps or excessive eye blinks were discarded and repeated. Acoustic distortion of the stimulus envelope resulting from sound transmission through the tube-phones was quantitatively examined, and could not account for the carrier-frequency dependent phase delays we observed in MEG recordings. The ~2-fold individual variation in carrier-frequency dependent phase delays found among our subjects (see Figure 6b) is also not compatible with an effect produced by the experimental equipment.

Data Analyses

The analyses were carried out using custom-written MATLAB programs (The Mathworks, Natick, MA). Statistical analyses also utilized SYSTAT (SPSS, Chicago, IL) and Statview (SAS, Cary, NC).

Measuring Phase Tracking

The following procedure was followed for each participant, MEG sensor and run. Data from each sensor were digitally resampled (RESAMP, Engineering Design) at 664 Hz prior to Fourier analysis in order to have 16 time points per 41.5 Hz cycle. This ensured that Fourier transforms which were an integer multiple of 16 points in length had a bin precisely centered on 41.5 Hz. Following resampling, data were discrete Fourier transformed (DFT) and the magnitudes and

phases of the 41.5 Hz Fourier coefficients were extracted. The population of phase angles computed for each sensor and run were rotated so that they were centered around 0, to avoid phase unwrapping discontinuities. This analysis was conducted independently at 30 DFT lengths for each sensor and run. The minimum analysis duration was 16 points per DFT (1/41.5 or ~24.1 ms, 2583 DFTs per sequence), and the maximum duration was 3360 points per DFT (~5 s, 12 DFTs per sequence). These 30 DFT lengths spanned two broad regions: 'short DFT lengths' [~24.1 ms to ~241 ms per DFT (16 points to 160 points in multiples of 16 points: 16, 32, 48, 64, 80, 96, 112, 128, 144, 160)], and 'long DFT lengths' [~480 ms to ~5 s per DFT [320 points to 3360 points in multiples of 160 points: 320, 480, 640, 800, 960, ..., 3360]]. At each DFT length, the correlation between the phase-time series and the resampled stimulus carrier frequency-time series was calculated (cf. Fig. 2a,b). Resampling of the stimulus was carried out to assure that 'time averages' of stimulus and brain response parameters were made in a precisely comparable fashion. The resampled stimulus carrier frequency-time series for a given DFT length was constructed by taking the mean stimulus carrier frequency during each DFT epoch, expressed in semitones with respect to 440 Hz. The phase-frequency correlation as a function of DFT length was termed the 'correlation contour' (cf. Fig. 2c). One correlation contour was generated for each sensor and run in the study. Because the number of phase and frequency values used to compute each correlation in the correlation contour differ at each DFT length (longer DFT lengths = fewer values), the criteria for significance of correlation also differed (cf. Fig. 2c, blue dotted line). This criterion value was computed based on a bootstrap using uniform random complex numbers to generate phase values.

For each participant, we also computed signal-to-noise ratios (SNRs) for each sensor and run. SNR was computed as the ratio of the average energy in the 41.5 Hz frequency bin to the average energy in the frequency bins within 5 Hz of this bin. This calculation was done at each DFT analysis length, yielding 30 SNR values for each sensor on each run. For short DFT lengths, frequency bin width was >2.5 Hz, so only one bin on either side of 41.5 Hz was used to compute noise levels. At the shortest DFT length, this quantity could not be computed, since the bin below 41.5 Hz also represented DC values. The shortest DFT length was therefore excluded from SNR analysis.

Identifying Phase Tracking Sensors and the Overall Quality of Tracking

For each participant, a 'tracking bank' of sensor locations was chosen based on how well the phase-time series from different channels correlated with the stimulus carrier frequency-time series across runs. To be included in a tracking bank, a sensor had to have a significant correlation between its phase-time contour and the stimulus frequency-time contour ($P < 0.05$) at more than one-half of the DFT analysis lengths on more than one-half of the stimulus presentations. That is, across the seven correlation contours computed for each sensor in a given individual, four or more of these contours had to have significant correlations at 15 or more of the 30 DFT lengths for the sensor to be included in that individual's tracking bank. (The particular DFT lengths which had significant correlations could be different from run to run.) The number of sensors in an individual's tracking bank ranged from 10 to 40 (mean \pm SEM: 23.2 ± 3.2 per participant). Subsequent analyses were carried out using only the sensors in the tracking bank.

To study tracking performance within a participant, correlation contours from tracking bank sensors were divided into three categories. Those contours which did not have correlations above criterion at 15 or more DFT lengths were designated 'nontracking contours', while those that did were designated 'tracking contours'. To further classify the correlation contours, each contour was averaged across all 30 DFT lengths to yield an average tracking value (for example, averaging the values of the black points in Figure 2c would yield one such value). For the 'tracking contours', these average tracking values were ranked from highest to lowest, and the median value was identified. Tracking contours whose average correlation was at and above the median value for that participant were designated as 'top 50% tracking contours', while the remainder were classified as 'bottom 50% tracking contours'. To give an idea of the numbers of contours in the different categories, a representative participant had 23 sensors in his/her

tracking bank (23×7 experimental runs = 161 contours), yielding 63 top 50% tracking contours, 62 bottom 50% tracking contours, and 36 nontracking contours.

To examine the overall quality of tracking at a particular sensor, the three categories of correlation contours were assigned numerical values (top 50% tracking contours were assigned a value of 2, bottom 50% tracking contours a value of 1, and nontracking contours a value of 0). For each participant, the seven resulting 'quality scores' at each tracking bank sensor were averaged to yield a 'mean quality of tracking' value. These mean values were divided into three categories. 'Poor tracking sensors' had mean values ≤ 0.8 (40% or less of the maximum value of 2); 'intermediate tracking sensors' had mean values >0.8 and <1.4 (between 41% and 69% of the maximum value); and 'good tracking sensors' had mean values ≥ 1.4 (70% or more of the maximum value).

Determining the Minimum DFT Length at which Phase Tracking is Observed

The following procedure was adopted for determining the shortest DFT duration at which significant phase tracking was observed. The formula for the probability of a single channel having r out of n runs with a significant correlation at the $P < 0.05$ level at one DFT length, taking into account multiple comparisons at 30 different DFT lengths, is $30 \times (0.05)^r (0.95)^{n-r} \times n! / [r!(n-r)!]$. For $n = 7$ runs, at least four significant runs would be needed to obtain a P value <0.05 (P in this case ~0.006). However, when this value is corrected for multiple sensor comparisons (max = 40 sensors in a tracking bank in our study, $0.006 \times 40 = 0.24$), it would no longer meet the criterion for significance. With five significant runs, the formula generates a P -value of 0.00018 ($0.00018 \times 40 = 0.007$), which remains below the 0.05 level of significance after taking multiple sensor comparisons into account. Each sensor in an individual's tracking bank was therefore screened for the shortest DFT length at which significant phase tracking was observed on five or more runs. This was chosen as the minimum DFT length for significant phase tracking at that sensor.

Phase Range Analysis

To determine the physical range over which aSSR phase varied in response to the changing carrier frequencies in the stimuli, it is not adequate to simply calculate the difference between the phase values associated with the minimum and maximum frequencies (220 and 880 Hz in this study). This arises from the fact that the observed range of a phase-time series depends on the DFT length used to derive that phase-time series. Longer DFT lengths result in a narrower phase range, due to averaging phase over increasingly larger sections of the ascending-descending phase pattern. Measures of phase range based on phase-time contours computed at one DFT length, say 720 ms DFTs, yield a different value than measures based on another DFT length, say 3.61 s DFTs (cf. Fig. 2a,b). Basing phase range estimates solely on phase-time contours computed at the shortest DFT lengths (~24 ms in this study) is also unreliable because at these short DFT lengths the phase-time contour is very noisy.

The procedure utilized here regards the observed decrement in phase range with increasing DFT length as a basis for estimating the true range of the phase-tracking signals. A set of 16 idealized phase-time contours was constructed at the shortest DFT length used (= 1 AM cycle, 2583 points long). These were ascending-descending patterns that mimicked the stimulus frequency-time contour. The distance between the highest and lowest points of each contour was set at one phase range value (smallest: $\pi/8$, largest: 2π). The phase ranges of successive idealized contours were $\pi/8$ apart (at the AM rate of 41.5 Hz, the difference between these successive steps corresponds to 1.51 ms). For each sensor and experimental run, the measured phase-time contour at each DFT length was compared to each idealized phase-time contour averaged over the same DFT length, and the absolute difference (sum of the absolute value of the difference between the measured and the idealized contour) was recorded, resulting in 16 difference values at each DFT length. Since the recorded phase contours and the idealized phase contours both shrank in phase range as DFT length increased, this procedure permitted identification of the idealized phase contour whose pattern of shrinking best matched the real data (minimum difference over all 30 DFT lengths). The phase range of this 'best fitting' idealized contour provided the estimate of the

phase range of that sensor on that run. This procedure was repeated for every run of every tracking bank sensor in each participant.

The values generated by this analysis were compared with an estimate of the cochlear delay between 220 and 880 Hz (Greenberg *et al.*, 1998). The present paper uses AM signals, whose spectral sidebands may cause the cochlear delay to differ from these estimated values. Greenberg *et al.* (1998 found that amplitude modulation caused cortical latency differences between tones of different carrier frequencies recorded with MEG to be diminished relative to their 'pure tone' values. This would suggest that the AM tones used in this study should result in smaller cochlear delays relative to pure tone stimuli. Pure tone cochlear estimates were used because they appeared to provide conservative values for comparison with the experimental data.

Simulation Model of Phase Tracking

The purpose of the simulation was to see if a simple model of neural response types could explain the observed pattern of correlation increase with increasing DFT analysis length. The logic underlying the model is that the aSSR waveform at each sensor where tracking is observed consists of a linear sum of two brain response components.

Component 1

Component 1 is a signal whose phase variation perfectly tracks the pitch of the stimulus. The relative strength of this signal is represented by the proportionality constant τ . Perfect phase tracking would arise from a heterogeneous population of cells strongly phase-locked to the AM envelope. These cells would consistently fire in a phase-locked fashion to the envelope of the stimulus. However, their firing phase relative to each other (i.e. the exact point during each AM cycle when the cells would fire) would vary according to the carrier frequency the cells are tuned to. Cells that respond to higher frequencies would respond relatively earlier than cells that are tuned to lower frequencies. The degree to which higher-frequency cells respond earlier than lower frequency cells was incorporated in the model by using actual phase ranges recorded for each individual subject. Thus the constant τ can be thought of as the proportion of cells generating the aSSR that are strongly phase-locked to the AM envelope AND whose relative phases vary consistently according to their carrier frequency tuning.

Component 2

Component 2 is a signal with uniform random phase variation. This would arise from cells that respond to the AM (and therefore contribute to the aSSR) but that do not have strong phase locking and/or do not have consistent relative phase variation among cells tuned to different carrier frequencies. The relative strength of this signal is represented by the proportionality constant $(1 - \tau)$.

The simulation was conducted as follows. For each participant, we used the phase range data from their top 50% tracking contours. For each phase range value, a simulated phase-tracking brain signal was generated according to the following equation:

$$\text{brain signal} = \{\tau \exp[j \times \text{pitches}(t) \times ((\phi_r(t))/2)]\} + \{(1 - \tau) \exp[j \times \text{noise}(t)]\}$$

where τ is the proportion of component 1 responses; $\text{pitches}(t)$ is the scaled pitch contour (ranges from -1 to 1); $\phi_r(t)$ is the phase range; and $\text{noise}(t)$ is the uniform random noise (ranges from $-\pi$ to π).

Each simulated brain signal corresponded to the shortest DFT analysis length of 24.1 ms, i.e. was 2583 points long. This was accomplished by making the scaled pitch contour in the above equation 2583 points long, with each point representing the average pitch of the musical scales during a short DFT interval (the noise signal had the same length). Once the tracking signal and the noise were added together, the resulting signal was analyzed for phase as a function of time and for the correlation of the phase time series with the pitch time series. Just as with real data, this analysis was conducted at 30 DFT lengths, yielding a simulated correlation contour. This entire process was repeated for each phase range value, and the resulting population of simulated correlation contours for each participant were then averaged to form a grand average correlation contour.

τ is the only free parameter in this model. For each participant this simulation was repeated 100 times for 100 different values of τ ranging from 0.01 to 1.00 in increments of 0.01. The value of τ that yielded the

best fit between model and data (sum of the absolute value of the difference between the data contour and the model contour) was chosen as the τ for that participant.

Results

aSSR Phase Tracking of Stimulus Carrier Frequency and its Spatial Distribution

Figure 2 illustrates the basic phenomenon of phase tracking, with a record from a single sensor obtained during one run in one individual. Figure 2a shows the phase-time contour and resampled pitch-time contour based on a DFT length of 480 points (~720 ms). Figure 2b shows data from the same sensor/run when analysis is based on a DFT length of 2400 points (~3.61 s). Using different DFT lengths results in a different number of phase and frequency values ($n = 86$ vs. 17 points, respectively). The criterion for significant correlation between phase and frequency contours increases with increasing DFT length, as shown by the dotted line in Figure 2c and described in Materials and Methods. Figure 2a,b represents two of the 30 DFT lengths analyzed for this channel: their numerical correlation values are given as insets and are graphically indicated in the overall 'correlation contour' of Figure 2c by dashed arrows. For this sensor (circled in Fig. 3c) and run (with an extremely good tracking performance), there is a strong similarity between the carrier frequency contour and the phase-time contour of the aSSR over the entire stimulus presentation period of more than one minute.

Figure 3a,b shows the spatial distribution of aSSR phase correlation with stimulus carrier frequency for one representative participant on the same experimental run. Figure 3a plots the correlation values for all sensors in the tracking bank at a relatively short analysis length (DFT duration = 96.4 ms), while Figure 3b shows the same data with a longer analysis length (DFT duration = 1.2 s). Figure 3c shows a composite of the tracking banks of all ten participants, indicating how many individuals had a particular sensor location in their tracking bank, with an indication of the mean quality of tracking at each of these sensor locations averaged over all 10 participants (see Materials and Methods). As in previous work with AM pure-tone signals (Patel and Balaban, 2000, 2001), the sensors showing aSSR phase correlations with stimulus carrier frequency were distributed in a roughly symmetrical pattern consistent with sources in both auditory cortices. There was a nonsignificant tendency for more tracking sites on the right side across all individuals (136 right, 96 left, $P = 0.077$, Fisher's exact test). The mean quality of tracking averaged over all subjects (poor, intermediate, good) for sites on the right and left did not differ (mean rank for right, left channels 33, 36, respectively, Mann-Whitney U -test, $U = 512$, $U' = 643$, $n = 33$ left sites, 35 right sites, $P = 0.42$).

Dependence of aSSR Phase Tracking on DFT Analysis Length

To examine phase tracking as a function of DFT length, correlation contours (such as that shown in Fig. 2c) were computed for each sensor and run of each participant's data. A mean contour was then calculated for each individual participant's tracking bank, using the top 50% tracking contours (see Materials and Methods). Figure 4a,b shows representative mean contours for two participants (black lines), one who had low variance and the other who had the highest variance. For

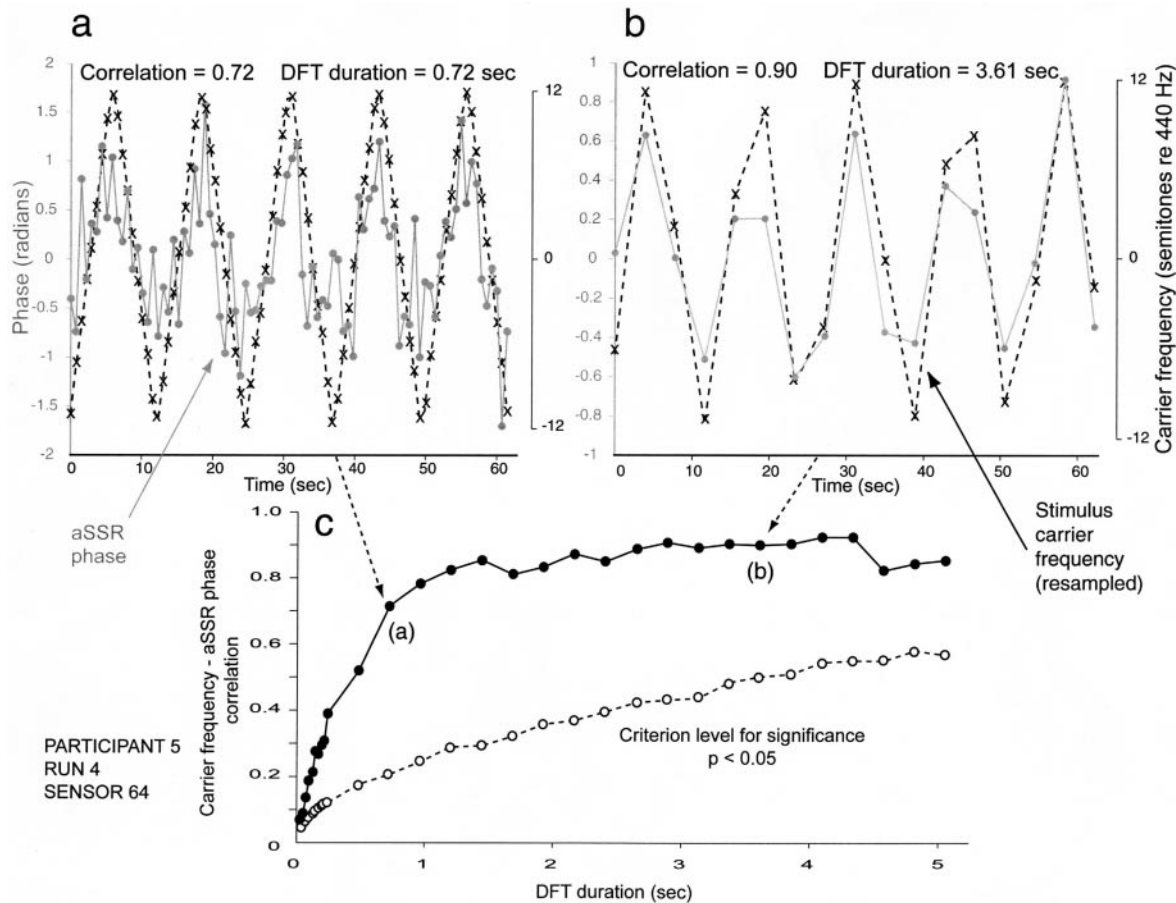


Figure 2. aSSR phase tracking at different analysis durations. The data represented in this figure were recorded from a single sensor (indicated by a circle in Fig. 3c), run, and participant. (a,b) Gray dots and solid gray lines are aSSR phase values (gray y-axis on the left), calculated from non-overlapping successive DFTs of the MEG signal recorded while the participant heard a ~1 min long carrier frequency pattern. This pattern, represented by the slow (-0.1 Hz) modulation of carrier frequency (denoted by black crosses and dotted black lines), consisted of 415 ms tones which ascended and descending in frequency in discrete steps (black y-axis on right). The carrier frequency pattern shows the average frequency during each DFT epoch. The y-axis ranges differ in (a) and (b) because longer analysis durations result in smaller phase range excursions (see Materials and Methods). Correlations improve with increasing analysis duration because of improving signal-to-noise ratios. (c) Black dots show the correlation between the aSSR phase time-series and the resampled stimulus carrier frequency time-series at 30 DFT durations. Together these points form a ‘correlation contour’ for this sensor and run in this participant. The points representing the data shown in (a) and (b) are indicated by dashed black arrows. The white points and dotted line in (c) represent the $P < 0.05$ criterion for significance at each DFT duration.

comparison, the average SNRs (gray lines) are also shown: these average SNR curves were computed from the same sensors and runs that yielded the grand average correlation contour. Also shown in Figure 4a,b are the $P \leq 0.05$ criteria for significance at each DFT length (gray dots). The correlation contours for all participants were similar and had asymptotes at DFT analysis lengths between 2 and 3 s. All participants showed mean correlation values at short DFT lengths (between 24 and 241 ms per DFT) that were significantly greater than chance, indicated by mean correlation values above the gray dots (error bars show 95% confidence intervals for the means).

Hemispheric differences in tracking overlooked by analyses with durations of half a second or longer (such as PET, fMRI or the results shown in Fig. 3b) may achieve significance when analyzed with shorter DFT durations (cf. Fig. 3a). We examined hemispheric differences in the top 50% tracking contours within each individual. Nine out of the ten participants had such contours on sensors over both sides of their heads. Average correlation contours were calculated separately for these right and left top 50% tracking contours for each participant. The difference between the right and left mean correla-

tion contours (right minus left) was then calculated for each participant at each DFT duration (9 participants \times 30 DFT durations, $R - L$). For each successive block of 10 DFT durations (90 values: 9 individuals \times 10 DFT durations), the significance of the $R - L$ difference was calculated using the Wilcoxon matched pairs sign-rank test, with Bonferroni correction for the three statistical comparisons (one for each block of 10 DFT durations). At the shortest 10 DFT durations (24–241 ms), there was a significant tracking advantage for the right side relative to the left (72/90 differences positive, $T^+ = 3285$, $P < 0.0001$). This difference was not shown at the intermediate 10 or the longest 10 DFT durations (intermediate: 482 ms – 2.65 s, 55/90 differences positive, $T^+ = 2599$, $P = 0.08$; long: 2.89 s – 5.06 s, 50/90 differences positive, $T^+ = 2415$, $P = 0.42$). Figure 3a,b shows the distribution of correlation values over the head of one subject at analysis durations of 96.4 ms and 1.2048 s, respectively, illustrating the difference in hemispheric asymmetry at short vs. intermediate analysis durations.

These analyses suggest (i) that changes in aSSR phase correlation with the stimulus at different DFT lengths are significantly better than chance even at short analysis durations with

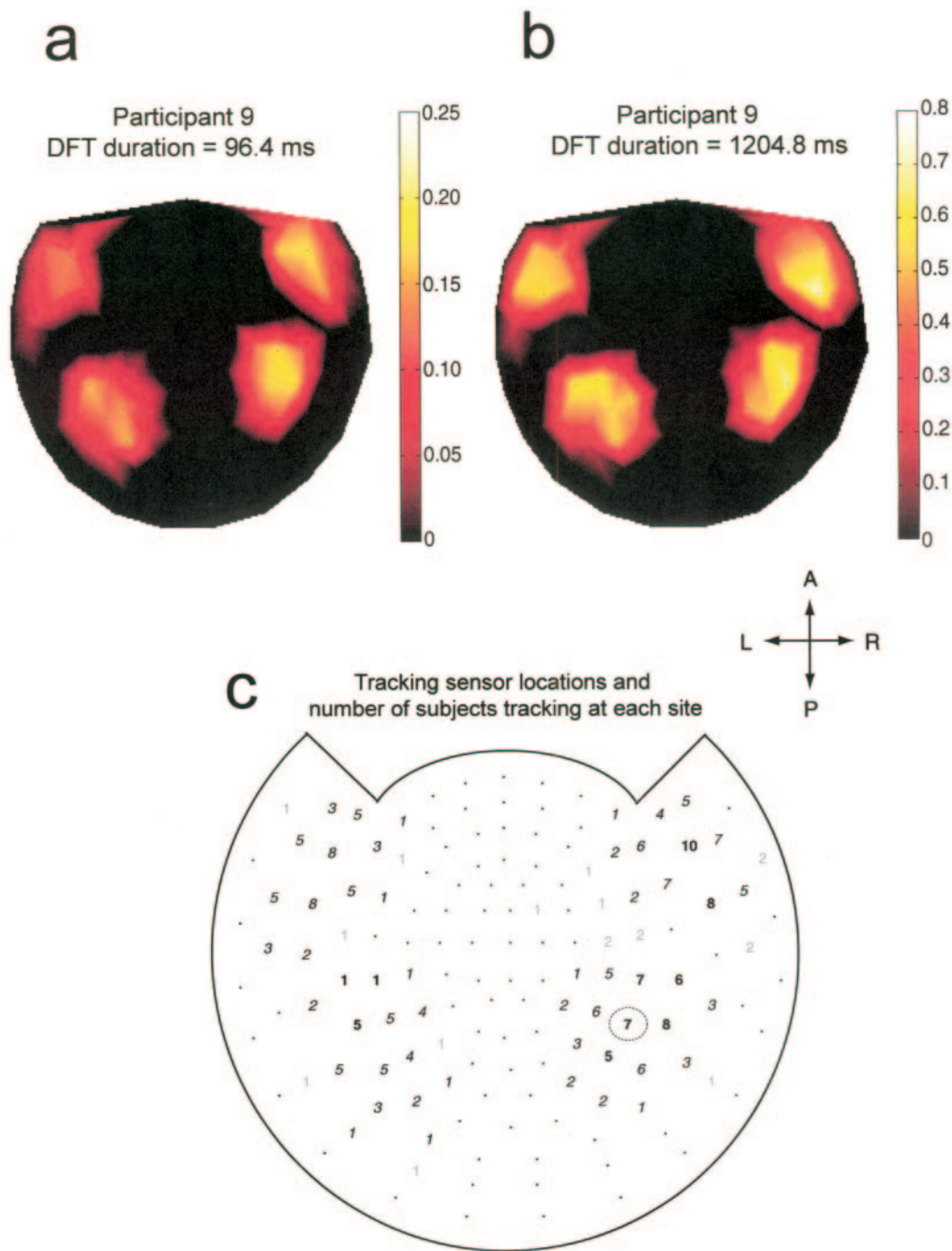


Figure 3. Individual and group data showing the spatial distribution and quality of phase tracking across the MEG sensor array. Panels (a,b,c) show a polar projection of the head surface as seen from above (A = anterior, P = posterior, L = left, R = right). (a,b) A smoothed spatial map of the aSSR phase contour correlations with stimulus carrier frequency at two different analysis durations for one experimental run of the same subject. The ranges of the two maps (given to the right of each diagram) are different because of the increase in the mean correlation value with increasing analysis duration (see Fig. 2). (c) Summary of tracking sensor locations and overall tracking quality across participants. Integers at each sensor location represent the number of participants who had that sensor in their tracking bank. Fonts represent overall tracking performance at that sensor averaged over all participants (normal gray = poor tracking sensors, italic black = intermediate tracking sensors, normal bold = good tracking sensors, see Materials and Methods). Small black dots indicate sensors that failed to meet the criteria for inclusion in the tracking bank. The dotted black oval in the right hemisphere shows the location of the sensor that generated the responses illustrated in Figure 2.

low SNRs; (ii) that phase tracking on the right side is significantly better at short DFT durations (~240 ms and less).

Shortest DFT Analysis Lengths with Significant aSSR Phase Correlations

A robust criterion was developed for deciding when aSSR phase-frequency correlations at a single sensor location repre-

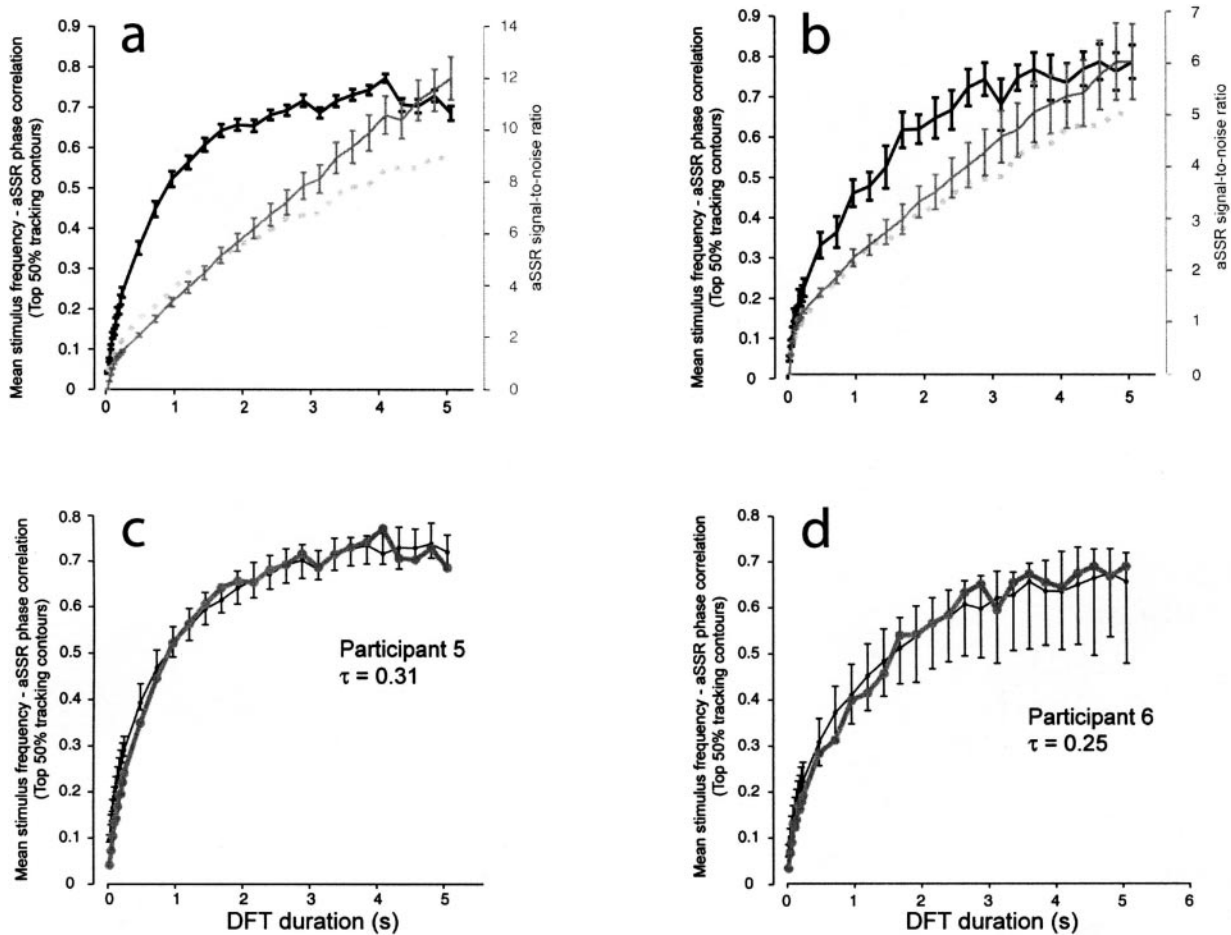


Figure 4. Changes in aSSR phase tracking with analysis duration. (*a,b*) Average correlation contours for two participants are shown in black, revealing a curvilinear increase in aSSR phase tracking with increasing DFT analysis length (data averaged over the top 50% tracking contours of each participant, see Materials and Methods). Left y-axis shows correlation values, x-axis shows DFT length. Error bars show 95% confidence intervals. Also shown are the average signal-to-noise ratio (SNR) of the aSSR (gray lines and right axis of each graph), computed from the same sensors and runs that yielded the average correlation contours. Error bars represent 95% confidence intervals for the means. Panels (*a,b*) show data from participants 5 and 6, respectively. (*c,d*) Results of a simulation model of phase tracking applied to data from the same two participants. The grand average correlation contour (thick gray line), and the correlation contour produced by the simulation of phase tracking (see Materials and Methods, thin black line with error bars) are plotted. The inset to each graph provides the value of τ yielding the best fit between the model and the data. τ represents the proportion of cells contributing to the aSSR directly involved in phase tracking of stimulus carrier frequency. Error bars indicate the extreme values observed in 1000 iterations of the simulation using the selected value of τ .

sent significant events in the face of multiple comparisons (a significant correlation at the $P < 0.05$ level at a given DFT analysis length on five or more experimental runs, see Materials and Methods). This criterion was uniformly applied to the tracking banks of all participants. The average shortest analysis length meeting the criterion for each participant ranged between 64 and 214 ms, and the number of tracking bank sensors where the criterion was met ranged between 10 and 39. Across participants, ~40% of tracking bank sensors met criterion at DFT lengths below 50 ms (68% met criterion below 100 ms). Finally, seven out of 10 participants had at least one sensor in the tracking bank that met the criterion at the shortest DFT length used in this study (24.1 ms), while the remaining three participants had at least one tracking bank sensor that met the criterion at the next DFT length (48.2 ms). Thus, all subjects had at least one sensor that showed consistent and significant phase tracking at analysis lengths of 50 ms or less.

Relation between Overall Tracking Performance and Tracking at Short DFT Analysis Lengths

The question of whether overall tracking performance is related to performance at short DFT lengths was addressed by examining correlation contours of each participant classified into three categories (top 50% tracking contours, bottom 50% tracking contours, and nontracking contours, see Materials and Methods). Figure 5 shows histograms of the minimum DFT length at which significant tracking was first observed in all correlation contours observed in the tracking banks of all participants, divided into tracking contours (top 50% contours, bottom 50% contours: Fig. 5*a*), and nontracking contours (Fig. 5*b*). The mean significant minimum DFT analysis lengths (± 1 SE) for the three categories of tracking performance (excluding data where no analysis lengths were significant) were 38.3 ± 0.9 ms for top 50% tracking contours, 58.3 ± 1.8 ms for bottom 50% tracking contours and 276.9 ± 32.8 ms for nontracking contours. These differences between tracking performance

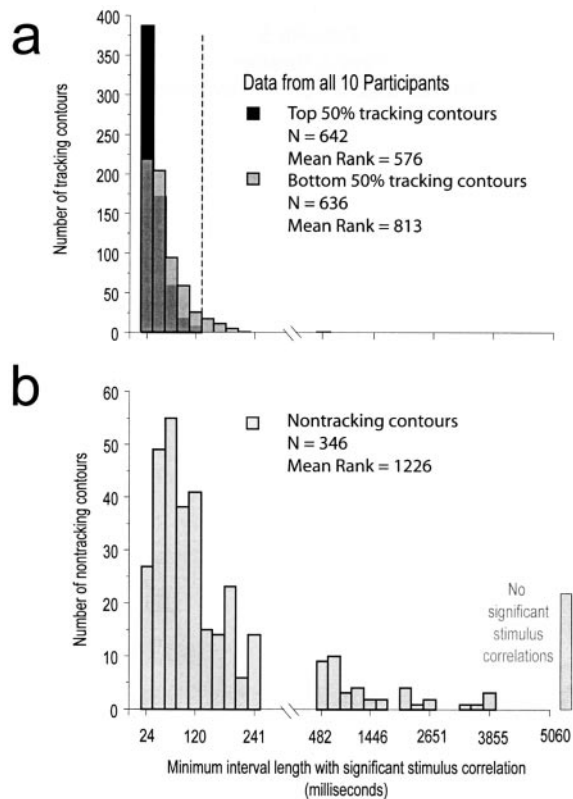


Figure 5. Histograms of the shortest DFT duration at which significant phase tracking is first observed for each correlation contour. Data from all correlation contours across all participants are included in this figure. Histograms are plotted separately for tracking and non-tracking experimental runs (*a,b*). Data from tracking runs (*a*) are further subdivided by tracking performance (top 50% tracking contours, bottom 50% tracking contours). The y-axis represents the number of correlation contours of each type (see Materials and Methods). The data from the top 50% tracking contours (black bars) are plotted behind those from the bottom 50% tracking contours (opaque gray bars). The vertical dotted line in (*a*) shows the upper limit of the top 50% tracking contour distribution.

categories were significant (Kruskal-Wallis ANOVA, $H = 426.5$, $P < 0.0001$, $n = 642, 636, 324$, all groups significantly different from each other at the 0.05 level in posthoc tests corrected for multiple comparisons). The same results were obtained for comparisons within each participant (all participants $H \geq 13.7$, $P \leq 0.001$; in 9/10 participants all groups were significantly different from each other in *post hoc* tests corrected for multiple comparisons; in the remaining participant the top 50% tracking contours were significantly different from the other two groups, which were not different from each other). Within each individual, the mean significant minimum DFT analysis lengths (± 1 SE) ranged from 31.2 ± 1.2 to 50.5 ± 6.8 ms.

These data suggest a functional relationship between how well the phase of neural activity at a particular sensor follows stimulus carrier frequency on any given recording, and the minimum analysis length at which it first shows a significant correlation with the stimulus. The best tracking contours (top 50%) also tend to show the lowest minimum DFT lengths for significant tracking. Figure 5*a* also demonstrates the heterogeneity in tracking responses in terms of their performance at short analysis lengths (cf. 'A model of phase tracking', below).

aSSR Phase Range

Figure 6*a* shows the mean phase range of the top 50% tracking contours for each individual, together with the 95% confidence

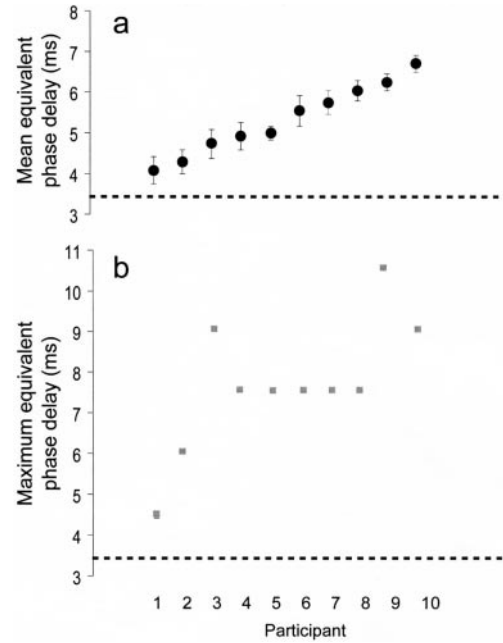


Figure 6. Mean phase range (*a*) and maximum phase range (*b*) of the observed aSSR for each subject, expressing the difference in aSSR phase between the lowest and highest carrier frequencies used in this study (220 and 880 Hz, respectively). Phase range is expressed as equivalent phase delay in ms; to convert to radians, multiply by $(2\pi \times 41.5)/1000$. The dotted line represents the expected delay value generated in the cochlea (3.41 ms; Greenberg *et al.*, 1998). Black data points in (*a*) represent the mean values for the top 50% tracking contours of each individual (individuals arrayed along the x-axis); error bars show the 95% confidence interval for the mean. Gray boxes in (*b*) represent the maximum delay value for each individual.

intervals for the mean (black circles and error bars). Phase range values have been converted to equivalent delay times at 41.5 Hz, in order to compare them with the expected cochlear delay between the high and low frequencies of 3.41 ms (Greenberg *et al.*, 1998). Figure 6*b* shows the maximum equivalent phase delay for each subject (gray boxes). The means and 95% confidence intervals of all 10 participants are above the expected cochlear delay (dashed line): these means ranged from 4.1 to 6.7 ms, with maxima ranging from 4.5 to 10.5 ms. Across subjects, mean equivalent phase delays were significantly associated with maximum equivalent phase delays (correlation = 0.72, $n = 10$, $P = 0.017$). These data demonstrate that aSSR phase tracking responses cannot be explained solely in terms of frequency-dependent delays arising in the cochlea; delays are expanded during neural processing between cochlea and cortex. The correlation between the mean and maximum delay values among subjects suggests that individual nervous systems produce individually distinctive ranges of delay values.

Relationship of aSSR Phase Tracking, aSSR Energy and aSSR Phase Range

Relationships among the quality of phase tracking, phase range, and aSSR signal energy at 41.5 Hz were examined in greater detail. Figure 7 shows a three-dimensional scatterplot of the relationships among these three variables, color-coded according to the three levels of tracking performance. These were quantitatively analyzed using partial Kendall's Tau correlation coefficients (Siegel and Castellan, 1988), which control

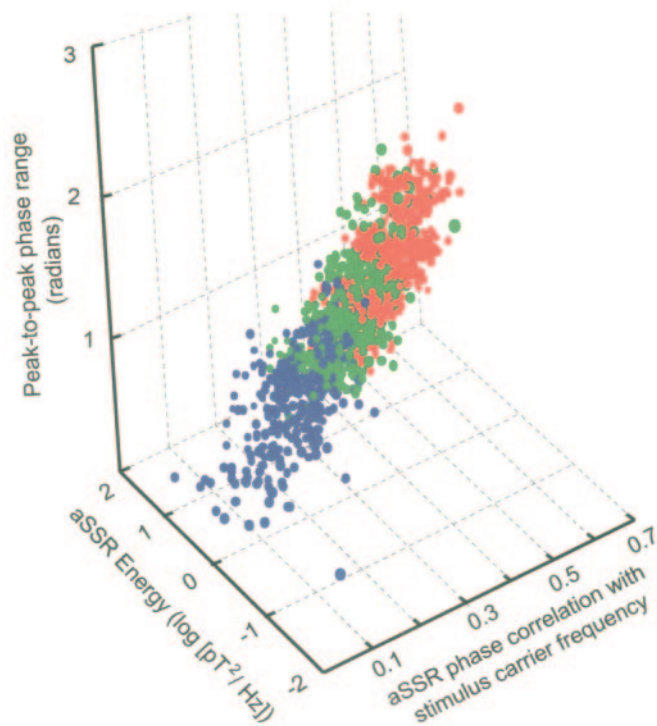


Figure 7. Relationship between aSSR energy, phase tracking, and phase range (data from the tracking banks of all participants combined; aSSR energy expressed in logarithmic units to more clearly represent the pattern of variation over the full dynamic range of signal intensities; pT = picotesla). Each point simultaneously represents (i) the average correlation value of a single correlation contour; (ii) the aSSR energy from that same sensor and run computed via a single minute-long DFT of the MEG signal; and (iii) the phase range computed for that same sensor and run. Data points are color coded as follows: red = top 50% tracking contours, green = bottom 50% tracking contours, blue = nontracking contours (see Materials and Methods).

for the interrelationships among the variables, and employing corrections for multiple statistical comparisons.

On average across participants, energy and tracking had a significant positive relationship (Kendall's Tau = 0.39, $n = 1624$, $P < 0.0001$ after Bonferroni correction); this same positive relationship was also found in 9/10 individuals (Kendall's Tau = 0.31–0.57, $n = 70$ –280, P all < 0.0012 after Bonferroni correction). Phase range and tracking also had a significant positive relationship in all individuals combined (Kendall's Tau = 0.54, $n = 1624$, $P < 0.0001$ after Bonferroni correction) and within all 10 individual individuals (Kendall's Tau = 0.50–0.66, $n = 70$ –280, P all < 0.0012 after Bonferroni correction). This indicates that sensors that tend to track better had relatively more signal strength than sensors that did not track well; and that sensors exhibiting larger phase ranges tended to track stimuli better than sensors having small response phase ranges. A positive relationship between sensor phase range and sensor phase correlation with stimulus carrier frequency is expected, because tracking sensors that have larger phase ranges can better match the details of stimulus carrier frequency variation, leading to higher correlations with the stimulus.

The positive relationships of tracking with phase range and energy also hold true when the analyses are limited either to tracking contours only, or to top 50% tracking contours only. The relationship between energy and tracking is manifest in 8/

10 individuals in the former case, and 6/10 individuals in the latter case; corresponding numbers for the relationship between phase range and tracking are 10/10 and 8/10 individuals, respectively. The relationships therefore cannot be an artifact of using data with a wider range of stimulus correlation values.

In contrast to these positive relationships of tracking with phase range and energy, there was a significant negative partial correlation between sensor energy and phase range, in all participants combined ($\tau = -0.42$, $n = 1624$, $P < 0.0001$ after Bonferroni correction) and within 9/10 individual participants ($\tau = -0.29$ to -0.56 , $n = 70$ –280, P all < 0.0012 after Bonferroni correction). Again, as with the positive relationships described above, this result does not change when analyses are limited either to tracking contours only, or to top 50% tracking contours only; in both cases, 9/10 participants still have significant partial negative correlations.

Given the positive relationship between tracking and sensor energy, it might have been expected that sensors with more aSSR energy would tend to have larger phase ranges. The fact that a significant pattern was obtained in the opposite direction suggests the separability of the phase and energy components of the aSSR response. That is, neural populations responsible for generating aSSR responses with large phase ranges are a subcomponent of all aSSR-responding cells.

A Model of Phase Tracking

The observation of significant phase tracking even at low SNRs, and the separability of the phase and energy components of the brain signals, prompted us to ask if the form of the correlation contours of phase tracking (Figs 2c, 4a,b) might tell us anything about the neural mechanisms of phase tracking.

To explore this issue we propose a model of aSSR phase tracking, embodied by a simulation of observed data (see explanation of the model in Materials and Methods). The model has one free parameter, τ , that represents the proportion of the neural response with firing perfectly phase-locked to the AM frequency, and with a firing delay that varies in a carrier-frequency dependent manner (the remainder of the response, $1 - \tau$, consisting of random 'phase noise').

Figure 4c,d show the best fitting curves (black) overlaid on grand average correlation contours (gray) for the two individuals whose data are shown in Figure 4a,b, together with an indication of the range of the best-fitting curve from 1000 iterations of the simulation using the indicated value of τ (black error bars). The curves from the remaining 8 subjects are very similar; τ values from all 10 individuals are concentrated over a small range, from 0.25 to 0.32, with a mean + SD of $0.278 + 0.026$. The observed range would likely be even smaller if the τ values were corrected for individual differences in phase range (the extent to which each individual's nervous system differentially 'magnifies' the frequency-dependent phase delay between cochlea and cortex). As discussed below, this estimate of τ is of interest because it favorably compares with cell population parameters observed in a recent neurophysiological study in the auditory cortex of unanesthetized primates (Liang *et al.*, 2002; see discussion below).

Discussion

There is a growing need in human auditory neuroscience for techniques that follow cortical activity as it varies over time in long sequences, with time resolution at the scale of tens to

hundreds of milliseconds. We believe that a dynamic approach to the auditory steady-state response (aSSR) provides one such technique, and that this approach serves as a valuable complement to the event-related approach to the aSSR. While the dynamic approach does not directly provide information relevant to the accurate localization of signal sources, it can reveal brain responses to time-varying acoustic stimuli not easily revealed in a static paradigm. Such responses provide information about how the auditory cortex follows patterns that vary over time, an issue of theoretical and practical significance in cognitive neuroscience (Poldrack *et al.*, 2001; Zatorre *et al.*, 2002). This is true both over periods of time that are sufficiently long for fMRI analysis (Janata *et al.*, 2002) and over shorter time periods, as explored here. Another advantage of the dynamic approach is that meaningful data can be gathered from single trials and individuals, facilitating research and clinical applications examining the variability of individual responses.

Consistent with previous research (Patel and Balaban, 2000), we found that the phase of the aSSR reliably tracked the carrier frequency contour of the tone sequences, with phase advancing for increasing carrier frequencies and delaying for decreasing carrier frequencies. This dependency of aSSR phase on carrier frequency is also consistent with research by other groups (Galambos *et al.*, 1981; John and Picton, 2000; Roß *et al.*, 2000), but had not previously been studied in a dynamic fashion. It is important to note that this does not imply that phase tracking is part of how the brain naturally follows tone sequences; after all, the presence of the aSSR is due to a stimulus manipulation – amplitude modulation – on the part of the experimenter. Rather, phase tracking is of neurobiological interest because it provides a method for probing aspects of auditory cortical responses to changing acoustic sequences.

A principal finding of the current study is that phase tracking can be observed even when analysis time windows are very short. This suggests that despite the unfavorable SNRs present at these short lengths, there is enough phase-locked information preserved in successive analysis windows to provide significant information about stimulus carrier frequency. While this method does not allow a precise estimate of how short the ‘unit of integration’ for phase-tracking cell populations is, it demonstrates that a significant amount of ‘phase-related’ carrier frequency information is present in signal intervals as short as 25–50 ms. An ancillary finding is that phase tracking is better in the right vs. the left hemisphere when data are analyzed at short (24.1–241 ms) but not at longer (482 ms–5 s) analysis durations, perhaps reflecting differences in the structural and/or functional properties of the right vs. left auditory cortex (cf. Zatorre *et al.*, 2002). This finding has significant implications for auditory hemispheric asymmetry research using methodologies such as PET and fMRI that may have integration times of ~0.5 s or longer.

A second main finding of interest concerns the physical range over which aSSR phase varies in response to changing carrier frequency. While the basic pattern of phase advance/delay with increasing/decreasing carrier frequency is consistent with frequency-dependent neural firing delays in the cochlea, the extent of these delays is too large to be explained solely by passive propagation of these relative delay values through the intervening nervous system. Rather, some expansion of these delays takes place between cochlea and cortex (cf. Greenberg *et al.*, 1998; Rupp *et al.*, 2002). Calcula-

tion of equivalent phase delays for each of our participants revealed that the degree of expansion was individually variable, with averages ranging from slightly but significantly greater than the expected cochlear delay to about twice the expected cochlear delay (Fig. 6a). There is nothing in the work presented here that suggests where in the pathway between cochlea and cortex the phase expansion may take place.

One possible common mechanism involved in this expansion might be the difference in how long it takes to integrate frequency information from a single cycle of a stimulus. If the cells producing cortical responses to different stimulus frequencies depend on input that at some point requires integration over some small number of stimulus cycles, there would be a frequency dependence in the ‘extra’ processing time that gets added to the cochlear delay. Relatively lower frequencies will have longer delays in their neuronal responses than relatively higher frequencies. It is unclear if individual differences in this integration time would be sufficient to account for the range of expansion values observed here, or if it will be necessary to invoke other anatomical or physiological mechanism(s) to fully explain the individual variation. Future experiments examining the phase ranges produced by stimulus sequences similar to those used here, but transposed to cover different absolute frequency ranges, could be used to both test for this common mechanism, and to see how well variation in this mechanisms might explain individual variation in the phase range of tracking responses. Research examining the connection between individual variation in carrier frequency/pitch perception and individual’s phase ranges could also prove to be instructive.

A third salient finding is that phase tracking is potentially explicable by a spatial sum of the activity of two cell populations, ‘tracking’ (proportion τ) and ‘non-tracking’ (proportion $1 - \tau$), involved in aSSR generation. Our estimate of τ (~25%) shows a marked resemblance to recent data from animal neurophysiology published just after the completion of the modeling work. Studying single unit responses to AM stimuli recorded in the primary auditory cortex of unanesthetized marmosets *Callithrix jacchus*, Liang *et al.* (2002 found a class of single units (BP for ‘bandpass’) that appear to have similar characteristics to the hypothesized ‘Component 1’ (tracking) cells in the simulation of phase tracking described above. Figure 12C of their paper (p. 2250) plots the percentage of recorded units that exhibit strong phase-locking at different AM frequencies. According to these curves, units responsive to AM rates of 41.5 Hz make up ~28 % of BP units and ~22% of all units recorded in their study. The criterion used for phase locking in the study (Rayleigh statistic >13.8) probably excludes cells that would contribute to MEG measurements; some of the BP cells may not have consistent relative phase response differences that vary with carrier frequency [as reflected in the ‘characteristic frequency’ (CF) response of the BP cells]; and humans may have different proportions of these cells in their auditory cortices in comparison to marmosets. Nevertheless, the close resemblance between the τ values calculated above and these data suggest that the phase tracking responses observed to AM pure tones could be predominantly driven by the behavior of one subpopulation of cortical units exhibiting phase-locked firing to the AM envelope (with a relative phase delay that varies with the carrier frequency tuning of the cell), together with non phase-locked activity at the AM rate produced by a variety of cortical units responsive to these sounds.

In conclusion, we suggest that phase tracking is due to a subpopulation (approximately 25%) of cells that generate the aSSR, and that these cells respond with very short temporal integration windows to changes in stimulus carrier frequency. Future work combining static and dynamic approaches to the aSSR could help elucidate the temporal response characteristics of the tracking cell populations to determine the shortest time intervals over which they can reflect changing carrier frequencies, whether cells contributing to phase tracking responses are more enriched in particular portions of the human auditory cortex (e.g. in core vs. belt areas of primary auditory cortex: Kaas *et al.*, 1999; Kaas and Hackett, 2000; Hackett *et al.*, 2001; Tian *et al.*, 2001; Wessinger *et al.*, 2001) and whether they are spatially segregated from non-tracking cells within these larger regions. Future work using the aSSR could also address the mechanisms underlying individual variation in phase tracking, whether the degree to which the auditory system magnifies cochlear frequency-specific phase delays during the perception of acoustic sequences depends on the nature of those sequences (Patel and Balaban, 2000) or the context in which they occur, and whether falling outside the 'typical' range is associated with perceptual and/ or cognitive disorders. More generally, we believe that aSSR dynamics can provide a window on activity changes in auditory cortical cell populations over relatively short time intervals, and thus can make a useful contribution to understanding how the brain follows natural stimulus sequences as they unfold over time.

Notes

We thank Lacey Kurelowech for technical assistance with MEG recordings, John Iversen for suggestions regarding the analysis, and two anonymous reviewers for comments that greatly improved the manuscript. A.P. was supported by an Esther J. Burnham fellowship and by Neurosciences Research Foundation as part of its program on music and the brain at The Neurosciences Institute.

Address correspondence to either A.D. Patel, The Neurosciences Institute, 10640 John Jay Hopkins Drive, San Diego, CA 92121, USA. Email: apatel@nsi.edu or to E. Balaban, Department of Psychology, McGill University, 1205 Dr Penfield Avenue, Montreal, Quebec, Canada H3A 1B1. Email: evan@psych.mcgill.ca.

References

Cabeza R, Kingstone A (2001) Handbook of functional neuroimaging of cognition. Cambridge, MA: MIT Press.

Fujiki N, Jousmaki V, Hari R (2002) Neuromagnetic responses to frequency-tagged sounds: A new method to follow inputs from each ear to the human auditory cortex during binaural hearing. *J Neurosci* 22(RC205):1-4.

Galambos R, Makeig S, Talmachoff PJ (1981) A 40-Hz auditory potential recorded from the human scalp. *Proc Natl Acad Sci USA* 78:2463-2647.

Galambos R, Makeig S (1988) Dynamic changes in steady-state responses. In: Dynamics of sensory and cognitive processing by the brain (Basar E, ed.), pp. 103-122. Berlin: Springer Verlag.

Greenberg S, Poeppel D, Roberts T (1998) A space-time theory of pitch and timbre based on cortical expansion of the cochlear traveling wave delay. In: Psychophysical and physiological advances in hearing (Palmer A, Summerfield Q, Rees A, Meddis R, eds), pp. 293-300. London: Whurr.

Gutschalk A, Mase R, Roth R, Ille N, Rupp A, Hahnel S, Picton TW, Scherg M (1999) Deconvolution of 40 Hz steady-state fields reveals

two overlapping source activities of the human auditory cortex. *Clin Neurophysiol* 110:856-868.

Hackett TA, Preuss TM, Kaas JH (2001) Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J Comp Neurol* 441:197-222.

Hari R, Hämäläinen M, Joutsiniemi S-L (1989) Neuromagnetic steady-state responses to auditory stimuli. *J Acoust Soc Am* 86:1033-1039.

Janata P, Burk JL, Van Horn JD, Leman M, Tillmann B, Bharucha JJ (2002) The cortical topography of tonal structures underlying Western music. *Science* 298:2167-2170.

John MS, Picton TW (2000) Human auditory steady-state responses to amplitude modulated tones: phase and latency measurements. *Hear Res* 14: 57-79.

Kaas JH, Hackett TA, Tramo MJ (1999) Auditory processing in primate-cerebral cortex. *Curr Opin Neurobiol* 9:164-170.

Kaas JH, Hackett TA (2000) Subdivisions of auditory cortex and processing streams in primates. *Proc Natl Acad Sci USA* 97:11793-11799.

Lewine JD, Orrison WW (1995) Magnetoencephalography and magnetic source imaging. In: Functional brain imaging (Orrison WW et al., eds), pp. 369-417. St Louis: Mosby.

Liang L, Liu T, Wang X (2002) Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. *J Neurophysiol* 87:2237 - 2261.

Nunez P (1981) Electric fields of the brain. New York: Oxford University Press.

Pastor MA, Artieda J, Arbizu J, Marti-Climent JM, Peñuelas I, Masdeu JC (2002) Activation of human cerebral and cerebellar cortex by auditory stimulation at 40 Hz. *J Neurosci* 22:10501-10506.

Patel AD (2003) A new approach to the cognitive neuroscience of melody. In: The cognitive neuroscience of music (Peretz I, Zatorre R, eds), pp. 325-345. Oxford: Oxford University Press.

Patel AD, Balaban E (2000) Temporal patterns of human cortical activity reflect tone sequence structure. *Nature* 404:80-84.

Patel AD, Balaban E (2001) Human pitch perception is reflected in the timing of stimulus related cortical activity. *Nat Neurosci* 4:839-844.

Plourde G, Stapells DR, Picton TW (1991) The human auditory steady-state potential. *Acta Otolaryngol* 491:153-160.

Poldrack RA, Temple E, Protopapas A, Nagarajan S, Tallal P, Merzenich M, Gabrieli, JD (2001) Relations between the neural bases of dynamic auditory processing and phonological processing: evidence from fMRI. *J Cog Neurosci* 13:687-97.

Roß B, Borgmann C, Draganova R, Roberts L, Pantev, C (2000) A high precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. *J Acoust Soc Am* 108:679-691.

Roß B, Picton TW, Pantev C (2002) Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field. *Hear Res* 165:68-84.

Rupp A, Uppenkamp S, Gutschalk A, Beucker R, Patterson RD, Dau T, Scherg, M (2002) The representation of peripheral neural activity in the middle-latency evoked field of primary auditory cortex in humans. *Hear Res* 174:19-31.

Siegel S, Castellán NJ Jr (1988) Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill.

Stapells DR, Linden D, Suffield JB, Hamel G, Picton TW (1984) Human auditory steady-state potentials. *Ear Hear* 5:105-113.

Tian B, Reser D, Durham A, Kustov A, Rauschecker, JP (2001) Functional specialization in rhesus monkey auditory cortex. *Science* 292:290-293.

Wessinger CM, VanMeter J, Tian B, Van Lare J, Pekar J., Rauschecker JP (2001) Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. *J Cogn Neurosci* 13:1-7.

Zatorre RJ, Belin P, Penhune V (2002) Structure and function of auditory cortex: music and speech. *Trends Cogn Sci* 6:37-46.