J Neurophysiol 107: 2033–2041, 2012.
First published October 5, 2011; doi:10.1152/jn.00310.2011.

Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: MEG evidence

Yadong Wang,[1,*] Nai Ding,[2,*] Nayef Ahmar,[2] Juanjuan Xiang,[2] David Poeppel,[3,4] and Jonathan Z. Simon[2,5]

[1]*Program in Neuroscience and Cognitive Science and* [2]*Department of Electrical and Computer Engineering, University of Maryland, College Park, Maryland;* [3]*Department of Psychology and* [4]*Center for Neural Science, New York University, New York, New York; and* [5]*Department of Biology, University of Maryland, College Park, Maryland*

Submitted 6 April 2011; accepted in final form 28 September 2011

Wang Y, Ding N, Ahmar N, Xiang J, Poeppel D, Simon JZ. Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: MEG evidence. *J Neurophysiol* 107: 2033–2041, 2012. First published October 5, 2011; doi:10.1152/jn.00310.2011.—Slow acoustic modulations below 20 Hz, of varying bandwidths, are dominant components of speech and many other natural sounds. The dynamic neural representations of these modulations are difficult to study through noninvasive neural-recording methods, however, because of the omnipresent background of slow neural oscillations throughout the brain. We recorded the auditory steady-state responses (aSSR) to slow amplitude modulations (AM) from 14 human subjects using magnetoencephalography. The responses to five AM rates (1.5, 3.5, 7.5, 15.5, and 31.5 Hz) and four types of carrier (pure tone and 1/3-, 2-, and 5-octave pink noise) were investigated. The phase-locked aSSR was detected reliably in all conditions. The response power generally decreases with increasing modulation rate, and the response latency is between 100 and 150 ms for all but the highest rates. Response properties depend only weakly on the bandwidth. Analysis of the complex-valued aSSR magnetic fields in the Fourier domain reveals several neural sources with different response phases. These neural sources of the aSSR, when approximated by a single equivalent current dipole (ECD), are distinct from and medial to the ECD location of the N1m response. These results demonstrate that the globally synchronized activity in the human auditory cortex is phase locked to slow temporal modulations below 30 Hz, and the neural sensitivity decreases with an increasing AM rate, with relative insensitivity to bandwidth.

magnetoencephalography; cortex

SLOW TEMPORAL MODULATIONS are prevalent in natural sounds (Singh and Theunissen 2003) and carry information critical to perceptual parsing. For example, in speech signals, temporal modulations of signal power below 10–15 Hz reflect the rhythm of syllabic scale articulatory gestures (Greenberg et al. 2003; Rosen 1992) and are crucial for speech intelligibility (Drullman et al. 1994; Shannon et al. 1995). Given the relevance for perception, the neural coding of slow temporal modulations has been studied extensively using various neurophysiological tools. In the mammalian primary auditory cortex, the firing of individual neurons is precisely phase locked to the stimulus modulation for modulation rates below 20 Hz but becomes nonstimulus synchronized at higher modulation rates (Eggermont 2002; Liang et al. 2002). The local field potential, which reflects dendritic activity synchronized over local neural networks, is also precisely synchronized to slow temporal modulations and can phase lock beyond 40 Hz (Eggermont 2002). Similarly, in the human auditory cortex, the local field potential is stimulus synchronized most strongly below 16 Hz for amplitude modulation (AM) noise (Liegeois-Chauvel et al. 2004) but can phase lock up to 100 Hz for click-train stimuli (Brugge et al. 2009). Furthermore, the hemodynamic response measured by functional MRI (fMRI), which reflects both phase-locked and nonphase-locked response power, is also most sensitive to AM below 16 Hz (Giraud et al. 2000). Therefore, at both the single neuron level and the local neural network level, it has been shown consistently that the mammalian auditory cortex encodes slow AM as a temporal code and that neural sensitivity decreases above 16 Hz, although the upper bound of the temporal code varies with the recording technique and the stimulus.

The manner in which slow modulations are encoded by large-scale, phase-locked neural activity, as measurable by magnetoencephalography (MEG) or EEG, is not agreed on universally. On the one hand, most studies use AM and frequency modulation (FM) at rates well above 10 Hz, and it is well established that the MEG/EEG response is most reliable when the modulation rate is near the relatively high rate of 40 Hz (Picton et al. 1987; Ross et al. 2000). On the other hand, there is a growing interest in the neural coding of temporal modulations of speech and other complex sounds, and these studies consistently show that MEG/EEG responses most strongly track only slow modulations, e.g., below 10 Hz (Ahissar et al. 2001; Aiken and Picton 2008; Howard and Poeppel 2010; Lalor et al. 2009; Luo and Poeppel 2007). It is uncertain whether this discrepancy results from a dichotomy in how complex natural sounds and simple AM sounds are encoded in large-scale synchronized activity or the lack of detailed characterizations of MEG/EEG responses to slow-rate AM.

In this study, we characterize the large-scale neural response to slow AM using MEG. The sustained MEG/EEG response, tracking a constant rate AM/FM, is typically referred to as the auditory steady-state responses (aSSR). Besides the 40-Hz range aSSR, the aSSR to temporal modulations below 20 Hz is also reported to be measurable (Alaerts et al. 2009; Ding and Simon 2009; Luo et al. 2006; Millman et al. 2009; Picton et al. 1987; Rees et al. 1986). In this low-frequency range, however, there is little consistency in descriptions of the shape of the modulation transfer function (MTF), i.e., how the aSSR power

* Y. Wang and N. Ding contributed equally to this work.

Address for reprint requests and other correspondence: J. Z. Simon, Dept. of Electrical and Computer Engineering, Univ. of Maryland, College Park, MD 20742 (e-mail: jzsimon@umd.edu).

changes as a function of stimulus modulation rate. Alaerts et al. (2009) reported that below 30 Hz, the EEG MTF for AM broadband noise shows two peaks, one near 10 Hz and the other at the lowest tested frequency, 4 Hz. In contrast, Ross et al. (2000) showed that the power of the MEG aSSR to an AM pure tone increases with increasing AM rate from 10 Hz to 40 Hz. Several other studies, using various types of temporal modulations, however, have suggested that the MEG/EEG aSSR is strongest below 5 Hz (Alaerts et al. 2009; Ding and Simon 2009; Millman et al. 2009; Picton et al. 1987; Rees et al. 1986).

The inconsistency of the shape of the MTF in the low-frequency region may be caused by several factors. One is that the roughly 1/f-shaped background activity in MEG/EEG recordings poses challenges for the detection and strength estimation of the aSSR. Consequently, various MEG signal analysis methods, differing in their noise susceptibility, may lead to different conclusions. In this study, we used a statistically optimal spatial filter to extract the aSSR from intrinsically noisy measurements and then estimated the spectral shape of the MTF using a method that removes the statistical bias that can plague power estimates under noisy conditions.

A second reason why MTF shapes have been reported so inconsistently is that the aSSR may be sensitive to the carrier bandwidth. Sounds with different bandwidths have been argued to be processed by different but overlapping areas in the human auditory cortex [Chevillet et al. 2011; Wessinger et al. 2001; also see Overath et al. (2012) for an alternative view in the companion paper]. Although the potentially distinct spatial positions of neural populations tuned to different bandwidths may not be resolvable by MEG/EEG, their response characteristics, e.g., the MTFs, are robustly reflected in the MEG/EEG recordings. To address the effect of carrier bandwidth on the aSSR, we used carriers with four different bandwidths, ranging from pure tones to 5-octave pink noise.

In summary, evidence from single-unit recordings in animals and invasive recordings in humans leads us to expect the aSSR measured by MEG to show the shape of a low-pass filter for frequencies below 10 Hz: auditory cortex should respond most strongly at the lowest frequencies, decreasing in response power as the frequency increases. In addition, because it has been suggested that the responses to stimuli of differing bandwidths are processed preferentially in different areas (or by different auditory circuits within the same area), we expect systematic dependence of the response location on the stimulus bandwidth. As seen below, there is strong evidence for the first expectation (low-pass character of response) and indirect evidence for the second (stimulus bandwidth dependence on location). Importantly, in a companion paper, Overath et al. (2012) report similar results for modulation rate and bandwidth throughout the brain, using a complementary technique (fMRI).

METHODS

Subjects. Fifteen subjects (nine female), all right handed (Oldfield 1971) and reporting normal hearing with no history of neurological disorder, listened passively to the acoustic stimuli while the MEG recordings were taken. The experimental procedures were approved by the University of Maryland Institutional Review Board, and written, informed consent was obtained from each participant. Subjects were paid for their participation. One female subject was removed from analysis due to interruption of her experiment by a power failure.

Stimuli. Sinusoidal AM sounds of 2,000 ms duration were presented to each subject. Twenty separate stimuli were created in MATLAB (MathWorks, Natick, MA) for five different modulation frequencies (1.5, 3.5, 7.5, 15.5, and 31.5 Hz) and four different carriers (pure tone, 1/3-octave pink noise, 2-octave pink noise, and 5-octave pink noise, all centered at 707 Hz) with 100% modulation depth. The stimuli were created offline, gated on and off using a 15-ms duration cosine-squared ramp, and saved in a 16-bit stereo wave format at a sampling rate of 44.1 kHz.

Experimental procedure. Stimulus delivery was carried out using Presentation (Neurobehavioral Systems, Albany, CA). All of the AM stimuli (five modulation frequencies $\times$ four carrier bandwidths) were presented 50 times/condition in pseudorandom order. Interstimulus intervals were randomized between 700 and 900 ms. The signals were delivered binaurally to the subjects' ears with 50 Ω sound tubing (E-A-RTone 3A, Etymotic Research, Elk Grove Village, IL) at a comfortable loudness of $\sim$70 dB sound-pressure level. Before the main experiment, 170 repetitions of a 50-ms, 1-kHz tone were represented to the subjects with the interstimulus interval randomized between 750 and 1,550 ms. MEG responses to these short tones were used to localize the source of the N1m component.

MEG recording. MEG recordings were conducted using a 157-sensor axial gradiometer whole-head system (Kanazawa Institute of Technology, Kanazawa, Japan). Three magnetometers are also included in the MEG system to measure the environmental magnetic field. The magnetic signals were bandpassed between 1 Hz and 200 Hz, notch filtered at 60 Hz, and sampled at the rate of 500 Hz. The 1-Hz high-pass filter's influence on the amplitude and phase of MEG recordings was corrected.

MEG signal processing. Time-shifted principle component analysis (de Cheveigné and Simon 2007) was applied to remove environmental noise collected by the magnetometers from each gradiometer. To reduce the impact of transient MEG responses [e.g., the N1m component (Näätänen and Picton 1987)], only data from 500 ms to 2,000 ms poststimulus were used for aSSR analysis and were zero padded to 1,000 samples. Modeling the aSSR as a sinusoid at the stimulus rate makes analysis in the frequency domain straightforward; the neural response at each harmonic of the stimulus rate was analyzed separately. It is worth pointing out that although the aSSR is analyzed in the frequency domain, we make no claim that it is generated by frequency domain analysis in the brain. The MEG recordings were transformed into the frequency domain using a 1,000-point (frequency resolution 0.5 Hz) discrete Fourier transform (DFT) without any smoothing window. Since all of the AM rates in this study are integer multiples of the frequency resolution of DFT, the sinusoidal aSSR is fully captured by the DFT coefficient at the stimulus AM rate. Therefore, the aSSR magnetic field at each AM rate, ω, is conveniently represented as a vector $S(\omega)$, consisting the DFT coefficients from all of the sensors.

Complex magnetic field. The scalp magnetic field consisting of the Fourier coefficients is complex, since it contains both the magnitude and phase information of the aSSR. Since electromagnetic signals travel at light speed, there is no phase lag between MEG recordings and neural source activity. Hence, if the aSSR is from a hypothetical single source, the MEG signals measured from all sensors should be in phase (or 180° out of phase where the field reverses direction). Mathematically, if the aSSR source is explained by an equivalent current dipole (ECD) located at $\boldsymbol{r}$, and the scalp magnetic field topography generated by this dipole, i.e., the lead field, is $\boldsymbol{a}(\boldsymbol{r})$, then the aSSR scalp magnetic field is $\boldsymbol{S}(\omega) = \alpha\exp(j\psi)\boldsymbol{a}(\boldsymbol{r})$, where α and ψ are the magnitude and phase of the aSSR. In this condition, the complex magnetic field can be decomposed into a real magnetic field $\boldsymbol{a}(\boldsymbol{r})$, which contains all structural information about the neural source and a single aSSR phase, ψ, which reflects the timing of the neural response. The outer product of $\boldsymbol{S}(\omega)$ and its Hermitian transpose, i.e.,

$\boldsymbol{S}(\omega)\boldsymbol{S}^{\mathrm{H}}(\omega) = \alpha^2\boldsymbol{a}(\boldsymbol{r})\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{r})$, is real, and the neural source-related information $\boldsymbol{a}(\mathbf{r})$ can be extracted by applying the eigendecomposition to matrix $\boldsymbol{S}(\omega)\boldsymbol{S}^{\mathrm{H}}(\omega)$. $\boldsymbol{a}(\boldsymbol{r})$ is referred to as the real-valued aSSR magnetic field of $\boldsymbol{S}(\omega)$.

In general, MEG responses are generated by multiple neural sources. If these sources are so close together that they are not spatially distinguishable by MEG, then they can still be modeled as a single ECD, whose location is the center of current–gravity of all of these neural sources. For example, suppose that there are two aSSR sources located very close to $\boldsymbol{r}$, but they have different orientations. Then, the aSSR scalp magnetic field becomes $\boldsymbol{S}(\omega) = \alpha_1\exp(j\psi_1)\boldsymbol{a}_1(\boldsymbol{r}) + \alpha_2\exp(j\psi_2)\boldsymbol{a}_2(\boldsymbol{r})$, where α_1, α_2, ψ_1, and ψ_2 are the magnitudes and phases of the two aSSR sources. $\boldsymbol{a}_1(\boldsymbol{r})$ and $\boldsymbol{a}_2(\boldsymbol{r})$ are the lead fields of the two sources, respectively (different, since their sources had different dipole orientations). If the two sources have the same response phase, $\psi_1 = \psi_2 = \psi$, then the aSSR magnetic field reduces to $\boldsymbol{S}(\omega) = \exp(j\psi)[\alpha_1\boldsymbol{a}_1(\boldsymbol{r}) + \alpha_2\boldsymbol{a}_2(\boldsymbol{r})]$, which can be modeled by a single ECD, whose orientation is a weighted sum of individual dipole orientations. In contrast, if the two sources have different response phases, i.e., $\psi_1 \neq \psi_2$, then $\boldsymbol{S}(\omega) = [\alpha_1\cos(\psi_1)\boldsymbol{a}_1(\boldsymbol{r}) + \alpha_2\cos(\psi_2)\boldsymbol{a}_2(\boldsymbol{r})] + j[\alpha_1\sin(\psi_1)\boldsymbol{a}_1(\boldsymbol{r}) + \alpha_2\sin(\psi_2)\boldsymbol{a}_2(\boldsymbol{r})]$. $\boldsymbol{S}(\omega)$ can no longer be modeled as a real-valued dipole with a single aSSR phase, and the outer product of itself with its Hermitian transpose becomes complex valued. In this case, however, $\boldsymbol{S}(\omega)$ can be modeled as arising from a single ECD with a complex-valued dipole moment, i.e., a fully complex dipole (Simon and Wang 2005). The fully complex dipole includes the special case of a real-valued dipole. Whether a fully complex dipole can be reduced to a real-valued dipole is characterized by the sharpness of the dipole (Simon and Wang 2005), which is zero for real-valued dipoles. Even when the aSSR is generated by more than two sources with different response phases, it can still be modeled by a fully complex dipole. The number of dipoles (greater than or equal to two), however, cannot be revealed by this analysis alone.

ECD modeling. A single ECD model was used to model the auditory response in each hemisphere. The lead field of each ECD was calculated based on an isotropic sphere model built for each subject using the software MEG Laboratory 2.001M (Yokogawa Electric, Japan; Eagle Technologies, Japan; Kanazawa Institute of Technology). An ECD is determined by its location and moment. The moment was estimated using the least-squares method (Mosher et al. 2003), whereas the position was estimated using a modified simplex search with clustering (Uutela et al. 1998). The goodness of fit (GOF) of an ECD model is evaluated as $1 - \|\boldsymbol{S}(\omega) - \boldsymbol{S}_{\mathrm{FIT}}(\omega)\|^2/\|\boldsymbol{S}(\omega)\|^2$, where $\|\cdot\|$ is the l_2 norm, and $\boldsymbol{S}_{\mathrm{FIT}}(\omega)$ is the magnetic field generated by the dipole model.

The ECD locations in the left and right hemispheres were chosen by maximizing the GOF in two steps. In the first step, the approximate location was determined separately in each hemisphere, using only sensors over the corresponding hemisphere. In each hemisphere, 20 random positions were selected as seeds for a simplex search, and the resulting locations were grouped into, at most, five clusters using K-means clustering (Uutela et al. 1998). In the second step, all sensors from both hemispheres were used simultaneously to determine the locations of bilateral ECDs, using the cluster centers of the ECD locations in both hemispheres as seeds. If an ECD were localized to be <2.5 cm away from the center of the head, it was viewed as a pathological solution and was removed from further analysis. The left/right ECD pair with the highest GOF was chosen as the final ECD estimate.

The same source localization procedure was applied to the N1m data. The MEG responses to all 170 repetitions of a 1-kHz tone pip were averaged. The N1m component was determined as the peak of the root-mean-square value over sensors between 70 and 120 ms after stimulus onset and was measured reliably from all subjects. The N1m magnetic field and the real-valued aSSR magnetic fields were modeled by real-valued dipoles, whereas the complex-valued aSSR magnetic fields were modeled by fully complex dipoles. Fully complex dipoles and real-valued dipoles were all localized using the same procedure.

Spatial filtering. Spatial filtering is a common technique for aSSR processing. It linearly combines the measurements from multiple MEG sensors to compute the response properties of a single component. The output of a spatial filter can be viewed as a "virtual sensor" (Millman et al. 2009). Here, we used the denoising source separation (DSS) spatial filter (de Cheveigné and Simon 2008), which is optimal for aSSR detection in Gaussian noise (see APPENDIX). The design of a DSS filter relies on the covariance matrix of the aSSR magnetic field and the covariance matrix of the background activity (see APPENDIX). Since the aSSR magnetic field is complex, the resulting DSS filter is also complex and is called the complex-valued DSS filter. If there is only one neural source of the aSSR, we can replace a complex-valued aSSR magnetic field with the corresponding real-valued aSSR magnetic field without any loss. The DSS filter designed using the real-valued aSSR magnetic field is real valued and is called the DSS filter with a single source assumption.

Significance testing. A spatial filter can help to detect neural signals. However, if the spatial filter is designed and evaluated on the same dataset, it will generate frequent false positives. Therefore, we used fivefold cross-validation to evaluate the statistical significance level of the aSSR, as estimated by spatial filter outputs. Specifically, the 50 total trials were divided into five sets. Four sets (40 trials), called the training sets, were used for spatial filter design, and the other set (10 trials), called the testing set, was used to evaluate the significance level. This process was repeated five times, with a different selection of the testing set on each repetition. The median of the five estimates of significance level was used as the final significance level. In other words, only if three out of five sets (or more) showed significant results ($p_{\mathrm{cv}} < 0.05$) was the response viewed as significant. If the five sets are statistically independent of one another, the probability that a response would be tested to be significant (median of $p_{\mathrm{cv}} < 0.05$), by chance, is smaller than 0.001 (based on the binomial distribution). Therefore, although we had 14 subjects, the probability that one or more subjects were judged to have a significant aSSR by chance is quite small.

The significance level of aSSR was evaluated on the testing set using a phase-projected bootstrap test. Specifically, after spatial filtering, the phase of the aSSR was first estimated from the training set. The aSSR in the testing set, a complex scalar after spatial filtering, was averaged over trials and then projected along the direction of the estimated aSSR phase (Picton et al. 2001). A bootstrap test was then applied to this phase-projected response, in which the 95% confidence interval of the null distribution of the phase-projected response was estimated based on the responses at this frequency to stimuli modulated at a frequency other than ω.

RESULTS

The spatial distributions of the aSSR magnetic fields from a representative subject are plotted in Fig. 1. A clear dipolar pattern is seen in each hemisphere, indicating that low-rate AM stimuli evoke measurable MEG responses. The power, phase, reliability, and source location of the aSSR are discussed in the following sections.

aSSR power and phase. Fig. 2 shows the power of the MEG response as a function of the stimulus AM rate and carrier bandwidth for each hemisphere. The response power is measured by the power of the ECD, which is estimated for each stimulus condition using the least-squares method (Mosher et al. 2003). The location of the ECD of the aSSR at each frequency is determined based on all stimulus conditions having the same AM rate (see *Source localization* below for details).

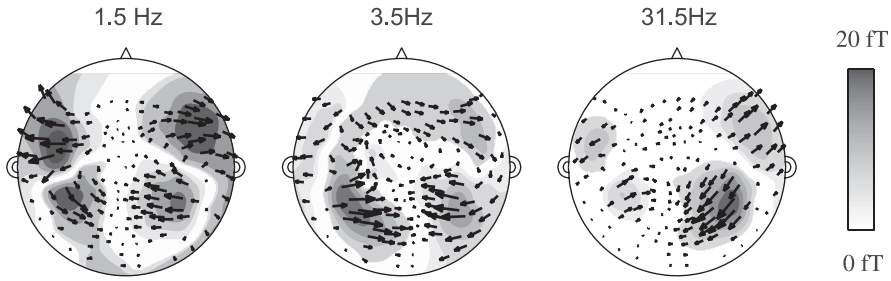

Fig. 1. The auditory steady-state response (aSSR) magnetic field distribution on a flattened head in response to amplitude modulation (AM) sounds with pure tone carrier (from a representative subject). The complex aSSR magnetic field measured at each magnetoencephalography (MEG) sensor at the stimulus AM rate is represented by an arrow, whose length and direction represent the magnitude and phase of the Fourier transform of the response at the stimulus frequency (Simon and Wang 2005).

The MEG response power is affected significantly by the stimulus AM rate in both hemispheres [AM rate $\times$ stimulus bandwidth two-way ANOVA; $F(4, 279) > 22$; $P < 10^{-4}$ for both hemispheres] and is affected significantly by the stimulus bandwidth in the right hemisphere [two-way ANOVA; $F(3, 279) = 3.4$; $P < 0.02$]. Further analysis shows that in the right hemisphere, the influence of stimulus bandwidth on aSSR amplitude is only significant at 1.5 Hz [one-way ANOVA; $F(3, 55) = 3.5$; $P < 0.03$] and marginally significant at 15.5 Hz [one-way ANOVA; $F(3, 55) = 2.4$; $P = 0.07$]. The MEG response power is stronger in the right hemisphere than left for 1.5 Hz AM with pure tone carrier and 31.5 Hz AM with all carriers [paired t-test; $t(13) = 2.2$; $P < 0.05$].

The MEG response has a low-pass profile: it decreases with increasing AM rate. This phenomenon by itself does not directly indicate a low-pass-shaped aSSR MTF, since the MEG response projected to the aSSR dipole location still contains both the aSSR and background activity. The power of the background activity attributed to the aSSR dipole location is shown in Fig. 2 and also has a low-pass profile (background activity at frequency ω is estimated by averaging over stimulus conditions whose stimulus rate is not ω). An unbiased estimate of aSSR power is obtained by taking the difference between MEG measurement power and background activity power (Kay 1988, see Chapter 11). The MTF based on the unbiased estimate of aSSR power still changes significantly as a function of modulation rate in both hemisphere [$F(4, 279) = 12$; $P < 10^{-4}$]. When fitted by a single linear function, the unbiased MTF in the left (right) hemisphere has a slope of -2.0 dB/octave (-1.4 dB/octave).

The phase of the aSSR evoked by AM stimuli with pure tone carriers is plotted in Fig. 3, as an example. Although all subjects are included in the aSSR power analysis, only subjects with significant aSSR are included in the phase analysis (see *Reliability of the aSSR* below). The reason is that for subjects who do not show significant aSSR, their response power will be consistently low, but their response phase will be randomly distributed between 0° and 360°, making the aSSR phase estimation unreliable. The data at 15.5 and 31.5 Hz are omitted, since there are not enough data in the frequency interval between them to determine the phase unwrapping (due to the 360° ambiguity in phase). The slopes of the phase-response function, namely, the apparent latency (Ross et al. 2000) or group delay of the aSSR, are 131 $\pm$ 29 and 147 $\pm$ 18 ms (mean $\pm$ SE) for the left and right hemispheres, respectively. The slope is calculated based on individual subjects who show significant responses at both 1.5 and 3.5 Hz (seven and nine subjects for the left and right hemisphere, respectively). Only two subjects show significant responses for both 1.5 and 7.5 Hz. The mean apparent latencies for the two subjects are 106.31 and 119.63 ms in the left and right hemispheres. The apparent latencies for the aSSR below 3.5 Hz are 89 $\pm$ 50 and 129 $\pm$ 59 ms (mean $\pm$ SE) for AM stimuli with a 1/3-octave pink noise carrier in the left and right hemisphere, and the apparent latency is 121 $\pm$ 20 ms (mean $\pm$ SE) for AM stimuli with a 2-octave pink noise carrier in the right hemisphere. The apparent latencies for other conditions are not consistently measurable, since few subjects show reliable aSSR phase at

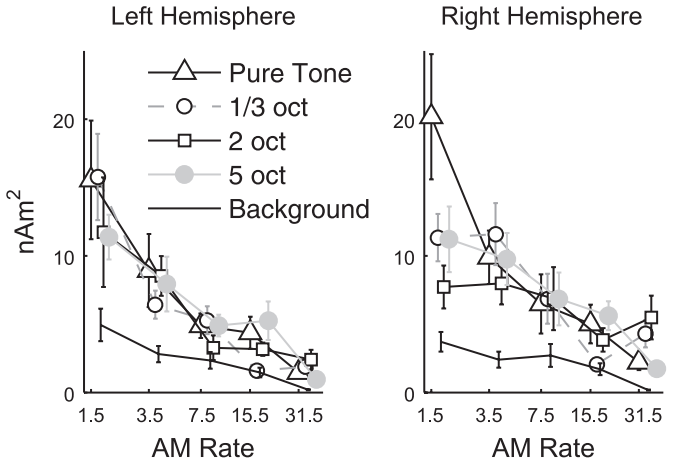

Fig. 2. The power of MEG response decreases as the stimulus AM rate increases. The response power is measured as the power of the equivalent current dipole (ECD) and averaged over subjects. Error bars represent ± 1 SE over subjects. The power of the background activity attributed to the aSSR ECD location is also shown in the figure as a solid black line.

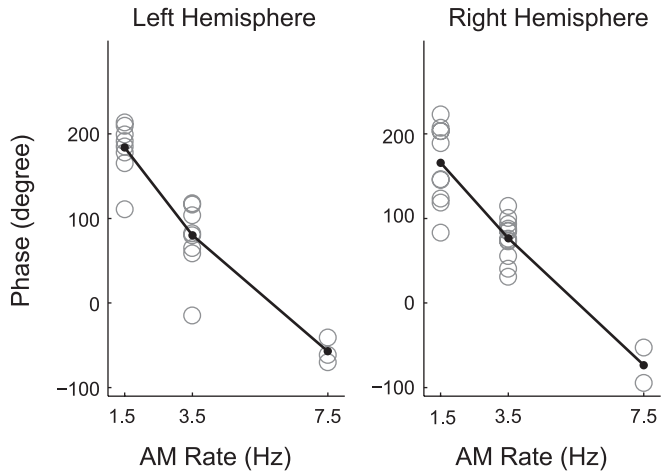

Fig. 3. The phase of the aSSR evoked by AM with a pure tone carrier. Each circle is the result of a single subject, and the line is the grand average. The phase is approximately linear with respect to the stimulus AM rate.

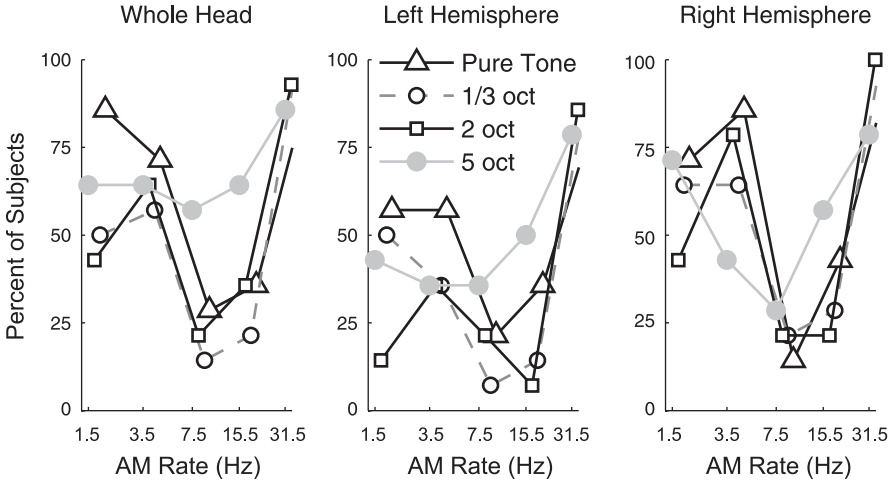

Fig. 4. The percentage of subjects showing a statistically significant aSSR. The aSSR is reliably measurable in most subjects at 1.5, 3.5, and 31.5 Hz.

both 1.5 and 3.5 Hz. At 15.5 Hz and 31.5 Hz, the phase of the aSSR is not affected significantly by stimulus bandwidth nor is it different between the two hemispheres (bandwidth × hemisphere two-way ANOVA; $P > 0.2$).

Reliability of the aSSR. The percentage of subjects showing a significant aSSR, based on a complex-valued DSS filter, is shown in Fig. 4. At the two lowest rates, 1.5 Hz and 3.5 Hz, the aSSR is measured reliably from most subjects, especially for the pure tone carriers. At 7.5 Hz and 15.5 Hz, few subjects show reliable responses, except for the widest bandwidth carrier, 5-octave pink noise. At 31.5 Hz, almost all subjects show measurable responses for all tested stimulus carriers. Table 1 illustrates the percent of subjects showing a significant neural response at each stimulus AM rate, averaged over conditions with different bandwidths.

To improve the detection performance, we combined responses under all stimulus conditions sharing the same modulation rate in the DSS filter design and applied the designed filter to process the response in individual conditions. The results of this method are shown in Table 1 and are manifestly better than the results based on DSS filters designed for individual conditions. This method is also used to test the significance of the harmonics of the aSSR, which are generally very weak except when the AM rate is 15.5 Hz (Table 2). The 15.5-Hz aSSR has strong second and third harmonics, presumably since the two harmonic frequencies, 31 Hz and 46.5 Hz, are relatively close to 40 Hz (Ross et al. 2000).

Table 1. *Percentage of subjects showing a significant aSSR, as the MEG signal is processed by different spatial filters*

AM Rate (Hz)	1.5	3.5	7.5	15.5	31.5
Individual conditions					
DSS filter (complex)	61	64	30	39	91
DSS filter (real/single source)	46	57	20	39	82
Combined conditions (same AM rate)					
DSS filter (complex)	71	75	32	53	98
DSS filter (real/single source)	63	71	34	50	100

The significance test used here is a phase-projected *t*-test. The results of the significance test are averaged over all conditions with the same amplitude modulation (AM) rate. The upper block shows the results when the spatial filters are designed based on individual stimulus conditions. The lower block shows the results when the spatial filters are designed based on combining all conditions with the same AM rate. The performance of the denoising source separation (DSS) filter, designed from a real-valued (single-source) magnetic field, is worse than the performance of the complex-valued DSS filter, indicating that the auditory steady-state response (aSSR) arises from multiple sources with different response phases. MEG, magnetoencephalography.

As discussed in METHODS, if the aSSR arises from a single neural source, the complex-valued aSSR magnetic field can be replaced by a real-valued magnetic field in spatial filter design. However, the performance of the DSS filter designed from a real-valued magnetic field (Table 1) is actually worse than the performance of the complex-valued DSS filter (Table 1), indicating that the aSSR arises from multiple sources with different response phases.

Additionally, to separately evaluate the power and the phase locking of the neural response, we applied bootstrap tests based on the power and intertrial phase coherence of the response separately (Table 3). The spatial filter used here is the real-valued DSS filter that combines all stimulus conditions with the same AM rate. The intertrial phase coherence (Fisher 1993) has much higher statistical power than does response power, indicating very strong phase locking in the response. For subjects who achieve significance with the phase coherence significance test, the phase coherence value does not vary much over conditions, with a value of 0.52 ± 0.09 (mean ± SD over conditions), which is much higher than that of background activity, 0.08 ± 0.01 (mean ± SD over conditions).

We emphasize that all of the results in this section are based on a fivefold cross-validation, as described in METHODS. Without the cross-validation, all of the results in Table 1 would saturate at >97%, an artifact of overfitting.

Source localization. The neural sources of the N1m and the aSSR are localized using real and fully complex dipole models, respectively. Only subjects showing a statistically significant neural response are included in the localization analysis. The GOF medians for the least-square ECD fits are 90% for the N1m and 75% for the aSSR. Only localization results with 80% or higher GOF are included for further analyses. An

Table 2. *Percentage of subjects showing significant responses at frequencies harmonically related to the stimulus AM rate*

AM Rate (Hz)	1.5	3.5	7.5	15.5	31.5
2nd harmonic	21	11	18	36	7
3rd harmonic	7	0	0	30	0

Table 3. *Percentage of subjects showing significant AM responses using separate power- and phase-based significance tests*

AM Rate (Hz)	1.5	3.5	7.5	15.5	31.5
Power	14	33	10	12	71
Phase coherence	44	53	17	35	98

unbalanced two-way ANOVA reveals that the aSSR source location and dipole orientation are not affected significantly by either stimulus rate or carrier bandwidth ($P > 0.1$). Since the aSSR source location is insensitive to stimulus conditions, all aSSR localization results are then combined to compare with the N1m source using a *t*-test. In the left (right) hemisphere, the aSSR source is 4.3 (7.6) mm more medially located than the N1m source [$t(60) > 2.2$, $P < 0.02$ for the left hemisphere; $t(67) > 2.2$, $P < 0.03$ for the right hemisphere]. No significant difference is revealed in the superior-inferior or anterior-posterior direction ($P > 0.2$).

To improve the accuracy of source localization, we averaged the aSSR over all stimulus bandwidths at the same frequency. At each frequency, only subjects showing significant aSSR in more than one condition are considered. The median GOF of the averaged aSSR is 83%. Nevertheless, no significant interaction between the aSSR source location and stimulus rate is found by an unbalanced one-way ANOVA. It is confirmed that the aSSR source location is significantly different from the N1m source location in the medial-lateral direction [8.2 mm, $t(37) = 2.54$, $P < 0.02$ in the left hemisphere; 5.8 mm, $t(37) = 2.58$, $P < 0.02$ in the right hemisphere]. The difference between the ECD location of aSSR and N1m is not caused by the difference between real- and complex-valued dipole localizations, since the aSSR source locations estimated by real and complex dipole models are not significantly different. [The GOF of real dipoles are higher—90% on average—as expected (Simon and Wang 2005).]

The sharpness of a fully complex dipole describes how well a complex-valued magnetic field can be approximated by a real-valued magnetic field (Simon and Wang 2005). The approximation is valid for sharpness values close to zero and is only possible if there is a single aSSR source or if all aSSR sources oscillate in phase. The sharpness values are summarized in Table 4. The sharpness of the 1.5-Hz aSSR is significantly larger than the sharpness of the 31.5-Hz aSSR in both hemispheres [paired *t*-test; $t(3) > 7.1$; $P < 0.01$], as is the sharpness of the 3.5-Hz aSSR in the left hemisphere [paired *t*-test; $t(5) > 4.2$; $P < 0.01$]. This suggests that for the lowest rates, the aSSR is not well approximated by a single, real-valued ECD.

DISCUSSION

Slow temporal modulations carry important information and are important building blocks of speech and other natural sounds (Rosen 1992; Shannon et al. 1995; Singh and Theunissen 2003). This study shows that these slow temporal modulations are neurally represented by large-scale, synchronized, phase-locked activity in a human auditory cortex. Unbiased response power estimation shows that the MTF of the low-frequency aSSR generally has a low-pass pattern and only weakly depends on the carrier bandwidth. Both of these findings are consistent with fMRI data investigating an analogous paradigm (Overath et al. 2012). There, too, one observes a low-pass pattern (of blood oxygen level-dependent responses) as a function of modulation frequency and relative independence from carrier bandwidth.

Power MTF of the MEG aSSR. The overall low-pass-shaped aSSR MTF found in this study is consistent with results from scalp EEG (Rees et al. 1986), intracranial EEG (Liegeois-Chauvel et al. 2004), and fMRI (Giraud et al. 2000; Schönwiesner and Zatorre 2009). It is also consistent with the earlier MEG studies on slow FM (Millman et al. 2009) and more complex temporal modulations (Ding and Simon 2009). This result demonstrates that the frequency dependence of neural sensitivity to the slow temporal modulations of natural sounds, repeatedly demonstrated using MEG and EEG (e.g., Abrams et al. 2008; Luo and Poeppel 2007), can be explained even by the neural coding scheme used for simple AM stimuli.

It should be emphasized that this low-pass profile of the aSSR MTF is not caused by the low-pass nature of background activity, since the aSSR power is estimated as the difference between the power of MEG measurement and the power of background activity. The aSSR power estimated this way is unbiased (Kay 1988). Although the aSSR power shows a low-pass pattern below 31.5 Hz, the signal-to-noise ratio (SNR) of the aSSR has a "U" shape and reaches its maximum at 31.5 Hz, because of the attenuation of background activity at high frequencies. The high SNR makes the aSSR at 31.5 Hz easy to detect. Above 31.5 Hz, previous studies have demonstrated consistently that the MEG aSSR shows maximal response power near 40 Hz and remains measurable up to at least 80 Hz (Gutschalk et al. 1999; Ross et al. 2000; Schoonhoven et al. 2003).

The stimulus carrier bandwidth does not affect the overall low-pass trend of MTF; however, it influences the detailed shape of the MTF, especially in the right hemisphere. The MTF decreases very quickly below 7.5 Hz when the stimulus is a pure tone but is roughly flat when the stimulus is a 5-octave broadband noise. This result is consistent with previous studies showing a strong 10-Hz aSSR using broadband noise carriers (Alaerts et al. 2009; Rees et al. 1986) but a weak 10-Hz aSSR using pure tone carriers (Rees et al. 1986; Ross et al. 2000). The core and belt auditory cortex have been argued to process preferentially pure tone and broadband noises, respectively (Wessinger et al. 2001). Although recent fMRI data do not replicate that profile (Overath et al. 2012) and indeed, directly challenge the hierarchical hypothesis articulated there, it is possible that these regions have different MTFs, and their differing response properties are ultimately reflected in the integrated MEG aSSR.

Cortical neural code for slow modulations. Recordings from the primary auditory cortex of human and other mammals show that modulations below 15 Hz are represented by a phase-locked temporal code (Eggermont 2002; Liang et al. 2002; Liegeois-Chauvel et al. 2004). Similarly, extracranial

Table 4. *Grand averaged sharpness of the fitted, fully complex aSSR current-equivalent dipoles*

AM Rate (Hz)	1.5	3.5	7.5	15.5	31.5
Left	0.20	0.13	0.05	0.11	0.06
Right	0.18	0.20	0.46	0.30	0.06

MEG/EEG recordings from human subjects have also been shown to be phase locked to slow modulations (Alaerts et al. 2009; Luo and Poeppel 2007; Millman et al. 2009). Here, we confirm these results by showing that the intertrial phase coherence of the aSSR is ~0.52, dramatically higher than the 0.08 of background activity.

Additionally, the analysis of the complex-valued magnetic field suggests that the phase-locked activity below 3.5 Hz originates from multiple brain regions. This is demonstrated by two separate methods. Firstly, the complex-valued magnetic field, based on a multiple source assumption, is better at detecting the low-rate aSSR below 3.5 Hz than the real-valued magnetic field, which is based on a single source assumption (e.g., Table 1). Secondly, the sharpness of the fitted, fully complex dipole is larger at 1.5 and 3.5 Hz than at 31.5 Hz (Table 4). Therefore, compared with the 31.5-Hz aSSR, the 1.5- and 3.5-Hz aSSR deviate more from the hypothesis that they are generated from a single source or multiple sources with the same response phase. Hence, the evidence that the aSSR in the low-frequency region (≤3.5 Hz) is generated from multiple sources is compelling. Since the sources of the low-frequency aSSR respond with different phases, it is possible that these sources are activated sequentially, which is also suggested by human intracranial recordings (Gourevitch et al. 2008). In comparison, Gutschalk et al. (1999) provide evidence that the aSSR in the 30- to 50-Hz range also derives from two sources.

The aSSR in the low-frequency range is more reliably detected in the right hemisphere than the left (Fig. 4), which is consistent with the hypothesis that the right hemisphere is dominant for representing slow modulations (Abrams et al. 2008; Poeppel 2003). The aSSR at the relatively high frequency, i.e., 31.5 Hz, is also lateralized to the right hemisphere, similar to the finding that the aSSR near 40 Hz shows a right-hemisphere dominance (Ross et al. 2005).

Neural sources of the low-rate aSSR. AM sounds activate many cortical networks, including both primary and nonprimary auditory cortex (Gourevitch et al. 2008; Liegeois-Chauvel et al. 2004). Given the spatial resolution of MEG, it is difficult to localize the aSSR neural source very precisely. The ECD location of a neural response is usually interpreted as the center of current–gravity of all of its sources (although, see Lütkenhöner and Mosher 2006). This study shows that the ECD of the low-rate aSSR is located 5–10 mm more medially than the ECD of N1m, consistent with earlier work using complex AM stimuli (Draganova et al. 2002). The aSSR ECD locations evoked by AM with different modulation rates (≤31.5 Hz) and carrier bandwidths are not distinguishable. Other studies have shown that the ECD of the 40-Hz MEG aSSR is also 5–10 mm more medially located than the ECD of N1m (Herdman et al. 2003; Ross et al. 2005) within the superior temporal plane, which includes primary auditory cortex. Therefore, the sources of the low-rate aSSR should also be near, possibly including, the primary auditory cortex, consistent with studies using complex stimuli, such as speech and AM, with a time-varying modulation rate (Ahissar et al. 2001; Lalor et al. 2009).

Although the low-rate aSSR and the 40-Hz range aSSR have similar ECD locations, they have quite different apparent latencies. The apparent latency for the 40-Hz aSSR is ~40 ms (Ross et al. 2000; Schoonhoven et al. 2003). In contrast, the apparent latency seen here for the aSSR below 7.5 Hz is between 100 and 150 ms, consistent with previous EEG studies (Alaerts et al. 2009; Lalor et al. 2009; Rees et al. 1986). The different latencies indicate that the aSSR in a low and 40-Hz frequency range may be generated by different neural networks, which is also suggested by human intracranial electrophysiological studies (Gourevitch et al. 2008). However, it is also possible that the latency of the aSSR neural sources is modulation-rate dependent. Since we have shown that the low-rate aSSR is generated from multiple neural sources with different response phases, the apparent latency value may only be an overall property of MEG responses and not directly related to the response latency of any neural source.

Importantly, the disparity between the low-rate aSSR source and the N1m source indicates that the low-rate aSSR evoked by sinusoidal AM is certainly not a simple superposition of the N1m response, nor is it likely to arise from the P2m component, which is less robust than the N1m, and with a more variable source location. The source location of P2m mainly differs from the source location of N1m in the anterior-posterior direction rather than medial-lateral direction (Altmann et al. 2008; Ross et al. 2007). Furthermore, the latency of the P2m is highly variable in different experiments, ranging from 150 to >200 ms (Altmann et al. 2008; Chait et al. 2004; Howard and Poeppel 2009; Ross et al. 2007). In addition, the energy of a sinusoidal AM is continuously and smoothly changing without any salient onset/offset cues in each modulation cycle, which makes it unlikely that a N1m–P2m complex is elicited by a specific feature in each AM cycle. In contrast, for a slow aSSR, generated by click trains, it is possible that some response components would result from the N1m–P2m complex.

In summary, this study characterizes the MEG response to AM, near and below 30 Hz. Although the low-frequency neural background activity is strong, the aSSR can still be measured reliably with appropriate spatial processing methods (e.g., DSS-generated filters) and significance tests (e.g., the phase-projected bootstrap test). The phase-locked aSSR is strongest for very low temporal modulations and decreases ~2 dB/octave with increasing AM rate. The low-frequency aSSR is generated from multiple neural sources, with their locations located more medially than the ECD of the N1m. This profile of neural sensitivity to slow temporal modulations, seen universally from the single unit level to the large-scale, synchronized neural network level, for both natural sounds and simple AM/FM, may serve as a fundamental principle of neural processing in auditory cortex.

APPENDIX

DSS. The linear DSS transform (de Cheveigné and Simon 2008; Särelä and Valpola 2004) uses two covariance matrices, $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$, and each DSS filter $\boldsymbol{w}$ can be expressed as the solution of the following generalized eigendecomposition problem (Fukunaga 1972).

$$\boldsymbol{R}_2\boldsymbol{w} = \lambda\boldsymbol{R}_1\boldsymbol{w} \quad (1)$$

In this study, $\boldsymbol{R}_1$ is chosen to be the covariance of background noise $\boldsymbol{R}_n$, and $\boldsymbol{R}_2$ is chosen to be the covariance of the aSSR magnetic field $\boldsymbol{S}(\omega)$: $\boldsymbol{S}(\omega)\boldsymbol{S}^H(\omega)$. With the use of $\boldsymbol{R}_1 = \boldsymbol{R}_n$, and $\boldsymbol{R}_2 = \boldsymbol{S}(\omega)\boldsymbol{S}^H(\omega)$ in *Eq. 1* and multiplying $\boldsymbol{R}_n^{-1}$ on both sides, we have, $\boldsymbol{w} = \alpha\boldsymbol{R}_n^{-1}\boldsymbol{S}(\omega)$ where $\alpha = \boldsymbol{S}^H(\omega)\boldsymbol{w}/\lambda$. Therefore, for aSSR detection, $\boldsymbol{w} = \boldsymbol{R}_n^{-1}\boldsymbol{S}(\omega)$ is the only DSS filter.

Likelihood ratio test for the aSSR in Gaussian noise. When the aSSR is assumed to be a sinusoid at frequency ω in the frequency domain, its detection may be formulated as the following hypothesis testing problem.

$$\begin{cases} \mathrm{H0} : \boldsymbol{M}(\omega) = \boldsymbol{N}(\omega) \\ \mathrm{H1} : \boldsymbol{M}(\omega) = \alpha \boldsymbol{S}(\omega) + \boldsymbol{N}(\omega), \ \alpha \neq 0 \end{cases} \quad (2)$$

The optimal significance test is a likelihood ratio test (Poor 1994), and the likelihood ratio function is $L[\boldsymbol{M}(\omega)] = P[\boldsymbol{M}(\omega) \mid \mathrm{H1}]/P[\boldsymbol{M}(\omega) \mid \mathrm{H0}]$. If we assume the background noise to be subject to the complex Gaussian distribution with zero mean and covariance $\boldsymbol{R}_\mathrm{n} = E[\boldsymbol{N}(\omega)\boldsymbol{N}^\mathrm{H}(\omega)]$ (Kay 1988, see Chapter 3), $L[\boldsymbol{M}(\omega)]$ is proportional to $\boldsymbol{M}^\mathrm{H}(\omega)\boldsymbol{R}_\mathrm{n}^{-1}\boldsymbol{S}(\omega)$. Therefore, the optimal decision statistic can be viewed as the output of a spatial filter $\boldsymbol{w} = \boldsymbol{R}_\mathrm{n}^{-1}\boldsymbol{S}(\omega)$, which is equal to the DSS filter.

Implementation of DSS filters. Evaluation of the DSS filters requires the aSSR magnetic field $\boldsymbol{S}(\omega)$ and the noise covariance matrix $\boldsymbol{R}_\mathrm{n}(\omega)$. $\boldsymbol{S}(\omega)$ was estimated by averaging the MEG measurements over trials, and $\boldsymbol{R}_\mathrm{n}(\omega)$ was estimated as the covariance matrix of the responses from conditions with the stimulus AM rate, not at ω. $\boldsymbol{R}_\mathrm{n}(\omega)$ is, in general, complex, but in practice, it was observed that using only the real part of $\boldsymbol{R}_\mathrm{n}(\omega)$ usually gave slightly higher statistical power. Therefore, only the real part of $\boldsymbol{R}_\mathrm{n}(\omega)$ was used for all analyses presented here.

ACKNOWLEDGMENTS

We thank Jeff Walker for excellent technical assistance and Minsuk Park and Harshavardhan A. Agashe for help with the analysis. We also acknowledge Alain de Cheveigné and Mary F. Howard for constructive comments on earlier versions of this manuscript.

Present addresses: Y. Wang, Ditech Networks, Mountain View, CA 94043; N. Ahmar, Ohio State University, Columbus, OH 43210; J. Xiang, Goldman Sachs, New York, NY 10282.

GRANTS

Support for this research was provided by the National Institute on Deafness and Other Communication Disorders Grants R01-DC-005660 and R01-DC-008342.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: Y.W., D.P., and J.Z.S. conception and design of research; Y.W. performed experiments; Y.W., N.D., N.A., J.X., and J.Z.S. analyzed data; Y.W., N.D., D.P., and J.Z.S. interpreted results of experiments; Y.W., N.D., and J.Z.S. prepared figures; Y.W., N.D., and J.Z.S. drafted manuscript; Y.W., N.D., D.P., and J.Z.S. edited and revised manuscript; N.D., D.P., and J.Z.S. approved final version of manuscript.

REFERENCES

Abrams DA, Nicol T, Zecker S, Kraus N. Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *J Neurosci* 28: 3958–3965, 2008.

Ahissar E, Nagarajan S, Ahissar M, Protopapas A, Mahncke H, Merzenich MM. Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proc Natl Acad Sci USA* 98: 13367–13372, 2001.

Aiken SJ, Picton TW. Human cortical responses to the speech envelope. *Ear Hear* 29: 139–157, 2008.

Alaerts J, Luts H, Hofmann M, Wouters J. Cortical auditory steady-state responses to low modulation rates. *Int J Audiol* 48: 582–593, 2009.

Altmann CF, Nakata H, Noguchi Y, Inui K, Hoshiyama M, Kaneoke Y, Kakigi R. Temporal dynamics of adaptation to natural sounds in the human auditory cortex. *Cereb Cortex* 18: 1350–1360, 2008.

Brugge JF, Nourski KV, Oya H, Reale RA, Kawasaki H, Steinschneider M, Howard MA, III. Coding of repetitive transients by auditory cortex on Heschl's gyrus. *J Neurophysiol* 102: 2358–2374, 2009.

Chait M, Simon JZ, Poeppel D. Auditory M50 and M100 responses to broadband noise: functional implications. *Neuroreport* 15: 2455–2458, 2004.

Chevillet M, Riesenhuber M, Rauschecker JP. Functional correlates of the anterolateral processing hierarchy in human auditory cortex. *J Neurosci* 31: 9345–9352, 2011.

de Cheveigné A, Simon JZ. Denoising based on spatial filtering. *J Neurosci Methods* 171: 331–339, 2008.

de Cheveigné A, Simon JZ. Denoising based on time-shift PCA. *J Neurosci Methods* 165: 297–305, 2007.

Ding N, Simon JZ. Neural representations of complex Temporal modulations in the human auditory cortex. *J Neurophysiol* 102: 2731–2743, 2009.

Draganova R, Ross B, Borgmann C, Pantev C. Auditory cortical response patterns to multiple rhythms of AM sound. *Ear Hear* 23: 254–265, 2002.

Drullman R, Festen JM, Plomp R. Effect of temporal envelope smearing on speech reception. *J Acoust Soc Am* 95: 1053–1064, 1994.

Eggermont JJ. Temporal modulation transfer functions in cat primary auditory cortex: separating stimulus effects from neural mechanisms. *J Neurophysiol* 87: 305–321, 2002.

Fisher NI. *Statistical Analysis of Circular Data*. New York: Cambridge University Press, 1993.

Fukunaga K. *Introduction to Statistical Pattern Recognition*. New York: Academic, 1972.

Giraud AL, Lorenzi C, Ashburner J, Wable J, Johnsrude I, Frackowiak R, Kleinschmidt A. Representation of the temporal envelope of sounds in the human brain. *J Neurophysiol* 84: 1588–1598, 2000.

Gourevitch B, Jeannes RLB, Faucon G, Liegeois-Chauvel C. Temporal envelope processing in the human auditory cortex: response and interconnections of auditory cortical areas. *Hear Res* 237: 1–18, 2008.

Greenberg S, Carvey H, Hitchcock L, Chang S. Temporal properties of spontaneous speech—a syllable-centric perspective. *J Phon* 31: 465–485, 2003.

Gutschalk A, Mase R, Roth R, Ille N, Rupp A, Hahnel S, Picton TW, Scherg M. Deconvolution of 40 Hz steady-state fields reveals two overlapping source activities of the human auditory cortex. *Clin Neurophysiol* 110: 856–868, 1999.

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

Howard MF, Poeppel D. Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension. *J Neurophysiol* 104: 2500–2511, 2010.

Howard MF, Poeppel D. Hemispheric asymmetry in mid and long latency neuromagnetic responses to single clicks. *Hear Res* 257: 41–52, 2009.

Kay SM. *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ: Prentice Hall, 1988.

Lalor EC, Power AJ, Reilly RB, Foxe JJ. Resolving precise temporal processing properties of the auditory system using continuous stimuli. *J Neurophysiol* 102: 349–359, 2009.

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

Liegeois-Chauvel C, Lorenzi C, Trebuchon A, Regis J, Chauvel P. Temporal envelope processing in the human left and right auditory cortices. *Cereb Cortex* 14: 731–740, 2004.

Luo H, Poeppel D. Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron* 54: 1001–1010, 2007.

Luo H, Wang Y, Poeppel D, Simon JZ. Concurrent encoding of frequency and amplitude modulation in human auditory cortex: MEG evidence. *J Neurophysiol* 96: 2712–2723, 2006.

Lütkenhöner B, Mosher editors JC. *Source Analysis of Auditory Evoked Potentials and Fields*. Philadelphia: Lippincott Williams & Wilkins, 2006.

Millman RE, Prendergast G, Kitterick PT, Woods WP, Green GGR. Spatiotemporal reconstruction of the auditory steady-state response to frequency modulation using magnetoencephalography. *Neuroimage* 49: 745–758, 2009.

Mosher JC, Baillet S, Leahy RM. Equivalence of linear approaches in bioelectromagnetic inverse solutions. In: *IEEE Workshop on Statistical Signal Processing*. St. Louis, MO, 2003, p. 294–297.

Näätänen R, Picton T. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24: 375–425, 1987.

Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9: 97–113, 1971.

Overath T, Zhang Y, Sanes DH, Poeppel D. Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: fMRI evidence. *J Neurophysiol*; published ahead of print February 1, 2012, doi:10.1152/jn.00308.2011.

Picton TW, Dimitrijevic A, John MS, Roon PV. The use of phase in the detection of auditory steady-state responses. *Clin Neurophysiol* 112: 1698–1711, 2001.

Picton TW, Skinner CR, Champagne SC, Kellett AJC, Maiste AC. Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone. *J Acoust Soc Am* 82: 165–178, 1987.

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

Poor HV. *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.

Rees A, Green G, Kay R. Steady-state evoked responses to sinusoidally amplitude-modulated sounds recorded in man. *Hear Res* 23: 123–133, 1986.

Rosen S. Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond B Biol Sci* 336: 367–373, 1992.

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

Ross B, Fujioka T, Tremblay KL, Picton TW. Aging in binaural hearing begins in mid-life: evidence from cortical auditory-evoked responses to changes in interaural phase. *J Neurosci* 27: 11172–11178, 2007.

Ross B, Herdman AT, Pantev C. Right hemispheric laterality of human 40 Hz auditory steady-state responses. *Cereb Cortex* 15: 2029–2039, 2005.

Särelä J, Valpola H. Denoising source separation. *J Mach Learn Res* 6: 233–272, 2004.

Schönwiesner M, Zatorre RJ. Spectro-temporal modulation transfer function of single voxels in the human auditory cortex measured with high-resolution fMRI. *Proc Natl Acad Sci USA* 106: 14611–14616, 2009.

Schoonhoven R, Boden CJR, Verbunt JPA, de Munck JC. A whole head MEG study of the amplitude-modulation-following response: phase coherence, group delay and dipole source analysis. *Clin Neurophysiol* 114: 2096–2106, 2003.

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M. Speech recognition with primarily temporal cues. *Science* 270: 303–304, 1995.

Simon JZ, Wang Y. Fully complex magnetoencephalography. *J Neurosci Methods* 149: 64–73, 2005.

Singh NC, Theunissen FE. Modulation spectra of natural sounds and ethological theories of auditory processing. *J Acoust Soc Am* 114: 2003.

Uutela K, Hamalainen M, Salmelin R. Global optimization in the localization of neuromagnetic sources. *IEEE Trans Biomed Eng* 45: 716–723, 1998.

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