

The Analysis of Simple and Complex Auditory Signals in Human Auditory Cortex: Magnetoencephalographic Evidence from M100 Modulation

Julian Jenkins III,^{1,2} William J. Idsardi,² and David Poeppel³

Objective: Ecologically valid signals (e.g., vowels) have multiple components of substantially different frequencies and amplitudes that may not be equally cortically represented. In this study, we investigate a relatively simple signal at an intermediate level of complexity, two-frequency composite tones, a stimulus lying between simple sinusoids and ecologically valid signals such as speech. We aim to characterize the cortical response properties to better understand how complex signals may be represented in auditory cortex.

Design: Using magnetoencephalography, we assessed the sensitivity of the M100/N100m auditory-evoked component to manipulations of the power ratio of the individual frequency components of the two-frequency complexes. Fourteen right-handed subjects with normal hearing were scanned while passively listening to 10 complex and 12 simple signals. The complex signals were composed of one higher frequency and one lower frequency sinusoid; the lower frequency sinusoidal component was at one of the five loudness levels relative to the higher frequency one: -20, -10, 0, +10, +20 dB. The simple signals comprised all the complex signal components presented in isolation.

Results: The data replicate and extend several previous findings: (1) the systematic dependence of the M100 latency on signal intensity and (2) the dependence of the M100 latency on signal frequency, with lower frequency signals (~100 Hz) exhibiting longer latencies than higher frequency signals (~1000 Hz) even at matched loudness levels. (3) Importantly, we observe that, relative to simple signals, complex signals show increased response amplitude—as one might predict—but decreased M100 latencies.

Conclusion: The data suggest that by the time the M100 is generated in auditory cortex (~70 to 80 msec after stimulus onset), integrative processing across frequency channels has taken place which is observable in the M100 modulation. In light of these data models that attribute more time and processing resources to a complex stimulus merit reevaluation, in that our data show that acoustically more complex signals are associated with robust temporal facilitation, across frequencies and signal amplitude level.

(Ear & Hearing 2010;xx;1-●)

INTRODUCTION

Magnetoencephalographic (MEG) studies have shown that both physical and perceptual attributes of auditory stimuli are reflected in the timing and magnetic field deflection of the major auditory-evoked component, called the N100m or M100 (Mäkinen et al. 2004; Salajegheh et al. 2004; Hari et al. 1980; Hari and Mäkelä 1988; Eulitz et al. 1995; Fujioka et al. 2003; Seither-Preisler et al. 2003; Kirveskari et al. 2006; Lütkenhöner et al. 2006). This field component, occurring ~100 msec after stimulus onset, has been shown to be sensitive to differences in

the frequency, waveform shape (e.g., sinusoidal versus sawtooth), and intensity of the stimulus (see Roberts et al. 2000 for review). For example, lower frequency signals (~100 Hz), a frequency range typical of the fundamental frequency of male voices, show longer latencies and decreased field deflections than higher frequency signals (~1000 Hz), a frequency range typical for the first formant in speech (Fant 1960). Crucially, this relation holds even when frequencies are equated for loudness level. However, this interaction is not static. The latency and amplitude in response to lower frequency signals can be decreased and increased, respectively, by increasing the intensity (amplitude) of the signal (Stufflebeam et al. 1998).

It has also been argued that the processing reflected in the M100 reflects perceptual attributes of a stimulus (Roberts et al. 2000). This implies that the perceptual attributes of signals are at least partially reflected in the latency and amplitude of the M100 component, insofar as the particular attribute can be extracted by ~80 msec, the time at which the M100 is generated. To that end, it is useful to investigate the link between loudness, a perceptual attribute of an auditory stimulus with a clear and obvious contribution from the physical structure (intensity) of the stimulus (although not a one-to-one relationship), and its effect on the latency and amplitude of the M100. In short, the question is to contrast the effect on the M100 of direct physical properties of the signal and properties derived from the signal.

The principal paradigm has been to play simple single signals (e.g., sinusoids and square waves) or single speech sounds (e.g., vowels) and to evaluate the timing, amplitude, and dipole localization of the M100 (Pantev et al. 1989; Eulitz et al. 1995; Diesch et al. 1996; Diesch & Luce 1997, 2000; Lütkenhöner et al. 2001; Obleser et al. 2003, Obleser et al. 2004a,b). On one hand, there have been numerous studies utilizing such signals that have been useful in characterizing response properties of the M100. At the other extreme, studies using ecological signals such as speech and music are acoustically and spectrally complex in a way that exceeds our current understanding. However, there is not yet a rich literature on acoustic stimuli that are of an “intermediate” nature. Therefore, the linking hypotheses between the processing of simple stimuli and the processing of ecologically natural stimuli remain a bit underspecified. Signals such as speech, music, and natural and artificial sounds contain not only mixtures of frequency components but also components that contribute in unequal ways to the overall structure of the signal, i.e., the power levels of the components may be heavily biased toward a single component or group of components.

There are two—somewhat independent—motivations to pursue a deeper understanding of the M100 and its properties.

¹Department of Biology, and ²Cognitive Neuroscience of Language Laboratory, University of Maryland, College Park, Maryland; and ³Department of Psychology, New York University, New York, New York.

First, there is a need to understand more about auditory neuroscience in the context of human auditory processing, and specifically the integration across auditory channels and critical bands, with a view to the cortical parsing of complex signals. The features of ecological signals contain components that must be integrated across a variety of critical bands and processing channels. The M100 may be particularly well suited to elucidate relatively early cortical aspects of neuronal encoding of complex sound processing (see Chait et al. 2006, 2007 for studies on the analysis of Huggins pitch and change detection, respectively). A second motivation derives from the fact that the M100 response and its predictable variability have been used increasingly in the investigation of auditory cognition. For example, the M100 latency dependency on frequency has been used to test theories of vowel perception (Poeppl 1997; Vihla et al. 2000; Vihla & Salmelin 2002, Tiitinen et al. 2005), virtual pitch (Monahan et al. 2008), and other phenomena. Insofar as the M100 is used to test models in auditory cognitive neuroscience, studies that investigate the factors modulating it merit further attention.

To put the study into context, the mid- and long-latency auditory-evoked potentials/auditory-evoked fields have been productively used in many studies investigating auditory perception. For example, investigation of the latency of evoked components has yielded important insights into auditory system development (Eggermont 1995), functional reorganization (Hirata et al. 1999; Brattico et al. 2003; Nikjeh et al. 2009; Okamoto et al. 2009), cochlear damage (Dietrich et al. 2001), and loudness perception and processing in both humans and nonhuman primates (Tucker et al. 2001; Tanji et al. 2010). Our study was designed to test specific hypotheses concerning loudness perception and signal component integration across critical bands. Because complex sounds contain multiple frequencies that vary in their relative power, it is important to get a more quantitative understanding of the response structure.

In this study, we used MEG, a technique well suited for the analysis of temporal information in auditory cortex (Roberts et al. 2000; Lütkenhöner & Poeppel 2010), to examine the effect on the M100 of variations in the frequency and amplitude of complex sinusoidal signals, i.e., signals composed of sinusoids of more than one frequency. Response attributes of the M100 for complex stimuli are compared with the M100 response attributes for the individual components of the complex stimuli. Intuitively speaking, the most straightforward hypothesis may be that the integrative processing (across frequency and time) required by complex signals may be associated with response prolongation, amplitude increases, and other factors reflective of the engagement of additional neural resources because of the increased complexity implicit in the perceptual analysis of structured signals. Interestingly, we observe the opposite pattern for response latency.

MATERIALS AND METHODS

Subjects

Fourteen right-handed (Oldfield 1971) normal-hearing adult subjects (nine females) underwent MEG scanning. One subject was excluded from the analysis because of an insufficient signal to noise ratio for all experimental conditions. Age range was 19 to 27 yrs, mean 21.8 yrs. Subjects were compensated (\$10/hr) for their participation. Presentation of stimuli and

biomagnetic recording was performed with the approval of the institutional committee on human research of the University of Maryland, College Park. Before the start of the experiment, informed written consent was obtained from each subject.

Threshold Testing

Threshold testing was conducted using the Hughson-Westlake paradigm (Carhart & Jerger 1959) while the subject was in the MEG shielded room. Subjects' hearing threshold was determined using 250, 500, 1000, and 2000 Hz tone pulses delivered via an Earscan 3 audiometer (Micro Audiometrics Corp., Murphy, NC). Potential subjects with >5 dB HL difference between ears or a hearing threshold >15 dB HL for any frequency were rejected.

Perceptual Loudness Estimation

Perceptual loudness was estimated by placing the subject in the MEG scanner with the experimental earphones (to simulate experimental conditions) before the start of the experiment. Each subject was asked to discriminate a pair of signals, with both members of a pair consisting of sinusoids. Subjects had to indicate which pair of signals had the same loudness (Moore 2004). The first signal in a pair was a 1000-Hz signal presented at 60 dB above threshold. The second signal was 127, 252, or 800 Hz signal at one of the three varying levels. The levels were chosen based on previous loudness studies and standards (ISO 226:2003), which give an estimate of the SPL for the loudness of sinusoids of varying frequencies to be matched to a 1000-Hz sinusoid. For all subjects, the levels of the lower frequency signals ranged from 50 to 75 dB SPL. The three levels were as follows: (a) 5 dB below the SPL most likely to provide a match to the 1000-Hz signal, (b) at the sound pressure level most likely to match that of the 1000-Hz signal, and (c) 5 dB above the sound pressure level most likely to match the 1000-Hz signal. All signals were of 1-sec duration and had a 7 msec \cos^2 onset and offset. Subjects were first asked to evaluate the loudness of the signal pairs independently as many times as needed and then to respond via button press when the two signals had the same perceptual loudness. If the signals were not matched in loudness then the level of the volume of the lower frequency signal was adjusted by the experimenter, and the subject was asked to repeat the task until the perceptual loudness was matched. The signals with the same perceptual loudness were carried over and used as the 0-dB signals for the experiment (see below).

Stimuli

Signals for the experiment were generated with MATLAB (v7, The Mathworks, Natick, MA). Four sinusoidal signals of the frequencies 127, 252, 800, and 1000 Hz were sampled at 44.1 kHz with 16-bit resolution. Signals were generated using the sine function not the cosine function. Duration was 400 msec with a 7 msec \cos^2 onset/offset ramp. Complex signals consisted of combinations of 127 Hz with 1000 Hz and 252 Hz with 800 Hz. Individual frequencies and their combinations were selected to be in nonoverlapping critical bands.

The two lower frequencies were presented at five different signal levels (with respect to the 800- and 1000-Hz signals 60 dB above the subject's threshold). A reference value of 1000 Hz was used because of the human auditory system's sensitiv-

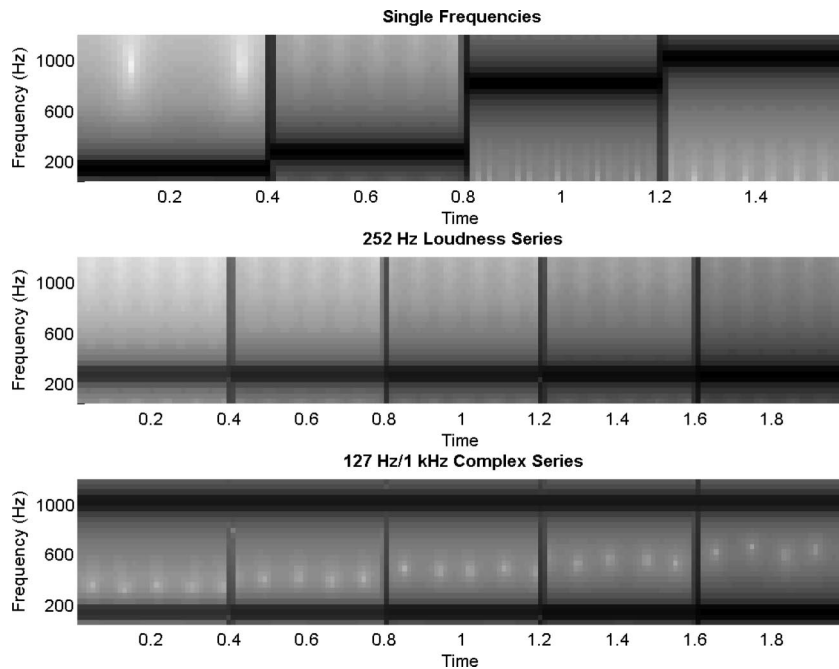


Fig. 1. Schematic spectrograms of stimuli. Top panel, Spectrograms of single sinusoids (left to right: 127, 252, 800, and 1000 Hz) at 0-dB loudness level. Middle panel, Loudness series for 252-Hz single sinusoid from -20 dB at the far left to $+20$ dB at the far right. Bottom panel, Loudness series for 127/1000 Hz complex sinusoid. Panels show full range of lower frequency component of stimuli from -20 to $+20$ dB levels.

ity to frequencies between 1000 and 2000 Hz, as well as the well known M100/N100m and associated parameters for this frequency. The relative values used were -20 , -10 , 0 , $+10$, and $+20$ dB. The two high-frequency signals, the 800- and 1000-Hz signals, were only presented at one loudness level (0 dB). The two complex signals were presented in mixtures of amplitudes for a total of 22 signals (12 single signals and 10 complex signals; see Fig. 1). The energy of the complex signals and simple signals was not identical because it was important to keep all signal components that constituted the complex signals identical to their simple signal counterparts to accurately gauge the effect on M100 latency and peak root mean square (RMS). Each stimulus was presented 100 times, pseudorandomly interleaved. Complex signals were more akin to “chords” rather than being mathematically additive; as such, special care was taken to make sure that both ears received the same signal being applied. The experimental materials were passively listened to; no response was required from the subjects. To maintain the vigilance of the subjects, a distracter task was incorporated into the experiment. Approximately Gaussian white noise (400-msec duration) was used as the target during the experiment and was pseudorandomly presented with the signals (17% of total). Subjects had to press a button in response to the noise target; these trials were excluded from analysis.

Delivery

All experimental stimuli were presented using a Dell Optiplex computer with a SoundMAX Integrated HD sound card (Analog Devices, Norwood, MA) via Presentation stimulus presentation software (Neurobehavioral Systems, Inc., Albany, CA). Stimuli were delivered to the subjects binaurally via Eartone ER3A transducers and nonmagnetic air-tube delivery (Etymotic, Oak Brook, IL). The interstimulus interval varied randomly in the time interval between 300 and 700 msec. The decision to use a relatively short interstimulus interval (ISI)

was based on practical considerations; the use of such an ISI ensured that subjects would not be in the scanner more than 1 hr, after which considerable discomfort may occur. Previous studies have investigated the impact of ISI on the peak and latency of the N1/N1m (Hari et al. 1982) with shorter ISIs exhibiting longer latencies. However, because it takes ~ 10 to 12 secs for the response to the previous stimulus to disappear completely, extending the ISI may be artificially inflating the M100 response. Subjects were instructed to stay as still as possible while positioned inside of the scanner; no mechanical measures were taken (such as a foam neck support) to maintain that position. However, the amount of deviation between the start of the experiment and its conclusion was quantified by taking measurements of the subjects’ head positions using marker coils.

Recording

Data were acquired using a 160-channel whole-head biomagnetometer with axial gradiometer sensors (KIT System, Kanazawa, Japan). Recording bandwidth was DC-100 Hz, with a 60-Hz Notch filter, at 1000-Hz sampling rate. Data were noise reduced using time-shifted principal component analysis (de Cheveigné & Simon 2007), trials averaged offline (artifact rejection ± 2.5 pT), bandpass filtered between 0.03 and 20 Hz, and baseline corrected over the 100-msec prestimulus interval. The relatively low upper frequency cutoff value is not unusual; auditory-evoked field studies generally have filter cutoff values of 20 to 30 Hz (see Tiitinen et al. 1993; Eulitz et al. 1995). In addition, time-shifted principal component analysis does not remove eye-blink artifacts, but trials with excessive frontal field deflections were excluded from the analysis.

Data Analysis: Peak RMS and Latency Analysis

Selection of maximally responsive auditory channels was performed in a pretest, using 250 and 1000 Hz sinusoidal signals of 400-msec duration presented at ~ 65 dB SPL. On

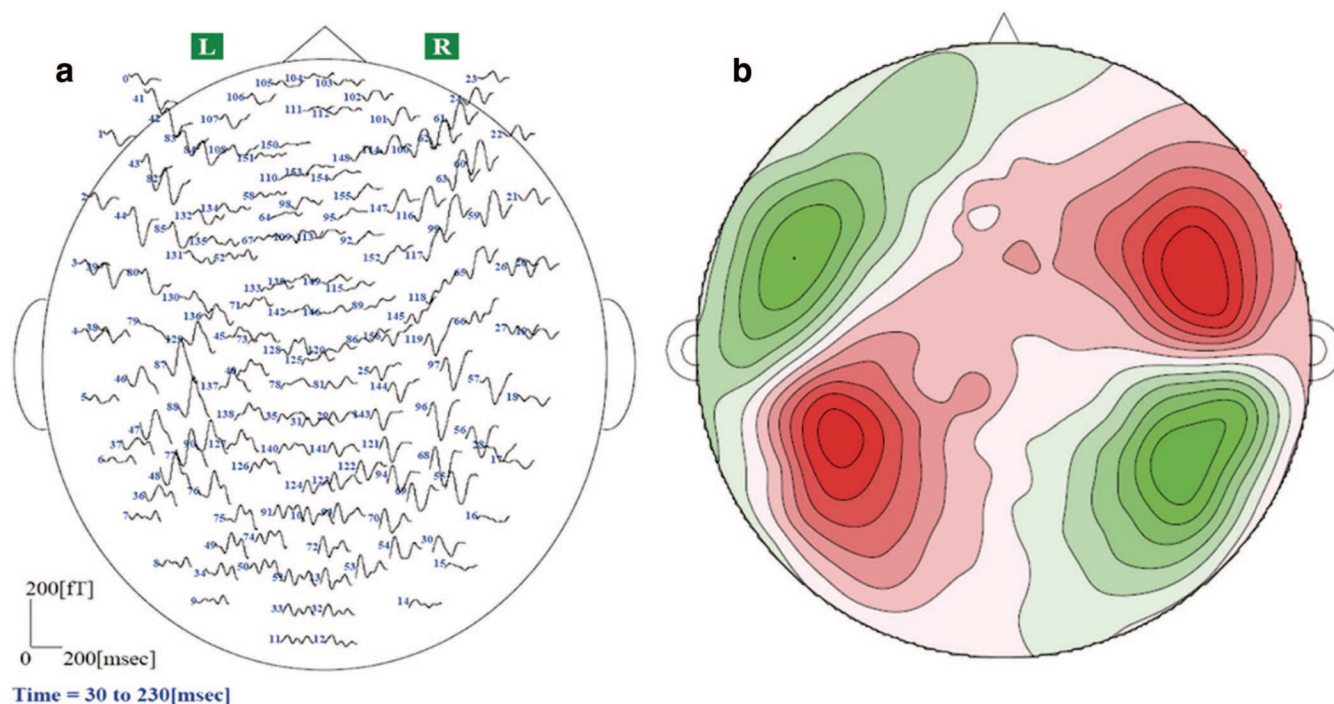


Fig. 2. a, Sensor configuration of whole-head biomagnetometer. The response of 157 data channels for one subject's response to a 252/800 Hz complex signal is shown. The peaks and troughs in the evoked magnetic field illustrate the typical distribution of an auditory-evoked response measured by axial gradiometers. The latency of M100 peak is 122 msec. As expected, channels along the midline, i.e., not overlaying temporal lobe cortical areas, do not reflect any interpretable evoked activity elicited by an auditory signal. b, Contour map of the response peak of the sensor data shown in (a). Magnetic field source is in red and sink is in green, depicting the canonical left vs. right dipolar configurations.

average, the 1000-Hz signal resulted in a stronger auditory response. Five channels from source and sink from each hemisphere (i.e., 20 channels total) with the maximum measured magnetic field deflection to the 1000-Hz signal were used for subsequent analysis. M100 peak RMS amplitude and latency (search window: 90 to 210 msec after stimulus onset) for each signal for each hemisphere were determined using the localizer channels. M100 peak RMS and latency data values for each hemisphere for each condition were averaged across the subjects (see Figs. 2 and 3 for visualization). The signal evaluation window ranged from 100-msec pretrigger to 600-msec posttrigger. M100 peak RMS time values for all data channels were collected and averaged across subjects for each stimulus and were plotted topographically to confirm the M100 response. In addition, when a subject's data did not show an auditory cortex magnetic field topography for a given condition, after averaging and filtering, that data (peak RMS and latency) was excluded from further analysis (Luo & Poeppel 2007). A total of 31 such exclusions were made.

Across-subject responses for all subjects showing a quantifiable M100 were characterized by collecting the individual RMS vectors into a matrix (according to condition) and then calculating the RMS of that matrix. The individual RMS values were calculated from the sensors selected from the pretest data for all time points of the observed response. The peak RMS value and latency for the across subject response was quantified in the same time window as the individual subject data. The motivation for using the RMS of the RMS (see Fig. 7) is as follows: (a) MEG is a technique well suited for within-subject analyses and (b) because the sensor positions are not in the same place for all subjects, unlike EEG where there are

consistent points of reference, grand averages are not as straightforward. Although some researchers do present grand-averaged data, the interpretation of the results still requires RMS. Because of the inherent difficulty of presenting grand averages of channels in MEG data because of the fact that channels are not in canonical positions, we use the RMS of RMS measure to present grand averages over a derived measure (RMS) that provides an index of regional activation. This has proven a valuable and valid measure in previous studies (Chait et al. 2006).

To assess the significance of the effect of auditory signal manipulation on the M100 latency and peak RMS, the values were collected and analyzed using repeated measures analysis of variance (ANOVA) (SPSS 16.0, SPSS Inc., Chicago, IL). Missing M100 values (see above) were replaced with the across-subject series mean for the particular signal type. A full factorial design was employed with amplitude and latency as the dependent measures and hemisphere (right hemisphere [RH] versus left hemisphere [LH]), signal type (simple versus complex), and loudness level as the factors. Three additional planned comparisons were performed using paired *t*-tests. The first compared simple low-frequency signals with complex signals (e.g., 127 Hz -20 dB to 127 Hz -20 dB/1000 Hz) for the M100 latency in the same hemisphere, the second compared the same kind of signal across hemispheres (e.g., 127 Hz 0 dB RH versus 127 Hz 0 dB LH), and the third analysis compared high-frequency signals with complex signals (e.g., 800-252 Hz -10 dB/800 Hz) within the same hemisphere. All statistical effects (repeated measures ANOVA and paired *t*-tests) are reported as significant at $p < 0.05$.

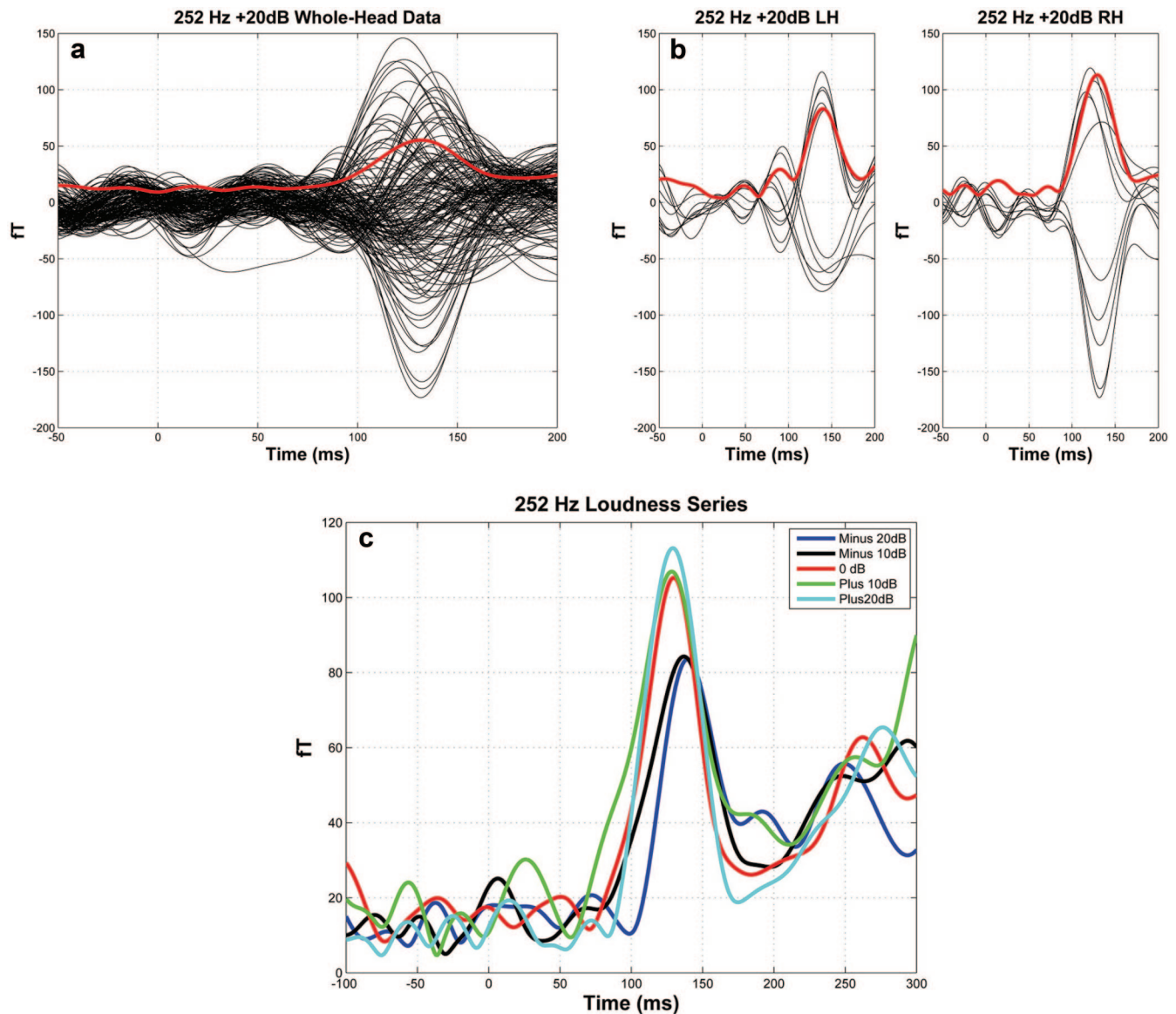


Fig. 3. a, Data from single subject for 157 channels in response to 252-Hz single signal at +20 dB presentation level. Black lines are magnetic field deflections from each channel; red line shows the root mean square (RMS) across all data channels. The data shown are averaged (88 trials) and band-pass filtered (0.03 to 20 Hz). This “global” view of the response (across hemispheres) illustrates the large and robust response profile elicited by the stimuli and provides a sense of the scale of the response under consideration in terms of timing and amplitude. b, Response broken down by hemisphere. For each hemisphere, the 10 best channels (maximum field deflection), 5 from the source and 5 from the sink, are selected and their RMS calculated. Data from single subject shown in (a) for 252 Hz +20 dB single signal. Left/right panels are the left/right hemispheres. c, Data from single subject in (a) and (b) for 252-Hz single signals at all loudness levels (–20, –10, 0, +10, and +20 dB) from the 10 right hemisphere channels with the maximum magnetic field deflection. Displayed are the RMS values with peaks corresponding to M100 for each signal. Signals with lower intensity levels (–20 and –10 dB) show delayed latencies and lower peak M100 amplitudes relative to other signals in the series. Latency values: 140 (–20 dB), 137 (–10 dB), 129 (0 dB), 128 (+10 dB), and 128 (+20 dB) msec.

Data Analysis: Dipole Source Estimation

Because (a) the M100 response should not solely be defined by its peak amplitude and latency and (b) the evidence that stimulus frequency may influence the sources of the M100, we conducted an analysis of the estimated position of the M100 sources. Dipole sources were estimated based on data from five subjects who exhibited quantifiable M100 responses for a majority of conditions and from whom we had robust digitized head-shape data. We did not have access to structural magnetic resonance images from our subjects and are consequently unable to perform individual anatomically constrained local-

izations. Single equivalent current dipole estimates with a goodness-of-fit <80% were excluded from subsequent statistical analyses. A simple spherical head model was used to determine the source of the M100 response (x , y , z axes) as well as the dipole angles (θ and ϕ) using 40 channels per hemisphere (20 source and 20 sink) with the greatest magnetic field deflection for a single moving dipole. Given their differences in spectral composition, we performed a single equivalent current dipole analysis to assess whether the different signals were associated with measurably different source localizations. The dipole modeling, performed on individual

subjects head shapes using a spherical model, implemented in the MEG160 platform, yielded x , y , and z coordinates of the dipole, angles ϕ and θ , as well as dipole moment in nAm. Statistical significance of the values of x , y , z , θ , and ϕ were assessed using a mixed effects ANOVA and Wilcoxon sign tests in R using the “languageR” statistical package (R Foundation for Statistical Computing, v 2.8.1; Baayen 2008). Wilcoxon tests were performed on the values of θ and ϕ both across and within subjects to assess any potential differences in the source orientation of the M100 response. The nonparametric tests were performed on single-frequency data values versus complex-frequency data values, e.g., 127 Hz versus 127/1000 Hz and 1000 Hz versus 127/1000 Hz. We employed the more conservative nonparametric tests because we could only use the data for 5 of 13 subjects for dipole localization. Because of such a small sample size out of the original subject pool, we thought it was more correct to assume that the data were not parametric.

RESULTS

The stimuli typically elicited a robust M100 response, although the quality of the M100 varied as a function of signal type. Figure 2 illustrates a typical evoked field distribution recorded from 157 channels for one subject presented with one complex signal. The displayed (averaged and filtered) data show the prominent magnetic field deflection in both right and LH channels. The peak (and trough) responses demonstrate the canonical distribution associated with the sensor configuration used in the current experiment (Fig. 2a). The spatial extent of the response, and the putative temporal lobe origin, is well captured by the absence of evoked activity along the midline channels. Calculating the topographic distribution of the response at the peak (Fig. 2b) provides evidence that the M100 recorded here yields the standard spatial response pattern observed for most auditory signals.

The butterfly plot depicted in Figure 3a shows the response to a 252 Hz signal at the +20 dB presentation level for one subject. Both the individual channel responses and the aggregate response across channels (RMS, red) point to the large peak slightly after 100 msec after stimulus onset. The quantitative analysis across experimental conditions that we perform here is based on a more selective group of channels. Figure 3b illustrates the dependent variable. The left and right panels show channels from the LH and RH, respectively. As described above, we selected five channels from the source and sink of each hemisphere (black lines) and calculated the RMS response (red) across the 10 selected channels in each hemisphere. The subsequent analyses here focus on the RMS peak derived from this visualization of the data.

Figure 3c shows the RMS for a 252-Hz signal presented at the five different loudness levels. The RMS curves show clearly the relationship between the peak response and loudness; in particular, for this subject and these data, one observes a clear grouping, with the two lowest loudness values being associated with significantly lower and later M100 peaks, replicating previous findings on loudness effects on the M100 (Stufflebeam et al. 1998).

Figure 4 summarizes the mean latency and peak M100 values across subjects for the 127-, 252-, 800-, and 1000-Hz signals at the equal loudness (0 dB) level. The 800- and

1000-Hz signals demonstrate decreased M100 latencies and increased peak M100 values relative to the 127- and 252-Hz signals. Although the peak RMS values across hemispheres show no significant difference (Fig. 4a), the RH shows significantly decreased latencies relative to the LH for all frequencies, replicating previous data.

Figure 5 summarizes the mean peak RMS and latencies for simple low-frequency and complex signals. Relative to simple signals, the complex signals show a significant increase in the peak RMS at all loudness levels for both sets of complex signals (127/1000 Hz and 252/800 Hz; Figs. 5c, d). Complex signals also show decreased latencies relative to simple signals; however, for the loudest simple signals and complex signals, the latency values are approximately the same (Figs. 5a, b). This decrease in M100 latency is contrary to the intuitive prediction that complexity would yield slower latency values because of increased cortical processing as a result of allocating more neuronal resources to parse the incoming signal. However, it has been shown in macaque auditory cortex that signal complexity decreases the latency of physiological responses (Lakatos et al. 2005). In addition, the time range of the M100 latency across loudness levels is not as great for the complex signals as opposed to the simple signals (~ 130 to 180 msec for simple signals, ~ 120 to 140 msec for complex signals). As with simple signals, the latency of the M100 is faster in the RH than in the LH. There were two significant main effects regarding M100 peak latency for the ANOVA that replicate findings from previous studies, namely those of hemisphere ($F[1,12] = 28.397$) and loudness level ($F[2.125,25.504] = 20.246$). The data show that the RH yields faster M100 peak latency values than the LH and that as the level of the signal increases, the M100 peak latency decreases (see Figs. 5a, b).

The core result of the ANOVA is the finding that complex signals yield faster M100 peak latencies than low-frequency simple signals ($F[3,36] = 23.730$). Across all loudness levels, the complex signals generated faster mean M100 values, although not all the mean latency values were significant (see below). There were two interaction effects: a two-way interaction effect of signal type \times loudness level ($F[12,144] = 4.726$) and a three-way effect of hemisphere \times signal type \times loudness level ($F[6.282,75.387] = 3.088$). The complex signals generated faster M100 latencies at low signal levels (-20 to 0 dB) but not for the higher signal levels. The three-way interaction indicates that the two-way interaction was confined to the RH.

In addition, the planned comparisons involving paired t -tests on the M100 peak latency for simple low-frequency signals versus complex signals (within hemisphere) indicated that on average, the simple signals, except at the highest loudness levels, generated slower M100 peak latency values relative to the complex signals. All values for the paired comparisons and their significance were greater than $t(12) = 3.46$ and $r = 0.707$ for the RH and $t(12) = 2.63$ and $r = 0.604$ for the LH.

Figure 6 summarizes the across subject group response in terms of the peak and latency of the RMS of individual subject RMS vectors. We observe that the pattern for the maximum RMS value and its latency hold for both the simple and complex tones, for all combinations explored. Crucially, the loudest and fastest latency simple signals (both low and high frequencies) peak at ~ 130 msec, except for the 800- and 1000-Hz signals in the RH, whereas the M100 response peaks

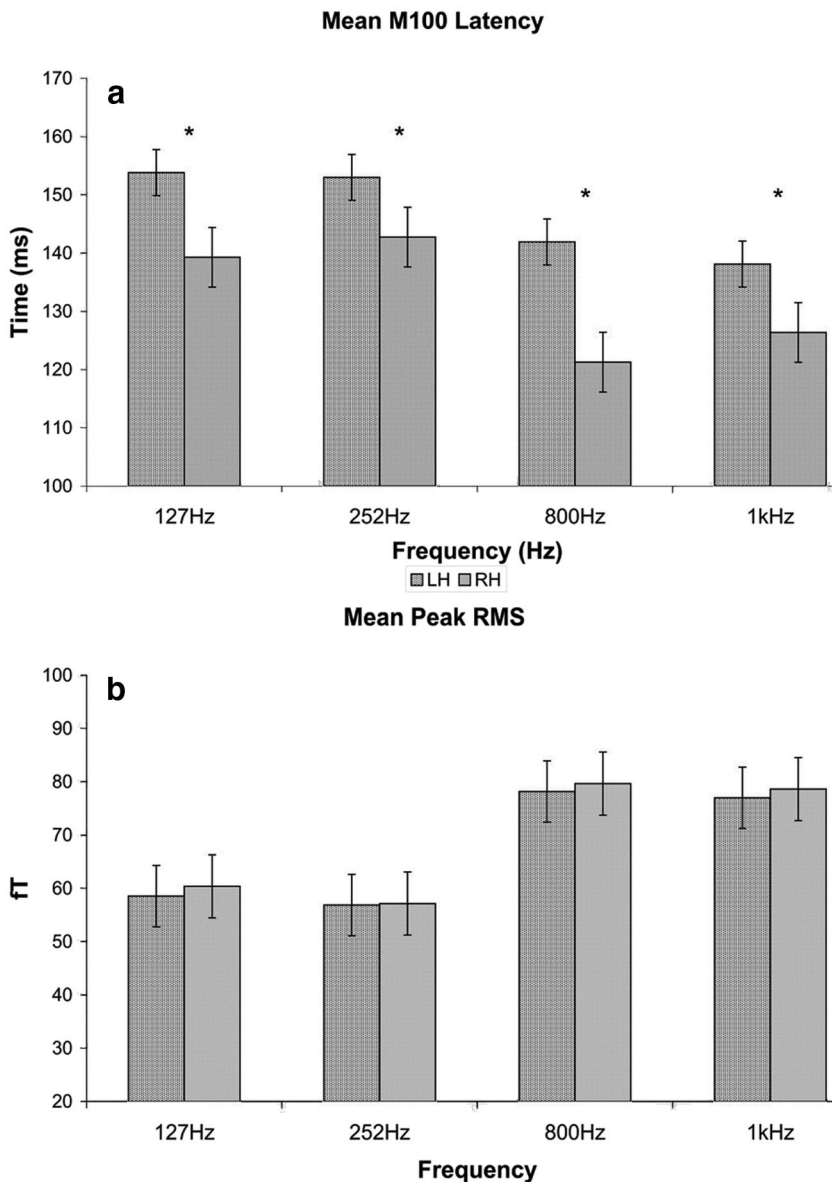


Fig. 4. a, Mean M100 latency across subjects for single signals at matched loudness levels (matched sensation level). Latencies for the right hemisphere are faster than those for the left hemisphere. The latency pattern replicates a number of previous reports that show a striking latency-frequency dependence independent of sensation level (see Roberts et al. 2000 for review). b, Mean peak root mean square (RMS) values for M100 across subjects for single signals at matched loudness levels. Peak RMS values are approximately the same in each hemisphere (i.e., effect size is about the same).

at ~120 msec. As with the mean peak latency ANOVA, there were several main significant results regarding the mean peak RMS value, those of signal type ($F[3,36] = 35.086$), loudness level ($F[4,48] = 34.644$), and the two-way interaction of signal type \times loudness ($F[12,144] = 6.729$). No paired comparisons between simple and complex signals across hemispheres were performed on the peak RMS value because there was significant overlap upon graphical inspection and that there was no significant difference between hemispheres resulting from the ANOVA analysis (see Figs. 5c, d).

The experiment shows that as signal level increased, the M100 peak RMS increased and the latency decreased. For 127- and 252-Hz single signals, the range of the M100 latency was about the same (~130 to 180 msec), with no obvious lateralization. The effect size for the M100 amplitude is the same for both the LH and RH for the 127- and 252-Hz signals. For both single and complex signals, for both hemispheres, the M100 latency seems to converge on the same approximate

value (~125 msec), indicating that the latency reaches a possible point of saturation.

Mean M100 latency (~180–130 msec) across subjects for 127- and 252-Hz single signals at all loudness levels, for both hemispheres, was approximately the same. As the signal amplitude increased, the M100 latency decreased and peak RMS value increased.

Additional ANOVAs were performed on the data without excluding any data. For the ANOVA performed on the latency of the M100, significant effects were found for hemisphere ($F[1,12] = 16.754$), signal type ($F[3,36] = 10.427$), and loudness level ($F[1.951, 23.407] = 5.527$). No significant two-way and three-way effects were found. For the ANOVA performed on the peak RMS of the M100, significant effects were found for signal type ($F[3,36] = 48.221$), loudness level ($F[2.194, 26.331] = 33.621$), and signal type \times loudness level ($F[12,144] = 10.651$). No other significant interactions were found. Analyses with and

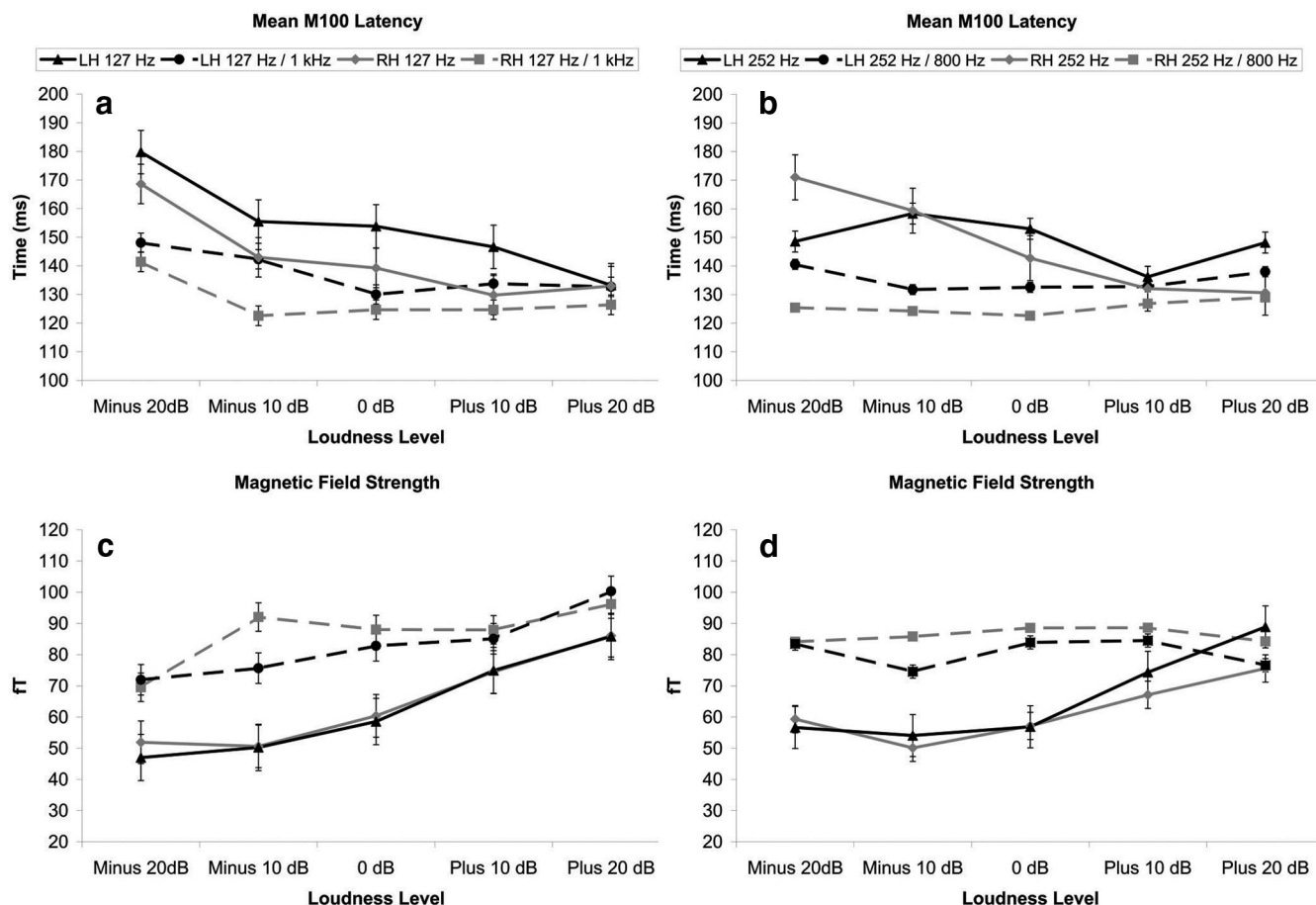


Fig. 5. a, 127-Hz single signal and 127/1000 Hz complex signal M100 latencies across subjects, for both hemispheres, plotted separately. Simple signals are indicated by the solid bold lines and complex signals by the dashed lines, the RH values are indicated by gray color and the LH by black color. Data are the subject means at each loudness level for single signals and mixture levels for complex signals. The means and SEs of the means presented exclude subject data that did not generate the M100 response. As loudness level/mixture level (signal intensity and/or power) increases (see Materials and Methods section for details), the M100 latency decreases. Data for single signals (solid lines) replicate findings from Stufflebeam et al. 1998 (latency decreases with increasing signal intensity and/or loudness). Complex signals (dashed lines) also show decreased latency relative to single signals. This pattern is contrary to the possible prediction that complex signals yield longer latencies. b, The same data for the 252-Hz single signals and 252/800 Hz complex signals. c, Mean peak RMS values across subjects for 127-Hz single signals (solid lines) and 127/1000 Hz complex signals (dashed lines). Relative to single signals, complex signals generate an increase in the M100 amplitude across all mixture levels. The effect size is the same in both the left hemisphere and right hemisphere. d, Analogous to the data in (c) for the 252-Hz single signals and 252/800 Hz complex signals.

without replacement both found significant effects of hemisphere and loudness level for M100 latency and signal type, loudness level, and type \times level for the RMS.

In no spatial dimension was there a systematic effect of stimulus on the localization. This is consistent with (i) the resolution of the method and model employed, and (ii) the fact that even expected spatial differences (e.g., between 250 and 1000 Hz) are not at all reliable (see Lütkenhöner 2003 for extensive data and discussion). The mixed-effects ANOVA performed on the dipole source data found only a significant interaction of hemisphere ($F[1,152] = 5112$), not surprising because of the independence of hemispheric responses. No other significant responses were found. The across-subject and within-subject Wilcoxon sign tests did not find any statistical differences in the orientation of the M100 source.

DISCUSSION

The major auditory-evoked component, the M100/N100m, exhibits variation in both the magnetic field amplitude and the

peak latency as a function of signal frequency and level. In this study, we expand on the dominant paradigm to investigate to what extent the relative level of components and complexity of a signal affect the magnetic field deflection and peak latency of the M100. The signals for the current experiment are an important intermediate level of spectral complexity between simple sinusoids and more ecologically valid signals such as vowels or musical tones. Typical MEG studies focusing on the M100 component use either simple sinusoidal signals or more complex speech tokens (e.g., synthetic vowels with three sinusoidal or approximately sinusoidal components) but do not examine how the individual simple components of a complex signal contribute to the overall M100 response of the complex signal. The data for the simple signals replicate previous findings, namely that as the signal level increases, the field deflection increases and the peak latency decreases and that the RH exhibits faster latencies than the LH (Stufflebeam et al. 1998; Lütkenhöner & Klein 2007; Howard & Poeppel 2009). The analysis was performed in sensor space, not source space,

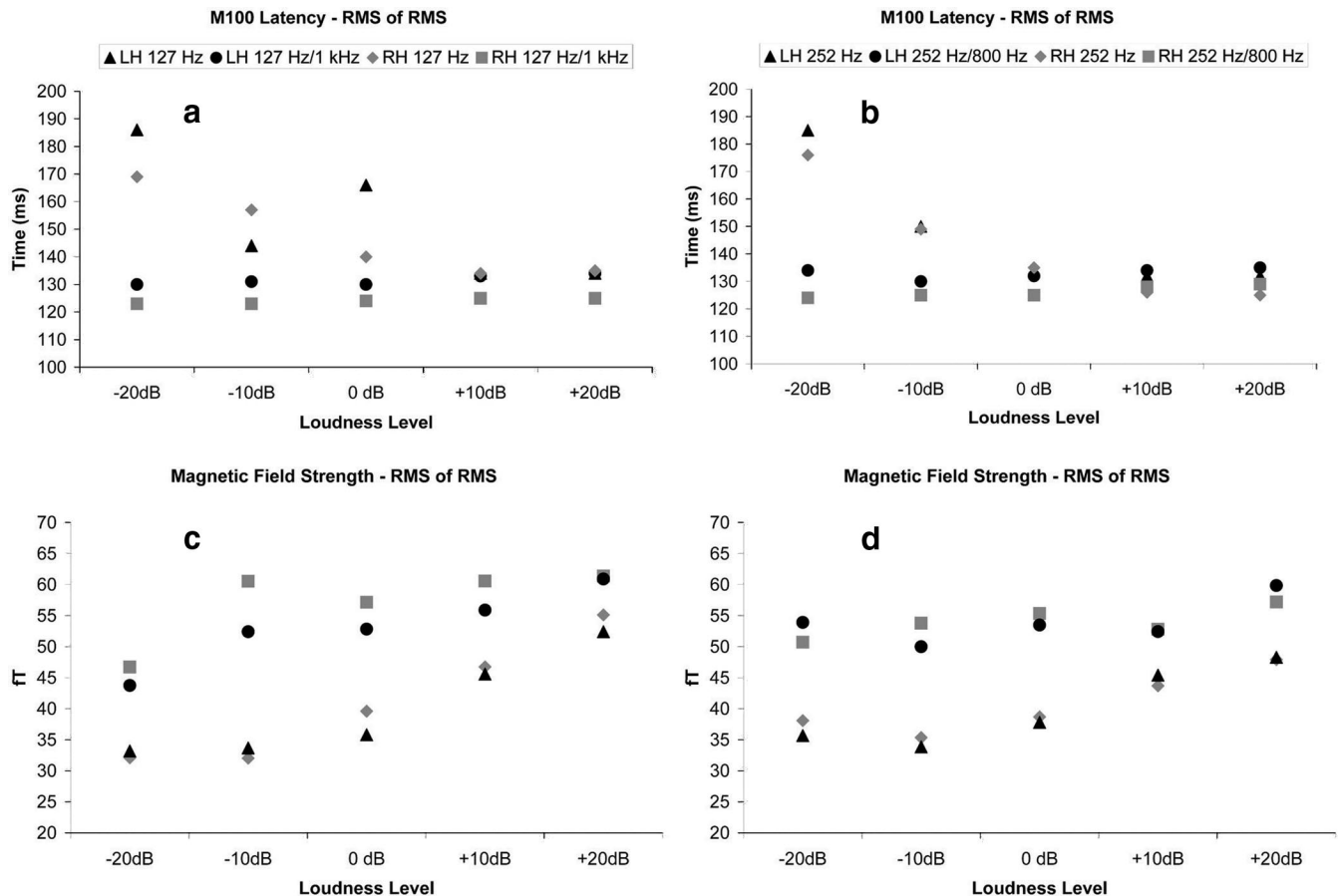


Fig. 6. a, 127-Hz single signal and 127/1000 Hz complex signal M100 latencies across subjects, for both hemispheres, plotted separately. Conventions used are the same as in Figure 5. Data are the RMS of subject RMS vectors collected for subjects demonstrating a quantifiable M100 for a specific condition. Across-subject RMS responses show the same response pattern as the mean values for the peak and latency of the M100. b–d, The same data as their counterparts in Figure 5, except that mean values are replaced with across-subject RMS values.

to stay as close as possible to the recorded data without making source configuration assumptions. However, the sources of the auditory M100 are well understood (Lütkenhöner & Steinsträter 1998), and hypotheses about localization are not at issue in this experiment. Although we did not obtain anatomic magnetic resonance images for subjects and can therefore not perform high-resolution anatomically constrained source localization on individual subjects, the contour maps for both simple and complex signals are indistinguishable from the patterns in the prior literature investigating the cortical sources underlying the M100. The cortical sources lie in the superior temporal lobe and are likely to include multiple distinct generators. The data we report here are highly likely to be associated with the same cortical generators. We were able to verify that there were no signal-driven source differences when using single equivalent current dipole modeling with a spherical head model derived from individual subject head shapes.

The core finding of this study is that complex signals demonstrate increased amplitude but decreased peak latency values. The finding that the magnetic field amplitude increases with complexity is not surprising: an increase in the amount of signal across channels would drive more neurons, and an increase in activity seems a reasonable hypothesis. What was unexpected is that complexity does not yield longer, slower latencies in cortical populations as a consequence of the

integration across critical bands. This could be for several reasons. First, it could be that the signals are parsed and integrated very early in the ascending auditory pathway and the results of this early analysis are reflected in the response of cortical neurons. The signals used contained components well separated in frequency (in different critical bands, no harmonic or octave relationships), and from a physiological and perceptual standpoint, the signals “appear” to be two separate sinusoids played simultaneously. Such a separation could theoretically aid in the analysis of the signals starting from the cochlea and proceeding to cortical neurons.

The data also suggest that certain frequencies may have a privileged analysis by the auditory system (i.e., the excitation thresholds for these frequencies are relatively low). This is evidenced by the finding that even at perceptually matched loudness levels, the 800- and 1000-Hz signals elicit greater response amplitudes and shorter M100 peak latencies than the 127- and 252-Hz signals. The preferential treatment (lower thresholds) given to these frequencies may contribute to the decreased M100 peak latencies of the complex signals used in the experiment. This view is supported by the across-subject data presented in Figure 7. Although the data presented in Figure 7 suggests that the latency of the complex signals is wholly determined by the higher frequency component, the results of the ANOVAs for latency found statistically signifi-

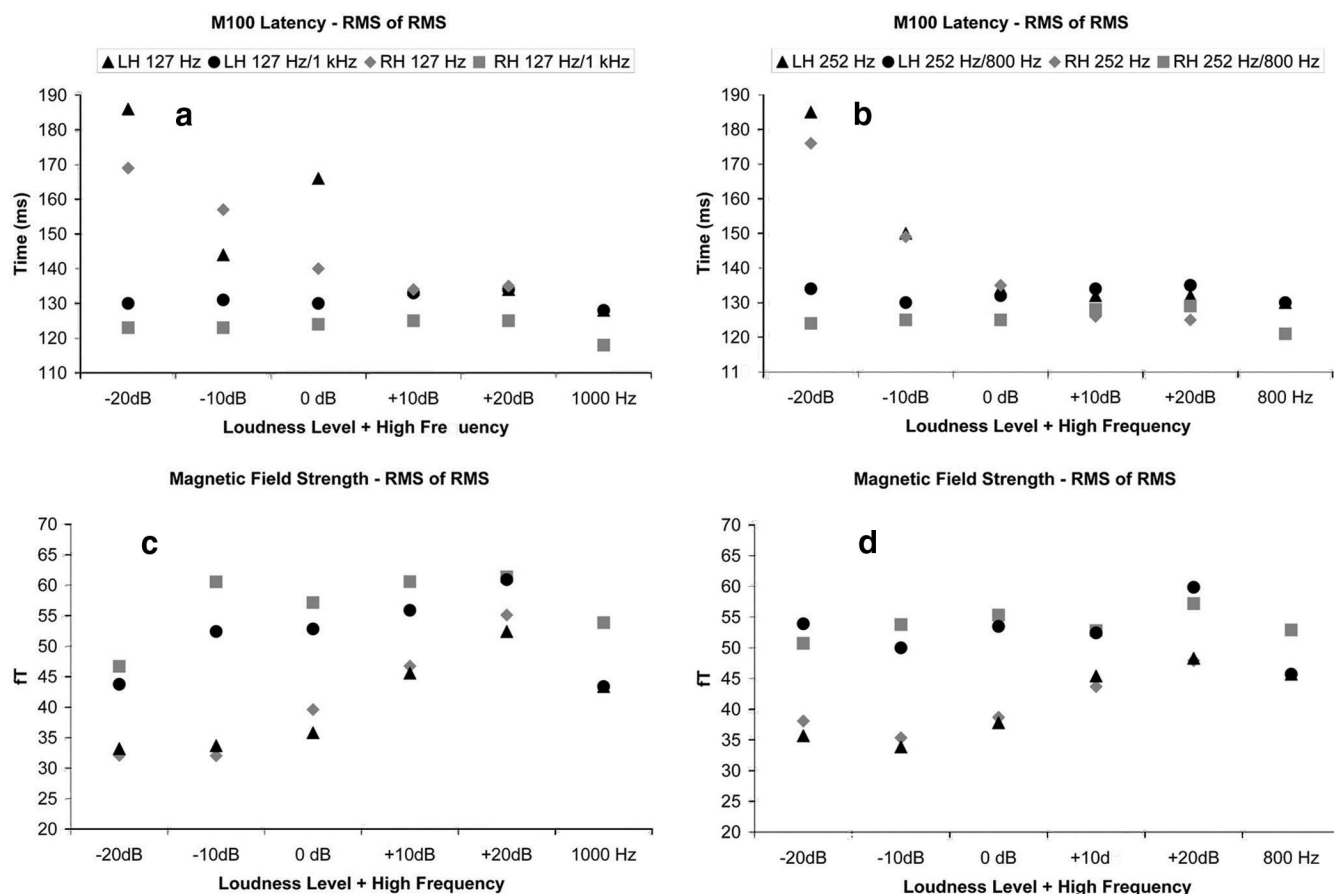


Fig. 7. a, 127/1000 Hz complex signal and 1000-Hz single signal M100 latency values for both hemispheres. Conventions used are the same as in Figure 5. Latency values are the RMS of subject RMS vectors collected for subjects demonstrating a quantifiable M100 for a specific condition. The latencies of the complex signals are closest to that of the 1000-Hz signal. b, The same data for the 252/800 Hz complex signals and 800-Hz single signal. c, Peak RMS values of across-subject RMS vectors for 127/1000 Hz and 1000-Hz single signals for both hemispheres. d, The same data for the 252/800 Hz complex signals and 800-Hz single signal.

cant effects for signal type and loudness level, a finding that cannot be completely discounted. Even when both the power and the perception of the complex signals was dominated by the lower frequency component, the latency of the complex signal is closer to that of the higher frequency component. This creates an interesting dissociation between perception and the physiological response (see below).

Another possible explanation for the data is related to the integration window of the M100 and the amount of sound energy located in that window. The physical structure of the signal and any modulations to the signal within the first 30 to 40 msec after onset can have profound effects on the latency, amplitude, and shape of the M100 (Gage et al. 2000). The complex signals have relatively more energy in this time period than the simple signals. The latency and amplitude facilitation may be analogous to that seen with pulse data (Sams et al. 1993).

An alternative scenario is that because the complex signals used here more closely resemble ecologically valid signals (e.g., vowels), cortical neurons are more “tuned” to them (Jacobsen et al. 2004), and this facilitates more robust and faster processing. It should be noted that the signals used contained only two frequency components as opposed to the three normally used with vowel tokens. Perceptually, this is a significant difference. Once three or more sinusoidal or ap-

proximately sinusoidal components are used, different perceptual effects can occur, based on the frequency grouping of the components as well as presentation of the signals (e.g., modulation and harmonic structure). As a kind of “halfway point” between simple sinusoids and the complex signals encountered normally in the environment, the signals shed light on the interactive effects of the processing reflected in the M100 response. The signals used also give evidence of the interactive effects associated with the processing (hemisphere, signal type, level, etc.), which would be harder to disentangle if the same method in this experiment was applied to more complex signals. However, this processing has a natural limit as evidenced by the apparent saturation of peak latency values (see Figs. 5a, b) whereby the values for both simple and complex signals converge onto the same approximate value.

Previous data suggest that different signal frequencies have different M100 source orientations, with low-frequency signals having the steepest angles in the sagittal plane and high frequencies the shallowest angles (Tiitinen et al. 1993). The results of the mixed-effects ANOVA and the Wilcoxon tests on the phase angles of the responses did not replicate this finding. This could be due to an inability to resolve these components because of the spatial limitations of MEG (Lütkenhöner & Steinsträter 1998; Lütkenhöner 2003). However, previous stud-

ies have found that signal level and context may affect the localization of the source (Okamoto et al. 2009; Tanji et al. 2010). The primary purpose of this experiment, however, was not a focus on source localization but rather quantifying direct, surface measurements. Furthermore, even though we were not concerned with differences in hemispheric processing for the M100 for this particular study, it is vital to consider the data from each hemisphere separately for several reasons. First, the M100 literature shows that the RH exhibits faster response latencies than the LH as well as a greater power at the peak. Second, there is evidence that the hemispheres evaluate different aspects of the auditory signal. Finally, the cytoarchitecture and the cortical folding between hemispheres differ. Thus, it would be safer to evaluate each hemisphere as a separate entity for statistical purposes.

One finding of particular interest is that of the dissociation of perception from the physiological response that we are quantifying. It was previously proposed that the M100 reflects perceptual attributes of the signal rather directly (Stufflebeam et al. 1998). In this experiment, the lower frequency components at their loudest levels (+10 and +20dB) completely dominated the perception of the complex signals. However, the latency and the peak RMS values were not closer to the values for the simple signals at those levels presented alone. The data suggest that the loudness of a signal is not entirely reflected in the M100 in the case of complex signals or at least that there is another mechanism that assists in assessing the loudness of signal and its individual components. At the very least, loudness cannot be “read out”—from a neural coding point of view—by decoding properties of the M100. This is consistent, of course, with stable loudness estimates developing over a period of about 200 msec (cf. Moore 2004).

Further studies into this topic should focus on the different grouping of signals, e.g., complex signals composed of 400- and 720-Hz components, or even components where frequencies are in the same critical band (Soeta & Nakagawa 2006). The facilitation of a complex signal and the proximity of its latency to that of a particular component may depend on the constituent frequencies, the position of those frequencies relative to hearing bandwidth, and their distance (logarithmic, linear, and log-linear) from each other.

ACKNOWLEDGMENTS

The authors thank Mary Howard for detailed comments on the article and Jeff Walker for excellent technical assistance.

This work was supported by NIH grant 2R01DC05660 (to D.P.) and training grant DC-00046 from the National Institute of Deafness and Communicative Disorders of the National Institutes of Health (to J.J. III).

Address for correspondence: Julian Jenkins III, M.S., Department of Biology, University of Maryland, 3263 Biology-Psychology Building, College Park, MD 20742. E-mail: julianj@umd.edu.

Received September 3, 2009; accepted February 1, 2010.

REFERENCES

Baayen, R. H. (2008). *languageR: Data Sets and Functions with “Analyzing Linguistic Data: A Practical Introduction to Statistics.”* R package version 0.953.

Brattico, E., Tervaniemi, T., Picton, T. W. (2003). Effects of brief discrimination-training on the auditory N1 wave. *Neuroreport*, 14, 2489–2492.

Carhart, R. & Jerger, J. F. (1959). Preferred method for clinical determination of pure-tone thresholds. *J Speech Hear Dis*, 24, 330–345.

Chait, M., Poeppel, D., de Cheveigné, A., et al. (2007). Processing asymmetry of transitions between order and disorder in human auditory cortex. *J Neurosci*, 27, 5207–5214.

Chait, M., Poeppel, D., Simon, J.Z. (2006). Neural response correlates of detection of monaurally and binaurally created pitches in humans. *Cereb Cortex*, 16, 835–848.

de Cheveigné, A. & Simon, J. Z. (2007). Denoising based on time-shift PCA. *J Neurosci Methods*, 165, 297–305.

Diesch, E., Eulitz, C., Hampson, S., et al. (1996). The neurotopography of vowels as mirrored by evoked magnetic field measurements. *Brain Lang*, 53, 143–168.

Diesch, E. & Luce, T. (1997). Magnetic fields elicited by tones and vowel formants reveal tonotopy and nonlinear summation of cortical activation. *Psychophysiology*, 34, 501–510.

Diesch, E. & Luce, T. (2000). Topographic and temporal indices of vowel spectral envelope extraction in the human auditory cortex. *J Cogn Neurosci*, 12, 878–893.

Dietrich, V., Nieschalk, M., Stoll, W. et al. (2001). Cortical reorganization in patients with high frequency cochlear hearing loss. *Hear Res*, 158, 95–101.

Eggermont, J. J. (1995). Evoked potentials as indicators of auditory development. *Int J Pediatr Otorhinolaryngol*, 32(suppl), S183–S186.

Eulitz, C., Diesch, E., Pantev, C., et al. (1995). Magnetic and electric brain activity evoked by the processing of tone and vowel stimuli. *J Neurosci*, 15, 2748–2755.

Fant, G. (1960). *Acoustic Theory of Speech Production*. The Hague, The Netherlands: Mouton.

Fujioka, T., Ross, B., Okamoto, H., et al. (2003). Tonotopic representation of missing fundamental complex sounds in the human auditory cortex. *Eur J Neurosci*, 18, 432–440.

Gage, N. M. & Roberts, T. P. L. (2000). Temporal integration: Reflections in the M100 of the auditory evoked field. *Neuroreport*, 11, 2723–2726.

Hari, R., Aittoniemi, K., Järvinen, M. L., et al. (1980). Auditory evoked transient and sustained magnetic fields of the human brain. Localization of neural generators. *Exp Brain Res*, 40, 237–240.

Hari, R., Kaila, K., Katila, T., et al. (1982). Interstimulus interval dependence of the auditory vertex response and its magnetic counterpart: Implications for their neural generation. *Electroencephalogr Clin Neurophysiol*, 71, 64–72.

Hari, R. & Mäkelä, J.P. (1988). Modification of neuromagnetic responses of the human auditory cortex by masking sounds. *Exp Brain Res*, 71, 87–92.

Hirata, Y., Kuriki, S., Pantev, C. (1999). Musicians with absolute pitch show distinct neural activities in the auditory cortex. *Neuroreport*, 10, 999–1002.

Howard, M. & Poeppel, D. (2009). Hemispheric asymmetry in mid and long latency responses to single clicks. *Hear Res*, 257, 41–52.

Jacobsen, T., Schröger, E., Alter, K. (2004). Pre-attentive perception of vowel phonemes from variable speech stimuli. *Psychophysiology*, 41, 654–659.

Kirverskari, E., Salmelin, R., Hari, R. (2006). Neuromagnetic responses to vowels vs. tones reveal hemispheric lateralization. *Clin Neurophysiol*, 117, 643–648.

Lakatos, P., Pincze, Z., Fu, K.-M., et al. (2005). Timing of pure tone and noise-evoked responses in macaque auditory cortex. *Neuroreport*, 16, 933–937.

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

Lütkenhöner, B. (2003). Single-dipole analyses of the N100m Are not suitable for characterizing the cortical representation of pitch. *Audiol Neurotol*, 8, 222–233.

Lütkenhöner, B. & Klein, J.-S. (2007). Auditory evoked field at threshold. *Hear Res*, 228, 188–200.

Lütkenhöner, B., Lammertmann, C., Knecht, S. (2001). Latency of auditory evoked field deflection n100m ruled by pitch or spectrum? *Audiol Neurotol*, 6, 263–278.

Lütkenhöner, B. & Poeppel, D. (in press). From tones to speech. In Schreiner, C. & J. Winer (Eds.), *Auditory Cortex*. New York: Springer.

Lütkenhöner, B., Seither-Preisler, A., Seither, S. (2006). Piano tones evoke stronger magnetic fields than pure tones or noise, both in musicians and non-musicians. *Neuroimage*, 30, 927–937.

- Lütkenhöner, B. & Steinsträter, O. (1998). High-precision neuromagnetic study of the functional organization of the human auditory cortex. *Audiol Neurotol*, 3, 191–213.
- Mäkinen, V., May, P., Tiitinen, H. (2004). Transient brain responses predict the temporal dynamics of sound detection in humans. *Neuroimage*, 21, 701–706.
- MATLAB [computer program] (2005). Version 7, SP14. Natick, MA: The Mathworks.
- Monahan, P. J., de Souza, K., Idsardi, W. J. (2008). Neuromagnetic evidence for early auditory restoration of fundamental pitch. *PLoS ONE*, 3, e2900.
- Moore, B. C. J. (2004). The Perception of Loudness. In *An Introduction to the Psychology of Hearing* (pp. 127–162). San Diego, CA: Elsevier, Academic Press.
- Nikjeh, D. D., Lister, J. J., Frischh, S. A. (2009). Preattentive cortical-evoked responses to pure tones, harmonic tones, and speech: Influence of music training. *Ear Hear*, 30, 432–446.
- Obleser, J., Lahiri, A., Eulitz, C. (2003). Auditory-evoked magnetic field codes place of articulation in timing and topography around 100 milliseconds post syllable onset. *Neuroimage*, 20, 1839–1847.
- Obleser, J., Lahiri, A., Eulitz, C. (2004a). Magnetic brain response mirrors extraction of phonological features from spoken vowels. *J Cogn Neurosci*, 16, 31–39.
- Obleser, J., Rockstroh, B., Eulitz, C. (2004b). Gender differences in hemispheric asymmetry of syllable processing: Left-lateralized magnetic N100 varies with syllable categorization in females. *Psychophysiology*, 41, 783–788.
- Okamoto, H., Stracke, H., Zwisterlood, P., et al. (2009). Frequency-specific modulation of population-level frequency tuning in human auditory cortex. *BMC Neurosci*, 10, 1.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Pantev, C., Hoke, M., Lütkenhöner, B., et al. (1989). Tonotopic organization of the auditory cortex: Pitch versus frequency representation. *Science*, 246, 486–488.
- Poeppel, D., Phillips, C., Yellin, E., et al. (1997). Processing of vowels in supratemporal auditory cortex. *Neurosci Lett*, 221, 145–148.
- R [computer program]. Version 2.8.1. Vienna, Austria: R Foundation for Statistical Computing; 2008.
- Roberts, T. P. L., Ferrari, P., Poeppel, D. (1998). Latency of evoked neuromagnetic M100 reflects perceptual and acoustic stimulus attributes. *Neuroreport*, 9, 3265–3269.
- Roberts, T. P. L., Ferrari, P., Stufflebeam, S. M., et al. (2000). Latency of the auditory evoked neuromagnetic field components: Stimulus dependence and insights toward perception. *J Clin Neurophysiol*, 17, 114–129.
- Salajegheh, A., Link, A., Elster, C., et al. (2004). Systemic latency variation of the auditory evoked M100: From average to single-trial data. *Neuroimage*, 23, 288–295.
- Sams, M., Hämäläinen, M., Hari, R., et al. (1993). Human auditory cortical mechanisms of sound lateralization: I. Interaural time differences within sound. *Hear Res*, 67, 89–97.
- Seither-Preisler, A., Krumbholz, K., Lütkenhöner, B. (2003). Sensitivity of the neuromagnetic N100m deflection to spectral bandwidth: A function of the auditory periphery? *Audiol Neurotol*, 8, 322–337.
- Soeta, Y., Nakagawa, S., 2006. Complex tone processing and critical band in the human auditory cortex. *Hear Res*, 222, 125–132.
- SPSS [computer program]. Version 16.0. Chicago, IL: SPSS Inc.; 2007.
- Stufflebeam, S. M., Poeppel, D., Rowley, H. A., et al. (1998). Perithreshold encoding of stimulus frequency and intensity in the M100 latency. *Neuroreport*, 9, 91–94.
- Tanji, K., Leopold, D. A., Ye, F. Q., et al. (2010). Effect of sound intensity on tonotopic fMRI maps in the unanesthetized monkey. *Neuroimage*, 49, 150–157.
- Tiitinen, H., Alho, K., Huotilainen, M., et al. (1993). Tonotopic auditory cortex and the magnetoencephalographic (MEG) equivalent of the mismatch negativity. *Psychophysiology*, 30, 537–540.
- Tiitinen, H., Mäkelä, A.M., Mäkinen, V., et al. (2005). Disentangling the effects of phonation and articulation: Hemispheric asymmetries in the auditory N1m response of the human brain. *BMC Neurosci*, 6, 62.
- Tucker, D. A., Dietrich, S., McPherson, D. L., et al. (2001). Effect of stimulus intensity level on auditory middle latency response brain maps in human adults. *J Am Acad Audiol*, 12, 223–232.
- Vihla, M., Lounasmaa, O.V., Salmelin, R. (2000). Cortical processing of change detection: Dissociation between natural vowels and two-frequency complex tones. *Proc Natl Acad Sci USA*, 97, 10590–10594.
- Vihla, M. & Salmelin, R. (2003). Hemispheric balance in processing attended and non-attended vowels and complex tones. *Brain Res Cogn Brain Res*, 16, 167–173.