ELSEVIER

Clinical Neurophysiology 115 (2004) 1885–1895

www.elsevier.com/locate/clinph

Tempo dependence of middle- and long-latency auditory responses: power and phase modulation of the EEG at multiple time-scales

Joel S. Snyder[a,*], Edward W. Large[b]

[a]*Department of Psychology, Cornell University, Ithaca, NY, USA*
[b]*Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, FL, USA*

Accepted 15 March 2004

Available online 10 May 2004

Abstract

Objective: We measured the influences of power and phase modulations of neuroelectric activity on auditory responses to pure-tone patterns with inter-onset intervals typical of music.

Methods: Tones were presented to 8 subjects at 10 different tempos from 150 to 3125 ms and with random intervals. We quantified time-frequency (TF) power with respect to a pre-tone-onset baseline and the TF phase coherence across trials. Peak-to-peak event-related potential (ERP) amplitude values for the middle and long-latency auditory responses were obtained for comparison.

Results: ERP amplitude, size of power modulation, and amount of phase coherence were larger at slower tempos for the long-latency response (LLR) but not for the middle-latency response (MLR). Multiple regression analysis indicated that for MLR and LLR, phase modulation was a better predictor of ERP amplitude than power modulation.

Conclusions: Phase modulation is a better predictor of ERP amplitude than power modulation for middle and long-latency auditory responses.

Significance: Lack of diminution of the MLR at fast tempos indicates its usefulness for studying early cortical processing of music and speech patterns.

Keywords: Power and phase modulation; Time-frequency representations; Auditory evoked response; Tempo; Music and speech processing

---

1. Introduction

Neurophysiological methods such as EEG and MEG are the most practical techniques for directly observing the temporal dynamics of human brain activity. Such techniques appear ideal for resolving activations of multiple neuronal populations during the course of perceptual, motor, and cognitive processing. The auditory evoked response has been particularly well described, with components measured at the scalp that arise from brainstem, mid-brain, and auditory cortical regions during the first 300 ms following tone onset (Näätänen and Winkler, 1999; Picton et al., 1974). Functional imaging studies show that auditory temporal patterns such as music and speech activate auditory cortical regions that play a role in auditory evoked response generation (Zatorre et al., 2002). Despite the excellent temporal resolution afforded by EEG and MEG, it is currently unclear how to best study early cortical processing of patterns such as musical sequences consisting of changing pitches and inter-onset intervals (IOI). However, the known influence of tempo (inversely related to IOI) on auditory responses suggests that early auditory cortical responses may offer the best opportunity to study such dynamic patterns.

The auditory long-latency response (LLR), occurring after 100 ms post-stimulus, is characterized by a large influence of tempo. Both the neuroelectric and the neuromagnetic N1 and P2 components of the LLR show rapid amplitude diminution as tempo increases (Carver et al., 2002; Hari et al., 1982; Sams et al., 1993). This is a limiting factor for the application of non-invasive neurophysiological measurement to on-line auditory processing of patterns such as music and speech, which unfold at a similar time-scale as the LLR. For example, the IOIs in music range

---

[*] Corresponding author. Address: Rotman Research Institute, Baycrest Centre for Geriatric Care, 3560 Bathurst Street, Toronto, Ont., Canada M6A 2E1. Tel.: +1-416-785-2500x3130; fax: +1-416-785-2862.

E-mail address: jsnyder@rotman-baycrest.on.ca (J.S. Snyder).

1388-2457/$30.00 © 2004 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.clinph.2004.03.024

mostly from 100 to 1200 ms IOI, the same range at which subjects are best able to tap in synchrony with metronomic patterns (Engström et al., 1996; Mates et al., 1994; Peters, 1989) and perceive changes in tempo (Drake and Botte, 1993; Friberg and Sundberg, 1995). This IOI range also marks the transition from transient to steady-state responses (SSR), with the middle of this range corresponding to the initial overlap between responses to adjacent tones (Carver et al., 2002). Investigations of transient responses and the SSR typically use stimulation rates much slower and much faster, respectively. This is due to methodological constraints designed to elicit responses with the largest signal-to-noise ratio, an important consideration. However, the behavioral significance of the 100–1200 ms IOI range suggests a need for its systematic investigation using EEG and MEG.

The middle-latency response (MLR), occurring within 100 ms latency, has a less well-understood tempo dependence than the LLR. This uncertainty is further complicated by the possibility of a partially independent gamma-band response (GBR) that has similar peak latencies. Several studies have shown diminution of the MLR at fast tempos (Erwin and Buchwald, 1986; Makeig, 1990), although the amount of decrease is much less dramatic than with the LLR. Furthermore, this decrease may occur only for the small, late peak (P1) at ~50 ms of the MLR (Carver et al., 2002; Erwin and Buchwald, 1986). There is conflicting evidence for tempo dependence of the GBR (Pantev et al., 1993; Ross et al., 2002). A recent report that provided an extremely fine sampling of the tempo range also found a narrow resonance effect in the MLR around 500 ms IOI (Carver et al., 2002), the center of the tempo perception range (Drake and Botte, 1993; Friberg and Sundberg, 1995). There are no reports for disappearance of the MLR at tempos typical of music and speech unlike with the LLR. The GBR and early peaks of the MLR may therefore provide an important tool for neurophysiological investigation of auditory pattern processing because of the more constant response across different IOIs. However, a more systematic characterization of the tempo dependence of the MLR seems appropriate before attempting such investigations with complex patterns such as music and speech.

An investigation of auditory brain dynamics using time-frequency (TF) methods (e.g. spectrograms) might additionally yield insight to the possible separate existence of the MLR and GBR. The evidence thus far points to similarities between the two responses, and to some important differences. The MLR and GBR contain similar positive and negative peaks within the first 100 ms post-stimulus, and behave similarly to paired-tone presentation (Müller et al., 2001). The MLR is generally obtained using a wide-band filter, while the GBR is obtained using a narrower-band filter such as 24–48 Hz (Pantev et al., 1991, 1993). The MLR displays tonotopic organization (Pantev et al., 1995) whereas tonotopy has not been found for the GBR (Pantev, 1995).

The TF dynamics of power and phase following tone onsets is also likely to help uncover the generating mechanisms for the MLR and LLR. Intracerebral optical imaging has shown that the pre-stimulus spatial patterns on the cortical surface combine linearly with the average evoked response to predict single-trial evoked responses (Arieli et al., 1996). More directly related to scalp recordings, several investigations have shown that the power and phase of pre-stimulus alpha activity (~10 Hz) relates to the size of subsequent event-related potentials (ERPs) in the auditory and visual domains (Brandt et al., 1991; Haig and Gordon, 1998; Jansen and Brandt, 1991; Kruglikov and Schiff, 2003; Makeig et al., 2002). The ease of recording the high-amplitude alpha wave has facilitated these demonstrations. However, it is likely that multiple bands of oscillatory activity underlie the appearance of other evoked responses that occur at faster time-scales such as the MLR.

A related line of research has focused on whether power or phase modulation is most predictive of evoked response size. Early reports claimed that post-stimulus power changed much less from pre-stimulus values than did phase even though power changes do indeed occur (Sayers et al., 1974). More recently, an analysis of the phase evolution showed an increasing uniformity of phase across trials during low-frequency evoked auditory response generation (Jansen et al., 2003). This phase-reorganizing effect mainly occurred in the 2–8 Hz range and predicted the differential response amplitude to paired tones. A more systematic comparison of power and phase modulation at a larger range of frequencies would help determine how much power vs. phase contribute to the MLR and the LLR. To accomplish this, we calculated independent measures of stimulus-related power and phase modulation in a broad-band of frequencies, and performed multiple regression analysis to compare the effects of power and phase while controlling for overlapping contributions.

The TF measures of power modulation and phase coherence that we use can be considered as induced and evoked responses (Pantev, 1995; Tallon-Baudry and Bertrand, 1999), respectively. *Induced* responses indicate changes in the amplitude of EEG activity following a sensory or cognitive event with various latencies. *Evoked* activity, on the other hand, indicates changes in the phase of EEG activity typically with relatively short latencies, suggesting that they bear the most similarity to sensory evoked responses. The particular measures of power modulation and phase coherence we used here are quite pure measures of stimulus-related power and phase modulations of EEG activity. The power modulation measure we use calculates an estimate of the baseline power spectrum on a single trial basis using the pre-stimulus period and then calculates the power in moving windows relative to the baseline, resulting in a high temporal resolution measure of power dynamics regardless of the phase of activity (Makeig, 1993). The phase coherence measure we use provides a measure of the similarity in phase structure of EEG activity across trials, regardless of the amplitude of the activity (Tallon-Baudry et al., 1996). Thus, high phase coherence occurring after stimulus onsets relative to the baseline period indicates a stimulus-related phase modulation of EEG activity. We calculated both power and phase measures with high temporal and frequency resolution, providing a detailed view of auditory stimulus-related power and phase dynamics. Plotting the results in a TF representation (i.e. a spectrogram) enables a visual evaluation of the temporal dynamics of EEG activity at multiple frequency scales, simultaneously.

2. Methods

2.1. Subjects and stimuli

We obtained written, informed consent from 8 young adults (24–37 years, mean 30 years; 5 males, 3 females) for participation in this experiment. Review committees at Cornell University and Florida Atlantic University approved our study protocol. We delivered auditory stimuli using a custom program written using MAX 3.6.2, running on a Macintosh G3 Powerbook. MAX sent musical instrument digital interface (MIDI) signals to an Akai S2000 MIDI synthesizer that converted the MIDI signals into tones. A Harmon/Kardon multi-media speaker system consisting of two speakers approximately 3 ft behind the subject's head played tones at a comfortable listening level. Auditory stimuli consisted of sine-wave tones (262 Hz, 55 dB SPL, 50 ms duration including a 10 ms rise time and no fall time). Subjects listened to isochronous tones in 10 different IOI conditions: 150, 250, 375, 500, 625, 750, 1000, 1250, 1875, and 3125 ms IOI, and a random condition with IOIs sampled from a flat distribution (375–750 ms).

2.2. Data collection

We recorded electrical brain activity using 84 channels from an EEG cap (Electro-Cap International, Inc., Eaton, OH) with ground placed on the right forehead and reference to linked-mastoids. Ground and reference electrode impedance were maintained at $<5\ \mathrm{k}\Omega$, and recording electrodes were $<10\ \mathrm{k}\Omega$. This was accomplished by cleaning the entire scalp with rubbing alcohol and applying conducting gel to each electrode site using a blunt needle before the experiment, and by applying gel between stimulus conditions as needed. EEG signals were amplified by a MicroAmp system (Sam Technology, Inc., San Francisco, CA), digitized at 256 Hz, and filtered with a passband of 0.05–100 Hz, sending the digitized EEG data to a Pentium III Dell Dimension XPS T450 computer running Manscan 4.1. In addition, a custom-designed micro-controller device received MIDI events and triggered stimulus event marking on the computerized EEG record.

Each EEG recording session lasted about 3 h, including preparation time and recording. During recording in a dimly lit room, subjects sat in a comfortable chair and fixated a cross on the wall in front of them. We instructed subjects to avoid eye and body movements during stimulus delivery, but not to try too hard to suppress eye-blinks. On-line monitoring of the EEG indicated that subjects were generally successful in avoiding movements. Each subject participated in all of the tempo conditions listed above in quasi-random order. To avoid boredom during listening, we alternated amongst the 5 fastest and 5 slowest IOI conditions. For each condition, 200 tones were presented in trains of 20 or 50 at a time with a few seconds of rest between trains. For the fastest tempos (150–1000 ms, and random), tones were presented in trains of 50, and for the slowest tempos (1250–3125), tones were presented in trains of 20. Between conditions, subjects had several minutes to rest and converse with the experimenter in order to maintain a high level of alertness.

2.3. Data processing

The EEG data were analyzed off-line using custom MATLAB scripts and functions from the wavelet, signal processing (*MathWorks*, Natick, MA), and EEGLAB toolboxes (Delorme and Makeig, 2002). We first extracted individual trials from the recording sessions in one-tone segments with half an IOI preceding the event marker and half an IOI following the event marker. Artifact rejection threshold was set at 50 μV, after baseline correction. Channels with fewer than half of recorded tone presentations left after artifact rejection were excluded from further analysis (mean number retained, 80.0 channels).

To calculate power modulation and phase coherence, we applied a TF analysis ('timef.m' from the EEGLAB toolbox). Power modulation refers to a mean increase in power (i.e. square of amplitude) of the ongoing EEG across trials whereas phase coherence refers to a similarity in phase of the ongoing EEG across trials. Therefore, if phase coherence increases from pre-stimulus to post-stimulus periods this implies stimulus-related phase modulation of the EEG. The first quantity calculated was post-stimulus increase of power (in dB) relative to a half-IOI pre-stimulus baseline from 0 to 60 Hz in steps of 4 Hz with a 125 ms sliding window and maximal overlap. This was performed for each individual trial and the resulting TF power modulations were then averaged across trials for each experimental condition. This measure of power modulation is a type of induced activity, in the sense that it disregards the phase of oscillatory activity (Makeig, 1993; Tallon-Baudry and Bertrand, 1999).

Phase coherence (in dimensionless units from 0 to 1) was calculated across trials independent of amplitude, and at the same TF resolution as the power modulation analysis (Tallon-Baudry et al., 1996). This indicated for each time-point and frequency how consistent the EEG phase was across trials within a condition for each subject. Thus, increases of this index following tone presentation indicated a stimulus-related increase in the phase consistency (i.e. phase modulation) across tone presentations. Significant power and phase values at the $P = 0.05$ level were determined using a bootstrap analysis with 200 replications (for more details, see 'timef.m' from the EEGLAB toolbox). In all TF representations, only significant power and phase values were displayed and all non-significant values were set to 0.

For each subject, average TF representations of significant power and phase modulation were taken across channels. As shown in Fig. 1, peak power and phase values were taken as the means across TF space for both the MLR (20–60 Hz, 0–75 ms) and LLR (0–20 Hz, 75–200 ms). ERPs were also calculated for the same sets of trials submitted to TF analysis, and averaged across channels. To isolate the MLR and the LLR, the mean ERPs across all tone presentations were filtered (bandpass at 10–60 Hz for the MLR, attenuated by 7.5 dB at 10 Hz and 6.6 dB at 60 Hz; lowpass at 20 Hz for the LLR, attenuated by 5.3 dB at 20 Hz). Peak-to-peak amplitude values were taken for the MLR and the LLR for comparison with power and phase modulation values (see Fig. 1). MLR amplitude was the difference between the minimum and maximum peaks in the latency range of 0–75 ms, while the LLR amplitude was the difference between the minimum and maximum peaks in the latency range of 75–200 ms.

3. Results

3.1. *Effect of tempo on amplitude, power, and phase dynamics of the MLR and LLR*

Fig. 2 shows the time series of averaged neuroelectric activity for one subject at each of the IOI conditions, with tone onset at 0 ms. An early auditory response (MLR) is clearly present across all of the conditions with large positive (Pa) and negative (Nb) peaks, and a small positive peak (Pb) within the first 100 ms post-tone onset. It shows no overall increase at slower tempos (i.e. larger IOIs). Peak amplitude for the random IOI condition is somewhat smaller than peak responses in the isochronous 375–625 conditions, but comparable in size to other IOI conditions. A late auditory response (LLR) is increasingly large at slower tempos, with large negative (N1) and positive (P2) peaks after 100 ms post-tone onset. The response in the random IOI condition is smaller than the slow tempo conditions, but similar in size to conditions with similar mean IOI content (500 ms).

Fig. 3 confirms these findings and separately visualizes stimulus-related TF power modulation (i.e. amplitude changes of EEG at each TF value) and phase coherence (i.e. phase consistency of EEG at each TF value) for the same subject as in Fig. 2. Each row of a TF representation indicates for a given frequency the amount of power modulation or phase coherence at each point in time. The top

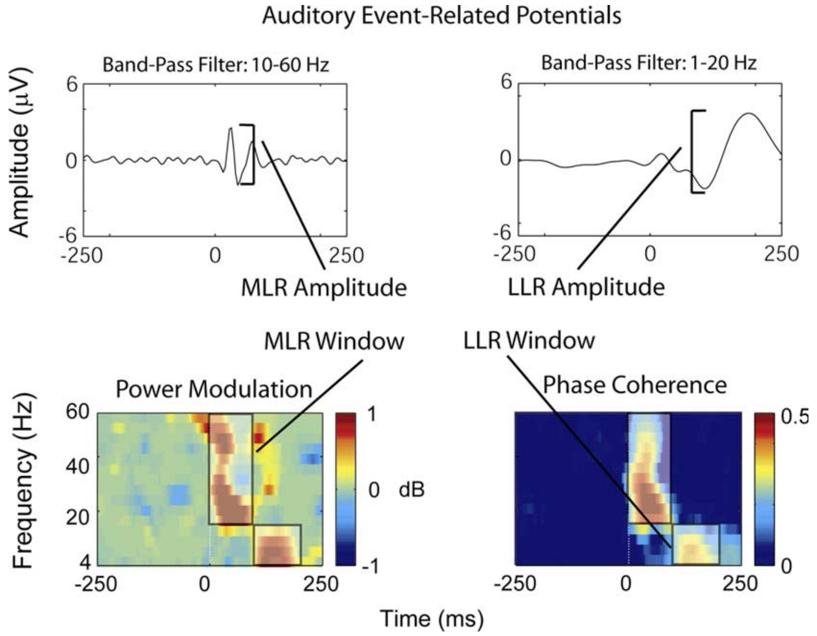

Fig. 1. Procedure for determining peak-to-peak amplitude of the event-related potential (μV), size of power modulation (dB relative to half-IOI pre-stimulus baseline), and amount of phase coherence (dimensionless units 0–1) for the middle-latency response (MLR) and the long-latency response (LLR). The peak-to-peak amplitudes of the MLR and LLR are the magnitudes of difference between the highest and lowest peaks from 0 to 75 and 75 to 200 ms latencies, respectively. For power modulation and phase coherence, the size of the response is the mean value in the time-frequency windows of 0–75 ms, 20–60 Hz for the MLR, and 75–200 ms, 0–20 Hz for the LLR.

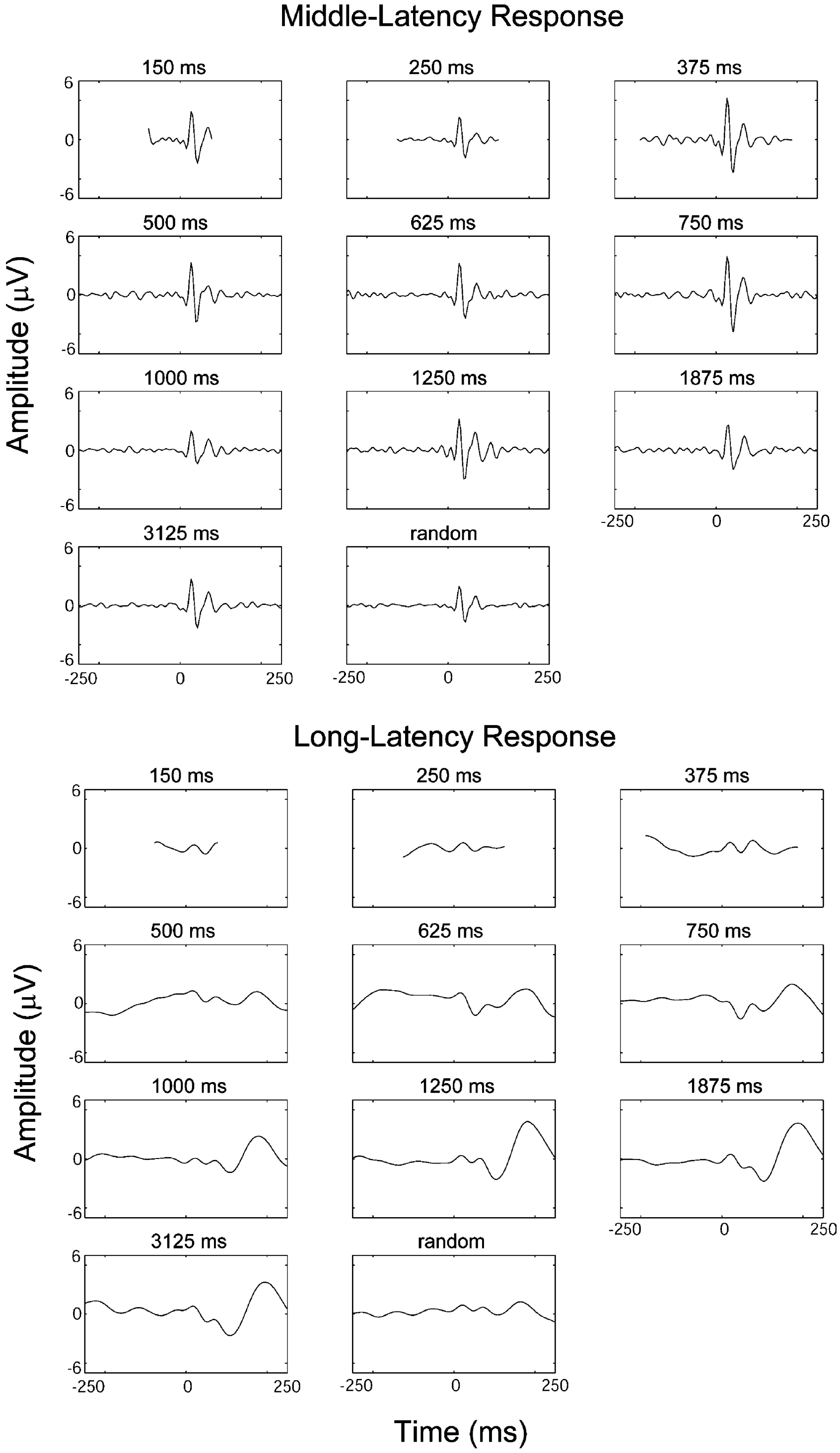

Fig. 2. Mean auditory event-related potentials across channels for subject 6 for each of the IOI conditions. For the first 10 conditions 200 tones were presented at a constant IOI of 150, 250, 375, 500, 625, 750, 1000, 1250, 1875, and 3125 ms. For the last condition, 200 tones were presented at random IOIs taken from a flat distribution of 375–750 ms. The MLR appears as a series of fast deflections in the first 100 ms post-tone onset, while the LLR appears as a series of slow deflections after 100 ms.

of Fig. 3 shows power modulations with respect to pre-stimulus baseline, and the bottom of Fig. 3 shows phase coherence across trials following tone onsets in all 11 conditions. In the fastest IOI conditions, power and phase modulation occurred mainly at high frequencies in the 20–60 Hz range and around the latency of the MLR, with no clear distinction between MLR and GBR TF ranges. At slower tempos, power and phase modulation increasingly occurred at low frequencies around the latency of the LLR. Finally, in several of the IOI conditions, high-frequency

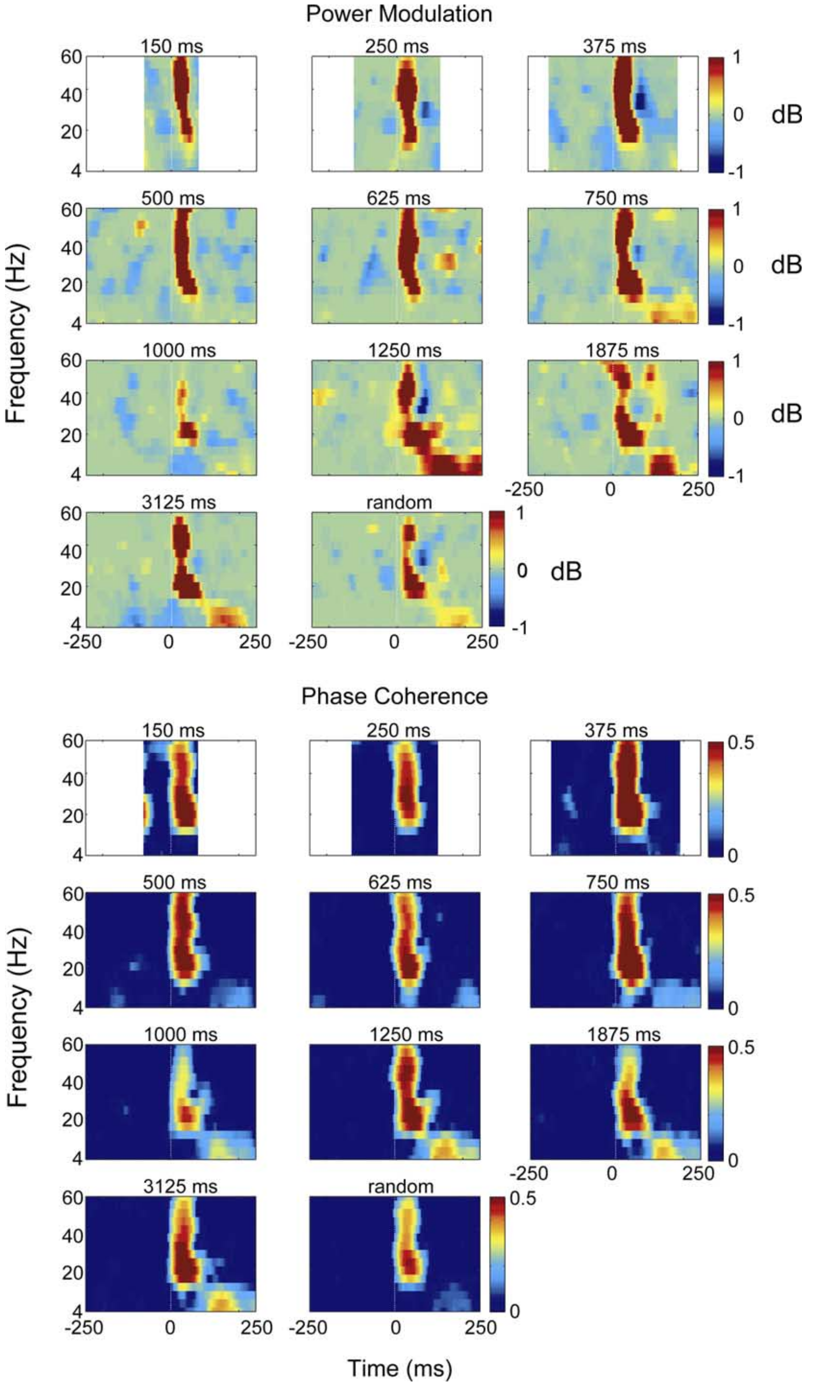

Fig. 3. Mean time-frequency (TF) representations of significant power modulation (dB relative to half-IOI pre-stimulus baseline) and significant phase coherence (dimensionless units 0–1) across channels for subject 6 for each of the IOI conditions. Non-significant values were set to 0. For both TF representation types, MLR appears as a broadband high-frequency activation in the first 100 ms post-tone onset, while the LLR appears as a narrower band low-frequency activation after 100 ms.

power increase corresponding to the MLR latency region was immediately followed by power decrease in the same frequency region. This effect was observed in the two subjects who showed the largest MLR.

We applied one-way repeated measures analyses of variance (ANOVA) to test for differences among the IOI conditions, with Greenhouse–Geisser corrections for possible sphericity violations. Separate ANOVAs tested for effects of IOI condition on ERP peak-to-peak amplitude, mean power modulation across the TF windows, and mean phase coherence across the TF windows, for the MLR and the LLR. The 150, 250, and 375 ms IOI conditions were not included in the LLR analyses because the full N1–P2 response did not occur in the extracted window. Planned simple contrasts with a Bonferonni correction compared the random IOI condition with the other IOI conditions.

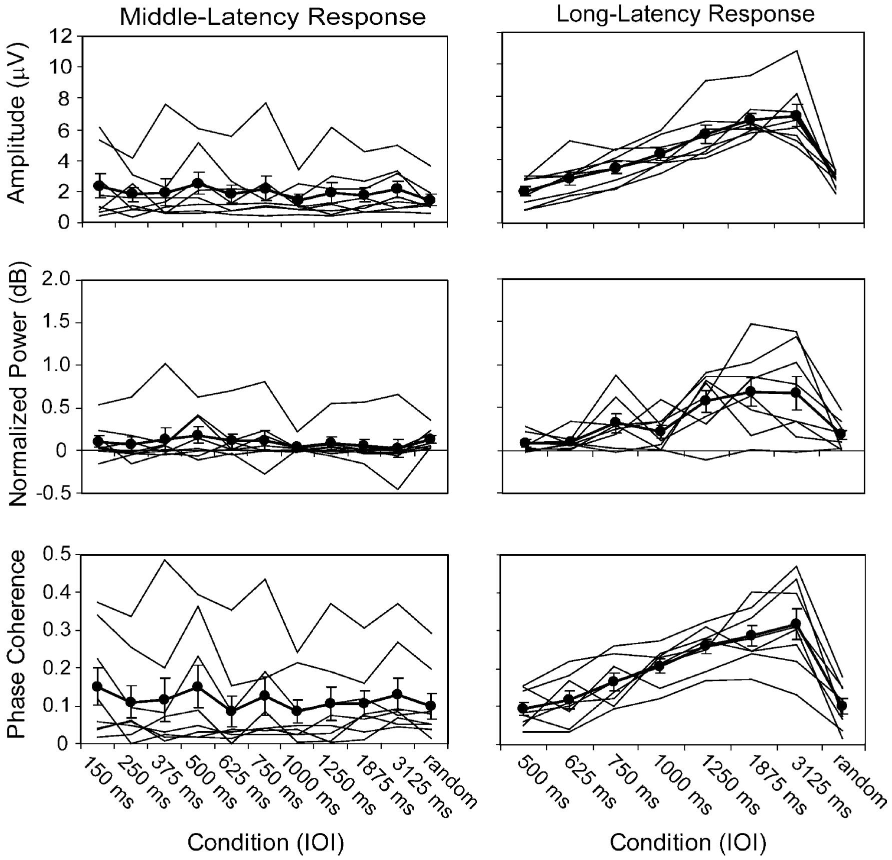

Fig. 4. Peak-to-peak amplitude (top), size of power modulation (middle), and amount of phase coherence (bottom) for each subject (thin lines) and mean (± SE, thick lines) in each of the IOI conditions for the middle-latency response (left) and for the long-latency response (right). These values were submitted to analyses of variance to test for effects of IOI condition and regression analyses to test for the relative contribution of power modulation and phase coherence in predicting peak-to-peak amplitude.

Fig. 4 shows the values for each IOI condition, as a mean across subjects and for each subject separately. Individual subjects varied more in the amplitude and IOI dependence of their MLRs than their LLRs.

For the MLR, there were no significant main effects of IOI condition for ERP amplitude ($F(10, 70) = 1.31$, $P = 0.30$), power modulation ($F(10, 70) = 1.33$, $P = 0.29$), and phase coherence ($F(10, 70) = 2.18$, $P = 0.10$). Planned contrasts showed that the random IOI condition did not differ from any of the other IOI conditions in MLR amplitude, power modulation, or phase coherence.

For the LLR, there were significant main effects of IOI condition for ERP amplitude ($F(7, 49) = 31.67$, $P < 0.001$), power modulation ($F(7, 49) = 7.49$, $P < 0.005$), and phase coherence ($F(7, 49) = 25.69$, $P < 0.001$). This is due to the increasing size of each of these quantities at slower tempos. Based on the planned contrasts, the random IOI condition showed significantly smaller ERP amplitudes than the 750, 1000, 1250, 1875, and 3125 ms conditions ($P < 0.005$), and smaller phase coherence than the 750, 1000, 1250, 1875, and 3125 ms conditions ($P < 0.005$). Thus, the random IOI condition differed from conditions containing larger mean IOI, consistent with the main effect of tempo condition.

3.2. Effect of power and phase modulation on peak-to-peak amplitude

The ANOVA results above and the corresponding data depicted in Fig. 4 suggest that phase modulation is a better predictor of ERP amplitude, for both the MLR and the LLR.

Table 1
Correlation matrices for peak-to-peak amplitude (μV), power modulation (dB), and phase coherence (dimensionless) of the middle-latency response (MLR) and long-latency response (LLR)

	ERP amplitude (μV)	Power modulation (dB)	Phase coherence
MLR			
ERP amplitude (μV)	1.0	0.834*	0.922*
Power modulation (dB)		1.0	0.843*
Phase coherence			1.0
LLR			
ERP amplitude (μV)	1.0	0.484*	0.709*
Power modulation (dB)		1.0	0.847*
Phase coherence			1.0

*Significant correlations at the $P < 0.001$ level with 86 degrees of freedom for MLR and 62 degrees of freedom for LLR.

Table 2
Regression table for power modulation and phase coherence as predictors of peak-to-peak amplitude of the middle-latency response (MLR) and long-latency response (LLR), showing slope (B), standard error of slope (SE B), normalized slope (β), t value and significance of β (t and P), and semi-partial correlation (sr) for each predictor variable

	B	SE B	β	t	P	sr
MLR						
Power modulation	1.42	0.54	0.20	2.64	0.010	0.11
Phase coherence	10.51	1.05	0.75	10.02	0.000	0.41
LLR						
Power modulation	−2.29	0.90	−0.41	−2.55	0.013	−0.22
Phase coherence	21.52	3.28	1.06	6.55	0.000	0.56

Table 1 shows the simple correlations (r) between the variables submitted to this analysis separately for the MLR and LLR. Power and phase modulation correlated significantly with ERP amplitude and with each other for both the MLR ($r = 0.83–0.92$, $P < 0.001$) and the LLR ($r = 0.48–0.85$, $P < 0.001$). We determined the relative contribution of phase and power modulation to ERP amplitude with multiple regression analysis (Darlington, 1990), as shown in Table 2. For the MLR, power and phase modulation significantly predicted 88 ERP responses across all 11 IOI conditions and 8 subjects ($R^2(2,85) = 0.86$, $P < 0.001$). The regression coefficients of phase modulation ($P < 0.001$) and power modulation were significant ($P < 0.025$). Semi-partial correlations (sr) indicated the unique contributions of phase and power modulation in predicting ERP size ($sr = 0.41$ and 0.11, respectively). For the LLR, power and phase modulation significantly predicted 64 ERP responses across 8 IOI conditions (500–3125 ms and random) and 8 subjects ($R^2(2,61) = 0.55$, $P < 0.001$). The regression coefficient of phase modulation was significant ($P < 0.001$) as was the coefficient of power modulation ($P < 0.025$), with unique contributions of $sr =$ 0.56 and −0.22, respectively. Thus, stimulus-related phase modulation predicted peak-to-peak ERP amplitude better than power modulation for early and late auditory responses. Power modulation significantly contributed to ERP amplitude for the late response but not the early response. Note that the unique contribution of power modulation was negative, demonstrating an inverse relation to ERP amplitude. This may be due to portions of the TF representation that underwent decreases in power after the initial increase that follows tone onsets.

4. Discussion

4.1. Tempo effects on auditory responses: differential refractoriness for MLR and LLR

We characterized the dependence of early (MLR) and late (LLR) auditory responses on tempo and structure of tone presentation in a behaviorally relevant IOI range. Both of these responses are sensitive to tone onsets and are therefore likely to be important indicators of sensory aspects of auditory pattern processing. We quantified the peak-to-peak amplitude, the size of power modulation, and the amount of phase coherence for tone onsets in isochronous patterns with IOIs from 150 to 3125 ms, and in patterns with IOIs taken randomly from a flat distribution (375–750 ms). We replicated previous studies by finding a larger LLR for slower tempos in each of the 3 measures of brain response (Carver et al., 2002; Hari et al., 1982; Sams et al., 1993). This dependence on tempo may reflect the maintenance of a sensory memory trace (Lü et al., 1992; Sams et al., 1993), or a refractory period (Budd et al., 1998) of large populations of neurons in auditory cortex (Godey et al., 2001).

The MLR showed no significant effects of IOI condition for any of the 3 response types we quantified, unlike with the LLR. Our quantification of the MLR was peak-to-peak amplitude of the largest positive (Pa) and negative (Nb) peaks, and is therefore consistent with a previous finding that Pa lacks tempo dependence (Erwin and Buchwald, 1986). However, previous studies of the MLR and GBR have found larger responses at slower tempos, though with smaller numbers of subjects and a coarser sampling of IOI (Makeig, 1990; Pantev et al., 1993). The relative constancy of MLR amplitude even at very fast tempos is consistent with a recent study that deconvolved the auditory SSR generated by clicks with IOIs from 19 to 31 ms to reveal an MLR-like waveform (Gutschalk et al., 1999). This response was localized to primary auditory cortex, confirming intracranial and source localization studies of the MLR (Godey et al., 2001; Liégeois-Chauvel et al., 1994; Yvert et al., 2001). The first stages of auditory cortical processing thus appear to show very little refractoriness, an important property for processing individual events in patterns such as those appearing in music and speech. This is in contrast to the LLR that may reflect more temporally integrated activity that codes for perceptual features such as pitch (Pantev et al., 1989).

Studies using higher spatial resolution techniques inform the generation of the differential response patterns across tempos of the MLR and the LLR. A recent intracranial study of two epilepsy patients recorded auditory evoked potentials directly from primary auditory cortex and from a secondary auditory area (Howard et al., 2000). In both patients, the primary auditory area showed very little degradation at fast tempos, whereas the secondary area declined markedly at fast tempos. Non-invasive studies using fMRI similarly find that brain areas are most responsive to lower stimulation rates at later stages in the ascending auditory system, from the brainstem to secondary cortical fields (Giraud et al., 2000; Harms and Melcher, 2002). These results are consistent with sources for the MLR and LLR in primary and secondary auditory cortical areas, respectively, with differential tempo dependence. High spatial resolution techniques can also reveal the extent to which temporal properties are represented spatially in individual brain structures. Despite the existence of single auditory cortical cells sensitive to IOI (deCharms et al., 1998), an fMRI study failed to find a topographic representation of IOIs (Giraud et al., 2000), as exists for pitch (Pantev et al., 1989). It is therefore currently unclear what principles of neural organization underlie auditory temporal representation.

4.2. Tempo effects on auditory responses: implications for studying pattern processing

Our study revealed non-linear response functions with some subjects showing peaks in the 375–625 ms IOI range, the center of an important tempo range for sensory-motor anticipation (Engström et al., 1996; Mates et al., 1994; Peters, 1989), and tempo discrimination (Drake and Botte, 1993; Friberg and Sundberg, 1995). This is consistent with a recent demonstration of large peaks in the P1 response at around 500 ms IOI, resulting from overlap with adjacent-tone responses (Carver et al., 2002). The lack of response diminution at fast tempos suggests the MLR as an important tool for studying pattern processing.

Presenting tones with randomly selected IOIs had little effect on response amplitude for the constant IOI conditions with similar mean IOI content for the MLR and the LLR. This suggests that tone-onset predictability did not have a substantial impact on response size. This is also consistent with the fact that neither the MLR nor the LLR showed diminution at the transition point of ~1200 ms past which people can no longer synchronize to metronome patterns (Engström et al., 1996; Mates et al., 1994).

Ongoing music and speech are both dynamic acoustic signals that evolve at multiple time-scales. Because of their rich temporal dynamics, it is logical that they would be ideal candidates for neuroelectric and neuromagnetic study. Recent studies showed that musical experience and musical stimulation affect auditory cortical responses at multiple stages of processing, including the MLR (Schneider et al., 2002), the LLR (Pantev et al., 1998, 2001), the SSR (Bhattacharya et al., 2001; Patel and Balaban, 2000), and the mismatch negativity (MMN, Tervaniemi, 2001). Thus, music demonstrates the importance of pathways containing these responses in processing complex auditory temporal patterns.

Due to their transient nature, the responses measured in the current experiment are particularly good tools for studying neural processing of onset characteristics. This includes features of individual sound events such as spectral content, pitch, intensity, and duration. But it may also include features that depend on a larger temporal context such as expectation for the timing and characteristics of events. Expectancy is fundamental to tapping in time with music (Snyder and Krumhansl, 2001; Toiviainen and Snyder, 2003), meter perception (Large and Jones, 1999), musical pitch processing (Krumhansl, 2000), and speech processing (Warren, 1984).

In addition to the importance of expectancy in auditory processing, it may also be especially amenable to non-invasive neurophysiological study because of its temporal nature, in comparison to other perceptual characteristics that likely depend on topographical representations in the brain (e.g. Janata et al., 2002; Pantev et al., 1995). Support for this includes the contingent negative variation and other slow brain responses that anticipate sensory and motor events (Walter et al., 1964), and the emitted responses that anticipate events in isochronous sequences at similar latencies as onset evoked responses (Janata, 2001; May and Tiitinen, 2001; Weinberg et al., 1970). Later emitted responses such as the MMN and P3 more likely represent a violation of temporal expectancy, rather than a direct manifestation of sensory expectancy (Janata, 2001; Yabe et al., 1998). Application of these paradigms to studying neural expectancy during music and speech processing would likely be profitable. For example, we recently studied the emitted response that occurs in the absence of expected tones in a simple binary metrical context, defined by the presence of alternating loud and soft tones (Snyder and Large, submitted). Power modulations in the TF range of the MLR occurred in the absence of loud and soft tones that were as large as when tones were present. We also recently found correlates of top-down modulation of the MLR to equi-intensity tone onsets according to metrical expectancy (Snyder et al., 2003), with larger responses to events occurring at metrically strong time points compared to weak time points. Continued research on more complex stimuli would likely yield important knowledge concerning the neural basis of temporal pattern processing, a goal proposed long ago by Lashley (1951).

4.3. Relation between the middle-latency response and the gamma-band response

The present results obtained no clear distinction in TF space between the MLR and the GBR. Instead, a large area of activation from 20 to 60 Hz occurred for both power modulation and phase coherence. The MLR and the GBR are usually obtained using wide-band and narrow-band filters such as 1–60 and 24–48 Hz, respectively (Pantev et al., 1993). Therefore, it seems appropriate to assume that the GBR is a subset of the MLR. This could explain the many similarities between the two responses as well as the differences. The MLR and GBR have similar peak latencies (Pantev et al., 1993), respond similarly to paired-tones (Müller et al., 2001), and are both localized in primary auditory cortex (Pantev et al., 1991, 1995; Yvert et al., 2001). The main difference is that the MLR shows tonotopic organization while the GBR reportedly does not (Pantev, 1995; Pantev et al., 1995). Despite the fact that dipole modeling assumes a point source of current flow in the brain, a relatively large surface area of cortex is activated by tone stimuli (Herdman et al., 2003). It is possible that the GBR takes up a subset of the area from which the MLR is generated. This subset may be an area that either does not have tonotopy or has tonotopy that is difficult to detect using scalp sensors. In support of the conclusion that the MLR and GBR represent similar responses, evoked and induced activity in the MLR TF range was recently found in a study of auditory rhythm processing, with no clear separation of the MLR and GBR in TF space (Snyder and Large, submitted).

4.4. Dynamics of power and phase coherence: unique information?

Submitting the peak-to-peak amplitude of MLR and LLR to regression analysis with power modulation and phase coherence as predictors demonstrated the dominant effect of phase coherence on ERP amplitude. Our results further indicate that wide frequency bands undergo transient phase reorganization after tone onsets. These results are consistent with previous claims that phase resetting of ongoing brain activity is a major determinant of the ERP (Jansen et al., 2003; Karakaş et al., 2000; Makeig et al., 2002; Sayers et al., 1974). Jansen et al. (2003) recently showed that phase dynamics of EEG activity in the 2–8 Hz band is mainly responsible for the auditory LLR. The present results extend this finding by showing that the auditory MLR also is based on phase reorganization, and by providing a quantitative comparison with power dynamics. The fact that the ERP mainly reflects phase organization suggests the need for further characterization of the different information about neural processing contained in power and phase dynamics. For example, the so-called *induced* brain activity that mainly reflects power dynamics shows different patterns than evoked activity in a wide range of perceptual, cognitive, and motor paradigms (Pantev, 1995; Tallon-Baudry and Bertrand, 1999). In particular, induced activity is extremely useful for detecting high-frequency brain activity that is not strictly time-locked to an event onset.

Acknowledgements

This research was supported by National Science Foundation grant BCS-0094229, awarded to EWL. We thank Dr J.A.S. Kelso who generously provided EEG facilities in support of this research, purchased under National Institute of Mental Health grant R37 MH42900-16. We also thank Bill McLean for technical assistance and Debora Yonezawa and Ted Zanto for help running subjects. Joel S. Snyder is now at the Rotman Research Institute, Baycrest Centre for Geriatric Care, Toronto, Canada.

References

Arieli A, Sterkin A, Grinvald A, Aertsen A. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science 1996;273:1868–71.

Bhattacharya J, Petsche H, Pereda E. Long-range synchrony in the γ band: role in music perception. J Neurosci 2001;21:6329–37.

Brandt ME, Jansen BH, Carbonari JP. Pre-stimulus spectral EEG patterns and the visual evoked response. Electroencephalogr Clin Neurophysiol 1991;80:16–20.

Budd TW, Barry RJ, Gordon E, Rennie C, Michie PT. Decrement of the N1 auditory event-related potential with stimulus repetition: habituation vs. refractoriness. Int J Psychophysiol 1998;31:51–68.

Carver FW, Fuchs A, Jantzen KJ, Kelso JAS. Spatiotemporal analysis of the neuromagnetic response to rhythmic auditory stimulation. Clin Neurophysiol 2002;113:1921–31.

Darlington RB. Regression and linear models. New York: McGraw-Hill; 1990.

deCharms RC, Blake DT, Merzenich MM. Optimizing sound features for cortical neurons. Science 1998;280:1439–43.

Delorme A, Makeig S. EEGLAB: matlab software for signal processing of electrophysiological brain data. http://sccn.ucsd.edu/eeglab/; 2002.

Drake C, Botte MC. Tempo sensitivity in auditory sequences: evidence for a multiple-look model. Percept Psychophys 1993;54:277–86.

Engström DA, Kelso JAS, Holroyd T. Reaction–anticipation transitions in human perception–action patterns. Hum Mov Sci 1996;15:809–22.

Erwin RJ, Buchwald JS. Midlatency auditory evoked responses: differential recovery cycle characteristics. Electroencephalogr Clin Neurophysiol 1986;64:417–23.

Friberg A, Sundberg J. Time discrimination in a monotonic, isochronous sequence. J Acoust Soc Am 1995;98:2524–31.

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 2000;84:1588–98.

Godey B, Schwartz D, de Graaf JB, Chauvel P, Liégeois-Chauvel C. Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. Clin Neurophysiol 2001;112:1850–9.

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 1999;110:856–68.

Haig AR, Gordon E. Prestimulus EEG alpha phase synchronicity influences N100 amplitude and reaction time. Psychophysiology 1998;35:591–5.

Hari R, Kaila K, Katila T, Tuomisto T, Varpula T. Interstimulus interval dependence of the auditory vertex response and its magnetic counterpart: implications for their neural generation. Electroencephalogr Clin Neurophysiol 1982;54:561–9.

Harms MP, Melcher JR. Sound repetition rate in the human auditory pathway: representations in the waveshape and amplitude of fMRI activation. J Neurophysiol 2002;88:1433–50.

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 2003;20:995–1005.

Howard MA, Volkov IO, Mirsky R, Garell PC, Noh MD, Granner M, Damasio H, Steinschneider M, Reale RA, Hind JE, Brugge JF. Auditory cortex on the posterior superior temporal gyrus. J Comp Neurol 2000;416:79–92.

Janata P. Brain electrical activity evoked by mental formation of auditory expectations and images. Brain Topogr 2001;13:169–93.

Janata P, Birk JL, Van Horn JD, Leman M, Tillmann B, Bharucha JJ. The cortical topography of tonal structures underlying Western music. Science 2002;298:2167–70.

Jansen BH, Brandt ME. The effect of the phase of prestimulus alpha activity on the averaged visual evoked response. Electroencephalogr Clin Neurophysiol 1991;80:241–50.

Jansen BH, Agarwal G, Hegde A, Boutros NN. Phase synchronization of the ongoing EEG and auditory EP generation. Clin Neurophysiol 2003;114:79–85.

Karakaş S, Erzengin ÖU, Başar E. A new strategy involving multiple cognitive paradigms demonstrates that ERP components are determined by the superposition of oscillatory responses. Clin Neurophysiol 2000; 111:1719–32.

Kruglikov ST, Schiff SS. Interplay of electroencephalogram phase and auditory-evoked neural activity. J Neurosci 2003;23:10122–7.

Krumhansl CL. Rhythm and pitch in music cognition. Psychol Bull 2000; 126:159–79.

Large EW, Jones MR. The dynamics of attending: how people track time-varying events. Psychol Rev 1999;106:119–59.

Lashley KS. The problem of serial order in behavior. In: Jeffress LA, editor. Cerebral mechanisms in behavior. New York: Wiley; 1951. p. 123–47.

Liégeois-Chauvel C, Musolino A, Badier JM, Marquis P, Chauvel P. Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. Electroencephalogr Clin Neurophysiol 1994;92:204–14.

Lü Z-L, Williamson SF, Kaufman L. Behavioral lifetime of human auditory sensory memory predicted by physiological measures. Science 1992; 258:1668–70.

Makeig S. A dramatic increase in the auditory middle latency response at very slow rates. In: Brunia CHM, Gaillard AWK, Kok A, editors. Psychophysiological brain research. Tilburg: Tilburg University Press; 1990. p. 60–4.

Makeig S. Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. Electroencephalogr Clin Neurophysiol 1993;86:283–93.

Makeig S, Westerfield M, Jung T-P, Enghoff S, Townsend J, Courchesne E, Sejnowski TJ. Dynamic brain sources of visual evoked potentials. Science 2002;295:690–4.

Mates J, Radil T, Müller U, Pöppel E. Temporal integration in sensorimotor synchronization. J Cogn Neurosci 1994;6:332–40.

May P, Tiitinen H. Human cortical processing of auditory events over time. NeuroReport 2001;12:573–7.

Müller MM, Keil A, Kissler J, Gruber T. Suppression of the auditory middle-latency response and evoked gamma-band response in paired-click paradigm. Exp Brain Res 2001;136:474–9.

Näätänen R, Winkler I. The concept of auditory stimulus representation in cognitive neuroscience. Psychol Bull 1999;125:826–59.

Pantev C. Evoked and induced gamma-band activity of the human cortex. Brain Topogr 1995;7:321–30.

Pantev C, Hoke M, Lutkenhoner B, Lehnertz K. Tonotopic organization of the auditory cortex: pitch versus frequency representation. Science 1989;246:486–8.

Pantev C, Makeig S, Hoke M, Galambos R, Hampson S, Gallen C. Human auditory evoked gamma-band magnetic fields. Proc Natl Acad Sci 1991;88:8996–9000.

Pantev C, Elbert T, Makeig S, Hampson S, Eulitz C, Hoke M. Relationship of transient and steady-state auditory evoked fields. Electroencephalogr Clin Neurophysiol 1993;88:389–96.

Pantev C, Bertrand O, Eulitz C, Erkindt C, Hampson S, Schuierer G, Elbert T. Specific tonotopic organization of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. Electroencephalogr Clin Neurophysiol 1995;94: 26–40.

Pantev C, Oostenveld R, Engelien A, Ross B, Roberts LE, Hoke M. Increased auditory cortical representation in musicians. Nature 1998; 392:811–4.

Pantev C, Roberts LE, Schulz M, Engelien A, Ross B. Timbre-specific enhancement of auditory cortical representations in musicians. NeuroReport 2001;12:169–74.

Patel AD, Balaban E. Temporal patterns of human cortical activity reflect tone sequence structure. Nature 2000;404:80–2.

Peters M. The relationship between variability of intertap intervals and interval duration. Psychol Res 1989;51:38–42.

Picton TW, Hillyard SA, Krausz HI, Galambos R. Human auditory evoked potentials I: evaluation of components. Electroencephalogr Clin Neurophysiol 1974;36:179–90.

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

Sams M, Hari R, Rif J, Knuutila J. The human auditory sensory memory trace persists about 10 sec: neuromagnetic evidence. J Cogn Neurosci 1993;5:363–70.

Sayers BM, Beagley HA, Henshall WR. The mechanism of auditory evoked EEG responses. Nature 1974;247:481–3.

Schneider P, Scherg M, Dosch HG, Specht HJ, Gutschalk A, Rupp A. Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. Nat Neurosci 2002;5:688–94.

Snyder J, Krumhansl CL. Tapping to ragtime: cues to pulse finding. Music Percept 2001;18:455–89.

Snyder JS, Large EW. Gamma-band activity reflects the metric structure of rhythmic tone sequences. Submitted for publication.

Snyder JS, Zanto TP, Large EW, Kelso JAS. Gamma-band activity during rhythmic processing: temporal structure, meter, and attention. Soc Neurosci Abstr 2003;29:486.3.

Tallon-Baudry C, Bertrand O. Oscillatory gamma activity in humans and its role in object representation. Trends Cogn Sci 1999;3:151–62.

Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J. Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. J Neurosci 1996;16:4240–9.

Tervaniemi M. Musical sound processing in the human brain: evidence from electric and magnetic recordings. In: Zatorre RJ, Peretz I, editors. The biological foundations of music. New York: The New York Academy of Sciences; 2001. p. 259–72.

Toiviainen P, Snyder JS. Tapping to Bach: resonance-based modeling of pulse. Music Percept 2003;21:43–80.

Walter WG, Cooper R, Aldridge VJ, McCallum WC, Winter AL. Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. Nature 1964;203: 380–4.

Warren RM. Perceptual restoration of obliterated sounds. Psychol Bull 1984;96:371–83.

Weinberg H, Walter WG, Crow HJ. Intracerebral events in humans related to real and imaginary stimuli. Electroencephalogr Clin Neurophysiol 1970;29:1–9.

Yabe H, Tervaniemi M, Sinkkonen J, Huotilainen M, Ilmoniemi RJ, Näätänen R. Temporal window of integration of auditory information in the human brain. Psychophysiology 1998;35:615–9.

Yvert B, Crouzeix A, Bertrand O, Seither-Preisler A, Pantev C. Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. Cereb Cortex 2001;11:411–23.

Zatorre RJ, Belin P, Penhune VB. Structure and function of auditory cortex: music and speech. Trends Cogn Sci 2002;6:37–46.