Spatiotemporal analysis of the neuromagnetic response to rhythmic auditory stimulation: rate dependence and transient to steady-state transition

Frederick W. Carver*, Armin Fuchs, K.J. Jantzen, J.A. Scott Kelso

Center for Complex Systems and Brain Sciences, Florida Atlantic University, 777 Glades Road, Boca Raton, FL 33431, USA

Abstract

Objective: Whole head magnetoencephalography was used to investigate the spatiotemporal dynamics of neuromagnetic brain activity associated with rhythmic auditory stimulation.

Methods: In order to characterize the evolution of the auditory responses we applied a Karhunen-Loève decomposition and k-means cluster analysis to globally compare spatial patterns of brain activity at different latencies and stimulation rates. Tones were presented binaurally at 27 different stimulation rates within a perceptually and behaviorally relevant range from 0.6 to 8.1 Hz.

Results: Over this range, we observed a linear increase of the amplitude of the main auditory response at 100 ms latency (N1m) with increasing inter-stimulus interval, and qualitative changes of the overall spatiotemporal dynamics of the auditory response. In particular, a transition occurred between a transient evoked response at low frequencies, and a continuous steady-state response at high frequencies.

Conclusions: We show the onset of temporal overlap between responses to successive tones that leads to this transition. Response overlap begins to occur near 2 Hz, marking the onset of a continuous perceptual representation. © 2002 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Auditory; Magnetoencephalography; Inter-stimulus interval; Rhythm; Principal components analysis; k-means cluster analysis

1. Introduction

The presentation rate of a series of simple tone or click stimuli is well known to systematically affect the human brain response recorded with electroencephalography (EEG) and magnetoencephalography (MEG). The auditory responses investigated by the majority of studies can be grouped into one of two general categories: transient and steady-state. Transient responses are so-named because they are elicited by a series of tones separated by sufficient time to allow for the neural response to each tone to return to baseline prior to the onset of the next tone. In addition, a randomized inter-stimulus interval is often employed in this type of study to avoid anticipation by the subject and to minimize the cumulative effect of any overlap of late and early responses across stimulus presentations. Since auditory evoked responses can last for over 300 ms (Picton et al., 1974), inter-stimulus intervals (ISIs) on the order of seconds are typically required for these studies. Steady-state

responses (SSRs), on the other hand, are recorded at much higher stimulation rates and with constant ISIs, so as to produce continuous activation between successive stimulus events. A large amount of the work on steady-state activity has focused on the 40 Hz steady state response, a near sinusoidal response elicited in humans by auditory stimulation near 40 Hz (Galambos et al., 1981; Stapells et al., 1984). Because of the large amplitude response in human EEG to stimulation at this frequency, this 40 Hz SSR has drawn considerable attention, both as a subject of study and as a research tool, with competing theories having been proposed to explain its origin (see for example Hari et al., 1989; Gutschalk et al., 1999; Pantev et al., 1996). Regan (1989) defines the 'ideal steady-state' as a response 'whose constituent frequency components remain constant in amplitude and phase over an infinitely long time period', implying a sustained oscillation at a single frequency or a small group of frequencies. The 40 Hz auditory SSR is a classic example of one such ideal response. In the present paper, we will use the more general definition of the SSR given in Regan (1982) since it provides a clearer definition of the border between transient and steady-state activity. By

this definition, the transition from a transient to a steady-state response occurs when the ISI becomes short enough to cause overlap between the transient responses to successive tones, thus producing a continuous evoked response. Since transient responses can last over 300 ms, the transition from transient to steady-state should occur at stimulation rates between 2 and 5 Hz. These rates have not been systematically investigated since they fall between those typically used to evoke transient or steady-state responses. In this experiment we used magnetoencephalography to systematically investigate the range of stimulation rates containing the hypothesized transient to steady-state transition. Our goal was to determine the onset of the SSR and describe changes in the evoked responses that occur as discrete cortical responses begin to interact.

Compared to the steady-state response, the origin of the transient response is relatively well understood. In EEG and MEG the response consists of a series of electric potentials or neuromagnetic fields that begin soon after tone onset. These were first recorded with EEG, and divided into 3 response epochs: the so-called early, middle and long latency responses (Picton et al., 1974). The early components, also known as brain-stem evoked responses, are subcortical in origin and occur within 10 ms of tone onset. These are followed by cortical middle latency responses (MLRs), which occur between 10 and 50 ms. Cortical components that peak after 50 ms are referred to as long latency responses (LLRs), and are characterized by longer durations and higher amplitudes than the MLRs. The most prominent long latency response is a negative wave near 100 ms, commonly referred to in the EEG literature as the N100 or N1. The term N1m refers to the corresponding magnetic field recorded at a similar latency using MEG. Because of its large amplitude, the N1/N1m complex has received the most experimental attention of the evoked responses (for reviews see Näätänen and Picton, 1987; Woods, 1995). This research has revealed that the N1/N1m consists of several separate components with different latencies and source locations (e.g. Loveless et al., 1996; for a discussion see Näätänen and Winkler, 1999). The main N1m component, which peaks at around 100 ms, is thought to originate in supratemporal auditory cortex along Heschl's gyrus (see Pantev et al., 1995 for specific tonotopy of the response).

Like most long latency responses, the amplitude of the N1/N1m is known to depend on the presentation rate of the auditory stimuli. Several experiments have studied the rate dependency of the N1/N1m response using inter-stimulus intervals on the order of one to several seconds (Hari et al., 1982; Lü et al., 1992; Sams et al., 1993). In general they report that the amplitude of the N1/N1m decreases with increased presentation rate. However, in these studies the stimulation rates were too slow to allow for a comprehensive investigation of how this amplitude decrease may be related to the transition from a transient to a steady state response. One intention of the current experiment was to study the effect of rate on the N1m over a behaviorally relevant higher range of stimulation rates. For this purpose, and for the purpose of observing the transient to steady-state transition, we used rhythmic tone stimulation from 0.6 to 8.1 Hz, a parameter range over which the N1m has not been thoroughly examined. Based on preliminary EEG work (Carver et al., 1999), we hypothesized that we would observe the near disappearance of the N1m response within this range of frequencies, which would be at a lower rate than predicted by previous models (Lü et al., 1992). This parameter range is scientifically important not only because it encompasses those rates at which transients become steady states, but also because it contains most of the rates over which isochronous stimuli are perceived as being rhythmic (Fraisse, 1982). In addition, qualitative transitions in the ability of subjects to coordinate rhythmic hand movements with a metronome are also known to occur within this range in both behavior (Kelso et al., 1990) and brain (Fuchs et al., 1992; Kelso et al., 1991, 1992; Mayville et al., 2001; Meyer-Lindenberg et al., 2002; Wallenstein et al., 1995).

The predicted low amplitude of the long latency responses at higher stimulation rates and the potential for interference between responses to successive tones makes it difficult to analyze the auditory responses within the present range of input parameters. Here we overcome these problems by employing analysis techniques that take advantage of the high degree of spatial information provided by a full-head 141 sensor MEG system. These methods include principal components analysis (also known as Karhunen-Loève decomposition) and k-means cluster analysis. Our results show a linear decrease of the N1m amplitude as the ISI decreases over the range of frequencies, with a predicted disappearance of the response at a stimulation rate between 6 and 10 Hz. The N1m attenuation also has effects on the spatiotemporal dynamics of the overall activity from responses close to the N1m, which become apparent as the amplitude of the N1m decreases at higher stimulation rates. We also show the development of interaction between activity elicited by successive tones, which marks the transition point between the transient and steady-state responses. As evidence for this transition, we show interference between the last long latency response at 500 ms and the evoked response to the following tone.

2. Methods

2.1. Subjects

Four subjects (one female and 3 males) participated in the experiment. Their ages ranged from 26 to 33 years. All participants reported being right-handed. The experiment was conducted in compliance with all standards of human research outlined in the Declaration of Helsinki as well as by the Institutional Review Board. Informed consent was obtained from each participant prior to MEG recording. All
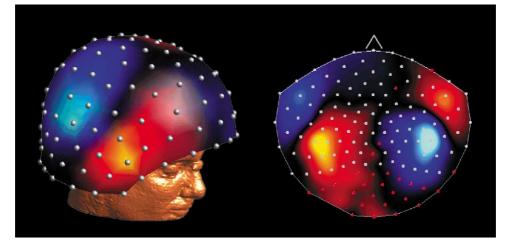
Fig. 1. (Left) Example of averaged magnetic activity from an N1m response interpolated with a spline of 3rd order and placed over a reconstructed MRI image. White dots represent the positions of the sensors relative to 3 fiduciary points (nasion, and left and right preauricular points). Light-blue/blue coloring indicates field lines entering the head; in yellow/red regions field lines exit. (Right) Two-dimensional polar projection of the same pattern of activity. The nose points to the top of the figure. Bilateral dipolar fields imply activity within both auditory cortices. Sensors colored red were not included in the analyses due to a high noise level.

subjects reported normal hearing. Audiometry was not performed.

2.2. Procedure

Subjects were seated inside a magnetically shielded room (Vacuum Schmelze, Hanau) with their heads held firmly within the dewar. Auditory stimulation was delivered binaurally through plastic headphones at an individually adjusted volume that each subject reported to be comfortable. Subjects were asked to attend to the auditory stimulation while fixing their gaze at a point two meters in front of them. They were further instructed to minimize all extraneous eye or body movements during the recordings. The experiment consisted of separate trials during which subjects listened to a series of tones presented at a constant rate of stimulation. Each series contained 40 1 kHz tones, with each tone lasting for 60 ms at a constant amplitude with infinite rise and fall times. Twenty-seven stimulation rates were used in the experiment, ranging from 0.6 to 8.1 Hz (see Fig. 2 for the complete list of rates). Step sizes between rates ranged from 0.1 to 0.8 Hz with larger steps at higher rates. Each rate was presented in at least 4 separate trials, providing a minimum of 160 cycles for averaging. The trial order was randomized across subjects.

2.3. Data acquisition

Neuromagnetic activity was recorded using a full-head magnetoencephalograph (CTF Inc., Port Coquitlam, BC) comprised of 141 SQuID (superconducting quantum interference device) sensors distributed homogeneously across the scalp. Fig. 1 (left) shows the approximate locations of the sensors relative to the head of a sample subject. A coordinate system for each subject's head was defined with respect to 3 fiduciary points: the nasion, and left and right

preauricular points. The 3-dimensional locations of these points in the sensor coordinate system were measured prior to each experiment. Conversion to third-order gradiometers was performed in firmware using a set of reference coils (Vrba et al., 1999). MEG and stimulus signals were bandpass (0.3–80 Hz) and notch filtered (50 and 100 Hz: the European line frequency and its harmonic) prior to digitization at a rate of 312.5 Hz.

2.4. Data processing

Separate ensemble averages of single stimulation cycles were calculated for each stimulation rate. Each cycle was centered at tone onset, and the duration was equal to the inter-stimulus interval. The MEG signals were manually inspected for artifacts before averaging, and any cycle with contaminations was discarded. The first two cycles from each trial were also excluded from the final average in order to avoid transients. Topographical mapping of the resulting average evoked fields was performed by a polar projection of the 3-dimensional sensor coordinates into two-dimensional space. A spline of 3rd order was used to interpolate activity between sensor positions. Fig. 1 (right) shows an example of this projection for an N1m response at 100 ms latency. Blue coloring in Fig. 1 represents magnetic field lines entering the head, with light blue indicating the strongest radial components of the field. Exiting field lines are indicated by red/yellow coloring, with yellow representing the highest amplitude. As expected for an N1m response, Fig. 1 shows dipolar-like activity patterns over both auditory cortices. Due to a high noise level, 30 sensors at mostly posterior locations were not included in the data analyses (see Fig. 1 for locations). This left a total of 111 channels, but did not create a significant loss of information
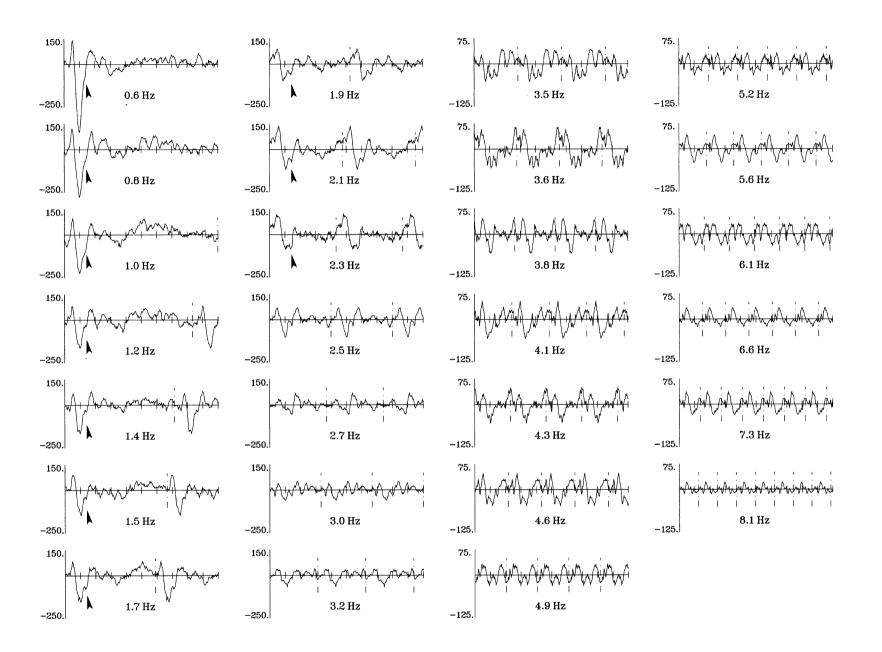
Fig. 2. Average time series for subject 2: 1 s of data at all stimulation rates from the channel that had the highest overall amplitude in the experiment for this subject. Each time series starts at tone onset, with dashed lines indicating the onset of succeeding tones. Note that the vertical femto-Tesla scale is different for the columns on the right in order to adjust for the lower amplitudes at higher rates. Hash marks on the horizontal axis indicate 100 ms intervals. Black arrows highlight a 150ms subcomponent of the N1m.

since most auditory events are seen over frontal, temporal, and central regions.

3. Results

3.1. Evolution of the response

Fig. 2 shows time series of the averaged auditory evoked responses from subject two. The single channel displayed at all stimulation rates was over right temporal cortex and recorded this subject's highest absolute amplitude of any sensor during the entire experiment. The time series show one second of averaged data starting at a tone onset. Predictably, the highest amplitude in the time series occurs during the N1m response at the lowest stimulation rate. As stimulation rate increases, the amplitude of the N1m systematically decreases. In fact, by 2.1 Hz the N1m is no longer the most dominant response. At this rate a response at 50 ms of opposite polarity has the highest amplitude. Another feature of the rate dependent N1m evolution is a subpeak of the response near 150 ms latency that only appears after the large 100 ms peak has begun to reduce (see arrows in Fig. 2). We will show evidence that this represents a separate subcomponent of the N1m. We will also quantify the rate dependence of the N1m and study how the response interacts with temporally adjacent responses. In terms of the overall character of the response, at low stimulation rates magnetic activity returns to base line level before the next tone onset. This is consistent with a transient response to each tone. However, as rate increases above 2 Hz, activity is still present at the onset of the next tone. At the higher stimulation rates, the time series show continuous activity between tones, indicative of a steady-state response. We will characterize the rate dependent transition between the transient and steady-state responses in terms of the onset of overlap between neural activity evoked by successive tones.

3.2. Amplitude of the N1m

We first analyzed how the amplitude of the dominant N1m auditory response depends on the rate of stimulation. The latency of the peak of the response was determined by the highest absolute amplitude in any sensor at 0.6 Hz, which turned out to be either 99.2 or 102.4 ms for all participants – a difference of only one sampling interval. In the following analyses we used these times as an estimate for the latency of the N1m at all stimulation rates because as rate increased the N1m became hard to distinguish from background activity and neighboring responses (see Fig. 2). In order to justify the use of one latency at all rates, we checked for any significant shift in latency at stimulation rates up to 2 Hz. This frequency was chosen as a cut-off since it was the highest stimulation rate at which the N1m was still identifiable based on amplitude alone. *t*-tests for each subject revealed that the mean of the latencies between 0.8 and 1.9 Hz were not significantly different from the

latencies originally determined at 0.6 Hz, indicating no significant temporal shift of the N1m.

As mentioned earlier, at the higher stimulation rates used in our study the amplitude of the brain signal at 100 ms was very low, which made the response difficult to observe in single channels. To overcome this problem, we treated the N1m response as a characteristic spatial pattern of activity based on all the sensors, and tracked the amplitude of this pattern across stimulation rate. For this purpose, we applied a Karhunen-Loève (KL) decomposition to the patterns at all stimulation rates at the latencies determined at 0.6 Hz (see Fuchs et al., 1992 for a discussion of KL decomposition). Fig. 3 shows the results of this procedure. For each subject, the first eigenvector of the decomposition captured the spatial pattern of the N1m response (see the topographical maps at left in Fig. 3). Note from the high eigenvalues (λ) that this mode accounts for over 80% of the variance in the signal for subjects 1 and 2, and about 65% for subjects 3 and 4.

The amplitude of the response at each stimulation rate was calculated by projecting the spatial pattern from each rate at the determined latency onto the largest eigenvector shown at left in Fig. 3. The plot of amplitude vs. inter-stimulus interval to the right of each eigenvector shows a strikingly linear relationship between amplitude and ISI, which is also supported by the high r^2-values obtained through linear regression. At higher stimulation rates, the N1m essentially vanishes. The *x*-intercept of the linear regression in each case provides a prediction for the ISI at which the response will reach zero amplitude. For each of the 4 subjects this occurs at an ISI between 98 and 162 ms, corresponding to stimulation rates between 6 and 10 Hz.

3.3. Extent of the N1m in time and frequency

At a stimulation rate of 0.6 Hz the N1m response has a rise and fall time that extends from approximately 70 to 130 ms past tone onset. From the KL analysis we know that the peak amplitude of the response reduced significantly as stimulation rate was increased. We now ask, first, if this reduction is accompanied by a temporal contraction of the N1m. In other words, does response duration decrease at higher stimulation rates? And second, if such contraction does occur, is it due to masking by other temporally adjacent responses of opposite polarity that were difficult to observe or suppressed at low stimulation rates?

To accomplish this task, it was necessary to characterize the brain response across a range of latencies as well as stimulation rates. Our goal was to find the region of activity corresponding to the N1m, and to distinguish this from activity associated with neighboring responses. We compared the brain activity between 35 and 160 ms from each stimulation rate, by compressing the data from each subject into a two-dimensional space with latency and stimulation rate making up the two axes. Each point in this space represents a spatial pattern of magnetic activity recorded at a specific latency and stimulation rate, i.e. a vector in a 111-dimensional space
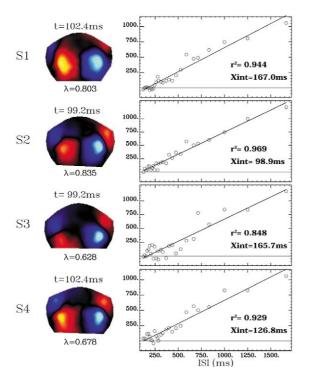
Fig. 3. The first mode of the KL decomposition of the peak N1m response patterns at all stimulation rates. The common latency at which the decomposition was performed was determined from the peak response time at the slowest rate. The high λ-values beneath each topographic map reflect the large portion of the signal accounted for by the first mode. A linear regression of the amplitudes of the first mode on stimulation ISI (solid line) revealed high r-squared values. The x intercept (Xint) represents the predicted ISI at which the N1m reaches zero amplitude. The y-axis scale is in arbitrary units produced by the decomposition procedure.

formed by the sensors. In order to find regions of high point density in this space (which correspond to well defined evoked responses in the latency-rate space), we applied a variant of the standard k-means clustering algorithm (Aldenderfer and Blashfield, 1984). The k-means algorithm finds the centers of regions of high point density (clusters) in the above vector space and assigns points in the neighborhood of these centers to the corresponding clusters. The procedure starts by choosing random initial cluster centers, and then assigning the vectors that are closest to each center to the corresponding cluster. The data within each cluster are then averaged to find the new cluster center. This process is repeated until it becomes stationary (i.e. the cluster membership does not change between iterations). The k-means algorithm requires a proximity measure between the objects to be clustered, for which we used the angle α between the high dimensional vectors.

$$\alpha = \arccos \frac{\vec{a} \cdot \vec{b}}{|\vec{a}||\vec{b}|}$$

where $\vec{a}$ and $\vec{b}$ are 111-dimensional vectors.

The angle measure was chosen as opposed to Euclidean distance because, similar to spatial correlation, angle depends only on shape and is independent of amplitude. Thus low and high amplitude patterns of the same shape will be clustered together. By varying the number of clusters from 1 to 7, we determined that two clusters produced the best segmentation of the data for all subjects. Using more than two clusters produced either redundant cluster centers or clusters with very few members. We modified the k-means algorithm so that after a few initial iterations only
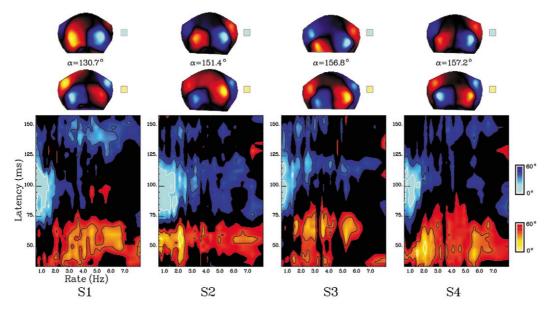

Fig. 4. Results from the cluster analysis procedure applied at all rates and latencies from 35 to 160 ms. The top two rows show the patterns corresponding to the cluster centers, and the angle α between them, for all subjects. Below these images are plots showing regions in a latency-rate space where patterns exist that have an angle less than 60° from the cluster centers. The color represents these angles, with shades of blue and red/yellow corresponding to angles from clusters one and two, respectively (see bars at right). A linear interpolation was applied between values at neighboring latencies and rates. Contour lines show the areas within 40° and 20° of the cluster centers (black and white lines, respectively).
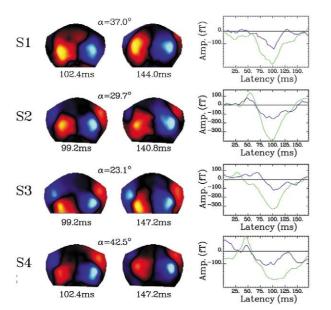
Fig. 5. (Left) Average topographic maps for 100 and 150 ms responses for each subject at a single latency for each response. Rates were included in the average only if they were within the N1m cluster center from Fig. 4. The exact latency used for each average is shown below each map, with the angle α between the spatial patterns indicated above. (Right) Time series showing the subpeaks of the N1m at 100 and 150 ms at stimulation rates of 0.6 Hz (green) and 1.7 Hz (blue) from a channel containing a large amplitude N1m response. The separate 150 ms response is visible at both rates, but becomes more apparent at the higher rates.

patterns with an angle of less than 45° from the old cluster center were used to calculate the new mean, thereby allowing us to focus on the most dominant activity patterns.

The spatial patterns of the final cluster centers are shown in Fig. 4 (top) for all subjects. Below these patterns are plots of the latency-rate space showing regions of similarity to the cluster centers with stimulation rate on the horizontal axis and latency from tone onset on the vertical axis. The colors represent angles less than 60° from the cluster centers, with shades of blue and red/yellow as indicated in the color bars (right) representing clusters 1 and 2, respectively. The colors were interpolated in order to fill out the space. We chose 60° as a threshold because it is less than half the angle between each pair of cluster centers and corresponds to a spatial correlation of 0.5 (see α values in the figure for the angles between the patterns). The top cluster center for each participant corresponds to the N1m response, with the dipolar structure similar to those from the KL procedure in Fig. 3. Consistent with this characterization, the blue region of similarity to this cluster includes latencies near 100 ms, at least for low stimulation rates. The red region of patterns in the second cluster occurs, in each case, earlier than the first cluster at a latency of about 50 ms. The pattern of neural activity represented by the second cluster center contains bilateral dipolar fields with a nearly opposite orientation to the N1m. This response more than likely corresponds to the first long latency component known as the P1/P1m in the
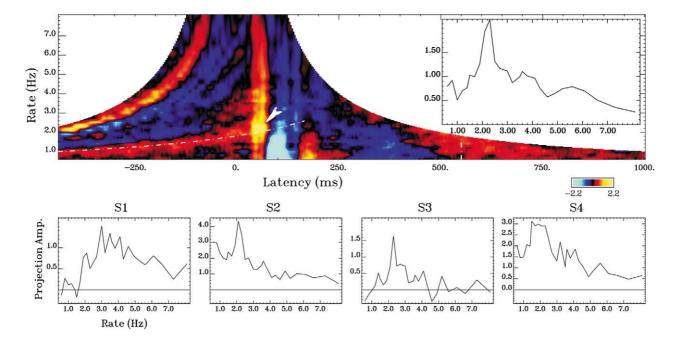

Fig. 6. Average of the projections of the recorded activity from all subjects onto the spatial patterns of their 50 ms responses from Fig. 4. (Top) An interpolated color image of the projection. The average amplitude of the 50 ms response vector is normalized to one, with the maximum amplitude of the projection colored yellow. Any amplitude less then the minimum value on the color scale is colored light blue. The plot is centered at tone onset, and the borders with the white region on the right and the left indicate the preceding and succeeding tone onsets. The dashed white line at right highlights the 500 ms response to the center tone. A dashed line at left tracks the 500 ms response to the previous tone, with an arrow indicating the region of interaction between the 500 and 50 ms responses. Inset: Amplitudes of the projection at the peak latency of the average 50 ms response exhibiting a pronounced peak when the 500 ms response interacts with the P1m around 2 Hz. Rate is shown on the *x* axis, and projection amplitude is on the *y* axis. (Bottom) Individual plots of P1m amplitudes for all subjects.

EEG/MEG literature (Näätänen and Winkler, 1999; Picton et al., 1974) and can also be seen in the single channel time series in Fig. 2. In the literature the same response has been grouped with the middle latency components and referred to as the Pb/Pbm or P50/P50m (Mäkelä et al., 1994; Yoshiura et al., 1996). For the purposes of this paper, we will consider it a long latency response, and call it the P1m.

The extent of the N1m response in the latency-rate space presents a complex pattern. The response activity near 100 ms does reduce at higher rates, although the clustering procedure found regions of similarity to the N1m at latencies other than 100 ms. The most prominent activity appears near 150 ms in the latency-rate plots of Fig. 4. Each subject shows coherent horizontal bands proximal to the first cluster at this latency. Since this activity is most coherent at higher frequencies, it could represent a latency shift of the N1m from 100 to 150 ms. Fig. 5 shows topographical images of the activity patterns at 100 and 150 ms latencies for each subject. These were generated by averaging the data from all stimulation rates that clustered with the N1m at these latencies. The results show that the 150 ms response has a spatial pattern differing by 30° or more from the main response at 100 ms for 3 of the 4 subjects. Because of this difference, the activity at 150 ms represents a separate response instead of a temporal shift of the 100 ms response. Evidence for a separate response is further enhanced by the single channel time series shown at right. For each subject, data are plotted at stimulation rates of 0.6 and 1.7 Hz from a channel exhibiting both a large N1m and a 150 ms response. At 0.6 Hz the main N1m deflection extends up to 150 ms, but at 1.7 Hz the response has diminished in both amplitude and temporal extent. At the higher frequency there is a distinct separate peak near 150 ms which exists alongside the dominant 100 ms response. This peak is also evident at lower rates (arrows in Fig. 2), suggesting that the main N1m component partially obscures the response at lower stimulation rates. This may explain why the clustering procedure shows little evidence of a separate 150 ms response at lower rates.

Earlier we hypothesized that a rate dependent reduction in the temporal extent of the 100 ms response would reveal other responses that were not clearly visible at low stimulation rates. This is apparent for the N1m subcomponent discussed earlier, but is also true for the earlier P1m response. In the latency-rate plots of subjects 1, 3, and 4, this response is either not apparent at the lower stimulation rates, or only occurs for a few milliseconds. As the temporal extent of the N1m reduces at higher rates, the 50 ms response extends into latencies previously occupied by the N1m. The increased temporal extent of the 50 ms response may indicate that the high amplitude 100 ms response masked the P1m at lower stimulation rates.

3.4. Overlap of responses to successive tones

How is the response to a given stimulus influenced by the response to the previous tone? Since the first major response in each subject is the P1m at 50 ms latency, we searched for evidence of late responses from the previous tone interfering with this response. We searched for this interference by using a projection of the entire dataset onto the spatial pattern of the 50 ms response. This proved to be a useful means of data compression by providing a single measure of similarity or dissimilarity to the P1m, thus emphasizing long latency responses that could positively or negatively interfere with the response. The measure also took full advantage of the high degree of spatial information available from all the sensors, which allowed for observation of low amplitude long-latency responses that were difficult to observe in single channels. Fig. 6 shows the average results for all subjects. The projection amplitudes A are from the formula

$$A = \frac{\vec{a} \cdot \vec{b}_{ij}}{\vec{a} \cdot \vec{a}}$$

where $\vec{a}$ is the vector representing the spatial pattern of the P1m, and $\vec{b}$ is the spatial pattern of activity recorded at latency i and stimulation rate j. The spatial pattern of the P1m used as a reference is from the clustering results shown in Fig. 4. The color plot in Fig. 6 shows the average of the projection data from all subjects at all stimulation rates. The intent is to show the evolution and onset of response overlap. Projection amplitudes are represented by a color scale, with colors interpolated between neighboring latencies and rates. Positive amplitudes are plotted in red/yellow indicating a high degree of similarity to the P1m; negative amplitudes are colored blue and indicate dissimilarity. The plot is centered at tone onset, with a full inter-stimulus interval presented before and after the tone at all but the very slow rates. Borders with the white regions at the left and the right indicate the previous and next tone onsets. After the center tone onset, the P1m and N1m appear as red and blue vertical stripes. The N1m appears as a negative deflection since the spatial pattern of the response is nearly opposite in polarity to the P1m (see Fig. 4). Several other long latency responses follow the N1m, including a broad region of positivity beginning near 500 ms, indicating a response with a similar spatial pattern to the P1m. The same long latency responses are seen emanating from the previous tone onset at left. A dashed line tracks the previous 500 ms response. As stimulation rate increases to 2 Hz, activity still present from the previous tone begins to overlap with the response to the central tone. When the 500 ms response overlaps the P1m, the spatially similar responses enhance each other. This is indicated by the increase in projection amplitude within the region of interaction (see arrow in Fig. 4).

The inset plot at the top in Fig. 6 highlights the affect of response overlap on the P1m. The graph shows the amplitude of the projection at the peak latency of the average P1m as a function of stimulation rate. This amplitude peaks around 2 Hz, which is when the 500 ms response has maximum overlap with the P1m. The plots at bottom show the individual P1m amplitudes for each subject. All subjects

show a spike in P1m amplitude near 2 Hz, with this being the maximum amplitude of the response in 3 out of 4 subjects. For the subjects who show a P1m response at all rates (2 and 4) the amplitude spike appears to interrupt a general trend of rate dependent reduction of the response. In the color plot, the region of interaction between the 500 ms response and the P1m also reveals a slight latency increase of the P1m and a delayed onset of the N1m. This implies that the constructive interference between the spatially similar 500 ms and P1m responses also coincides with destructive interference with the dissimilar N1m.

4. Discussion

The intention of this experiment was to characterize the cortical auditory response to rhythmic stimuli within a parameter range that is relevant for studies of perception and behavior. The relatively simple yet systematic design used in this study may serve as a foundation for more complex studies in areas such as music and speech perception and sensorimotor coordination. Several studies have investigated the rate dependency of the N1m at rates slower than used in this experiment (Lü et al., 1992; Mäkelä et al., 1993; Sams et al., 1993). Among these, Lü et al. (1992) used ISIs covering the range from 0.8 to 16 s. In contrast to the linear dependence found here, Lü et al. (1992) modeled the ISI dependency of the main component with a function of the form $A(1 - e^{-(t-t_0)/\tau})$, with amplitude A, ISI t, time of decay onset t_0, and lifetime τ. This function becomes asymptotic at high ISIs, accounting for the fact that the amplitude increase levels off with at ISIs above 10 s. Sams et al. (1993) used a similar exponential function to model the rate dependence of the N1m using ISIs ranging from 0.75 to 12 s. We found that the N1m is linearly dependent on inter-stimulus interval over the range of stimulation rates from 0.6 to 8.1 Hz. The linear dependence of the N1m in the present experiment should not be surprising, however, since the ISIs used here are in the linear range of the previous exponential models, before the function approaches an asymptote at high ISIs.

One of the predictions of the Lü et al. (1992) exponential model of the N1m is that the component disappears at an ISI equivalent to the duration of the stimulus. Thus, the evoked response is considered refractory with respect to tone offset. To our knowledge, our experiment is the first to test this prediction at ISIs shorter than the 0.8s used in Lü et al. (1992). For all subjects the N1m disappears at higher ISIs than the tone offset at 60 ms, and in fact for 3 of the subjects the ISI is more than twice that value (Fig. 3). It is possible, however, that the increase of the zero amplitude times was due to interference from the 50 ms response. The present results (Fig. 4) show that the 50 ms response for 3 of our subjects eventually extends very close to 100 ms. These data suggest that as the N1m reduces in amplitude, the 50 ms response may become a greater part of the signal at 100 ms

and thus interfere with the remaining low amplitude N1m. Interference from activity originating from the previous tone could also have caused difficulty in observing the N1m at higher stimulation rates. Overlap between responses to successive tones was present in all subjects at stimulation rates above 2 Hz.

In an EEG study using stimulation rates from 0.5 to 10 Hz, Erwin and Buchwald (1986) found that the 50 ms P1 response reduces in amplitude with increased stimulation rate. For the subjects showing a P1m at all rates in Fig. 6, the response shown is also attenuated by increasing rate, although the general pattern is interrupted by a large amplitude spike near 2 Hz from interference with the spatially similar 500 ms response to the previous tone. Obtaining an accurate model of the rate dependence of the P1m over this range of stimulation rates is made difficult by this potential for interference. One possible method for overcoming this problem is the maximum length sequence (MLS) method employed by Picton et al. (1992). The MLS consists of a specially constructed sequence of inter-stimulus intervals designed to average out the effect of response overlap. Using this method, the authors confirmed that the P1 is attenuated by increasing rate. Besides interference from previous responses, the P1m may also be affected by overlap from the large amplitude N1m response at low stimulation rates. This was apparent in the results of the clustering procedure (Fig. 4) in at least 3 subjects. Interference from either the N1m or activity induced by the previous tone may explain why some studies report an increase rather than a decrease of P1m amplitude with increased stimulation rate (see for example Alain et al., 1994).

In this study, a 150 ms response was also obscured by the main N1m component at low stimulation rates. Because of its temporal and spatial proximity to the N1m, this response is best described as a subcomponent of the main N1m response. Multiple subcomponents of the N1/N1m have been observed in previous EEG and MEG studies at latencies ranging from 75 to 170 ms (for reviews see Näätänen and Picton, 1987; Woods, 1995). However, due to its relative insensitivity to radial sources, fewer components are typically observed with MEG. Some of the studies of N1m rate dependence report a secondary N1m response, although none at a latency of 150 ms (Lü et al., 1992; Sams et al., 1993). Mäkelä et al. (1993) found that as stimulation rate increased to 1 Hz, a single source was no longer adequate to model the N1m. This coincides with the pattern of emergence observed here in which a second component became apparent at higher rates. However, few studies of N1m rate dependence have used rates in the range of the present study. Although not a study of rate dependence, Loveless et al. (1996) used pairs of stimuli in a similar range of ISIs as studied here and found a similar N1m subcomponent. The authors used ISIs ranging from 70 to 500 ms, and showed that the last tone in the pair produced separate N1m components at 100 and 150 ms. As shown here, the second component was significantly stronger at shorter ISIs.

By the definition used here, the onset of overlap between the transient responses to successive tones marks the beginning of the continuous steady-state response (Regan, 1982). Here we found this onset at a stimulation rate of 2 Hz. The steady-state response, however, is traditionally recorded at much higher stimulation rates (~10–80 Hz). As reported earlier, at 40 Hz the steady-state response forms a near sinusoidal oscillation at the stimulation frequency, which is a much simpler response than the time series found here at 2 Hz (see Figs. 2 and 6). The source of the high-frequency steady-state response is still under debate, with argument centering on whether the response results from overlap of transient responses, or phase locking of ongoing cortical rhythms, or some combination of the two (Azzena et al., 1995). Due to their rate dependence, long latency responses are not considered to make up a large part of the response (Hari et al., 1989). This contrasts with the onset of the low-frequency SSR seen here, which comes in the form of overlapping long latency responses. The SSR can also be thought of as an oscillatory response that has a resonant frequency near 40 Hz (Galambos et al., 1981). Using amplitude modulation of a simple tone, Picton et al. (1987) showed that the SSR actually has two resonant regions – one from 20 to 50 Hz, and another from 2 to 5 Hz. Based on these results, it may be possible to divide the steady-state response into distinct low- and high-frequency regimes.

Several studies have recorded middle latency responses at stimulation rates from 8 to 10 Hz to test the idea that the 40 Hz response results from overlap of these responses (Azzena et al., 1995; Gutschalk et al., 1999; Hari et al., 1989). The rates used here bordered this range and provided strong evidence that LLRs from previous tones could exist within the latency range of the middle latency response to the present tone (between 10 and 50 ms). We show evidence for a 150 ms N1m subcomponent that is observable up until 8.1 Hz (Fig. 4). At a stimulation rate of 8 Hz, a 150 ms response lands precisely at 25 ms past the next tone onset. Therefore, this response could interact with the MLRs recorded at these rates, and cause discrepancies when trying to account for the 40 Hz response by simple summation. If residual LLRs can be seen at 8.1 Hz, then the possibility exists that long latency responses contribute to the SSR at 40 Hz as well. In accordance with this idea, Santarelli et al. (1995) found responses in the latency range of the LLRs after the last click of a 40 Hz click train, suggesting that long latency responses are evoked by each click.

The present results suggest several issues for future study. Among these is determining the characteristics of the observed activity at 500 ms past tone onset. Standard transient responses, recorded with long randomized ISIs, typically do not contain a coherent response at such a long latency (Picton et al., 1974). Possibly, 500 ms activity only occurs during rhythmic stimulation. One method to clarify this issue would be to present stimuli with slightly variable ISIs. If the 500 ms response is induced by rhythmic stimuli only, it should disappear with enough variability in

the tone sequence. The present experiment could also be repeated, but with a slightly higher range of ISIs. The 500 ms response might disappear as the ISI increases beyond the theoretical limit of rhythm perception. Fraisse (1982) reports that this upper limit is near 1800 ms, which is close to the longest ISI used in this experiment. Future work might also determine if the onset of continuous activity between tones near 2 Hz has any significance for the processing of auditory information. Results from previous behavioral studies suggest that this possibility may deserve further investigation (see Gura, 2001). For example, if subjects tune a metronome to a tempo that sounds the most natural, they consistently prefer a rate very close to 2 Hz (Fraisse, 1982). In addition, the just noticeable difference between two different rates of isochronous tone stimulation is lowest within a region centered at 2 Hz (Drake and Botte, 1993). Auditory stimulation at 2 Hz is significant in sensorimotor coordination experiments as well. When subjects flex their index finger between successive tones (syncopate), they find the task increasingly difficult as the stimulation rate is increased. Beyond a critical point they spontaneously switch to synchronizing movement with the metronome (Kelso et al., 1990). In unpracticed subjects, this transition typically occurs at movement rates approaching 2 Hz and may be extended to higher frequencies with practice (Jantzen et al., 2001). The change from a transient to a steady-state response occurs when discrete cortical activity overlaps to form a continuous sensory/perceptual representation. The aforementioned perceptual and behavioral phenomena may well be linked to this qualitative transition in the nature of the auditory response.

Acknowledgements

References

Alain C, Woods DL, Ogawa KH. Brain indices of automatic processing. Neuroreport 1994;6:140–144.

Aldenderfer MS, Blashfield RK. Cluster analysis, Newbury Park: Sage Publications, 1984.

Azzena GB, Conti G, Santarelli R, Ottaviani F, Paludetti G, Maurizi M. Generation of human auditory steady-state responses (SSRs). I: stimulus rate effects. Hear Res 1995;83:1–8.

Carver FW, Fuchs A, Mayville JM, Davis SW, Kelso JAS. Systematic investigation of the human brain's response to rhythmic auditory stimulation. Dyn Neurosci 1999;VII.

Drake C, Botte MC. Tempo sensitivity in auditory sequences: evidence for a multiple-look model. Percept Psychophys 1993;54:277–286.

Erwin RJ, Buchwald JS. Midlatency auditory evoked responses: differential recovery cycle characteristics. Electroenceph clin Neurophysiol 1986;64:417–423.

Fraisse P. Rhythm and tempo. In: Deutsch D, editor. The psychology of music, New York: Academic Press, 1982. pp. 149–180.

Fuchs A, Kelso JAS, Haken H. Phase transitions in the human brain: spatial mode dynamics. Int J Bifurc Chaos 1992;2:917–939.

Galambos R, Makeig S, Talmachoff PJ. A 40-Hz auditory potential recorded from the human scalp. Proc Natl Acad Sci USA 1981;78:2643–2647.

Gura T. Rhythm of life. New Sci 2001;171:32–35.

Gutschalk A, Mase R, Roth R, Ille N, Rupp A, Hähnel 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;110:856–868.

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. Electroenceph clin Neurophysiol 1982;54:561–569.

Hari R, Hämäläinen M, Joutsiniemi SL. Neuromagnetic steady-state responses to auditory stimuli. J Acoust Soc Am 1989;86:1033–1039.

Jantzen KJ, Fuchs A, Mayville JM, Deecke L, Kelso JAS. Neuromagnetic activity in alpha and beta bands reflect learning-induced increases in coordinative stability. Clin Neurophysiol 2001;112:1685–1697.

Kelso JAS, DelColle JD, Schöner G. Action-perception as a pattern formation process. In: Jeannerod M, editor. Attention and performance XIII. Hillsdale, NJ: Erlbaum, 1990. pp. 139–169.

Kelso JAS, Bressler SL, Buchanan S, DeGuzman GC, Ding M, Fuchs A, Holroyd T. Cooperative and critical phenomena in the human brain revealed by multiple SQuIDs. In: Duke D, Pritchard W, editors. Measuring chaos in the human brain, Teaneck, NJ: World Scientific, 1991. pp. 97–112.

Kelso JAS, Bressler SL, Buchanan S, DeGuzman GC, Ding M, Fuchs A, Holroyd T. A phase transition in human brain and behavior. Phys Lett A 1992;169:134–144.

Loveless N, Levänen S, Jousmäki V, Sams M, Hari R. Temporal integration in auditory sensory memory: neuromagnetic evidence. Electroenceph clin Neurophysiol 1996;100:220–228.

Lü ZL, Williamson SJ, Kaufman L. Human auditory primary and association cortex have differing lifetimes for activation traces. Brain Res 1992;572:236–241.

Mäkelä JP, Ahonen A, Hämäläinen M, Hari R, Ilmoniemi M, Kajola M, Knuutila J, Lounasmaa OV, McEvoy L, Salmelin R, Salonen O, Sams M, Simola J, Tesche C, Vasama JP. Functional differences between auditory cortices of the two hemispheres revealed by whole-head neuromagnetic recordings. Hum Brain Mapp 1993;1:48–56.

Mäkelä JP, Hämäläinen M, Hari R, McEvoy L. Whole-head mapping of middle-latency auditory evoked magnetic fields. Electroenceph clin Neurophysiol 1994;92:414–421.

Mayville JM, Fuchs A, Ding M, Cheyne D, Deecke L, Kelso JAS. Event-related changes in neuromagnetic activity associated with syncopation and synchronization timing tasks. Hum Brain Mapp 2001;14:65–80.

Meyer-Lindenberg A, Ziemann U, Hajak G, Cohen L, Berman KF. Transitions between dynamical states of differing stability in the human brain. Proc Nat Acad Sci 2002;99:10948–10953.

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

Näätänen R, Winkler I. The concept of auditory stimulus representation in cognitive science. Psychol Bull 1999;125:826–859.

Pantev C, Bertrand O, Eulitz C, Verkindt C, Hampson S, Schuirer G, Elbert T. Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. Electroenceph clin Neurophysiol 1995;94:26–40.

Pantev C, Roberts LE, Elbert T, Roß B, Wienbruch C. Tonotopic organization of the sources of human auditory steady-state responses. Hear Res 1996;101:62–74.

Picton TW, Hillyard SA, Krausz HI, Galambos R. Human auditory evoked potentials. I: evaluation of components. Electroenceph clin Neurophysiol 1974;36:179–190.

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

Picton TW, Champagne SC, Kellett AJC. Human auditory evoked potentials recorded using maximum length sequences. Electroenceph clin Neurophysiol 1992;84:90–100.

Regan D. Comparison of transient and steady-state methods. Ann N Y Acad Sci 1982;388:45–71.

Regan D. Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine, New York: Elsevier, 1989.

Sams M, Hari R, Rif J, Knuutila J. The human auditory sensory memory trace persists about 10 s: neuromagnetic evidence. J Cogn Neurosci 1993;5:363–370.

Santarelli R, Maurizi M, Conti G, Ottaviani F, Paludetti G, Pettorossi VE. Generation of human auditory steady-state responses (SSRs). II: addition of responses to individual stimuli. Hear Res 1995;83:9–18.

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

Vrba J, Cheung T, Taylor B, Robinson SE. Synthetic higher-order gradiometers reduce environmental noise, not the measured brain signals. In: Yoshimoto T, Kotani M, Kuriki S, Karibe H, Nakasato N, editors. Recent advances in biomagnetism, Sendai: Tohoku University Press, 1999. pp. 105–108.

Wallenstein GV, Kelso JAS, Bressler SL. Phase transitions in spatiotemporal patterns of brain activity and behavior. Physica D 1995;20:626–634.

Woods DL. The component structure of the human auditory evoked potential. Electroenceph clin Neurophysiol Suppl 1995;44:102–109.

Yoshiura T, Ueno S, Iramina K, Masuda K. Human middle latency auditory evoked magnetic fields. Brain Topogr 1996;8:291–296.