

## Temporal dynamics of pitch in human auditory cortex

Alexander Gutschalk,<sup>a,\*</sup> Roy D. Patterson,<sup>b</sup> Michael Scherg,<sup>a</sup>  
Stefan Uppenkamp,<sup>b,c</sup> and André Rupp<sup>a</sup>

<sup>a</sup>Department of Neurology, University of Heidelberg, Heidelberg 69120, Germany

<sup>b</sup>Centre for the Neural Basis of Hearing, Department of Physiology, University of Cambridge, Cambridge CB2 3EG, UK

<sup>c</sup>Medizinische Physik, University of Oldenburg, Oldenburg, Germany

Received 30 September 2003; revised 16 January 2004; accepted 20 January 2004

Available online 13 April 2004

Recent functional imaging studies have shown that sounds with temporal pitch produce selective activation in anterolateral Heschl's gyrus. This paper reports a magnetoencephalographic (MEG) study of the temporal dynamics of this activation. The cortical response specific to pitch was isolated from the intensity-related response in Planum temporale using a 'continuous stimulation' paradigm in which regular and irregular click trains alternate without interruption. The mean interclick interval (ICI) was 6, 12, 24, or 48 ms; the train length was 720 ms. The auditory sustained field serves as a level-dependent baseline that enhances the signal-to-noise ratio over previous techniques.

The onset of pitch was accompanied by a prominent transient field, followed by a strong sustained field, both of which were associated with sources in lateral Heschl's gyrus. The sustained field rose from baseline about 70 ms after the onset of temporal regularity, asymptoted at about 450 ms, and commenced its return to baseline about 70 ms after pitch offset. The peak of the transient field occurred between 130 and 190 ms after regularity onset depending on the ICI.

The latencies of the cortical pitch response are substantially longer than might be anticipated from temporal models of pitch perception. This finding suggests that the temporal integration associated with periodicity processing occurs in a subcortical structure, and that the cortical responses reflect subsequent processes involving the measurement of pitch values and changes in pitch.

© 2004 Elsevier Inc. All rights reserved.

*Keywords:* Pitch; Auditory cortex; Auditory evoked fields; Magnetoencephalography

### Introduction

There has recently been a series of brain imaging studies of temporal pitch processing in cerebral cortex. The studies focus on broadband stochastic sounds with varying degrees of temporal

regularity, which have the advantage that both the pitch and pitch strength can be varied without affecting the average distribution of energy over frequency and time. These sounds make it possible to isolate the neural response associated with the perception of pitch, or a change in pitch, from the general response to the onset and presence of a sound. The initial studies employed regular interval sounds (RIS) in which the temporal fine structure of a random noise is regularized by making a copy of the noise, delaying it, and adding it back to the original noise, repeatedly (Yost et al., 1996). Positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) (Griffiths et al., 1998, 2001; Patterson et al., 2002) were used to show that there was selective activation for pitch in a region of Heschl's gyrus (HG), anterior and lateral to primary auditory cortex, and the level of activation increased with the degree of temporal regularity (i.e., the pitch strength). Selective activation for melody appeared in regions beyond HG in Planum polare and the superior temporal gyrus.

These studies prompted Gutschalk et al. (2002) to measure the sustained fields (SF) produced by click trains using magnetoencephalography (MEG), and to show that regular click trains, which are perceived to have a sustained temporal pitch, produce a markedly stronger SF than irregular click trains that do not have a pitch. The sensor waveform produced by a regular click train is illustrated by the bold gray line in the left panel of Fig. 1, while that produced by an irregular click train is shown by the thin gray line. (The figure introduces the fields and sources that are the subject of this paper along with the terminology.) After click train onset, the auditory evoked field (AEF) begins with a succession of transient fields (TF) consisting mainly of the peaks P1m, N1m, and P2m. The SF overlaps with these peaks at its onset around 300 ms, after which it rises to steady state around 400 ms, where it stays until the stimulus ends. The analysis of the data revealed that the sensor waves can be explained by the combination of an 'anterior source' in lateral HG and a 'posterior source' in Planum temporale (PT). Source waves for regular and irregular click trains are illustrated in the right-hand column of Fig. 1. The strength of the SF in the anterior source varied with the degree of temporal regularity in the click train, but the source was insensitive to sound level over a wide range. The SF of the posterior source varied with level but was insensitive to temporal regularity. The double dissociation of responses made it possible to separate these

---

\* Corresponding author. Department of Neurology, University of Heidelberg, Im Neuenheimer Feld 400, 69120 Heidelberg, Germany. Fax: +49-6221-56-5258.

E-mail address: alexander\_gutschalk@med.uni-heidelberg.de (A. Gutschalk).

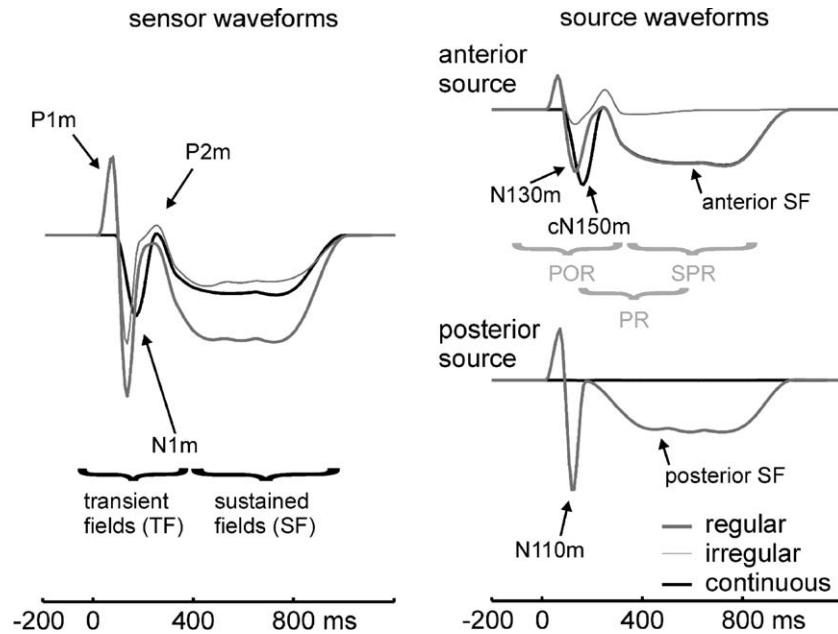


Fig. 1. Summary of the nomenclature used in this paper to describe the onset and sustained responses to click trains. For simplicity, off-responses are excluded from this scheme. Following Picton et al. (1995), we simply assign consecutive integers to the peaks of the transient fields when referring to the sensor waveforms, or when the generator is not known (left-hand panel). The P (positive) or N (negative) refers, by convention, to the polarity of the wave at the vertex electrode in EEG. The peaks in the source waveforms (right-hand panel) are designated by their latency in ms (Scherg, 1990). N and P now refer to the polarity of the wave at the cortical surface, which largely coincides with the vertex polarity in the case of auditory cortex. This specification allows us to distinguish the source components (e.g. N130, N110) that merge to form the N1m wave of the sensor data. The prefix 'c' indicates that data were obtained with the continuous paradigm. In these (black) waveforms, no P1m is observed and activity in the posterior source is avoided. The negative transient (N130m, cN150m) is referred to as the 'pitch onset response' (POR; Krumbholz et al., 2003). Together with the 'sustained pitch response' (SPR), it forms what might be referred to as the 'pitch response' (PR).

relatively close sources with confidence, and once separated, it became clear that the generators of the N1m are also affected by the regularity of the click train. To differentiate the components of the N1m, we refer to them by their typical peak latency (e.g., N130m and N110m), as suggested by Picton et al. (1995).

Subsequently, Krumbholz et al. (2003) used MEG to demonstrate that the perception of the onset of pitch was accompanied by a transient, surface-negative magnetic field whose source was also in lateral HG. This transient field was isolated by appending a spectrally matched noise to the front of a RIS. The pitch onset occurs without a concomitant change in level, which avoids the production of the TFs associated with changes in intensity (Biermann and Heil, 2000; Mäkelä et al., 1988). The latency of the peak of the TF was in the range 100–200 ms and it varied inversely with the pitch of the sound. Moreover, the amplitude of the TF varied directly with the pitch strength, indicating that the TF is, indeed, associated with the onset of the perception of pitch. They called this TF a pitch onset response (POR).

In this paper, we use MEG and continuous alternation of regular and irregular click trains to investigate the temporal properties of the SF and TF *simultaneously*, to compare the locations of their sources, and their relationship to other components of the AEF. If the posterior source is not affected by regularity, the SF will be the same during the regular and irregular intervals, and we can eliminate the posterior SF from the analysis by setting the measurement baseline during the irregular interval. The activity during the regular click train can then be associated with the anterior source. The AEF evoked during the regular interval of a continuously alternating click train is illustrated by the black waves in Fig. 1.

#### Models of pitch perception and models of the AEF

In the imaging studies involving RIS (Griffiths et al., 1998, 2001; Krumbholz et al., 2003; Patterson et al., 2002), the auditory image model (AIM) of Patterson et al. (1992, 1995) was used to simulate the internal representation of the sound, and to generate hypotheses about where different aspects of the processing might be in the auditory pathway. A brief description of the AIM is presented here to illustrate how time-domain auditory models can be used to interpret the transient and sustained components of the AEF and their relationships.

Time-domain models like AIM simulate the spectral analysis performed in the cochlea by the basilar membrane and outer hair cells using a bank of 75–100 band-pass auditory filters.<sup>1</sup> It is this process that creates the tonotopic dimension of auditory processing. Then, in each of the frequency channels defined by a filter, the model simulates the neural transduction performed by the inner hair cells and primary auditory fibers, using a unipolar, compressive firing mechanism that effectively records the times of the peaks in the wave flowing from that particular filter. Examples of the basilar membrane motion and neural activity patterns produced

<sup>1</sup> The low pitches produced by the broadband sounds used in the studies referred to here (e.g., Griffiths et al., 1998; Gutschalk et al., 2002) cannot be explained by the spectral analysis performed in the cochlea. The pitch appears to be based on the temporal information in the auditory nerve, which is concentrated and extracted in the ascending auditory pathway somewhere between the brain stem and auditory cortex (Griffiths et al., 2001; Langner and Schreiner, 1988; Winter et al., 2001).

in response to a natural vowel with a temporal pitch of 125 Hz are presented in Patterson et al. (1995, Fig. 2).

In the next stage of processing, the system is assumed to evaluate the time-interval information in each frequency channel by computing something like an interspike interval histogram (IIH) for the channel, and it is here that the responses to regular and irregular click trains become markedly different. The IIH evoked by a regular click train shows a pronounced peak at the period of the train in virtually every frequency channel; whereas, there are no enduring peaks in any of the IIHs for an irregular click train. The array of IIHs is called an ‘auditory image’ and it is this representation that gives the model its name. The brain imaging studies with RIS were partly motivated by the desire to delimit the location of this representation in the auditory pathway, since it plays a central role in this time-domain model of auditory perception. Examples of the neural activity patterns and auditory images produced in response to different samples of RIS are presented in Griffiths et al. (1998, Figs. 1 and 2, 2001, Fig. 1), Patterson et al. (2002, Fig. 1) and Krumbholz et al. (2003, Figs. 1 and 2).

In the MEG study of Gutschalk et al. (2002), AIM was used to illustrate the neural activity patterns and auditory images produced by regular and irregular click trains (their Fig. 1). There is a vertical ridge of activity in the auditory image of the regular click train, centered on its ICI (5 ms), and no corresponding feature in the auditory image of the irregular click train. Gutschalk et al. (2002) conclude that the pitch-related SF observed in lateral HG represents a process that follows the construction of the auditory image, and that it might represent the integration of pitch information across frequency channels, and/or the calculation of the specific pitch value. Patterson et al. (2002) argue that the fMRI activity produced by RIS in lateral HG leads to similar conclusions.

AIM preserves the temporal information necessary to explain the transient components of the AEF as well as the SF, as would virtually any time-domain model of pitch perception, and so these models also provide a basis for interpreting the dynamics of the AEF. Patterson et al. (1992) argued that the dynamics of auditory perception indicate that activity in the auditory image builds up and decays exponentially, with a half life of 30 ms—a time constant that is much longer than those observed in the filtering and neural transduction processes. The black curve in Fig. 2 shows the height of the vertical ridge in the auditory image, as a function of time, in response to a regular click train with a duration of 720 ms and an ICI of 12 ms (a pitch of 83 Hz). The ordinate is inverted so that the curve goes down when pitch strength *increases*, to facilitate comparison with the MEG response in lateral HG. The height of the pitch ridge increases (in the negative direction) over the first 100 ms of the sound and then asymptotes to a fixed level that is sustained for the duration of the sound. When the regular click train terminates, the height of the ridge decays back to baseline at the same rate as it grew when the pitch came on. This suggests that there should be a source that produces a SF with the time course resembling the black curve in Fig. 2, in the region of the auditory pathway associated with the auditory image. AIM does not immediately predict that there should be transient fields like those that appear in the AEF. It would be relatively simple, however, to extend the model to include a process that monitors change in the SF, and the output of such a process would be expected to resemble the derivative of the SF function, which is shown by the dashed line in Fig. 1. Thus, the dashed line shows the polarity and form of the TFs that might be expected to accompany the predicted SF.

It is clear from the outset that such a simple model of auditory processing is highly unlikely to explain the complicated dynamics of the AEF. Nevertheless, as we show in the discussion, the model does

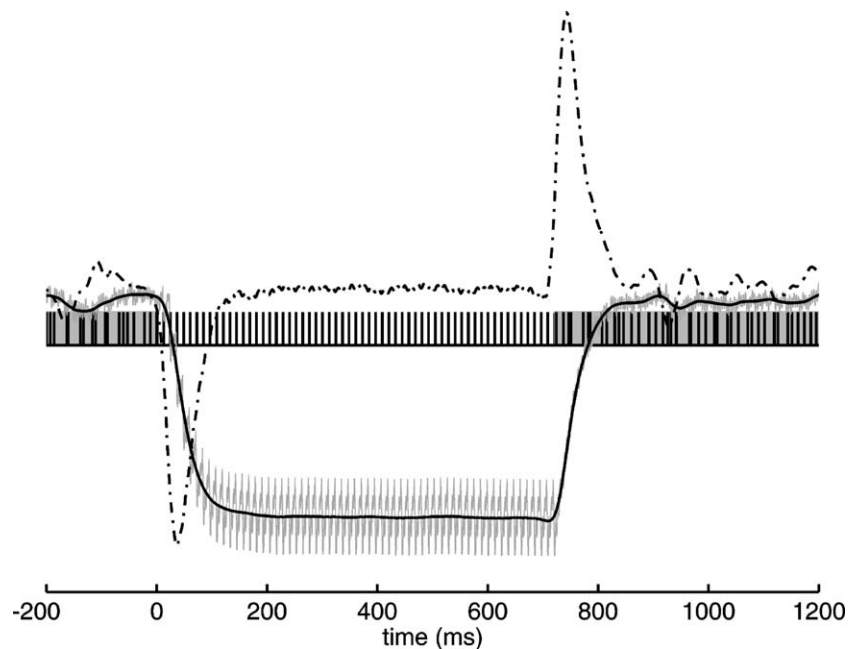


Fig. 2. Temporal development of regularity representation in the auditory image model (Patterson et al., 1995). The stimulus is a regular click train with an ICI of 12 ms and a length of 720 ms that is embedded in an irregular click train (shaded) with the same click density. To derive a flat level during the irregular trains, the simulation for 10 different irregular trains was averaged. The gray curve represents the strength of the auditory image buffer at a delay of 12 ms summed over all channels. The black line is a lowpass-filtered version of the same curve (20 Hz, Butterworth, zero phase shift). To get an estimate for a transient response, we calculated the first derivative of the black curve, which is represented by the dashed graph.

constrain the interpretation of the transient and sustained components of the AEF and enables us to use MEG data to study auditory processing in considerable detail. Specifically, comparison of the TFs and SFs evoked by regular and irregular click trains appears to confirm that, if there is an auditory image like that proposed in AIM, then it is in a subcortical structure (or maybe in A1), and the processing in lateral HG represents subsequent processing of pitch-related features after they appear in the auditory image.

## Materials and methods

The SF involves at least two sources (Gutschalk et al., 2002); the TF involves multiple generators that overlap the SF in time (e.g., Gutschalk et al., 1998; Loveless et al., 1996; Lü et al., 1992; Scherg et al., 1989). In a volume conductor like the brain, adjacent current sources blend and, as a result, spatiotemporal source analysis cannot normally separate more than two or three processes within the confines of auditory cortex (Scherg, 1990). Thus, it is important to use an experimental paradigm that limits the number of sources involved as much as possible. The SF observed by Gutschalk et al. (2002) in Planum temporale depends on stimulus intensity *but not* on stimulus regularity. Accordingly, we can assume that continuous stimulation without changes in bandwidth or intensity fixes this component of the SF, enabling us to use it as a baseline. Moreover, the paradigm allows us to avoid the part of the onset response that depends on the acceleration of peak pressure (Biermann and Heil, 2000; Krumbholz et al., 2003; Mäkelä et al., 1988).

This paper reports data obtained in two experimental sessions. Experiment 1 introduces the continuous presentation paradigm of regular and irregular click trains and compares it with the presentation of stimuli in silence. Experiment 2 uses the continuous stimulation paradigm to investigate the temporal dynamics of the regularity specific auditory evoked fields and its dependence on the click rate.

### Listeners

Eighteen listeners (10 female, 8 male) participated in experiment 1, and 12 listeners (6 female, 6 male) participated in experiment 2. The mean age was 28.7 and 28.0 years in experiments 1 and 2, respectively. Five of the listeners participated in both experiments. All of the listeners were right handed and had no history of peripheral or central hearing disorder. They provided informed consent before participating in the experiments.

### Stimuli

Regular and irregular click trains of 720-ms duration were presented at an intensity of 28 dB relative to threshold for click trains with a 12-ms interclick interval (ICI). The sampling rate was 48 kHz and the duration of the individual clicks was the minimum 1 sample (20.83  $\mu$ s). The stimuli were presented diotically using ER-3 transducers (Etymotic Research, Inc.) with 90-cm plastic tubes and foam ear pieces. These transceivers act like bandpass filters with an effective band width from approximately 500–3000 Hz.

In experiment 1, the ICI for the regular click trains was 12 ms; the irregular trains had the same average rate of clicks but the position of each click in the series was uniformly jittered  $\pm 6$  ms. Two conditions were compared: First, 300 regular and irregular click trains were presented with an intertrain interval (ITI) selected

at random from 600 to 1000 ms (Fig. 2). The order of the regular and irregular trains was randomized; that is, the click train was presented once or twice before switching to the other condition. In the second condition, the regular and irregular click trains were presented in alternation without any silent interval in between (Fig. 3B). The sessions contained either 200 repetitions (first six listeners) or 400 repetitions.

In experiment 2, the click trains were presented continuously, alternating from regular to irregular every 720 ms. The ICI was 6, 12, 24, or 48 ms; the intensity of the clicks was held constant in all conditions. Click trains were either perfectly regular or jittered uniformly over  $\pm$  one-half of the ICI; that is, over  $\pm 3, 6, 12,$  or  $24$  ms for the respective conditions. This is the maximum range of jitter that avoids the problem of successive intervals overlapping. The important properties of this method of jitter are that it produces an irregular click train with the same energy and average click rate as the corresponding regular click train, but without the buzzy pitch produced by the regular click train. These jitter conditions are used to calculate the nonpitch baseline. Every 25 repetitions, the ICI changed to the next condition; 200 repetitions of each condition were presented ( $8 \times 25$ ). For the regular click trains, the precise length from the first to the last click was 714, 708, 696, and 672 ms for ICIs of 6, 12, 24, and 48 ms, respectively.<sup>2</sup> The mean length of the irregular trains had the same duration for each ICI.

The 12-ms ICI was also used in two additional experiments which are not reported here in detail but which were performed in the same recording session. Together, the 12-ms conditions of all three experiments provide an average with 600 replications for this standard condition. This grand average was used to establish the source locations for experiment 2, since it has the best signal to noise ratio. The average was also used for a more detailed analysis of the transient and sustained fields presented in this paper.

### Recording and data processing

During the recording session, listeners watched a self-selected silent movie in an attempt to maintain their vigilance at a relatively steady level. The magnetoencephalogram was recorded continuously with a Neuromag-122 whole-head MEG system (Neuromag, Helsinki). The sampling rate was 1000 Hz with a 330-Hz lowpass filter and no highpass filter (direct coupled). The data were averaged offline with the BESA2000 software (MEGIS, Grädfelfing); artifact-contaminated epochs were rejected by an automatic gradient criterion and by visual inspection. In the noncontinuous condition, the length of the average epoch was from 400 ms before to 1600 ms after stimulus onset. The baseline was derived from the interval 100 ms before stimulus onset. In the continuous condition, the baseline was derived from the final 100 ms of the irregular stimulus, preceding regularity onset. The data were lowpass-filtered at 20 Hz (zero-phase shift Butterworth filter, 24dB/oct.) if not otherwise indicated.

T1-weighted magnetic resonance images (MRI) were obtained from each listener on a 1.5-T, Siemens Symphony MRI-scanner. Scans were performed in 176 sagittal slices yielding an isotropic voxel size of 1 mm<sup>3</sup>. Dipole locations were coregistered on

<sup>2</sup> The *alternation interval* is always 720 ms. The interclick interval at the transition belongs half to the regular and half to the irregular train, or to neither, or to both, which is why the *CT duration* and the *alternation length* are sometimes different.

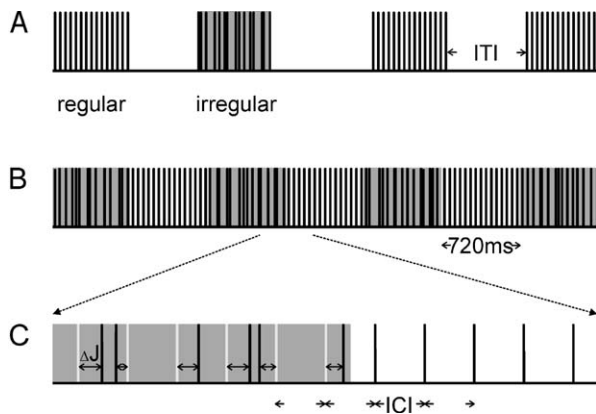


Fig. 3. Summary of the stimuli used in experiment 1. (A) The noncontinuous condition comprised regular and irregular click trains that were repeated with a variable intertrain interval (ITI) of 0.6–1.0 s. (B) In the continuous condition, the regular and irregular trains were presented in constant alternation between regular and irregular. (C) Within the irregular click trains, clicks were jittered around the regular position by  $\Delta I$ , which was in the range of  $\pm 1/2$  ICI. The intervals where clicks can appear in the irregular condition are shaded in gray.

individual scans and the positions were then transformed into the coordinate system of Talairach (Talairach and Tournoux, 1988) for further evaluation using Brainvoyager (Brain Innovation B.V., Maastricht). To determine whether source locations were independent, the transformed coordinates of the dipoles were tested for spatial separation in both hemispheres using the general linear model for repeated measures. The significance levels were not adjusted for multiple comparisons.

#### Source analysis

Spatiotemporal, dipole-source analysis (Scherg, 1990; Scherg and Von Cramon, 1985) was used to separate the anterior and posterior sources of the SF. A spherical head model was used and the position of the sphere was aligned to the individual head surface. The dipoles for the posterior SF source could only be fitted to the irregular click trains of the *noncontinuous* condition. This is the only condition that can be contrasted with a silent baseline, and does not evoke sustained, pitch-related activity. The dipoles for the anterior SF source could be fitted either to the regular-minus-irregular difference in the noncontinuous condition, or the regular trains in the continuous condition, where the baseline is set during the irregular click trains. Both conditions contrast regular and irregular click trains, so the remaining sustained field is the component evoked by temporal regularity (i.e. the pitch of the stimulus). In all cases, SF dipoles were fitted in the interval from 400- to 720-ms post-stimulus onset. The dipolar sources for the TFs were derived from the same three conditions as for the SFs. To suppress overlapping sustained activity, the data were highpass filtered at 3 Hz (zero-phase shift Butterworth filter, 12 dB/oct.). The first negative deflection in the AEF that peaks after 100–150 ms (N1m) was then fitted in a 30-ms interval around that peak. Dipole locations were considered acceptable, if they were within 15 mm of the listener's auditory cortex, defined as the gray matter of HG and PT.

The spatial filter used to extract source waveforms was obtained by combination of the lead field vectors of the posterior and anterior SF dipoles (or the respective combination of N1m dipoles). To model drift and other low frequency artifacts, a

principal component analysis (PCA) was computed over the interval from 1340 to 1440 ms, post-stimulus onset (i.e., the repetition of the baseline interval). The PCA component that accounted for the largest variance in this interval was then added to the spatial filter for the respective condition. In the case of eye movement artifacts, blinks were averaged, and another PCA component explaining the blinks was included in the spatial filter (Berg and Scherg, 1994; Ille et al., 2002).

In experiment 2, the dipoles were fitted to the grand average of the 12-ms condition. The dipoles derived from this condition were then used as a spatial filter for the analysis of the grand average as well as for all four ICIs. Two source models were used: one consisted of the SF dipoles on their own; the other contained the SF and N1m dipoles combined. Further details of the source analysis are provided along with the experimental data below.

## Results

### Overview

The MEG responses evoked by regular and irregular click trains provide several forms of supporting evidence for the hypothesis that there is a pitch processing region in the anterior part of auditory cortex (HG), and a separate region in the posterior part of auditory cortex (PT) that is highly active in response to the same stimuli at the same time, but which is not concerned with the presence or absence of temporal regularity in the click trains. The results are presented in three parts: (1) The data from the first experiment are used to locate dipoles for the SF and TF sources in both the anterior and posterior regions of auditory cortex (Fig. 4), and the SF and TF source waveforms are analyzed (Fig. 5). (2) The

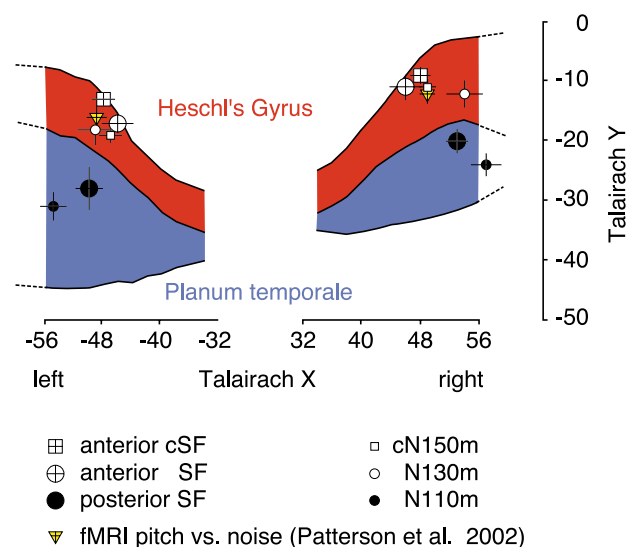


Fig. 4. Dipole sources with standard errors for experiment 1. The dipoles are projected onto an axial plain through HG and PT with the sulcal borders reported by Leonard et al. (1998) as an average from 53 normal adult brains. The dashed lines are an extrapolation to  $x$  values of  $\pm 60$  mm, where it may extend in some subjects, depending on the individual brain shape. Note that the small displacement in the  $z$  direction (Table 1) between the dipole locations and the location obtained with fMRI by Patterson et al. (2002) does not appear in the axial section shown here.

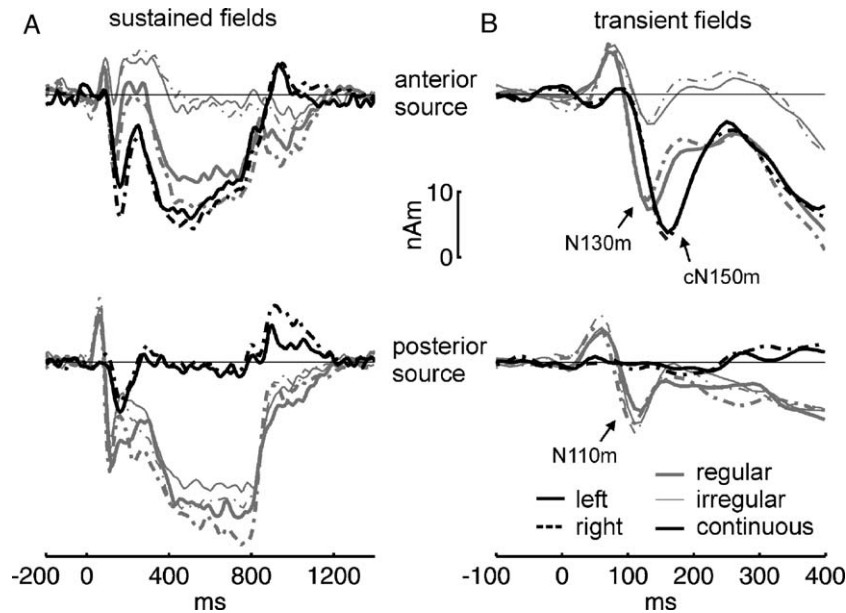


Fig. 5. Grand average source waveforms for the SFs (A) and the TFs (B); note the expanded time scale for the TFs in (B). Black lines represent the continuous condition;  $t = 0$  ms marks the start of the regular train (baseline set to the irregular trains); gray lines show the noncontinuous data, where the thick lines represent the regular and the thin lines the irregular condition. Source waves for the left (solid) and right (dashed) hemispheres are superimposed in both A and B.

results of the two-dipole model to the continuous data of the anterior source are then compared to those from a four-dipole model with separate sources for the transient and sustained fields (Fig. 6), to determine the degree to which they can be separated. (3) The data from the second session are used to describe the effect of ICI on the SF and the TF in the anterior region of auditory cortex (Fig. 7), again with a view to determining whether they represent the same mechanism.

#### Dipole sources for the SF and TF and their source waves

The mean dipole positions for the SF sources and TF sources from both continuous and noncontinuous stimulation are presented in Fig. 4 on an axial plane through HG and PT with the borders suggested by Leonard et al. (1998). There are four dipoles in a cluster in the anterior part of auditory cortex that are sensitive to the temporal regularity of the click trains, and there are two dipoles in the posterior part of auditory cortex associated with a source that is not sensitive to temporal regularity.

#### Sustained field sources

The SF dipoles for the anterior source were estimated from regular-minus-irregular differences in the noncontinuous condition and from the regular-irregular contrast in the continuous conditions. On average, the dipoles are in the anterolateral part of HG in both hemispheres; those for continuous stimulation (cSF) are slightly anterior and lateral to those for noncontinuous stimulation, but the difference is not significant ( $F_{2,32} = 1.0$ ; n.s.). The SF dipoles for the posterior source are in the anterior part of PT. The spatial separation of the anterior and posterior sources is highly significant (anterior source fitted to noncontinuous difference condition:  $F_{2,32} = 11.97$ ,  $P < 0.0001$ ; anterior source fitted to continuous-stimulation condition:  $F_{2,34} = 17.97$ ,  $P < 0.0001$ ). Dipoles were successfully fitted to all but one of the 18 listeners. For the exception, CP, no dipole could be fitted in the left

hemisphere for the noncontinuous difference condition. Table 1 presents the Talairach coordinates for comparison with the MEG data of Gutschalk et al. (2002) and the fMRI data of Patterson et al. (2002). Individual dipole locations are more variable; of the accepted 107 dipoles, 91 project close (up to 5 mm) to the gray matter of HG or PT, 14 dipoles within a limit of 6–10 mm, and 2 in a limit of 11–15 mm. When deviations in z-directions are neglected, 25 of the posterior SF dipoles project to PT and 11 to HG. For the anterior sources, 66 dipoles project to (or slightly anterior to) HG and 5 dipoles to PT.

The grand average source waveforms are shown in Fig. 4. The spatial filter in this case was composed of the posterior sources fitted to the irregular trains in the noncontinuous condition and the anterior sources fitted to the continuous regular trains. In the anterior region, the SFs for the regular trains are large and of similar size for the two paradigms; there is no SF for the noncontinuous, irregular click trains. In the posterior region, the SFs for the noncontinuous condition are even larger than in the anterior region, but there is no significant difference between the SFs for regular and irregular click trains. In the continuous condition, the baseline is set during the irregular click trains. No SF is observed for the regular trains, which is compatible with the finding that the SF in the posterior region is not sensitive to regularity, as reported by Gutschalk et al. (2002).

#### Transient field sources

At its onset, the SF overlaps with a TF consisting mainly of the peaks P1m and N1m (cf. Figs. 1 and 5). The spatial filter fitted for the SF separates these peaks and shows that the behaviour of the N1m resembles that of the SF (Fig. 5A). In the anterior source, the N1m is much larger for the regular click trains than it is for the irregular click trains; in the posterior source, the amplitude is essentially the same for regular and irregular click trains. The latency of the negative peak in the anterior source is slightly longer

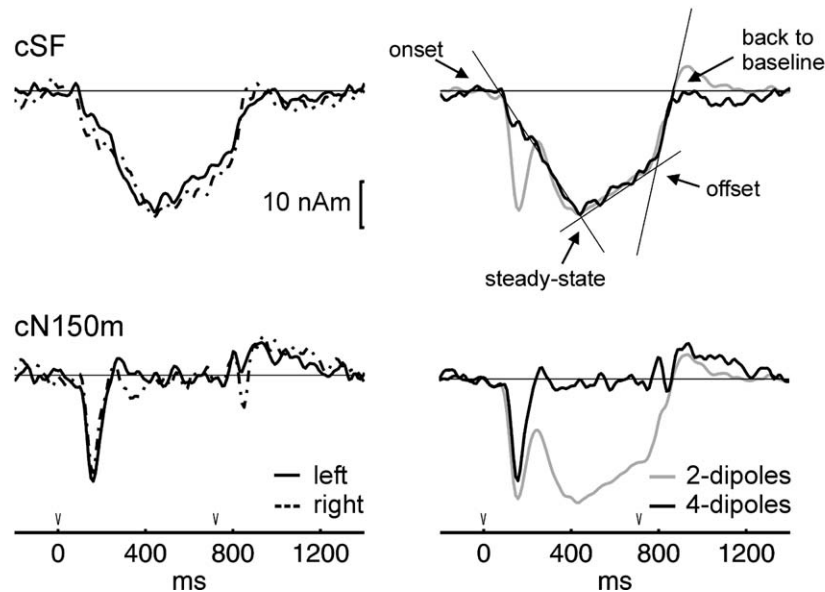


Fig. 6. Separation of the transient and sustained fields in the continuous paradigm; grand average source waveforms. Left column: separated, anterior sustained field and cN150m source waves; solid and dashed lines for the left and right hemispheres, respectively. Right column: the black lines show the average of the left and right waves in the left-hand column; the gray lines show the source waves derived with the two dipole model fitted to the sustained field on its own. Regression lines were fitted for the rising, steady and falling portions of the sustained field. Onset and offset of the regular train are marked by small arrows on the time scale.

(around 130 ms) than in the posterior source (around 110–120 ms). When the stimulation is *continuous* and the change that induces the deflection is the onset of pitch rather than the onset of energy with pitch, the negative deflection is much larger in the anterior source. Moreover, the peak of the negative deflection is delayed to around 150 ms after regularity onset, and in most listeners, there is no P1m before the negative deflection. The negative deflections will be distinguished in the following by abbreviations that code the delay of the peak: the two derived from the noncontinuous data will be called the N110m (posterior) and the N130m (anterior), and the one derived from with continuous stimulation will be called the cN150m, which encodes both the peak time (150 ms) and the paradigm (c). The nomenclature is summarized in Fig. 1.

To investigate the sources of the TF (and especially the N1m) more closely, and to compare them with the SF sources, we fitted dipoles to the following three conditions: (1) the N110m in the irregular condition, (2) the N130m in the regular–irregular difference, and (3) the cN150m in the regular, continuous condition. Dipoles could be fitted for the N110m and the cN150m in all listeners. The N130m in the difference condition (regular minus irregular) could be fitted with one dipole in each auditory cortex in 15 of the listeners, but not in the remaining three listeners.

On average, the N110m dipoles fitted to the irregular condition are in PT a little lateral and posterior to the source of the posterior SF (Fig. 4). The position of the N110m is not statistically different from that of the posterior SF ( $F_{2,34} = 0.91$ ; n.s.). The dipoles fitted to the cN150m and the dipoles fitted to the N130m are in lateral HG close to those of the anterior SF source; the dipoles of the N130m are slightly lateral to those for the cN150m. The spatial difference between the N130m and the cN150m is not significant ( $F_{2,28} = 0.53$ ; n.s.). Comparison of the cN150m and the cSF reveals a significant difference ( $F_{2,34} = 3.4$ ;  $P < 0.05$ ), although the two sources are less than a centimeter apart. The location of the

N110m source is significantly different from that of the N130m source ( $F_{2,28} = 11.08$ ;  $P < 0.001$ ) as well as from the cN150m ( $F_{2,34} = 18.67$ ;  $P < 0.0001$ ). Of the 104 acceptable dipoles, 77 project to the gray matter of HG or PT within a distance of 5 mm, 20 within a distance of 6–10 mm, and 7 within a distance of 11–15 mm.

The grand average source waves derived with the TF source model are shown in Fig. 5B; note the change of time scale from A. The anterior source was fitted to the cN150m, the posterior source to the N110m. As expected, the N110m is very similar for the regular and irregular stimuli in the noncontinuous condition, while no significant activity is observed in this source with continuous stimulation. The cN150m is now mainly associated with the anterior source and no longer appears in the posterior N110m source. The N130m at the onset of the regular click train (in the noncontinuous condition) has a similar shape as the cN150m. It is the case, however, that the wave has a more rapid onset; the mean peak latencies are listed in Table 2. With regard to the P1m in the noncontinuous data, the source waves for both the transient and sustained fields show that the latency of the P1m is shorter in the posterior source (around 60ms) than in the anterior source (around 75ms); this is the opposite of the negative peaks, N110m and N130m, where the posterior source has the shorter delay. When the P1m is fitted with a separate dipole, its mean location is in between the anterior and posterior sources. In the single subject data, however, the location is typically close to one or the other, which suggests that there are multiple sources involved in the P1m as well. However, there is no significant difference in the P1m derived from the regular and irregular conditions, and so this remains difficult to access.

In summary, in both the anterior and posterior regions, the location of the negative part of the TF is very similar to that of the SF source, and they have the same stimulus specificity. A more detailed analysis of the separation of the cN150m and the (anterior) cSF is the topic of the next subsection.

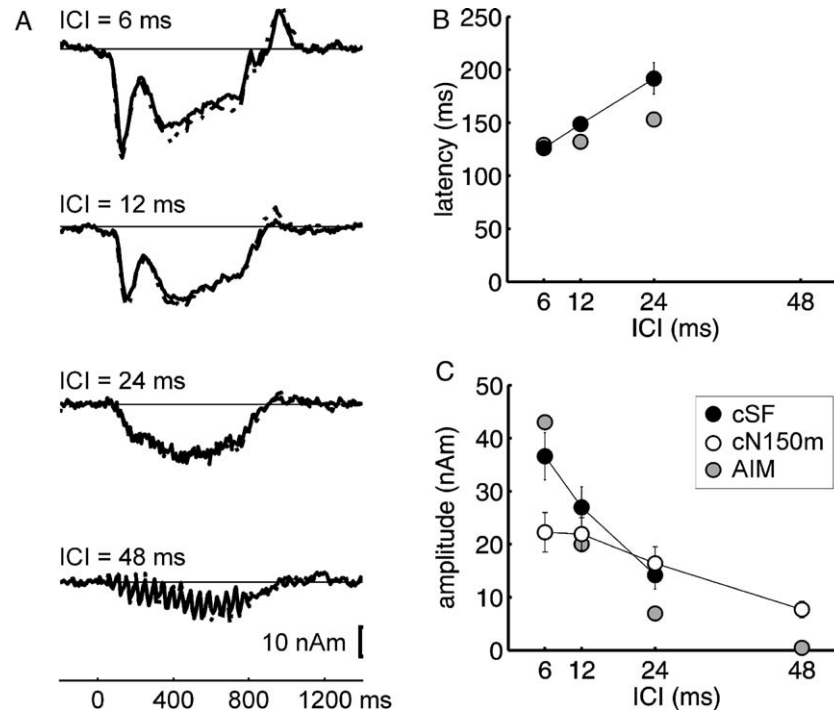


Fig. 7. (A) Source waveforms of the ICI experiment (experiment 2). Left (solid) and right (dashed) hemispheres waveforms are superimposed; the 60 Hz lowpass filter leaves the time-locked steady state response intact that is seen for ICIs of 24 and 48 ms. (B) Peak latencies of the cN150m and (C) amplitudes of the cN150m and the cSF. The cN150m is represented by black circles, the mean cSF amplitude from 400 to 720 ms by the white circles. Gray circles represent predictions for amplitude and latency derived with the auditory image model. Latencies represent the turning point of the rising sustained representation plus a fixed delay of 100 ms. The amplitudes represent the heights of the sustained representation relative to the irregular click train baseline (cf. Fig. 2). The amplitude was arbitrarily scaled to 20 nAm for the 12-ms ICI condition.

#### Separation of the transient and sustained source waves

Although the sources of the cSF and cN150m are near, they do not appear to be completely identical. The 12-ms condition that was presented 600 times to 12 listeners in session 2 provides a more sensitive measure of the spatial difference. With one exception, all of the dipoles were within 5 mm of the gray matter of HG or PT. For the exception (subject MK), the right cN150m dipole was located 9 mm above lateral HG. Of the 24 dipoles fitted for the cSF, 23 were in or slightly anterior to HG, one projected to the anterolateral PT. Most dipoles were in the intermediate or lateral part of the anterior half of HG. Some of the lateral dipoles were slightly anterior to the first transverse sulcus, which is sometimes not clearly discriminable in lateral HG. For the cN150m, 22 dipoles projected to HG and 2 to PT. Table 1 shows that on average the cN150m dipoles are located slightly posterior to the cSF dipoles in these data, as in the data from the first session. In the analysis of variance, the spatial difference between the location of the cSF source and the cN150m source was significant ( $F_{2,22} = 5.06$ ;  $P < 0.05$ ). However, when the hemispheres were examined separately, the difference was only significant in the left hemisphere ( $F_{2,22} = 8.82$ ;  $P < 0.01$ ; interaction of source  $\times$  hemisphere:  $F_{2,22} = 4.1$ ;  $P < 0.05$ ). It is also the case that the cN150m dipoles were tilted forward  $7^\circ$  more than the cSF dipoles ( $F_{2,22} = 7.44$ ;  $P < 0.01$ ). So there is probably some displacement between the transient and sustained fields, but the two sources would appear to be closely related.

The two dipoles fitted for the cSF were then combined with those for the cN150m to form a spatial filter consisting of four

dipoles. The grand average source waves derived with the four-dipole model are plotted in the left-hand column of Fig. 6; the two hemispheres are presented separately but there were essentially no differences. The latency of the cN150m from the four-dipole model is 150.6 ms ( $\pm 2.0$  ms), which is very similar to the 154.3 ( $\pm 3.1$ ) from the two-dipole model. The data for the two hemispheres were averaged and plotted in the right-hand column of Fig. 6 for comparison with the waves from the two-dipole model. The comparisons show that the two-dipole model produces a function that is essentially the sum of the cN150m and cSF. The positive inflection following the cN150m would appear to be mostly due to the blending of the transient and the sustained fields rather than an overlapping P2m, since a small P2m is only observed in few subjects and it does not show up in the grand average source waves (black line). Finally, the cN150m dipole picks up a transient response after the offset of regularity. This off-response consists of a small P1m and a small N1m that are only separated in the four-dipole model. They are followed by a prominent P2m that can also be observed to follow the cSF in the source waves of the two-dipole model. The latency of these waves are P1m = 85.9 ms ( $\pm 9.6$  ms standard deviation), N1m = 132.8 ms ( $\pm 13.4$  ms) and P2m = 219.9 ms ( $\pm 22.6$  ms). The P2m-off was not observed in the noncontinuous condition, where the SF is followed instead by continued negative source activity (cf. Fig. 5A). However, the SFs do not simply fade out in the noncontinuous condition either, so there is probably a transient off-response involved in this case as well.

The three straight lines in the upper right panel are regression lines fitted to the source waves to characterize the onset (150–450

Table 1  
Dipole locations in the space of Talairach and Tournoux (1988)

Condition	Talairach coordinates (x, y, z; mean $\pm$ SD)					
	Left auditory cortex			Right auditory cortex		
<i>Experiment 1 (n = 18)</i>						
N110m	-55 $\pm$ 8	-31 $\pm$ 10	11 $\pm$ 13	57 $\pm$ 9	-24 $\pm$ 8	15 $\pm$ 11
N130m <sup>a</sup>	-49 $\pm$ 10	-18 $\pm$ 9	12 $\pm$ 13	54 $\pm$ 9	-12 $\pm$ 8	8 $\pm$ 9
cN150m	-47 $\pm$ 7	-19 $\pm$ 5	8 $\pm$ 9	49 $\pm$ 5	-11 $\pm$ 6	10 $\pm$ 6
Posterior SF	-50 $\pm$ 8	-28 $\pm$ 15	12 $\pm$ 8	53 $\pm$ 6	-20 $\pm$ 8	12 $\pm$ 10
Anterior SF <sup>b</sup>	-46 $\pm$ 9	-17 $\pm$ 7	7 $\pm$ 8	46 $\pm$ 9	-11 $\pm$ 9	6 $\pm$ 5
Anterior SF, continuous	-48 $\pm$ 7	-13 $\pm$ 5	9 $\pm$ 10	48 $\pm$ 6	-9 $\pm$ 6	8 $\pm$ 9
<i>Experiment 2 (n = 12)</i>						
Anterior sustained field	-48 $\pm$ 5	-15 $\pm$ 2	5 $\pm$ 6	47 $\pm$ 5	-10 $\pm$ 6	6 $\pm$ 5
cN150m	-47 $\pm$ 6	-20 $\pm$ 3	4 $\pm$ 7	47 $\pm$ 4	-11 $\pm$ 7	8 $\pm$ 6
<i>Gutschalk et al. (2002)</i>						
Posterior source	-47 $\pm$ 11	-24 $\pm$ 7	14 $\pm$ 10	50 $\pm$ 9	-20 $\pm$ 5	18 $\pm$ 9
Anterior source	-51 $\pm$ 10	-18 $\pm$ 5	10 $\pm$ 7	47 $\pm$ 8	-13 $\pm$ 5	11 $\pm$ 8
<i>Patterson et al. (2002)</i>						
fMRI, fixed pitch vs. noise						
Original MNI coordinates	-55 $\pm$ 4	-13 $\pm$ 5	2 $\pm$ 3	57 $\pm$ 3	-9 $\pm$ 5	-2 $\pm$ 3
Estimated Talairach <sup>c</sup>	-49	-16	1	49	-12	-3

<sup>a</sup> n = 15.

<sup>b</sup> n = 17.

<sup>c</sup> cf. <http://www.mrc.cbu.cam.ac.uk/Imaging/mnispace.html>.

ms), steady state (450–800 ms), and offset (800–860 ms) of the cSF. Projection of the onset line back to baseline indicates that the onset latency is about 60 ms, which confirms that the cSF onset overlaps with the cN150m. The transition from the rising to the ‘steady’ portion of the cSF occurs at about 440 ms. The transition from the steady to the falling portion of the cSF occurs about 75 ms after regularity offset, indicating that the offset latency is just a little longer than the onset latency. The projection of the offset line intersects the baseline at about 150 ms after the offset of regularity, so the offset of the SF is more abrupt than the onset, as is indicated by the slopes of the first and third lines.

#### Variation of the transient and sustained fields with click rate

Grand average source waves for click trains with ICIs from 6 to 48 ms are shown in Fig. 7A; the dipoles used in the spatial filter were fitted to the sustained field of the 12-ms condition described above. The cSF amplitude decreases monotonically with click rate (i.e., increasing ICI), and it is still greater than zero when the ICI is 48 ms. The results are consistent with those observed in the anterior source previously (Gutschalk et al., 2002). For five of the listeners who also participated in experiment 1, the dipoles for the posterior source could be included in the spatial filter. This enabled us to confirm that there was no sustained field at the location of the posterior source in all ICI conditions.

The latency of the cN150m increases monotonically with the ICI, while its amplitude decreases. The latencies for ICIs of 6, 12, and 24 ms are 126, 150, and 188 ms, respectively. For convenience, we will continue to refer to this wave as the cN150m. It was not possible to separate the cN150m from the cSF when the ICI was 48 ms. Also, when the four-dipole model was used, no TF could be distinguished. With increasing delay, the cN150m-wave

moves closer to the cSF and the inflection between them disappears. The fusion of the cSF with the cN150m and the variability between subjects combined to preclude estimates of the cSF latency for all ICIs.

The latency of the cN150m is plotted as a function of ICI in Fig. 7B, along with estimates derived from the auditory image model introduced in Fig. 2 (Patterson et al., 1995). The model correctly predicts that the latency of the response will increase with increasing ICI; however, the latency of the cN150m grows at a greater rate than predicted by AIM. The amplitude of the cSF and the cN150m is plotted in Fig. 7C, along with estimates from AIM. The model predicts that the amplitude will decrease as ICI increases, but as with the latency, the rate of decrease is less than predicted by AIM. The amplitude of the cN150m clearly decreases faster than that of the cSF.

## Discussion

In the first part of the discussion, we will briefly address the advantages of the continuous stimulation paradigm, and then focus on the source analysis and the anatomical generators of the transient and sustained AEFs. The final part of the discussion is

Table 2  
Latencies of transient source waveforms (n = 18)

Peak	Dipole source	Latency (ms) $\pm$ standard error		
		Irregular	Regular	Continuous
P1m	posterior	73.8 $\pm$ 1.6	76.2 $\pm$ 3.5	–
P1m	anterior	57.4 $\pm$ 4.9	59.7 $\pm$ 2.2	–
N1m	posterior	111.8 $\pm$ 2.9	112.2 $\pm$ 1.9	–
N1m	anterior	131.0 $\pm$ 5.2	132.5 $\pm$ 3.9	154.3 $\pm$ 3.1

concerned with the temporal dynamics of the evoked pitch response and the implications for the physiology of pitch perception.

#### *Continuous stimulation paradigm*

The data from the first experiment show that the anterior SF, elicited by pitch stimuli, can be recorded with continuous and noncontinuous stimulation. In the continuous condition, the subtraction is implicitly performed by setting the baseline during the nonpitch condition. Continuous stimulation provides several advantages over noncontinuous stimulation: First, the time to acquire an equal number of averages is shorter, as no silent epochs are required. Second, the signal-to-noise ratio for the continuous condition is better by a factor of the square root of two, compared with a subtraction condition having the same number of replications. In other words, only half the number of replications is needed to achieve the same signal-to-noise ratio. Third, the part of the onset response that is evoked by the onset of stimulus energy is avoided (Jones et al., 1998; Krumbholz et al., 2003; Mäkelä et al., 1988). This is especially important for ICIs in the lower pitch range, where the latency of the cN150m overlaps with the P2m (in the range 150–200 ms). Finally, the reduction in the number of sources simplifies the analysis of the remaining sources; the fewer the number of sources that need to be modeled, the better the signal-to-noise ratio in the source waveforms.

#### *Sources of the transient and sustained fields*

The results of the source analysis indicate that anterior and posterior generators can be separated in auditory cortex for the transient onset fields as well as for the sustained fields. The locations identified for the TFs and SFs are similar; the spatial difference between them was only significant in the continuous stimulation condition. Moreover, the functional properties suggest that the TF and SF from corresponding sites belong together.

Both the anterior SF and the cN150m/N130m are associated with temporal pitch. Their amplitude decays as the ICI increases over a similar range, although the amplitude of the cN150m decays faster than the cSF with increasing ICI. This difference, and the temporal overlap of the source waves, suggests that they are not generated by the identical population of neurons. The dipole locations for the anterior SF and the cN150m/N130m are very close to those reported by Patterson et al. (2002) using fMRI and a pitch-minus-noise contrast. The MEG sources are displaced 5–10 mm above the fMRI sources (cf. Fig. 4; Table 1), but this is probably due to technical differences in the imaging techniques. The spatial extent of the source is quite focal in the fMRI (Patterson et al., 2002). With respect to the close vicinity in which the dipole locations for the cN150m and the cSF are found, one would expect them to be largely generated within this same auditory area. Recent single-cell data from Marmosets support the view that the transient and sustained responses come from the same area inasmuch as neurons with both sustained and transient responses appear in both core and belt areas (Barbour and Wang, 2003; Lu et al., 2001). With reference to human anatomy, the pitch-evoked responses are generated in a region lateral to primary auditory cortex (i.e., koniocortex) (Braak, 1978; Hackett et al., 2001). Recent cytoarchitectonic studies (Morosan et al., 2001; Rademacher et al., 2001) report that this area has properties similar to the core region. However, data obtained with a variety of histochemical staining techniques (Wallace et al., 2002) suggest

that the area in lateral HG is more like a nonprimary belt region. Together, the fMRI and MEG findings suggest that the area in lateral HG contains a pitch processing center. Krumbholz et al. (2003) have recently suggested the functional label ‘pitch onset response’ (POR) for the cN150m. Since the anterior sustained field appears to be part of a closely related process, an appropriate functional label would be the ‘sustained pitch response’ (SPR). More generally, the two together would be the ‘pitch response’ (PR; cf. Fig. 1).

The amplitude of the posterior SF was linked to stimulus intensity in our previous study (Gutschalk et al., 2002). While the dependence of the N1 on stimulus intensity has been reported to be less clear than that of the sustained potential (Näätänen and Picton, 1987; Picton et al., 1978b), the N1m (in this case probably equivalent to our N110m) was found to be modeled well by the acceleration of peak pressure at stimulus onset (Biermann and Heil, 2000). This part of the sound onset response disappears when temporal regularity is turned on without a concomitant change in stimulus intensity (see also Krumbholz et al., 2003; Mäkelä et al., 1988). Apart from sensitivity to stimulus intensity and insensitivity to temporal regularity, the defining characteristics of the posterior source are rather unspecific. In fMRI studies, broadband sounds are shown to activate the majority of auditory cortex independent of whether they produce a pitch (Patterson et al., 2002). When the stimulus is a modulated sinusoid and the intensity is varied, medial Heschl’s gyrus was found more sensitive to stimulus intensity than the surrounding belt (Brechmann et al., 2002; Hart et al., 2002). This might initially suggest that the posterior source sums the activity of primary and secondary areas throughout auditory cortex, and that the location in Planum temporale is just the center of this large area. However, human intracranial recordings do not reveal transient activity beyond latencies of 80 ms within the confines of medial HG. Moreover, major contributions to the N1m are recorded from PT (Liegeois-Chauvel et al., 1994). Thus, the fMRI data of Hart et al. (2002) and Brechmann et al. (2002) may rather relate to earlier components such as the P1m, which was largely attenuated in continuous stimulation like the N110m. With respect to histological segmentation in auditory cortex (Braak, 1978; Galaburda and Sanides, 1980; Rivier and Clarke, 1997; Wallace et al., 2002), we would therefore expect that the ‘posterior source’ in PT is more focal, and reflects neural activity from auditory fields of the posterior ‘belt’ (also called ‘parakoniocortex’), with very limited contribution from medial HG.

It was not possible to locate the source of the triphasic response observed at the transition from the regular to irregular sound due to its small amplitude. It remains to be determined whether it is concerned with the offset of pitch, or whether it is related to the onset of the irregular click train. Since it is not an intensity change response, it would appear to be associated with some aspect of timbre or stimulus change like the response described by Jones et al. (1998).

#### *Dynamics of the pitch response and temporal models of pitch perception*

In the auditory image model, at the onset of temporal regularity, the activity at the time interval corresponding to the pitch increases from baseline at the start of the second cycle of regularity and rises to steady state over about 100 ms when the pitch is 50 Hz or more (cf. Fig. 2). In the MEG data, the sustained part of the pitch response appears with a latency in the range 60–80 ms. The

latency of the earliest response in primary auditory cortex of humans is 15–19 ms (Liegeois-Chauvel et al., 1991), and in more lateral parts of HG it is in the range 25–40 ms. So the onset of the SPR is not much longer than might be expected from AIM. However, AIM would predict that the field would asymptote in the range of 150 ms, which is much less than the observed latency of 450 ms. This is a significant difference in the sense that if the half life of the auditory image were extended to fit the MEG data, the model would be too sluggish to segregate successive auditory events as we hear them.

AIM does not include a separate transient processor, but the POR might be explained in terms of the temporal derivative of the sustained wave (cf. Fig. 2). In this case, the latency of the POR would be directly proportional to the ICI. The observed latency for the POR (cN150m) increases monotonically with ICI, but it is four times the ICI. The same relation was reported by Krumbholz et al. (2003) for their RI sounds. Forss et al. (1993) had earlier reported that the latency of the N1m complex increases with ICI (Forss et al., 1993); however, the growth of the latency with ICI was much smaller, indicating that different sources were involved. Mäkelä et al. (1988) reported that a subcomponent of the N1m appeared at the onset of a square wave when it was preceded by an intensity-matched noise. The response was not accompanied by positive deflections and so it is reasonable to assume that they were observing the POR. The period length in that study ranged from 0.5 to 5 ms. The latency decreased from 120 ms for the 5-ms period to a plateau of 100 ms for periods of 2 ms or less (Mäkelä et al., 1988). These latencies are similar to those that would be extrapolated from our click train data.

The SPR was still observed at the lowest click rate, 21 Hz in experiment 2. In our previous study, the lowest rate for the SPR was between 10 and 13 Hz (Gutschalk et al., 2002). Krumbholz et al. (2003) recorded a POR with their lowest rate, 16 Hz. No POR was recorded at 21 in the current study, perhaps because the interstimulus interval (ISI) was rather short (0.72 s). It was 5 s (or 2 s from noise onset) in Krumbholz et al. (2003). The POR fuses with the SPR at long ICIs, and so it can only be detected when it is larger than the SPR. The pitch strength becomes steadily weaker below 50 Hz, and at rates below 30 Hz, rate discrimination becomes difficult (Krumbholz et al., 2000) and the stimuli can no longer be used to detect errors in melodies (Pressnitzer et al., 2001). Nevertheless, the sounds are still reported to have a weak pitch and the clicks are not perceived as single events until the ICI is about 100 ms. Thus, the presence of the cN150m and the anterior SF at rates in the region of 20 Hz is compatible with the suggestion that both arise from a pitch processing region in HG.

On the basis of fMRI data, Patterson et al. (2002) proposed that the extraction of time-interval information from the neural firing pattern in the auditory nerve, and the construction of the time-interval histograms that form the auditory image, occurs at a subcortical stage, while the specific pitch values and the pitch salience are determined in lateral Heschl's gyrus. This hypothesis was partly based on physiological evidence indicating that time-interval processing begins in subcortical regions (Griffiths et al., 2001; Langner and Schreiner, 1988; Winter et al., 2001). Recently, Penagos et al. (2003) have reported an increase in activity in lateral HG when tone complexes include resolved harmonics. This would support the view that integration of pitch information across frequency channels occurs in this region of HG. Budd et al. (2003) have also reported enhanced activity in lateral HG when there is an interaural correlation between noise samples presented

at the two ears. Although these stimuli do not produce a pitch, the activation supports the interpretation that there is cross channel integration of temporal fine-structure information in this region.

The dynamics of the PR revealed by MEG provides independent support for the hypothesis that the activity in HG represents pitch processing after the auditory image. Firstly, the pitch response in cortex begins far too late to represent the time-interval processing underlying the construction of the auditory image. Secondly, the rate of onset is too slow to be a direct representation of the growth of the vertical ridge in the auditory image at pitch onset. Thirdly, the response at pitch-offset is fundamentally different from that at pitch onset, which would not be the case if it were a direct measure of ridge height in the auditory image. Thus, it seems more likely that the pitch response represents the integration of time-interval information across frequency channels, and/or the calculation of the specific pitch value and its strength.

### Acknowledgments

The MRI data were acquired in the Department of Neuroradiology, University of Heidelberg, Germany. We are grateful to Prof. Klaus Sartor and Dr. Sabine Heiland, who provided access to the MRI facilities and generous technical support. The research was supported by the Deutsche Forschungsgemeinschaft (Ru 652/1–3) and the UK Medical Research Council (G9900369, G9901257).

### References

- Barbour, D.L., Wang, X., 2003. Auditory cortical responses elicited in awake primates by random spectrum stimuli. *J. Neurosci.* 23, 7194–7206.
- Berg, P., Scherg, M., 1994. A multiple source approach to the correction of eye artifacts. *Electroencephalogr. Clin. Neurophysiol.* 90, 229–241.
- Biermann, S., Heil, P., 2000. Parallels between timing of onset responses of single neurons in cat and of evoked magnetic fields in human auditory cortex. *J. Neurophysiol.* 84, 2426–2439.
- Braak, H., 1978. The pigment architecture of the human temporal lobe. *Anat. Embryol.* 152, 141–169.
- Brechmann, A., Baumgart, F., Scheich, H., 2002. Sound-level-dependent representation of frequency-modulations in human auditory cortex: a low noise fMRI study. *J. Neurophysiol.* 87, 423–433.
- Budd, T.W., Hall, D.A., Goncalves, M.S., Akeroyd, M.A., Foster, J.R., Palmer, A.R., Head, K., Summerfield, A.Q., 2003. Binaural specialisation in human auditory cortex: an fMRI investigation of interaural correlation sensitivity. *NeuroImage* 20, 1783–1794.
- Forss, N., Mäkelä, J.P., McEvoy, L., Hari, R., 1993. Temporal integration and oscillatory response of the human auditory cortex revealed by evoked magnetic fields to click trains. *Hear. Res.* 68, 89–96.
- Galaburda, A., Sanides, F., 1980. Cytoarchitectonic organization of the human auditory cortex. *J. Comp. Neurol.* 190, 597–610.
- Griffiths, T.D., Büchel, C., Frackowiak, R.S.J., Patterson, R.D., 1998. Analysis of temporal structure in sound by the brain. *Nat. Neurosci.* 1, 422–427.
- Griffiths, T.D., Uppenkamp, S., Johnsrude, I., Josephs, O., Patterson, R.D., 2001. Encoding of the temporal regularity of sound in the human brainstem. *Nat. Neurosci.* 4, 633–637.
- Gutschalk, A., Scherg, M., Picton, T.W., Mase, R., Roth, R., Ille, N., Klenk, A., Hähnel, S., 1998. Multiple source components of middle and late latency auditory evoked fields. In: Hashimoto, I., Kakigi, R. (Eds.), *Recent Advances in Human Neurophysiology*. Elsevier, Amsterdam, pp. 270–278.
- Gutschalk, A., Patterson, R.D., Rupp, A., Uppenkamp, S., Scherg, M.,

2002. Sustained magnetic fields reveal separate sites for sound level and temporal regularity in human auditory cortex. *NeuroImage* 15, 207–216.
- Hackett, T.A., Preuss, T.M., Kaas, J.H., 2001. Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J. Comp. Neurol.* 44, 197–222.
- Hart, H.C., Palmer, A.R., Hall, D.A., 2002. Heschl's gyrus is more sensitive to tone level than non-primary auditory cortex. *Hear. Res.* 171, 177–190.
- Ille, N., Berg, P., Scherg, M., 2002. Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *J. Clin. Neurophysiol.* 19, 113–124.
- Jones, S.J., Longe, O., Vaz Pato, M., 1998. Auditory evoked potentials to abrupt pitch and timbre change of complex tones: electrophysiological evidence of 'streaming'? *Electroencephalogr. Clin. Neurophysiol.* 108, 131–142.
- Krumbholz, K., Patterson, R.D., Pressnitzer, D., 2000. The lower limit of pitch as determined by rate discrimination. *J. Acoust. Soc. Am.* 108, 1170–1180.
- Krumbholz, K., Patterson, R.D., Seither-Preisler, A., Lammertmann, C., Lütkenhöner, B., 2003. Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cereb. Cortex* 13, 765–772.
- Langner, G., Schreiner, C.E., 1988. Periodicity coding in the inferior colliculus of the cat: I. Neuronal mechanisms. *J. Neurophysiol.* 60, 1799–1822.
- Leonard, C.M., Puranik, C., Kuldau, J.M., Lombardino, L.J., 1998. Normal variation in the frequency and location of human auditory cortex landmarks. Heschl's Gyrus: where is it? *Cereb. Cortex* 8, 397–406.
- Liegeois-Chauvel, C., Musolino, A., Chauvel, P., 1991. Localization of the primary auditory area in man. *Brain* 114, 139–151.
- Liegeois-Chauvel, C., Musolino, A., Badier, J.M., Marquis, P., Chauvel, P., 1994. Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroencephalogr. Clin. Neurophysiol.* 92, 204–214.
- Loveless, N., Levänen, S., Jousmäki, V., Sams, M., Hari, R., 1996. Temporal integration in auditory sensory memory: neuromagnetic evidence. *Electroencephalogr. Clin. Neurophysiol.* 100, 220–228.
- Lü, Z.L., Williamson, S.J., Kaufman, L., 1992. Human auditory primary and association cortex have different lifetimes for activation traces. *Brain Res.* 572, 236–241.
- Lu, T., Liang, L., Wang, X., 2001. Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nat. Neurosci.* 4, 1131–1138.
- Mäkelä, J.P., Hari, R., Leinonen, L., 1988. Magnetic responses of the human auditory cortex to noise/square wave transitions. *Electroencephalogr. Clin. Neurophysiol.* 69, 423–430.
- Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., Zilles, K., 2001. Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. *NeuroImage* 13, 684–701.
- Näätänen, R., Picton, T.W., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24, 375–425.
- Patterson, R.D., Robinson, K., Holdsworth, J.W., McKeown, D., Zhang, C., Allerhand, M., 1992. Complex sounds and auditory images. In: Cazals, Y., Demany, L., Horner, K. (Eds.), *Auditory Physiology and Perception*. Pergamon, Oxford, pp. 429–446.
- Patterson, R.D., Allerhand, M., Giguere, C., 1995. Time-domain modelling of peripheral auditory processing: a modular architecture and a software platform. *J. Acoust. Soc. Am.* 98, 1890–1894.
- Patterson, R.D., Uppenkamp, S., Johnsrude, I.S., Griffiths, T.D., 2002. The processing of temporal pitch and melody information in auditory cortex. *Neuron* 36, 767–776.
- Penagos, H., Oxenham, A.J., Melcher, J.R., 2003. Effects of harmonic resolvability on the cortical activity produced by complex tones. *Assoc. Res. Otolaryngol. Abstr.* 301, 75–76.
- Picton, T.W., Woods, D.L., Proulx, G.B., 1978. Human auditory sustained potentials: II. Stimulus relationships. *Electroencephalogr. Clin. Neurophysiol.* 45, 198–210.
- Picton, T.W., Lins, O.G., Scherg, M., 1995. The recording and analysis of event-related potentials. In: Boller, F., Grafman, J. (Eds.), *Handbook of Neuropsychology*, vol. 10. Elsevier, Amsterdam, pp. 3–73.
- Pressnitzer, D., Patterson, R.D., Krumbholz, K., 2001. The lower limit of melodic pitch. *J. Acoust. Soc. Am.* 109, 2074–2084.
- Rademacher, J., Morosan, P., Schormann, T., Schleicher, A., Werner, C., Freund, H.-J., Zilles, K., 2001. Probabilistic mapping and volume measurement of human primary auditory cortex. *NeuroImage* 13, 669–683.
- Rivier, F., Clarke, S., 1997. Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: evidence for multiple auditory areas. *NeuroImage* 6, 288–309.
- Scherg, M., 1990. Fundamentals of dipole source analysis. In: Grandori, F., Hoke, M., Romani, G.L. (Eds.), *Auditory Evoked Magnetic Fields and Electric Potentials*, *Advances in Audiology*, vol. 6. Karger, Basel, pp. 40–69.
- Scherg, M., Von Cramon, D., 1985. Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. *Electroencephalogr. Clin. Neurophysiol.* 62, 32–44.
- Scherg, M., Vajsar, J., Picton, T.W., 1989. A source analysis of the human auditory evoked potentials. *J. Cogn. Neurosci.* 1, 336–354.
- Talairach, P., Tournoux, J., 1988. *A Stereotactic Coplanar Atlas of the Human Brain*. Thieme, Stuttgart.
- Wallace, M.N., Johnston, P.W., Palmer, A.R., 2002. Histochemical identification of cortical areas in the auditory region of the human brain. *Exp. Brain Res.* 143, 499–508.
- Winter, I.M., Wiegrebe, L., Patterson, R.D., 2001. The temporal representation of the delay of iterated rippled noise in the ventral cochlear nucleus of the guinea-pig. *J. Physiol.* 537, 553–566.
- Yost, W.A., Patterson, R.D., Sheft, S., 1996. A time domain description for the pitch strength of iterated rippled noise. *J. Acoust. Soc. Am.* 99, 1066–1078.