

Neuromagnetic Evidence for a Pitch Processing Center in Heschl's Gyrus

K. Krumbholz^{1,2}, R.D. Patterson², A. Seither-Preisler¹,
C. Lammertmann¹ and B. Lütkenhöner¹

¹Institute of Experimental Audiology, University Clinic Münster, Münster, Germany and ²Center for the Neural Basis of Hearing, Department of Physiology, University of Cambridge, Cambridge, UK

There have been several attempts to use the neuromagnetic response to the onset of a tonal sound (N100m) to study pitch processing in auditory cortex. Unfortunately, a large proportion of the N100m is simply a response to the onset of sound energy, independent of whether the sound produces a pitch. The current study describes a novel stimulus paradigm designed to circumvent the energy-onset response and thereby isolate the response of those neural elements specifically involved in pitch processing. The temporal resolution of magnetoencephalography enables us to show that the latency and amplitude of this pitch-onset response (POR) vary with the pitch and pitch strength of the tone. The spatial resolution is sufficient to show that its source lies somewhat anterior and inferior to that of the N100m, probably in the medial part of Heschl's gyrus.

Introduction

Pitch is important in virtually all aspects of hearing; it is the basis of melody in music and prosody in speech. Recent fMRI studies indicate that there are specialized neural assemblies for pitch processing in the antero-lateral part of Heschl's gyrus (Griffiths *et al.*, 1998, 2001; Patterson *et al.*, 2002). The purpose of the current study was to use the superior temporal resolution of magnetoencephalography (MEG) to investigate the dynamics of pitch processing in this region of auditory cortex. Previous neuromagnetic studies on pitch processing have tended to focus on the prominent N100m response that peaks ~100 ms after the onset of a sound and corresponds to the negative deflection N100 in electroencephalographic recordings (Pantev *et al.*, 1989, 1996; Langner *et al.*, 1997). Forss *et al.* (Forss *et al.*, 1993) have shown that the latency of the N100m elicited by a regular click train is inversely related to the pitch of the sound, which led Crottaz-Herbette and Ragot (Crottaz-Herbette and Ragot, 2000) to propose that the cortical elements at the source of the N100m are involved in pitch processing. However, on reviewing a wide range of studies, Näätänen and Picton (Näätänen and Picton, 1987) concluded that an N100m can be elicited by the onset of almost any kind of sound, irrespective of its physical or perceptual properties. So, while it is the case that the latency of the N100m varies with pitch, it is also the case that it varies with the intensity and spectral composition of the sound (Roberts *et al.*, 2000; Lütkenhöner *et al.*, 2001; Seither-Preisler *et al.*, 2002). This means that any components of the N100m associated with pitch are fundamentally confounded with components that reflect responses to other stimulus features, such as loudness and timbre. In the present study, we describe a novel sound that enables us to avoid confounding the pitch-onset response (POR) with the sound-onset response (SOR), and so isolate the response to the onset of pitch information.

Bilsen (Bilsen, 1966) and later Yost (Yost, 1996) have shown that it is possible to manipulate the temporal structure of a random noise on the millisecond timescale and increase the regularity of time intervals between local waveform peaks,

thereby introducing a pitch into the perception of the sound *without* changing the energy or producing harmonically spaced peaks in the tonotopic distribution of the neural activity elicited by the sound. The current study shows that this regular-interval (RI) sound makes it possible to segregate the cortical response to the onset of sound energy from that associated with the processing of temporal regularity, and thus to segregate the source associated with the processing of pitch in auditory cortex.

Griffiths and colleagues have used RI sounds and functional brain imaging to confirm the common hypothesis that there is a hierarchy of pitch processing in the auditory pathway beginning in sub-cortical structures (Griffiths *et al.*, 2001) and extending up through Heschl's gyrus out onto planum polare (PP) and planum temporale (PT) (Griffiths *et al.*, 1998). In the most recent study (Patterson *et al.*, 2002), they showed that the antero-lateral part of Heschl's gyrus is particularly sensitive to the contrast between RI sounds and noise, and they concluded that this region was concerned with the extraction of pitch information from representations created in sub-cortical structures. They also inverted the contrast to try and identify regions where noise produced more activation than tonal sounds and, intriguingly, found none whatsoever, anywhere in the auditory pathway. The importance of lateral Heschl's gyrus in pitch processing has also been emphasized by Gutschalk *et al.* (Gutschalk *et al.*, 2002) who contrasted the MEG responses to regular and irregular click trains (CTs) with varying sound levels. They found a double dissociation involving a source in lateral Heschl's gyrus that was sensitive to CT regularity but not to CT level and a source in PT that was sensitive to CT level but not to CT regularity.

In previous studies with RI sounds, the different stimulus conditions were presented separately in discrete trials with silence between them; in this case, the MEG onset response is dominated by the N100m. This paper introduces a new paradigm, in which a continuous sound is constructed from a segment of noise and a segment of RI sound with the same energy and a very similar spectral profile. Perceptually, the sound comes on with a hiss characteristic of random noise and then it changes to a musical note with a distinct pitch and a timbre rather like a 'cracked' bassoon. The effect of the manipulation is limited to the temporal microstructure of the sound; the neural tonotopic representation and its gross temporal structure are essentially unchanged. This is illustrated in Figure 1. Figure 1*a* shows the waveform of a noise that becomes a RI sound at 2000 ms and Figure 1*b* shows the simulated neural response to the stimulus at the output of the cochlea. Each horizontal line in Figure 1*b* shows the spike probability in an individual auditory nerve fiber as a function of time. The ordinate shows the fiber's best frequency. The transition from the noise to the RI sound is not accompanied by marked changes either in the waveform (Fig. 1*a*) or the neural response (Fig. 1*b*). In particular, the

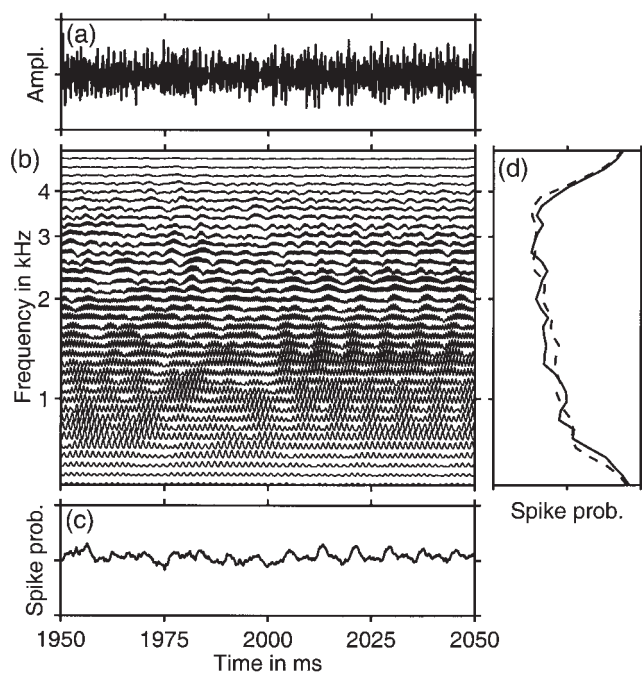


Figure 1. Waveform (a) and simulated neural activity pattern (b) of a noise that becomes a RI sound with a pitch of 125 Hz at 2000 ms. The RI sound was constructed by delaying a sample of random noise by 8 ms ($1/125$ Hz), adding it back to the original noise and iterating the process 16 times. Each horizontal line in (b) represents the simulated spike probability (Patterson *et al.*, 1995) of an individual auditory nerve fiber as a function of time (abscissa) and fiber best frequency (ordinate). Panel (c) shows spike probability, averaged over all fibers; there is no discontinuity at 2000 ms. The dashed and solid lines in (d) show spike probability averaged over time separately for the noise and RI sections of the sound, respectively; there are no harmonically spaced peaks in the RI summary (solid line).

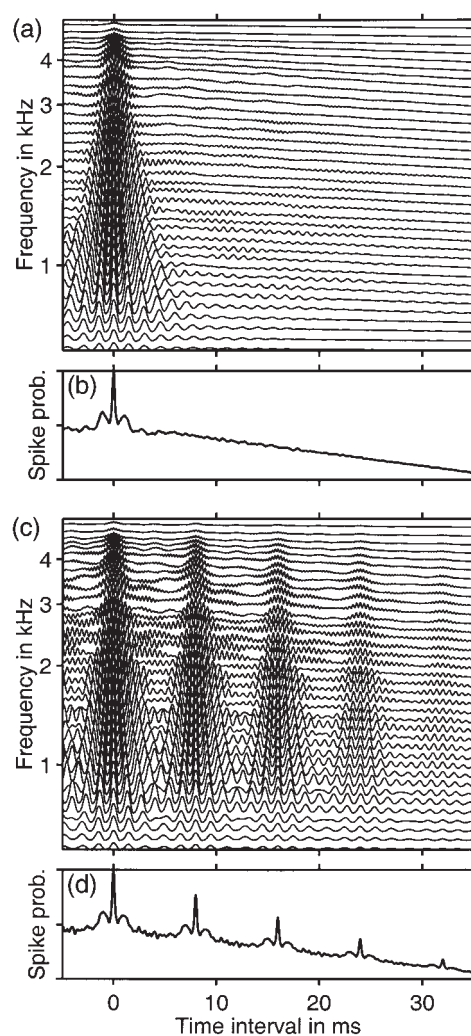


Figure 2. Time-interval histograms (Patterson *et al.*, 1995) of the simulated neural response to a noise (a) and a RI sound (c). Each horizontal line in (a) and (c) shows the distribution of time intervals between spikes in the corresponding channel of the primary neural response in (b), either before (a) or after (c) the transition to the RI sound at 2000 ms. Panels (b) and (d) show the time-interval histograms averaged across frequency channels for the noise (a) and the RI sound (c).

transition does not produce a discontinuity in the activity averaged over frequency (Fig. 1c), or activity averaged over time (Fig. 1d).

The temporal regularity that distinguishes the RI sound from the noise can be revealed by computing time-interval histograms from the neural activity patterns before and after the transition from noise to RI sound. Each horizontal line in Figure 2a shows the distribution of time intervals between neural spikes from the corresponding channel of the simulated neural pattern (Fig. 1b) in response to the noise; Figure 2c shows the distributions for the RI sound. The microstructure of noise is completely random, so the distribution of time intervals is uniform in all frequency channels (the concentration at 0 ms in Fig. 2a simply indicates the presence of activity in the channel). In contrast, the temporal regularity in the RI sound produces a concentration of time intervals at 8 ms and integer multiples thereof. The pattern is present in all channels, so these time intervals appear as vertical ridges in this representation (Fig. 2c). The ridges represent the pitch-related information in RI sounds. They produce peaks in the average time-interval histogram (Fig. 2d). The vertical location of the first peak (8 ms) corresponds to the reciprocal of the perceived pitch (125 Hz). The height of the peak increases with the degree of temporal regularity in the stimulus, which also increases the strength, or salience, of the pitch. In time-domain models of auditory processing, it is assumed that the auditory system transforms the fragile spike-timing information in auditory nerve firing (Fig. 1b) into a more stable, time-interval representation of the kind shown in Figure 2. In this representation, time interval is presented as a spatial dimension,

similar to frequency, and for convenience, it is reasonable to think of the time-interval representation as the pattern of activity in a two-dimensional array of neurons, possibly located at the level of the inferior colliculus in the midbrain.

Despite the smooth transition from noise to RI sound in the average statistics of the stimulus, there is a prominent perceptual change at the transition and it is accompanied by a prominent deflection in the magnetic field. In this paper we report a systematic investigation of this novel POR, and we compare its latency to the time required to form a stable estimate of the pitch of RI sounds perceptually. There is also a striking asymmetry inasmuch as the reverse transition from RI sound to noise produces essentially no deflection in the MEG response. In visual research, the continuous-stimulus paradigm has been used to isolate the response of cells that are specialized in the processing of visual motion (Cornette *et al.*, 1998; Ahlfors *et al.*, 1999), and to segregate different functional subcomponents of this motion-onset response (MOR) (Niedeggen and Wist, 1999). There is an interesting parallel between the auditory POR and

the visual MOR inasmuch as both are highly asymmetric; the offset of visual motion, like the offset of pitch, produces little or no response.

In the current experiments, the neuromagnetic response to the transition from a noise to a RI sound was measured as a function of stimulus parameters that control the pitch of the RI sound and its salience, to determine whether the amplitude and/or latency of the magnetic response reflect pitch and/or pitch strength, and whether the location of the source is the same as that of the N100m. The stimulus in each trial of these experiments consisted of two segments: a 2000 ms 'standard' segment intended to produce an onset response, followed by a 1000 ms 'test' segment intended to produce a 'change of information' response. In the first two experiments, the standard was a random noise and the test stimulus was a RI sound. In the third experiment, the standard and test sounds were reversed, so the standard was a RI sound and the test was a noise. The RI sounds were produced from a random noise by a delay-and-add process (Yost, 1996). Imagine a broadband noise with infinite duration. It is possible to impart a temporal regularity to the noise, by delaying a copy of the noise by d ms, adding it back to the original, and repeating the process n times. The sound has a pitch (in kHz) corresponding to the reciprocal of the delay (in ms). Each cycle of the delay-and-add process is referred to as an iteration. Iteration increases the degree of regularity in the waveform by increasing the probability of time intervals at the delay, and so the number of iterations, n , determines the strength, or salience, of the pitch percept. When n is 2, the tonal component of the sound is weak compared to the noise component; when n is 8 or more, the tonal component dominates the perception.

Materials and Methods

Stimuli and Listeners

The sounds used in the current experiments were presented at 65 dB hearing level and they were filtered to remove energy below 0.8 kHz and above 3.2 kHz. The sounds were produced by a speaker (compressor driver type) outside the magnetically shielded measuring room and delivered to the listener's right ear via 6.3 m of plastic tubing with an inner diameter of 16 mm. The passband in the transfer function of the plastic tubes approximately corresponded to the passband of the stimuli (0.8–3.2 kHz). Each stimulus was presented 100 times during the course of the experiment and the order of the conditions was randomized. The inter-trial interval was 5 s. The standard and test sounds were gated on and off with 5 ms cosine-squared ramps. At the transition from standard to test sound, the ramps overlapped so that the envelope of the composite stimulus remained flat (see Fig. 1a). Eight listeners participated in the first two experiments, where the test stimulus was a RI sound and the standard was a random noise. Nine listeners participated in the third experiment, six of whom had participated in the first two experiments; in the third experiment, the standard was a RI sound and the test sound was a noise. All listeners had normal audiological status and no history of neurological disease. Informed consent was obtained from each listener and the experimental procedures were approved by the Ethics Commission of the University of Münster.

Neuromagnetic Recordings

The magnetic fields were recorded over the listener's left hemisphere using a 37-channel first-order gradiometer system (Biomagnetic Technologies) in a magnetically shielded room. The data were acquired with a sampling rate of 297.6 Hz, filtered online between 0.1 and 100 Hz, and stored in 4 s stimulus-related epochs. The listeners were asked to stay awake and they were allowed to watch soundless video-films during the experiments.

Data Analysis

The 100 data epochs acquired for each stimulus condition were averaged and low-pass filtered at 20 Hz using a zero-phase-shift filter. Epochs with

amplitudes larger than 3 pT were considered artifactual and rejected. The sources of the N100m and the POR were analyzed with a single fixed dipole model assuming a spherical volume conductor. The center of the volume was estimated by approximating the scalp underneath the measuring coils by a sphere. Dipole parameters were derived using a maximum likelihood estimation procedure (Lütkenhöner, 1998a,b; Lütkenhöner *et al.*, 2003). The estimation of the time-invariant dipole parameters was restricted to a time window of 40 ms around the maximum in the root-mean-square (RMS) amplitude of the respective deflection. In order to analyze the N100m and P200m responses, the traces were baseline-corrected to the 100 ms period of silence just before stimulus onset. In the first and second experiments, the standard was always a noise, so the traces for all trials in each experiment were averaged, and the averaged traces analyzed to determine the location of the source. The baseline for the POR was the 100 ms segment of noise just before the transition to the RI sound. Sources were fitted separately for the POR in each stimulus condition, i.e. each combination of delay and number of iterations, because the latency of the POR depended on these parameters. Representative dipole parameters for the POR were produced by taking the median over the parameters for individual stimulus conditions. In one of the eight listeners who participated in the first two experiments, the signal-to-noise ratio of the responses was so low that many of the conditions did not yield a stable dipole solution, so this listener was discarded from further analysis.

Psychophysical Pitch-discrimination Experiment

A psychophysical pitch-discrimination experiment was performed to measure the time required to form a stable estimate of the pitch of RI sounds and compare it to the latency of the POR. Four listeners with no history of hearing impairment or neurological disease participated in this experiment. The experiment was carried out in a sound-insulated room. The stimuli were RI sounds with 16 iterations and varying delays, d . They were gated on and off with 2.5 ms cosine-squared ramps and presented binaurally to the listeners through headphones (AKG K 240 DF). The pitch-discrimination threshold (PDT) was measured as a function of the duration of the RI sounds, using an adaptive two-alternative, forced-choice procedure. In each trial, two RI sounds were presented with a silent gap of 700 ms. The delays of the two RI sounds differed slightly and the listener had to indicate, which of the two sounds had the higher pitch, namely, the shorter delay. The duration and the mean delay of the two RI sounds were fixed throughout each threshold run. The delay difference between the two RI sounds was decreased by a factor, v , after three consecutive correct responses and increased by the same factor after each incorrect response, tracking the delay difference that yields 79% correct responses (Levitt, 1971). The factor was 1.5 and 1.3 up to the first and second reversals of the delay difference, and was reduced to 1.15 for the rest of the 10 reversals that made up each threshold run. Each threshold estimate is the geometric mean of the delay differences at the last eight reversals. Three to five threshold estimates were gathered for each stimulus condition, that is, each combination of the mean delay and stimulus duration, and averaged.

All stimuli were presented with a constant overall energy; when the stimulus duration was 512 ms, the intensity level was ~ 59 dB SPL. The shortest and the longest stimulus durations tested were 16 and 1024 ms corresponding to intensity levels of 74 and 56 dB SPL, respectively.

Results

The Cortical Response to the Onset of Pitch in a Continuous Sound

The main result of this study is illustrated in Figure 3 for one representative listener. The left column shows the evoked magnetic fields at the onset of a noise (Fig. 3a) and a RI sound (Fig. 3c); the right column shows the response to the transition from one sound to the other at 2000 ms. The onset responses to the noise and RI sound (Fig. 3a,c) have essentially the same latency and amplitude, and the value of the latency is a little less than 100 ms (vertical, dash-dotted lines) indicating that these are classic N100m responses. The transition from noise to RI

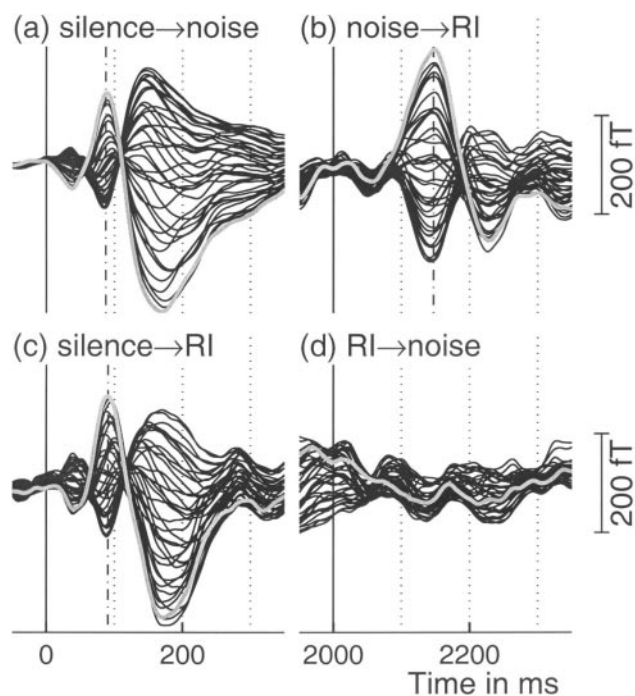


Figure 3. Neuromagnetic fields evoked by the onset of a noise (a) and a RI sound (c) at 0 ms, and by the transition from a noise to a RI sound (b) and from a RI sound to a noise (d) at 2000 ms. The RI sound had a delay of 8 ms and was generated with 16 iterations of the delay-and-add process, so it produced a strong pitch at 125 Hz, which is just below the note 'C' one octave below 'middle C' on the piano keyboard. Whereas the transition from noise to RI sound produces a large response (b), the transition from a RI sound to noise produces essentially no response whatsoever. The data are from one representative listener. Each panel shows a compilation of the 37 measurement channels, averaged over 100 presentations of the respective stimulus. The data were low-pass filtered at 20 Hz, and baseline corrected to the 100 ms period of silence just before the onset of the stimulus at 0 ms. The POR had the same polarity as the N100m as illustrated by the gray line in (a) and (b), which highlights one specific channel.

sound (Fig. 3b) produces an enhanced response, referred to as the POR, with a much longer latency (~150 ms). In contrast, the transition from RI sound to noise (Fig. 3d) produces *no discernible response* whatsoever, despite the fact that it produces a perceptual change that is just as salient as the transition from noise to RI sound.

The Amplitude and Latency of the POR Vary with the Pitch and Pitch Strength

The amplitude and latency of the POR varied with the pitch and pitch strength of the RI sound. In the first experiment, the delay, d , was fixed at 16 ms, corresponding to a pitch of 62.5 Hz, and the number of iterations, n , was varied from 2 to 32 in doublings; in the second experiment, n was fixed at 16 and d was varied from 4 to 64 ms in octave steps. For each stimulus condition in each experiment, a single equivalent dipole model was used to estimate the strength and location of the source of the magnetic field during the POR. Figure 4a,b shows the average dipole moments for seven listeners plotted as a function of time relative to sound onset. Figure 4a shows that the number of iterations, which determines the salience of the pitch, has a large effect on the amplitude of the POR. Figure 4b shows that the delay, which determines the pitch, affects both the amplitude and the latency of the POR. The condition labeled 'noise' was a control, where the transition was from one sample of noise to another. As expected, this condition produced no discernible response.

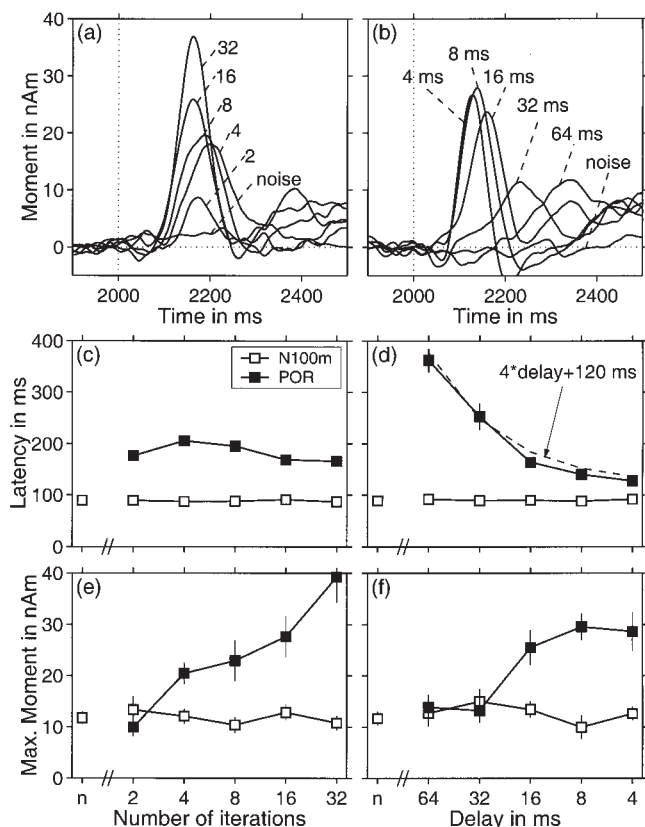


Figure 4. Upper panels: average dipole moments as a function of time in response to the transition from a noise to a RI sound, when the delay was fixed at 16 ms and the number of iterations was varied from 2 to 32 (a), and when the number of iterations was fixed and the delay was varied from 4 to 64 ms (b). The condition labeled 'noise' was a control, where the transition was from one sample of noise to another. The dipole moment is plotted as a function of time relative to stimulus onset, the transition from noise to RI sound was at 2000 ms. Middle and lower panels: The filled symbols show the latency (c, d) and amplitude (e, f) of the POR as a function of the number of iterations (c, e) and the delay (d, f) of the RI sound. For comparison, the open symbols show the latency and amplitude of the N100m response to the onset of noise in the respective stimulus condition, and in the noise control condition, labeled 'n' in each panel. The small vertical lines show the standard error of the mean and in many cases they are smaller than the size of the symbol. The dashed line in (d) represents an empirical description of the POR latency, given by 120 ms plus four times the delay of the RI sound.

The latency and amplitude of the peak of the POR were determined for all of the dipole moment functions of each individual listener. The average latency values for the two experiments are presented by filled symbols in the middle panels of Figure 4; the average amplitude values are presented by filled symbols in the lower panels. For a comparison, the open symbols show the latency and amplitude of the N100m response to the onset of noise in the respective stimulus condition, and in the noise control condition, labeled 'n' in each panel. Whereas the number of iterations, n , has only a small effect on POR latency (Fig. 4c), the delay, d , has a pronounced effect on latency (Fig. 4d); it increases from ~130 ms when the delay is 4 ms to over 350 ms when the delay is 64 ms. The dashed line in Figure 4d shows that the relationship is largely linear; the latency along the dashed line is four times the delay plus 120 ms. The amplitude of the POR also increases roughly linearly with each doubling in the number of iterations (Fig. 4e). The amplitude increases abruptly as the delay decreases from 32 to 16 ms (Fig. 4f). When the RI sound is high-pass filtered at 800 Hz, as in this experiment, the

lower limit of pitch for this stimulus is between 16 and 32 ms (Krumbholz *et al.*, 2000; Pressnitzer *et al.*, 2001). When the delay is 32 or 64 ms, the temporal regularity is perceived as flutter or repetition of a nondescript noisy feature. This suggests that a prominent POR is associated with the presence of pitch. Closer examination of the latency data (Fig. 4*d*) indicates that there may be a discontinuity in the gradient of the latency-delay function between 16 and 32 ms; the function is considerably steeper for delays greater than 16 ms.

The statistical significance of the effects of number of iterations and delay on the latency and amplitude of the T_R was verified by submitting the individual latency and amplitude data to a one-way ANOVA with repeated measures. Scheffé's *post hoc* test showed that the significant ($P < 0.0001$) main effect of delay on the T_R amplitude (Fig. 4*f*) was due to significant differences between the amplitudes for delays of 64 and 32 ms and those for 16, 8 and 4 ms ($P = 0.0036$). The differences within each of these two groups were insignificant ($P = 0.6478$). An analysis of covariance applied to the T_R latencies for different delays confirmed that there was a significant difference between the gradients of the latency-delay function (Fig. 4*d*) for delays below and above 16 ms ($P < 0.0001$).

In the third experiment, the standard segment of the stimulus was a RI sound with 16 iterations and a delay of 4, 8 or 16 ms, and the test sound was a random noise. None of the transitions from a RI sound to a noise produced a measurable transient response in any listener (see Fig. 3*d*).

The Location of the Source of the POR

The presence of a strong magnetic response to the transition from noise to tone, and the absence of a response to the transition from tone to noise, suggest that the N100m and the POR are independent neural responses generated by largely different neural populations. This conjecture was supported by the analysis of the locations of the equivalent current dipoles for the N100m and the POR. On average, the POR dipole was 12.4 mm more anterior, 6.0 mm more medial and 10.9 mm more inferior than the N100m dipole. The orientations of the POR and the N100m dipoles, on the other hand, were essentially equal.

Each of the three Cartesian coordinates of the individual dipole locations for the N100m and the POR was submitted to a one-way ANOVA with repeated measures. The analysis showed that the anterior and inferior shifts of the POR dipole relative to the N100m dipole (12.4 and 10.9 mm) were both highly significant ($P < 0.0001$ and $P = 0.0031$); the medial shift (6.0 mm) was also significant, albeit with a slightly larger value of P ($P = 0.0171$).

Figure 5 shows the proportion of the field explained by these current dipoles in two time ranges, one about the noise onset at 0 ms (left column), the other about the transition from noise to RI sound at 2000 ms (right column); the data are from one representative listener with an intermediate signal-to-noise ratio. The gray shading in Figure 5 shows the RMS amplitude of the measured field from all 37 gradiometer channels. The black shading shows the RMS amplitude of the deviation of the measured field from the field predicted by the current dipole; for convenience, the RMS deviation will be referred to as the 'residual' field of the dipole. Figure 5*b* shows that the magnetic field (gray shading) in the time range associated with the POR, marked by vertical dashed lines, is much larger than the residual field of the POR dipole (black shading), indicating that the POR dipole produces a good fit to the field of the POR. Figure 5*a* shows that the same dipole does not provide a good fit to the N100m; between the vertical dashed lines, marking the time

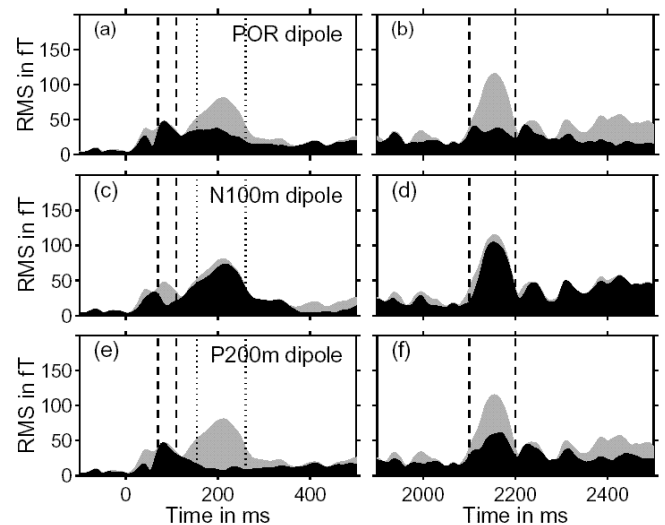


Figure 5. Proportion of the measured field that can be explained by three separate dipoles in two 600 ms time ranges, one about the noise onset at 0 ms (left column), the other about the transition from noise to RI sound at 2000 ms (right column). The data are from one representative listener with an intermediate signal-to-noise ratio. The gray shading shows the RMS amplitude of the measured field. The black areas show the RMS amplitude of the deviation between the measured field and the fields predicted by the POR dipole (a, b), the N100m dipole (c, d) and the P200m dipole (e, f). The latency ranges for the N100m and the POR are marked by pairs of vertical dashed lines; the latency range for the P200m is marked by a pair of dotted lines.

range for the N100m, the residual field is as large as the field itself. The situation is essentially reversed for the N100m dipole shown in the middle row of Figure 5; the N100m dipole produces a good fit in the time range of the N100m response (between the dashed lines in Fig. 5*c*), and a poor fit in the time range of the POR (between the dashed lines in Fig. 5*d*).

The response to the onset of sound energy is actually triphasic (peaks with inverted polarities at ~50, 100 and 200 ms; see Fig. 3*a,c*). There is a large positive peak after the N100m, which is referred to as P200m. A pair of vertical dotted lines marks the time range for the P200m in the left column of Figure 5; the residual field of the P200m dipole is shown in the bottom row of Figure 5. The P200m dipole produces a good fit to the field in the time range of the P200m (Fig. 5*e*), as would be expected. Moreover, the P200m dipole produces a relatively good fit to the POR (Fig. 5*f*), and the POR dipole produces a relatively good fit to the P200m (Fig. 5*a*). In contrast, the N100m dipole does not produce a good fit to either the P200m (Fig. 5*c*) or the POR (Fig. 5*d*). Taken together, these results suggest that the location of the source of the POR is similar to that of the P200m, and they both differ from the location of the source of the N100m. The fact that the POR and the P200m have similar topographies does not mean, however, that the neural processes underlying the two deflections are equivalent or have the same functional significance. The POR and the P200m have different latencies and opposite polarities. Moreover, the POR is specific to the onset of pitch in a sound, whereas the P200m can be elicited by tone and noise alike (see Fig. 3).

Lütkenhöner and Steinsträter (Lütkenhöner and Steinsträter, 1998) performed a high-precision measurement of the source locations of the N100m and P200m responses in a single listener, using sinusoids with varying frequencies as stimuli; the sources were then co-registered with a three-dimensional reconstruction of the listener's auditory cortex. Their results suggest that the N100m arises mainly from planum temporale, whereas the

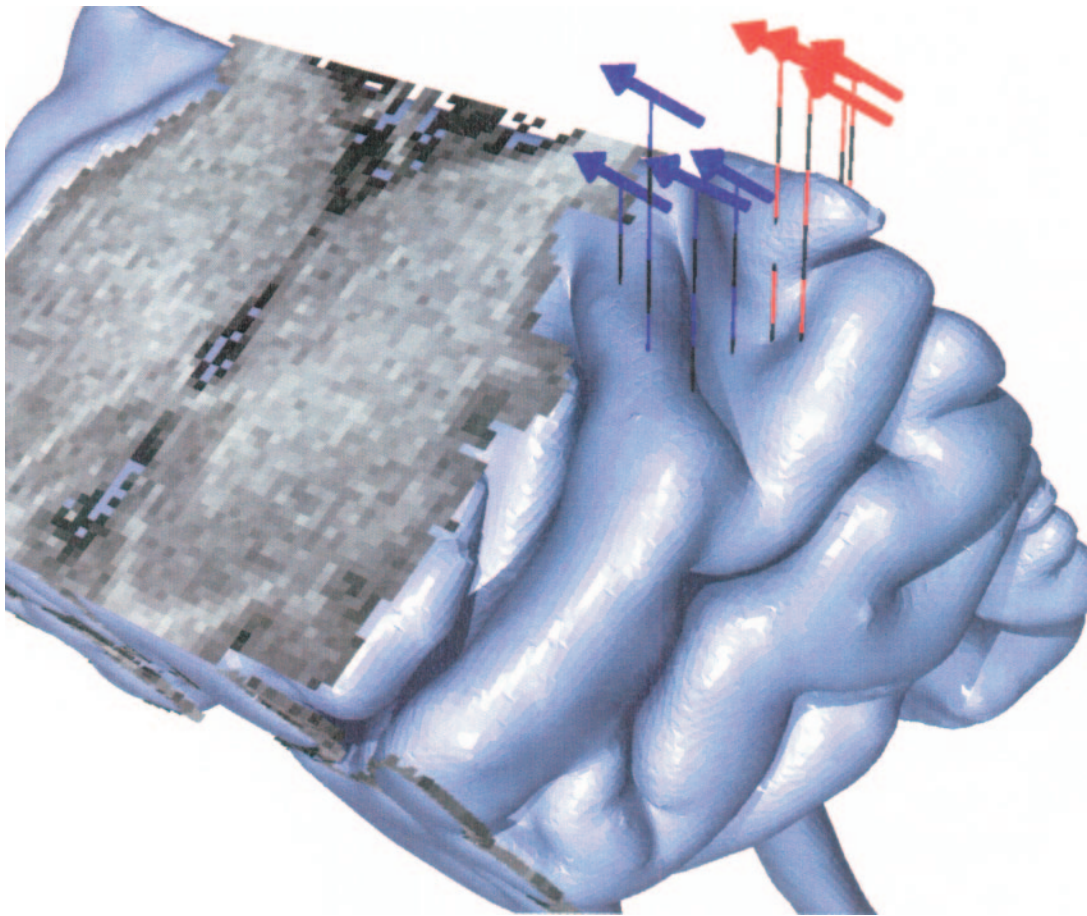


Figure 6. Source locations of the POR (blue) and the N100m (red) for a single listener, estimated from four measurement sessions and projected into a three-dimensional reconstruction of the listener's left temporal lobe. The dipoles are shifted upwards by 3 cm from the actual position of the dipole to prevent them from being partially hidden under the cortical surface. Each color bar on the vertical source markers is 5 mm in height.

source of the P200m reflects activity centered on Heschl's gyrus, anterior and inferior to the source of N100m. In order to determine whether the same is true for the POR, additional measurements were obtained from a listener with a large signal-to-noise ratio. The standard segment of the stimulus was a noise and the test segment was a RI sound with an 8 ms delay and 16 iterations that produces a strong pitch and thus a large POR. Four separate measurement sessions were performed, during each of which the stimulus was presented 420 times. Figure 6 shows a three-dimensional reconstruction of the listener's left temporal lobe derived from magnetic resonance images. The vertical lines with red arrows show the equivalent current dipoles for the N100m from the four measurement sessions; the vertical lines with blue arrows show the comparable dipoles for the POR. Despite the variability, it is clear that the POR dipoles are anterior and inferior to the N100m dipoles. The location of the N100m dipole is consistent with Lütkenhöner and Steinsträter's assumption that the N100m receives major contributions from planum temporale. The location of the POR dipole appears to be on Heschl's gyrus in a position similar to the dipole location that Lütkenhöner and Steinsträter reported for the P200m.

The Latency of the POR and the Perceptual Integration Time for Pitch

The results from the previous sections indicate that the POR reflects the activity of those neural elements in auditory cortex

that are involved in pitch processing. The function relating POR latency to the delay of the RI sound (Fig. 4d) shows that the neural elements at the source of the POR integrate pitch-related information over about four times the delay before generating a response. The functional imaging data of Griffiths *et al.* (Griffiths *et al.*, 2001) show that the processing of temporal pitch information is organized hierarchically in the auditory system. In this section, we report a psychophysical experiment designed to measure the *perceptual* integration time for pitch, that is, the time required to form a stable pitch estimate. The purpose was to try and determine the point in the pitch hierarchy represented by the POR, by comparing the latency of the POR to the perceptual integration time for pitch.

In the experiment, listeners were required to indicate which of two RI sounds had the higher pitch, and PDT was defined to be the minimum difference in delay required for statistically reliable discrimination. For each of four different delays of the RI sounds, ranging from 4 to 32 ms in octave steps, the PDT was measured as a function of stimulus duration. The data are presented in Figure 7; the parameter is the delay. The figure shows that threshold decreases rapidly as duration increases from ~ 4 to 8 times the delay of the RI sound. When the sounds were shorter than four times the delay, it was not possible to measure a stable threshold. This suggests that the auditory system has to integrate over a duration of at least four times the delay to derive a rough estimate of the pitch for these sounds – a period that is comparable to the POR latency. At the same time,

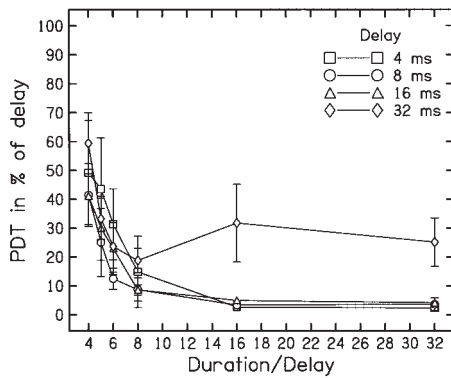


Figure 7. Average PDT for RI sounds with delays of 4, 8, 16 and 32 ms, plotted as a function of the *normalized* duration of the stimuli, that is, duration divided by the respective delay. The PDT is the difference between the two delays at threshold expressed as a percentage of the geometric mean of the delays. The data points show the average PDT of four listeners and the error bars show the standard error of the mean. The RI sounds were generated with 16 iterations of the delay-and-add process.

the auditory system appears to be able to integrate over a period of up to eight times the delay to attain a more precise pitch estimate. Beyond eight times the delay, the PDT asymptotes and the value of the asymptote is considerably lower for the 4, 8 and 16 ms delays than it is for the 32 ms delay. This is probably because a RI sound with a 32 ms delay does not produce a precise pitch when filtered as in the current experiment (Krumbholz *et al.*, 2000).

Discussion

The present study describes a transient neuromagnetic response, referred to as the POR, which can be elicited by the transition from a noise to a tone even when there is no concurrent change in sound energy. The transition from a tone to a noise, on the other hand, produces no discernable transient response, despite the fact that it is perceptually obvious when it occurs. This suggests that the cortical generators of the POR are associated with the neural processing of pitch-related information in sounds. This hypothesis is corroborated by the finding that the latency and amplitude of the POR vary with the pitch and the pitch strength of the tone. In contrast, the N100m responses to the onset of a tone and to the onset of a noise have essentially the same shape (see Figs 3*a,c*). It is also the case that the location of the source of the POR differs from that of the N100m. Together these findings suggest that the neural generators of the POR and the N100m are functionally independent.

The comparison of the physiological and perceptual data suggests that the neural elements at the source of the POR are involved in extracting an initial estimate of the pitch of a sound. The latency of the POR corresponds to the time that is required to determine that the sound has a unique pitch. At the same time, the POR occurs prior to the time required to refine the pitch value to the point where it could be used for melodic pitch perception (Krumbholz *et al.*, 2000; Pressnitzer *et al.*, 2001). The POR seems to represent a source, or sources, on medial Heschl's gyrus, adjacent to a larger region in the antero-lateral half of Heschl's gyrus where functional imaging studies have shown that activation is highly correlated with the degree of regularity in RI sounds (Griffiths *et al.*, 1998, 2001; Patterson *et al.*, 2002). In addition, a recent MEG study (Gutschalk *et al.*, 2002) with click trains has shown that regular click trains produce much more activity than irregular click trains with the

same average click rate in medial Heschl's gyrus. With regard to the hierarchy of pitch processing, these findings support the hypothesis that pitch is extracted and refined in centers progressing laterally along Heschl's gyrus and on out into adjacent areas.

Notes

Research supported by the Deutsche Forschungsgemeinschaft (Lu342/4-2), the UK Medical Research Council (G9901257) and the Austrian Academy of Sciences (APART 524).

Address correspondence to Dr Katrin Krumbholz, IME, AG Kognitive Neurologie, Forschungszentrum Jülich, 52 425 Jülich, Germany. Email: k.krumbholz@fz-juelich.de

References

- Ahlfors SP, Simpson GV, Dale AM, Belliveau JW, Liu AK, Korvenoja A, Virtanen J, Huotilainen M, Tootell RB, Aronen HJ, Ilmoniemi RJ (1999) Spatiotemporal activity of a cortical network for processing visual motion revealed by MEG and fMRI. *J Neurophysiol* 82:2545–2555.
- Bilsen FA (1966) Repetition pitch: monaural interaction of a sound with the repetition of the same, but phase-shifted sound. *Acustica* 17:295–300.
- Cornette L, Dupont P, Spileers W, Sunaert S, Michiels J, Van Hecke P, Mortelmans L, Orban GA (1998) Human cerebral activity evoked by motion reversal and motion onset. A PET study. *Brain* 121:143–157.
- Crottaz-Herbette S, Ragot R (2000) Perception of complex sounds: N1 latency codes pitch and topography codes spectra. *Clin Neurophysiol* 111:1759–1766.
- Forss N, Mäkelä JP, McEvoy L, Hari R (1993) Temporal integration and oscillatory responses of the human auditory cortex revealed by evoked magnetic fields to click trains. *Hear Res* 68:89–96.
- Griffiths TD, Buchel C, Frackowiak RS, Patterson RD (1998) Analysis of temporal structure in sound by the human brain. *Nat Neurosci* 1:422–427.
- Griffiths TD, Uppenkamp S, Johnsrude I, Josephs O, Patterson RD (2001) Encoding of the temporal regularity of sound in the human brainstem. *Nat Neurosci* 4:633–637.
- Gutschalk A, Patterson RD, 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.
- Krumbholz K, Patterson RD, Pressnitzer D (2000) The lower limit of pitch as determined by rate discrimination. *J Acoust Soc Am* 108:1170–1180.
- Langner G, Sams M, Heil P, Schulze H (1997) Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: evidence from magnetoencephalography. *J Comp Physiol A* 181:665–676.
- Levitt H (1971) Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 49:467–477.
- Lütkenhöner B (1998a) Dipole source localization by means of maximum likelihood estimation: I. Theory and simulations. *Electroencephalogr Clin Neurophysiol* 106:314–321.
- Lütkenhöner B (1998b) Dipole source localization by means of maximum likelihood estimation: II. Experimental evaluation. *Electroencephalogr Clin Neurophysiol* 106:322–329.
- Lütkenhöner B, Steinsträter O (1998) High-precision neuromagnetic study of the functional organization of the human auditory cortex. *Audiol Neurootol* 3:191–213.
- Lütkenhöner B, Lammertmann C, Knecht S (2001) Latency of auditory evoked field deflection N100m ruled by pitch or spectrum? *Audiol Neurootol* 6:263–278.
- Lütkenhöner B, Krumbholz K, Lammertmann C, Seither-Preisler A, Steinsträter O, Patterson RD (2003) Localization of primary auditory cortex in humans by magnetoencephalography. *Neuroimage* 18: 58–66.
- Näätänen R, Picton T (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.
- Niedeggen M, Wist ER (1999) Characteristics of visual evoked potentials generated by motion coherence onset. *Cogn Brain Res* 8:95–105.
- Pantev C, Hoke M, Lütkenhöner B, Lehnertz K (1989) Tonotopic organ-

- ization of the auditory cortex: pitch versus frequency representation. *Science* 246:486–488.
- Pantev C, Elbert T, Ross B, Eulitz C, Terhardt E (1996) Binaural fusion and the representation of virtual pitch in the human auditory cortex. *Hear Res* 100:164–170.
- Pantev C, Oostenveld R, Engelien A, Ross B, Roberts LE, Hoke M (1998) Increased cortical representations of musicians. *Nature* 392:811–814.
- Patterson RD, Allerhand M, Giguère C (1995) Time-domain modeling of peripheral auditory processing: a modular architecture and a software platform. *J Acoust Soc Am* 98:1890–1894.
- Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD (2002) The processing of temporal pitch and melody information in auditory cortex. *Neuron* 36:767–776.
- Pressnitzer D, Patterson RD, Krumbholz K (2001) The lower limit of melodic pitch. *J Acoust Soc Am* 109:2074–2084.
- Roberts TP L, Ferrari P, Stufflebeam SM, Poeppel D (2000) Latency of the auditory evoked neuromagnetic field components: stimulus dependence and insights toward perception. *J Clin Neurophysiol* 17:114–129.
- Seither-Preisler A, Krumbholz K, Lütkenhöner B (2002) MEG-correlates of pitch and spectrum in the auditory cortex. *Proceedings of the 13th International Conference on Biomagnetism, Jena, Germany*, pp. 122–124.
- Yost WA (1996) Pitch strength of iterated rippled noise. *J Acoust Soc Am* 100:3329–3335.