

Research paper

# From noise to pitch: Transient and sustained responses of the auditory evoked field

A. Seither-Preisler<sup>a,\*</sup>, Roy D. Patterson<sup>b</sup>, K. Krumbholz<sup>c</sup>, S. Seither<sup>a</sup>, B. Lütkenhöner<sup>a</sup>

<sup>a</sup> Department of Experimental Audiology, ENT Clinic, Münster University Hospital, Kardinal-von-Galen-Ring 10, D-48149 Münster, Germany

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

<sup>c</sup> MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD, UK

Received 27 January 2006; received in revised form 22 April 2006; accepted 27 April 2006

Available online 11 July 2006

## Abstract

In recent magnetoencephalographic studies, we established a novel component of the auditory evoked field, which is elicited by a transition from noise to pitch in the absence of a change in energy. It is referred to as the ‘pitch onset response’. To extend our understanding of pitch-related neural activity, we compared transient and sustained auditory evoked fields in response to a 2000-ms segment of noise and a subsequent 1000-ms segment of regular interval sound (RIS). RIS provokes the same long-term spectral representation in the auditory system as noise, but is distinguished by a definite pitch, the salience of which depends on the degree of temporal regularity. The stimuli were presented at three steps of increasing regularity and two spectral bandwidths. The auditory evoked fields were recorded from both cerebral hemispheres of twelve subjects with a 37-channel magnetoencephalographic system. Both the transient and the sustained components evoked by noise and RIS were sensitive to spectral bandwidth. Moreover, the pitch salience of the RIS systematically affected the pitch onset response, the sustained field, and the off-response. This indicates that the underlying neural generators reflect the emergence, persistence and offset of perceptual attributes derived from the temporal regularity of a sound.

© 2006 Elsevier B.V. All rights reserved.

**Keywords:** Pitch; Spectral bandwidth; Temporal coding; Auditory cortex; RIS; MEG; Pitch onset response; Sustained field

## 1. Introduction

To identify objects in the external world, it is necessary to segregate relevant features from meaningless background noise. Patterns exhibiting some degree of regularity

are considered worthy of further analysis. In the auditory system, the temporal regularity of incoming sounds is analyzed on different time scales. Periods on the order of seconds give rise to the sensation of rhythm, while shorter periods on the order of 10 ms give rise to the sensation of pitch (Krumbholz et al., 2000; Pressnitzer et al., 2001).

The investigation of pitch-related neural activity in auditory cortex provides a methodological challenge. For conventional stimuli, like sinusoids and complex tones, it is not possible to vary pitch without varying the spectrum of the sound. It is effectively possible, however, using a regular interval sound (RIS), which is generated from random noise by delaying a sample of noise, adding it back to the original, and iterating this delay-and-add process a number of times (Bilsen, 1966; Yost, 1996). The parameters of the RIS may be chosen such that it elicits the same long-term tonotopic representation in the auditory system as the ori-

*Abbreviations:* AEF, auditory evoked field; AI, auditory image; BB, broadband; GOF, goodness-of-fit; MEG, magnetoencephalography; NAP, neural activity pattern; NB, narrowband;  $n_{it}$ , number of iterations; POR, pitch onset response; RIS, regular interval sound; SF, sustained field; SF<sub>noise</sub>, sustained field evoked by noise; SF<sub>RIS</sub>, sustained field evoked by regular interval sound; SPR, sustained pitch response (evoked by regular click train); RMS, root-mean-square (effective value)

\* Corresponding author. Current address: Cognitive Science Section, Department of Psychology, University of Graz, Universitätsplatz 2/III, A-8010 Graz, Austria. Tel.: +43 316 380 8544; fax: +43 316 380 9806.

E-mail address: [annemarie.seither-preisler@uni-graz.at](mailto:annemarie.seither-preisler@uni-graz.at) (A. Seither-Preisler).

ginal noise, although it produces a strong pitch. The temporal regularity of the RIS is encoded in the interspike-interval statistics of the auditory nerve (Krumbholz et al., 2003). The reciprocal of the RIS delay,  $d$ , determines the pitch value, the number of iterations,  $n_{it}$ , determines the salience of the pitch percept.

In previous magnetoencephalographic studies, we used a continuous stimulation paradigm, in which a segment of noise flows into a segment of RIS with the same temporal and spectral envelope and power (Krumbholz et al., 2003; Seither-Preisler et al., 2004), to identify a component of the auditory evoked field (AEF) that responds to the transition from noise to pitch. The latency of this pitch onset response (POR) systematically increased with  $d$  and, to a smaller extent, decreased with  $n_{it}$ , while the amplitude of the POR increased with  $n_{it}$ . Krumbholz et al. (2003) estimated the time constants of the POR generators by comparing the pitch–latency function of the POR with pitch-discrimination threshold determined psychoacoustically for RISs of different durations. When the sounds were shorter than  $4d$ , it was not possible to measure a stable threshold. This suggests that the auditory system has to integrate over a duration of at least  $4d$  to derive a rough estimate of the pitch for these sounds. To attain a more precise pitch estimate, the auditory system appears to be able to integrate over a period of up to  $8d$ . Krumbholz et al. (2003) suggested that the latency of the POR reflects a constant transmission delay of about 120 ms, plus four times the RIS delay.

The POR was attributed to a cortical source with a center of ‘gravity’ in the medio-lateral part of Heschl’s gyrus, somewhat anterior and inferior to the source of the N100m, which is the dominant response to stimulus onset (Krumbholz et al., 2003). Reducing the time interval between stimulus onset and transition from noise to RIS resulted in a decreased POR amplitude, suggesting that the response to energy onset may cause a degree of refractoriness in the POR (Seither-Preisler et al., 2004). Thus, at least one of the N100m generators seems to be related to temporal pitch extraction. A subsequent study (Seither-Preisler et al., 2006) corroborated this conclusion.

In some respects, the POR is similar to the acoustic change complex (N100-P200) that is elicited by qualitative transitions in concatenated sounds. The complex has been observed for changes in energy (Ostroff et al., 1998; Martin and Boothroyd, 2000) as well as spectral envelope (Martin and Boothroyd, 2000). It has also been found when a noise changed into a harmonic complex tone of the same bandwidth, but with a definite pitch (Martin and Boothroyd, 1999). The authors speculated that the response might reflect an increase in the amount of neural phase locking in the auditory system. The differences in spectral fine structure of the two segments of the stimulus meant that it was not possible to interpret the results solely in terms of temporal pitch processing.

The present study was designed to investigate the sensitivity of different AEF components to noise and temporally

encoded pitch. Whereas previous studies (Krumbholz et al., 2003; Seither-Preisler et al., 2004) were largely confined to transient responses, the current paper also includes the sustained field response to *ongoing* RIS stimulation. The sustained response would appear to reflect processing that is functionally related to the *persistent* perception of RIS pitch. The relevant response is isolated by comparing the transient and the sustained responses to both noise and RIS.

## 2. Design of experiment: model

A series of simulations of the representation of RIS in the auditory system were performed in the course of designing the experiments reported in this study. In what follows, we present the simulations for the stimuli that were selected for the MEG experiment. The results of the simulations are not suitable for quantitative comparison with MEG data, but they do support qualitative comparisons that provide a good basis for the Discussion.

### 2.1. Methods

The simulations were based on a recent version of the Auditory Image Model (Patterson and Irino, 1998). A compound bandpass filter simulated the transfer characteristic of the outer and middle ear; it is the sum of four parallel, 2nd order bandpass filters (100–1300 Hz, gain of –12 dB; 350–6500 Hz, gain of 1.5 dB; 1800–5200 Hz, gain of 5 dB; 7500–14000 Hz, gain of –11 dB). Basilar membrane motion was simulated with a dual-resonance, nonlinear filter-bank (Lopez-Poveda and Meddis, 2001; Meddis et al., 2001) with 70 frequency channels covering the frequency range between 400 and 6400 Hz. The inner hair cell receptor potentials were simulated using the parameters recommended by Shamma et al. (1986). The spike probability at the synapse between inner hair cell and auditory-nerve fibre was simulated using the model of Sumner et al. (2002); there were 500 high-spontaneous-rate fibers per channel. The output of this stage is a three-dimensional representation of auditory nerve activity (‘neural activity pattern’ or NAP), reflecting the strength and the timing of neural firing along the tonotopic frequency axis (Patterson et al., 1995).

The pitch extraction stage was simulated with the Auditory Image Model (Patterson et al., 1995) using strobed temporal integration (Patterson et al., 1992). This time-domain model of pitch processing assumes that the auditory system transforms the fragile spike-timing information in the NAP into a more stable time-interval representation referred to as the auditory image (AI). The processing is assumed to occur somewhere between the cochlear nucleus and auditory cortex. Interspike-interval histograms are computed from the larger peaks in the NAP, separately for each frequency channel, to produce the AI; the mechanism performs a subset of the calculations involved in autocorrelation – a subset which preserves the temporal asymmetries that we perceive as timbre differences in temporally asymmetric sounds

(Patterson and Irino, 1998). Then, the interspike-interval histograms are averaged across the tonotopic dimension of the AI to produce a summary AI. If the stimulus contains temporal regularity, a peak will arise in the summary AI; the time interval of the peak predicts the pitch of the sound, and the peak height relative to the background predicts the salience of the pitch.

The height of the first peak in the autocorrelation function of a sound is closely related to the perceived pitch salience (Yost et al., 1996). Shofner and Selas (2002) have recently shown that the subjective percept follows Stevens's power law,  $\Psi = k\Phi^\beta$  ( $\Psi$ : magnitude of the sensation,  $\Phi$ : intensity of the stimulus,  $k$ : constant factor,  $\beta$ : exponent). The exponent  $\beta$  is typically greater than 1, indicating that pitch salience grows faster than the height of the first peak in the autocorrelation function. The Auditory Image Model (Patterson et al., 1995) focuses on the initial stages of pitch extraction, and so it does not include the nonlinear growth of pitch strength described by Shofner and Selas (2002).

## 2.2. Stimuli

A 2000-ms segment of random noise was concatenated with a 1000-ms segment of RIS; they had the same power and were passed through the same bandpass filter. The noise and the RIS were gated on and off with 5-ms cosine squared ramps. At the transition from noise to RIS, the ramps overlapped so that the envelope of the composite stimulus remained flat in this region. The delay of the RIS was 16 ms (a pitch of 62.5 Hz); the number of iterations,  $n_{it}$ , was 1, 4, or 16. There were two spectral bandwidths, as shown in the upper panels of Fig. 3: a 1-octave band (lower and upper cutoff frequencies: 1131 and 2262 Hz, respectively) and a 4-octave band (lower and upper cutoff frequencies: 400 and 6400 Hz, respectively). Both bands were centered around 1600 Hz. For convenience, they will be referred to as the narrowband (NB) and the broadband (BB) conditions. All of the stimuli were normalized to the same RMS level, which was set to 60 dB SPL.

## 2.3. Simulations

Fig. 1a shows the time course of the NAP (averaged across all frequency channels) in response to the BB-stimulus having a RIS segment with  $n_{it} = 16$ . A large phasic response to stimulus onset is followed by a region of steady activity that persists until stimulus offset. No discontinuity is observed at the transition from noise to RIS (dashed line), indicating that the two stimulus segments had the same intensity.

Fig. 1b shows spectral profiles of the NAPs for the two bandwidths. The abscissa is the tonotopic axis of the cochlea, and the ordinate represents the auditory-nerve spike probability in each frequency channel averaged over the first 500 ms of the signal. The spectra for noise (solid line) and RIS ( $n_{it} = 16$ ; dotted line) look very similar,

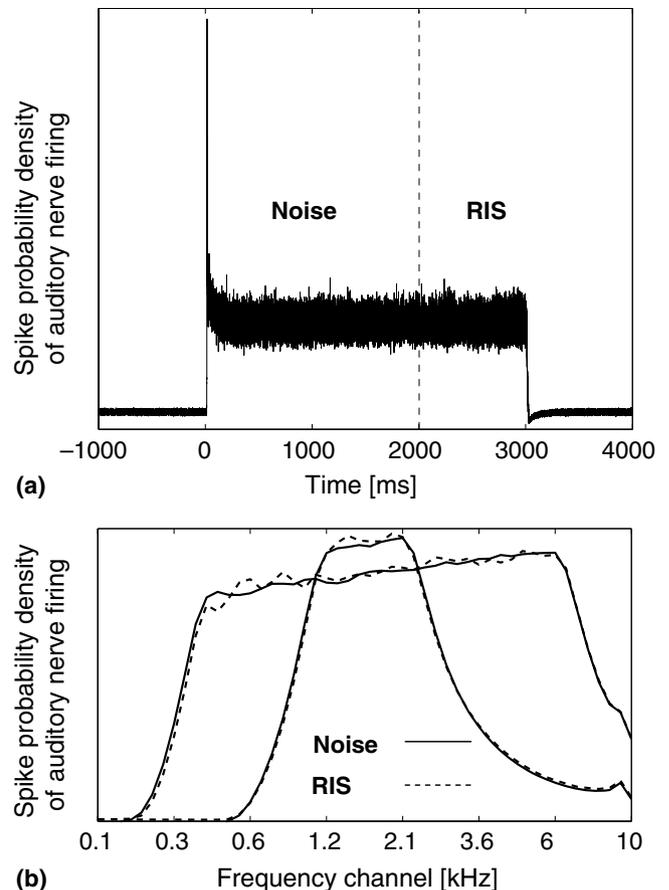


Fig. 1. (a) Firing activity of the auditory nerve in response to the broadband stimulus (BB) comprising four octaves and having a RIS segment with  $n_{it} = 16$ . There is a large phasic response to the energy onset of the noise, followed by an elevated activity level that persists throughout the stimulus. No discontinuity is observed at the transition from noise to RIS (dashed line). (b) Internal spectral representations of noise and RIS ( $n_{it} = 16$ ) at the level of the auditory nerve. The abscissa represents the tonotopic frequency axis of the cochlea and the ordinate is the spike probability density of auditory nerve firing at the respective frequency channels. As a consequence of peripheral compression, the overall neural activity level, indicated by the area underneath the curves, is higher for the BB stimuli. No systematic differences are observed between the representations of noise (dashed lines) and RIS (solid lines), indicating that the perceptual difference between both stimulus types cannot be a function of the internal stimulus spectrum.

although the curves for the noise appear slightly smoother. The small ripples in the RIS spectra are not systematic and a different RIS sample would show ripples at other frequencies. Thus, there is no harmonic structure that could function as a spectral template for extraction of a missing fundamental at the RIS delay (Terhardt et al., 1982). Rather, RIS-induced pitch sensations have to be attributed to a temporal analysis of neural interspike intervals, as postulated by Cariani (1999), Meddis and O'Mard (1997), and Patterson et al. (1992). The total amount of neural activity, as indicated by the area underneath the curves, is clearly greater for the BB stimulus, although the total stimulus intensity is the same for the two bandwidths. This is a consequence of the compressive properties of basilar membrane

motion (Yates et al., 1990), inner hair cell potentials, and auditory nerve firing (Meddis, 1986, 1988): For narrowband signals, the stimulated hair cells are driven into saturation at a lower intensity, because the total energy is concentrated on a smaller tonotopic array. As a consequence, for intensities above 30 dB SPL (Moore and Glasberg, 1996; Zwicker and Fastl, 1990), a bandwidth increase beyond a critical value ('critical bandwidth') enhances the total amount of neural activity, resulting in a growth in perceived loudness (Moore et al., 1997). Fig. 1b also shows that activity increases with channel frequency because the bandwidth of the auditory filter increases with channel frequency.

Fig. 2 shows the summary AIs that provide the basis for the pitch predictions. As a consequence of peripheral compression, the total activity level is lower for the NB stimuli (left column) than for the BB stimuli (right column). This is the case for noise (upper row) as well as RIS (lower three rows). The summary AIs for the RIS stimuli show a peak at 16 ms (dashed lines), reflecting the RIS delay, and the height of the peak relative to the stochastic background activity predicts the relative strength of the perceived pitch. The peak is somewhat larger in the BB conditions, which is an indirect consequence of peripheral compression (the RIS-specific interspike

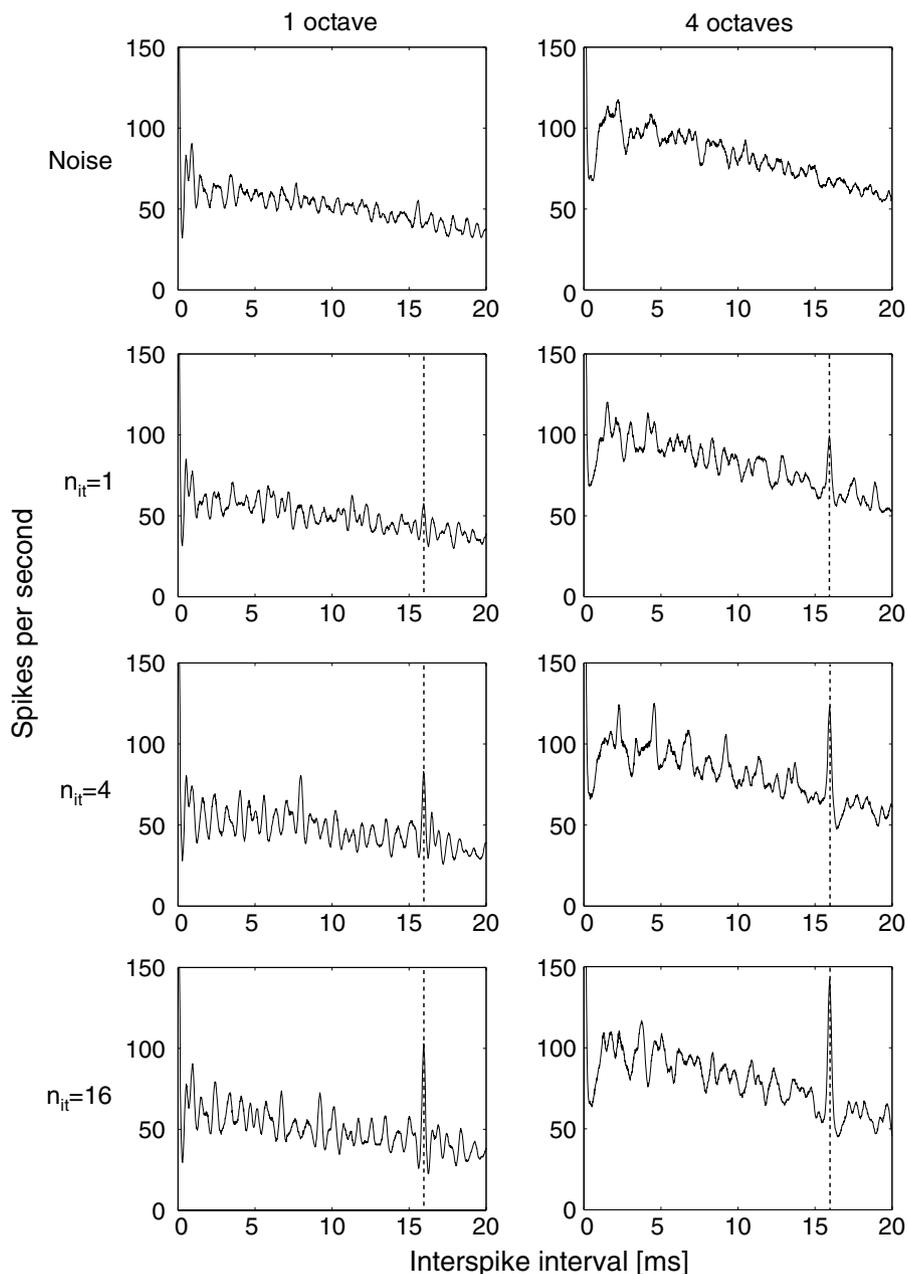


Fig. 2. Summary auditory images (summary AIs), representing the basis for pitch estimation, for the two noise and six RIS segments. The peak at 16 ms (dashed lines), signifying a prominent interspike interval at the RIS delay corresponding to 62.5 Hz, increases progressively with  $n_{it}$  (from top to bottom) and is relatively smaller for the NB stimuli (left column) than for the BB stimuli (right column). This indicates that the pitch salience of the RIS increases both with  $n_{it}$  and spectral bandwidth.

interval is derived from a higher overall activity). The peak increases, of course, with  $n_{it}$ .

#### 2.4. Predictions derived from the simulations

The above simulations suggest that the pitch salience will be greatest for the broadband RIS with the largest  $n_{it}$  – a prediction that was confirmed by informal listening. This stimulus is expected to elicit the largest POR, as the amplitude of this component increases with the salience of the pitch (Krumbholz et al., 2003; Seither-Preisler et al., 2004). According to Gutschalk et al. (2004), a sub-component of the sustained field, located on lateral Heschl's gyrus, responds to the temporal regularity of click trains. If the same process is applied to RIS stimuli (as would be expected), the sustained field should exhibit a similar correlation with pitch salience. It seems likely that an increase of bandwidth and  $n_{it}$  will result in an increase in the amplitude of the RIS-induced sustained field ( $SF_{RIS}$ ). For the noise-induced components (N100m and  $SF_{noise}$ ), the simulation suggests sensitivity to the bandwidth of the noise segment, since peripheral compression reduces within channel differences relative to bandwidth differences.

### 3. Experimental methods

#### 3.1. Subjects

Twelve right-handed subjects (six male, six female) participated in the experiment; they had normal audiological status with no history of neurological disease. After having explained the nature and purpose of the investigation, consent was obtained. The experiment was conducted in accordance with the Ethics Commission of the University of Münster and the Declaration of Helsinki.

#### 3.2. Stimuli

The stimuli in the MEG experiments were identical to those in the simulations, except that the sound delivery system acted as a low-pass filter. To avoid electronically induced magnetic artifacts, the stimuli were played out using a compressor-driver-type speaker outside the magnetically shielded chamber and delivered to the listener's right or left ear via 6.3 m of plastic tubing. The filtering caused by this tubing is illustrated in the bottom row of Fig. 3. Energy transmitted through the tubing is progressively attenuated for frequencies above 1.3 kHz, and becomes negligible above 4 kHz. As a consequence, the effective bandwidth of the BB stimuli is reduced to about three octaves. However, the filtering does not alter the fact that the BB stimuli are attenuated in power spectral density relative to NB stimuli. The attenuation corresponds to 7.01 dB before filtering and to 6.51 dB after filtering. The effect on the overall RMS values of the stimuli is almost negligible as well: after filtering, the stimulus intensities differ by only 1.05 dB (BB stimuli slightly more intense). This level difference is of the order

of the perceptual limits (Zwicker and Fastl, 1990), and it is unlikely to produce a measurable effect on the cortical AEF.

#### 3.3. Neuromagnetic recordings

Before the actual experiment, hearing threshold was determined for the NB stimulus having a RIS segment with  $n_{it} = 1$ . In the experiment itself, all six stimuli were presented at an intensity 60 dB above that threshold. The stimulus material was presented in five runs in randomized order, with a total of 100 trials per condition. The silent interval between two successive stimuli was 2 s (identical to the duration of the noise segment before RIS onset).

The magnetic field was recorded over the auditory cortex contralateral to the stimulated ear by means of a 37-channel first-order gradiometer system (Biomagnetic Technologies, San Diego). The sensor array was positioned as closely as possible to the head and manually centered about 1.5 cm superior to the point T3/T4 of the 10–20 system for electrode placement. The data were collected in two sessions, one for each hemisphere.

The listeners were asked to stay awake, and they watched soundless video films during the experiments. Data were sampled at 1041.7 Hz using 16-bit analog-to-digital converters (recording bandwidth 0.1–100 Hz) and stored on hard disk for further analysis.

#### 3.4. Data analysis

An epoch was considered to be contaminated by artefacts (eye blinks, body movements, external disturbances) if a data value exceeded the pre-stimulus baseline (the mean value in the 200-ms time window before stimulus onset) by 3 pT in any of the 37 channels. Such epochs were discarded. The remaining data were averaged separately for each stimulus condition. The signals were finally low-pass filtered with a cut-off frequency of 20 Hz (second-order butterworth filter).

An example of the data for a single representative subject is presented in Fig. 4: It shows the neuromagnetic responses from all 37 sensors to the BB stimulus having a RIS segment with  $n_{it} = 16$ . The transitions from noise to RIS at 2000 ms, and from RIS to silence at 3000 ms, are indicated by vertical lines. About 100 ms after noise onset, there is a clear N100m, which is followed by the sustained  $SF_{noise}$ . About 170 ms after RIS onset, there is a large POR, which is succeeded by the sustained  $SF_{RIS}$ . About 90 ms after stimulus offset, an off-response is observed.

A least-squares fitting algorithm based on a model with a single fixed dipole in a sphere (Lütkenhöner et al., 2003b) was used to estimate the neuromagnetic sources of the N100m (induced by noise-onset) and the POR (induced by the transition to the RIS), separately for each hemisphere. The N100m fit was based on the average of all stimulus conditions, while the POR fit was based on the condition eliciting the strongest response (BB,  $n_{it} = 16$ ). This procedure resulted in two source waveforms, model-

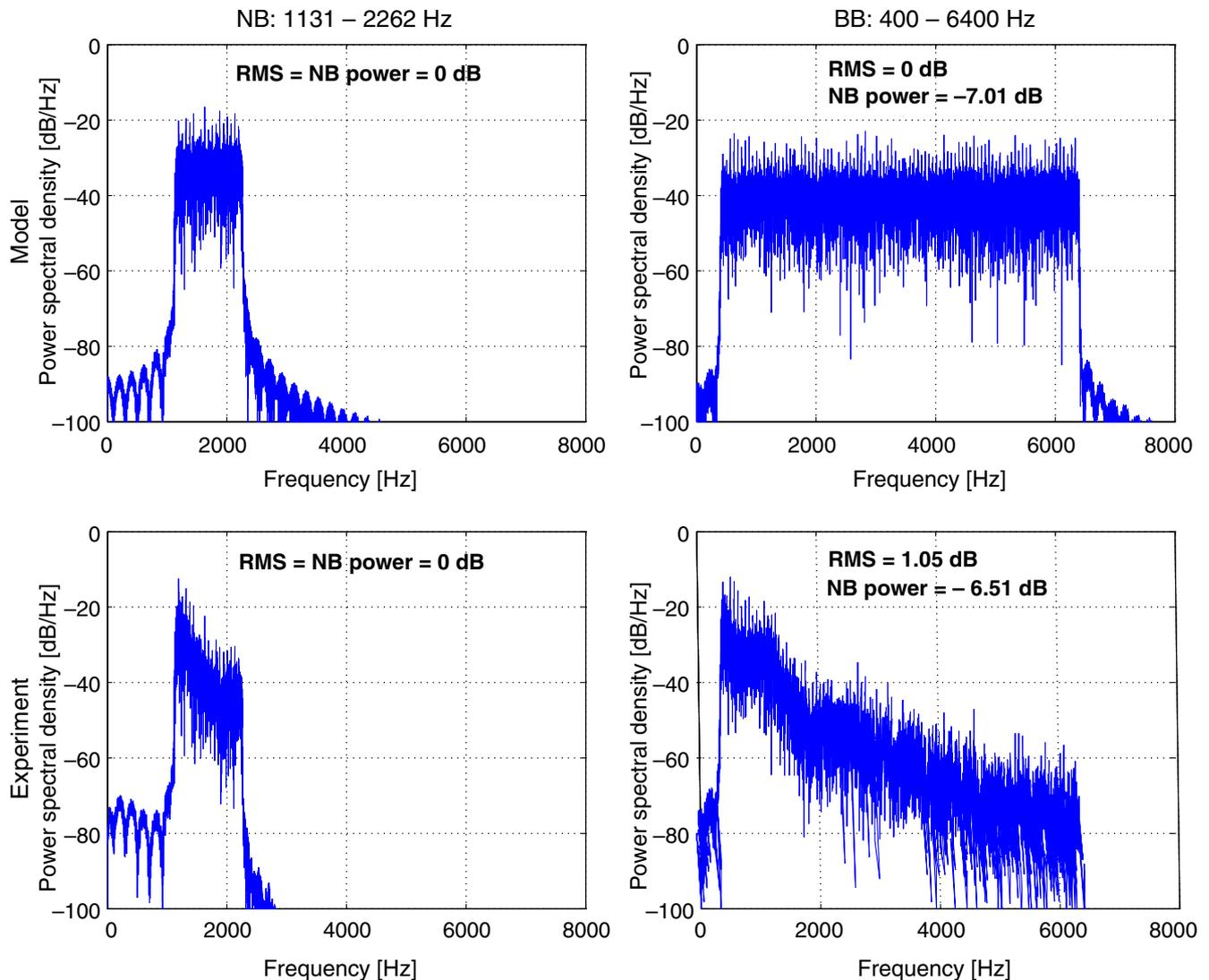


Fig. 3. Upper panels: spectra of the NB and BB stimuli used in the simulations. Bottom panels: spectra after passing the stimuli through the sound delivery system.

ing the neural activity at specific sites of auditory cortex: one for the N100m source and another one for the POR source. Data were excluded from the subsequent statistical analysis, if the percentage of the magnetic field variance accounted for by the dipole model (goodness-of-fit; GOF) was below 90%. Most subjects exhibited a GOF above 95%. Subject S2 was excluded from the statistical analyses based on the N100m fit (GOF for left and right hemisphere: 83.9% and 77%). The right hemisphere of subject S8 (GOF: 85.7%) was excluded from the analyses based on the POR fit. We refrained from fitting the sources of the  $SF_{\text{noise}}$ ,  $SF_{\text{RIS}}$ , and off-response, because the amplitudes were relatively small.

In the next step, the amplitudes and latencies of the transient components and the amplitudes of the sustained components were derived from one of the two source waveforms. The baseline level was set to the mean activity in the 200-ms time window before stimulus onset (for the N100m,  $SF_{\text{noise}}$ ,  $SF_{\text{RIS}}$ , and off-response) or RIS onset

(for the POR). The N100m and off-response are both elicited by rapid transitions in sound energy and their sources are closely related (Hari et al., 1987). Although this justifies the use of the N100m source waveform, for establishing the parameters of the off-response, the off-parameters were derived from the POR-source waveform as well, for comparison. Because of the need to compare POR,  $SF_{\text{noise}}$  and  $SF_{\text{RIS}}$  in a consistent way, the parameters of these components were derived from the source waveform obtained in the POR fit. The amplitudes and latencies of the transient AEF components were established from the dipole maxima in the respective time windows (N100m: 90–110 ms after stimulus onset; POR: 100–200 ms after RIS onset; off-response: 50–110 ms after stimulus offset). The amplitudes of the  $SF_{\text{noise}}$  and  $SF_{\text{RIS}}$  were defined as the mean dipole moments 1000–2000 ms after noise onset and 500–1000 ms after RIS onset, respectively. These relatively late time windows are basically free of transient response components.

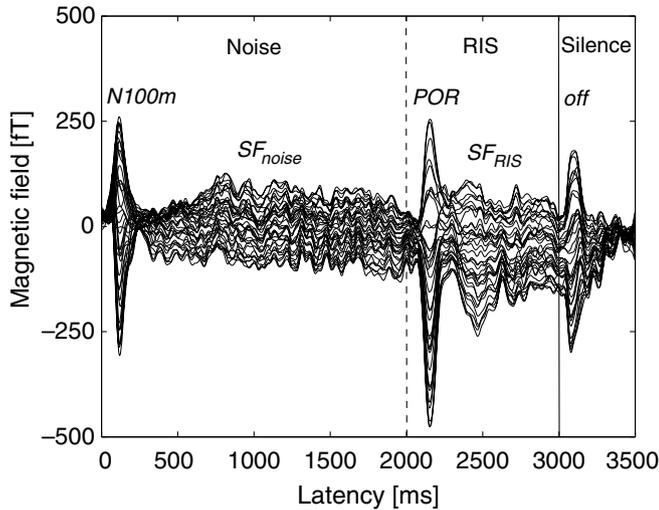


Fig. 4. Neuromagnetic response of an exemplary single subject (BB stimulus,  $n_{it} = 16$ ). The curves for all 37 measurement channels were overlaid. The transitions from noise to RIS at 2000 ms and from RIS to silence at 3000 ms are indicated by a, respectively.

### 3.5. Statistical analyses

For the noise-induced components, two-way, repeated-measures analyses of variance were performed for the independent variables ‘Hemisphere’ and ‘Bandwidth’, and for the dependent variables ‘AEF dipole moment’ (N100m and  $SF_{noise}$ ) and ‘AEF latency’ (N100m), respectively. As the POR was evoked in only some of the RIS conditions, the relative number of POR occurrences was analyzed by means of non-parametric paired-sign tests. Moreover, for conditions evoking a clear POR, the dipole moments and latencies were analyzed by means of paired *t*-tests. For the  $SF_{RIS}$  and off-response, three-way analyses of variance were calculated for the independent variables ‘Hemisphere’, ‘Bandwidth’, and ‘Iterations’, and the dependent variables ‘AEF dipole moment’ ( $SF_{RIS}$  and off-response) and ‘AEF latency’ (off-response), respectively. The analyses were done with StatView 5.0.1 (SAS Institute Inc., Cary, NC).

## 4. Experimental results

### 4.1. Time course of magnetic field strength

The source waveforms for the POR fit and the N100m fit were very similar, except that the relative amplitude of the POR was larger when directly assessed from the POR fit. This was especially true for the condition with the greatest pitch salience (BB,  $n_{it} = 16$ ). Fig. 5 shows the source waveforms for the POR fit (a: all conditions) and the N100m fit (b: condition with the greatest pitch salience), averaged over all subjects and both hemispheres (with appropriate GOF values). The long horizontal dotted lines mark the baseline levels (200 ms prior to stimulus onset). For all conditions, noise onset elicits a negative N100m, which is fol-

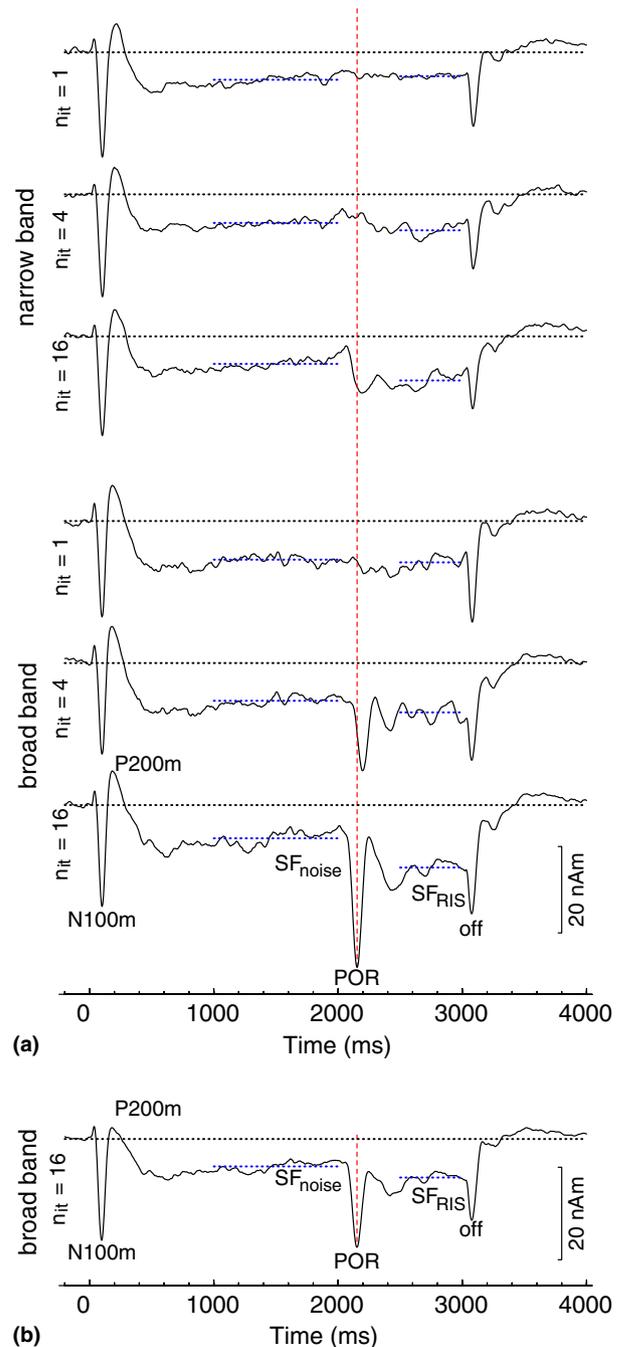


Fig. 5. Source waveforms: (a) POR fit (all conditions), (b) N100m fit (condition with the greatest pitch salience: BB,  $n_{it} = 16$ ). The curves represent the group average from all subjects and both hemispheres (with appropriate GOF values); the transition from noise to RIS is marked by a vertical dashed line. Long horizontal dotted lines mark the baseline levels, 200 ms prior to stimulus onset. Short horizontal dotted lines mark the mean dipole moments of the sustained activity ( $SF_{noise}$ : 1000–2000 ms after noise onset;  $SF_{RIS}$ : 500–1000 ms after RIS onset).

lowed by a positive deflection, the P200m. Thereafter, there is a small but consistent negative field relative to the prestimulus baseline. This is the  $SF_{noise}$  and it continues until the noise ends (horizontal dotted line: mean dipole moment for the time interval 1000–2000 ms after noise onset). About 170 ms after the transition from noise to RIS

(marked by the vertical dashed line), another negative deflection emerges in conditions where the pitch is relatively strong (the NB condition with  $n_{it} = 16$ , and the BB conditions with  $n_{it} = 4$  and  $n_{it} = 16$ ). This is the POR and its amplitude is directly related to the salience of the pitch produced by the RIS. Note that the same stimuli provoke an increase in the negative sustained field,  $SF_{RIS}$  (horizontal dotted line: mean dipole moment for the time interval 500–1000 ms after RIS onset), relative to  $SF_{noise}$ . Shortly after stimulus offset, a transient off-response can be observed for all conditions.

4.2. Statistical results

4.2.1. N100m and SF in response to noise

While the RIS segment varied with  $n_{it}$ , the preceding noise segment did not. Therefore, the noise-induced components were only analyzed with regard to hemisphere and the bandwidth of the noise. The results of the 2-way analysis of variance are presented in Fig. 6. With regard to the transient components, there were no significant hemispheric differences, for the latency or the amplitude (dipole moment) of the N100m. The mean N100m latency was significantly shorter for the BB condition (Fig. 6b), but the difference was small (2 ms). With regard to the sustained field, there were no significant hemispheric differences, but the mean dipole moment was significantly larger for the BB condition (Fig. 6a, right) and the difference was relatively large (1.6 nAm; 25%). The lack of a

Table 1

Statistical results for the number of POR occurrences

Condition	Mean proportion of POR occurrences (%)
<i>Iterations</i>	
$n_{it} = 1$	4.2
$n_{it} = 4$	47.9
$n_{it} = 16$	70.8
<i>p-Values for paired-sign tests</i>	
$n_{it} = 1$ vs. $n_{it} = 4$ : 0.001**	
$n_{it} = 1$ vs. $n_{it} = 16$ : 0.0005***	
$n_{it} = 4$ vs. $n_{it} = 16$ : 0.03*	
<i>Bandwidth</i>	
NB	18.1
BB	63.9
<i>p-Value for paired-sign test</i>	
NB vs. BB: 0.001**	
<i>Hemisphere</i>	
Left	40.3
Right	41.7
<i>p-Value for paired-sign test</i>	
Left vs. right: 0.69	

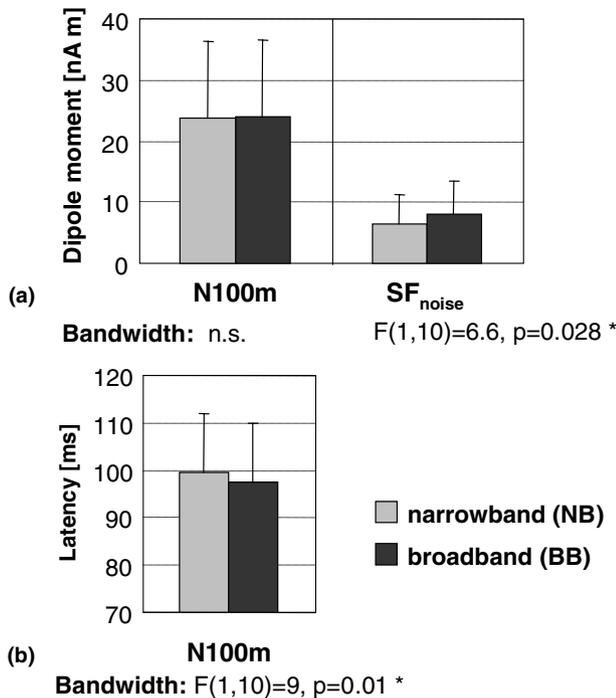


Fig. 6. Means and standard deviations of the dipole amplitudes and latencies for N100m (left) and  $SF_{noise}$  (right). Bandwidth had a significant effect on the latency (b), but not on the mean amplitude (a, left) of N100m. The dipole moment of  $SF_{noise}$  (a, right) was significantly higher at the BB condition.

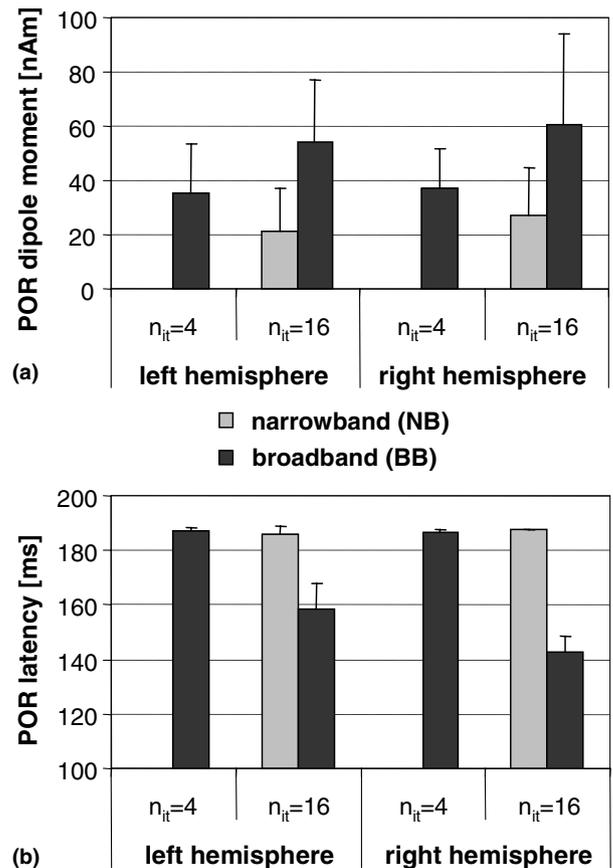


Fig. 7. Mean POR amplitudes (a) and latencies (b). Only conditions in which a POR was elicited in a sufficient number of subjects (NB:  $n_{it} = 16$ ; BB:  $n_{it} = 4$  and 16) are considered. Error bars show standard deviations.

bandwidth effect on the N100m amplitude is addressed in Section 5 ('Noise evoked neural activity: Effects of spectral bandwidth') in terms of inter-individual variability.

#### 4.2.2. POR and SF in response to RIS

For the broadband RIS with  $n_{it} = 16$ , there was a prominent POR in both hemispheres of all subjects. The proportion of subjects exhibiting a POR decreased substantially with decreasing bandwidth or number of iterations. There were no hemispheric differences regarding the presence or absence of the POR. A summary of the statistical results is provided in Table 1.

In the following, the amplitudes and latencies of the POR will be analyzed only for those three conditions which elicited a POR in a sufficient number of cases (NB:  $n_{it} = 16$ ; BB:  $n_{it} = 4$  and 16). The results are presented in Fig. 7 (a, amplitudes; b, latencies). There were no hemispheric differences either for the amplitude or the latency of the POR.

When  $n_{it}$  increased from 4 to 16 (BB stimulus), the mean dipole moment increased by 10.1 nAm (38.4%;  $t(10) = -3.9$ ,  $p = 0.003$ ) in the left hemisphere and by 12.7 nAm (47.7%;  $t(9) = -4$ ,  $p = 0.003$ ) in the right hemisphere. At the same time, the mean latency decreased by 30 ms ( $t(10) = -11.7$ ,  $p < 0.0001$ ) in the left hemisphere, and by 37 ms ( $t(9) = -12.2$ ,  $p < 0.0001$ ) in the right hemisphere. A comparison of the two bandwidths ( $n_{it} = 16$ ) shows that the BB stimuli produced larger mean dipole moments (left

hemisphere: +15.3 nAm (72.5%),  $t(4) = -3.3$ ,  $p = 0.028$ ; right hemisphere: +12.2 nAm (45%),  $t(3) = -3$ ,  $p = 0.05$ ) and shorter latencies (left hemisphere: -25 ms,  $t(4) = -6.2$ ,  $p = 0.0035$ ; right hemisphere: -45 ms,  $t(3) = -14.3$ ,  $p = 0.0007$ ) than the NB stimuli.

The dipole moment for SF<sub>RIS</sub> increased significantly with increasing  $n_{it}$  and spectral bandwidth. The results are graphically illustrated in the left panel of Fig. 8a. Relative to SF<sub>noise</sub>, SF<sub>RIS</sub> was larger by 2.1 nAm (31.9%), on the average. Fig. 9 illustrates the relative activity change from SF<sub>noise</sub> to SF<sub>RIS</sub>, for the three  $n_{it}$ -conditions (data averaged for both hemispheres). There were no hemispheric differences.

#### 4.2.3. Off-response

For the N100m-fit, three subjects (S2, S3, S11) did not show a consistent off-response and were excluded from the analysis of this component. For the POR-fit, subject S8 was excluded, as well, because of the low GOF value for the right hemisphere. As the results were very similar, they are illustrated only for the N100m-fit (Fig. 8). There was a significant effect of  $n_{it}$  on the latency (N100m-fit:  $F(2,8) = 6.2$ ,  $p = 0.01$ ; POR-fit:  $F(2,7) = 14.8$ ,  $p < 0.001$ ), but not on the amplitude of the off-response (N100m-fit:  $F(2,8) = 0.4$ ,  $p = 0.7$  n.s.; POR-fit:  $F(2,7) = 0.2$ ,  $p = 0.9$  n.s.). Moreover, the NB conditions resulted in significantly smaller dipole moments (N100m-fit:  $F(1,8) = 24.8$ ,

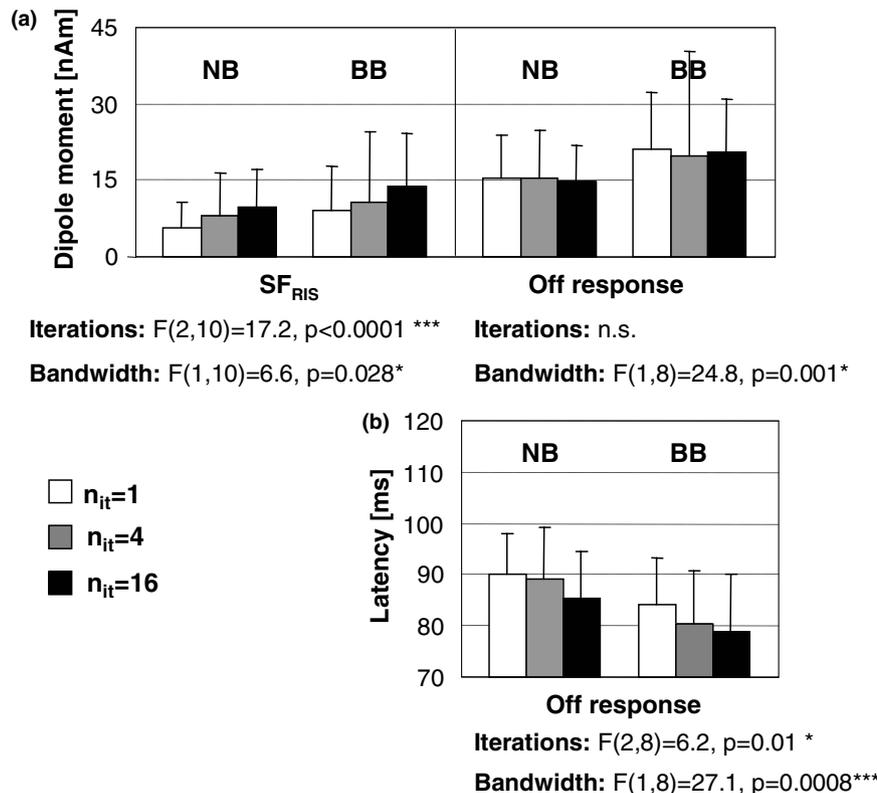


Fig. 8. Means and standard deviations of the dipole amplitudes and latencies for the RIS-induced SF<sub>RIS</sub> (left) and off-response (right). Both the  $n_{it}$  and the bandwidth of the RIS had a significant effect on the dipole moment of SF<sub>RIS</sub> (a, left). The results for the off response relate to the N100m fit (results for POR-fit: see text). The latency of the off-response varied significantly both with the  $n_{it}$  and the bandwidth of the RIS (b), while the dipole moment was only affected by the bandwidth (a, right).

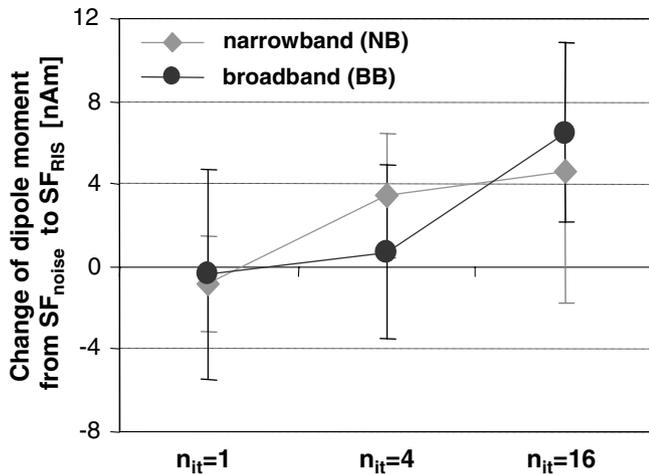


Fig. 9. Change of neural activity level from ongoing noise stimulation ( $SF_{noise}$ ) to ongoing RIS stimulation ( $SF_{RIS}$ ); error bars: standard deviations. While RIS segments with a poor pitch salience ( $n_{it} = 1$ ) do not noticeably change the neural activity level, as compared to random noise, RIS segments with an intermediate ( $n_{it} = 4$ ) to high ( $n_{it} = 16$ ) pitch salience progressively increase this level, at both spectral bandwidths.

$p = 0.001$ ; POR-fit:  $F(1,7) = 21.5$ ,  $p = 0.002$ ) and longer latencies (N100m-fit:  $F(1,8) = 27.1$ ,  $p = 0.0008$ ; POR-fit:  $F(1,7) = 66.9$ ,  $p < 0.0001$ ) than the BB conditions. As for all other comparisons, there was no significant hemispheric effect.

## 5. Discussion

### 5.1. Pitch-related neural activity

#### 5.1.1. Pitch onset response

As predicted by the model (Fig. 2), the pitch salience of the RIS had a substantial impact on the emergence, the amplitude, and the latency of the POR. The effect of  $n_{it}$  on the POR has already been described in Krumbholz et al. (2003). In that study,  $n_{it}$  was varied in doublings from 2 to 32. The POR amplitude increased monotonically across the  $n_{it}$ -range, whereas its latency decreased by about 35 ms when  $n_{it}$  increased from 4 to 16. The results of the present study are consistent with these findings.

A novel aspect of the current study is the finding that spectral bandwidth affects both the amplitude and the latency of the POR. The finding is consistent with the Auditory Image Model which predicts that pitch-induced response amplitudes should increase and latencies should decrease with pitch strength. In the simulated interspike-interval histograms (Fig. 2), the peak at the RIS delay of 16 ms is greater for the BB stimuli. Thus, the larger POR associated with the broadband RIS appears to be a direct correlate of pitch strength. According to the simulations, the decrease in POR latency with increasing spectral bandwidth or  $n_{it}$  may be explained as follows: BB stimuli produce more overall activity than NB stimuli, and consequently there are also more interspike intervals at the RIS delay. Therefore, the critical time for establishing a reliable

auditory image will be shorter. Such an interpretation is consistent with the finding of Rupp et al. (2003) that for short RIS segments embedded in noise, the minimal duration for RIS detection decreases with increasing pitch salience ( $n_{it}$ ).

The latency effect might also be expected from the known behaviour of the AEF. Transient AEF components reflect the simultaneous activity of an ensemble of neurons that form a functional entity. Broadly speaking, a high degree of synchrony not only increases the amplitude of the response, it also decreases the latency of the peak of the deflection (Hari, 1990). So, RIS segments with strong pitch salience might be expected to produce stronger cortical activation and shorter POR latencies, than RIS segments with less pitch salience. However, it should be noted that such a link between AEF amplitude and latency is not inevitable, and there are examples, where the parameters do not covary (e.g. Lütkenhöner et al., 2001; Seither-Preisler et al., 2003). Irregularities are to be expected, especially if the underlying source consists of several subcomponents. Depending on the signs, amplitudes, and latencies of the subcomponents, the dependence of the net deflection on the stimulus may be difficult to predict, and difficult to interpret (Lütkenhöner et al., 2006). In such cases, the parameters of the AEF deflection may be only loosely related to the stimulus, and an increase in dipole moment might be associated with an increasing, decreasing, or constant latency.

#### 5.1.2. RIS-induced sustained field

On average, the magnitude of  $SF_{RIS}$  was 31.9% larger as compared to  $SF_{noise}$ . This effect was not due to a slow drift in the magnetic field, but rather to an abrupt change in the baseline after pitch onset (c.f. Fig. 5). Remarkably, the dipole moment for  $SF_{RIS}$  increased systematically with  $n_{it}$  and spectral bandwidth. This implies that  $SF_{RIS}$  receives contributions from neural generators that are involved in the extraction of the pitch of ongoing sound signals. The result is consistent with the findings of Gutschalk et al. (2002), who recorded the SF in response to regular and irregular click trains presented at different rates and intensities. They identified two sources of the SF: a posterior one with a center of gravity in planum temporale, sensitive to intensity, and an anterior one with a center of gravity in lateral Heschl's gyrus, sensitive to temporal regularity. In a subsequent study, Gutschalk et al. (2004) used a discontinuous stimulation paradigm with regular and irregular click trains separated by pauses, and a continuous stimulation paradigm in which regular and irregular click trains alternated without interruption. The location and functional significance of the posterior and anterior source were largely confirmed, both for transient and sustained AEF components. The posterior source accounted for the regularity-unspecific N100m and subsequent SF. The sources of the two AEF components were located in the auditory fields of the posterior belt area ('paraconiocortex'), as specified by Braak (1978), Galaburda and Sanides (1980),

Rivier and Clarke (1997), and Wallace et al. (2002). The anterior source was located on lateral Heschl's gyrus, close to the focus of activity previously found for RIS-stimulation in an fMRI-experiment (Patterson et al., 2002). This source accounted for the transient response to the onset of temporal regularity (discontinuous stimulation: N130m, continuous stimulation: N150m) and for the subsequent SF. Analogous to the nomenclature of Krumbholz et al. (2003) regarding the response to pitch onset (POR), the regularity-sensitive SF was termed the 'sustained pitch response' (SPR). The source locations of the N130m, N150m, and SPR were very similar, though not identical. At least in the left hemisphere, the N150m was slightly posterior and more forward tilted as compared the SPR, suggesting that these components reflect activity from the same neural network, but from different cell populations.

The source location and pitch–latency function of wave N150m, evoked by the transition from irregular to regular click trains, largely resemble those of the POR, evoked by the transition from noise to RIS. This suggests that the transient POR/N150m and the sustained SF<sub>RIS</sub>/SPR are closely related and generated by a pitch processing unit in lateral Heschl's gyrus.

### 5.1.3. Off response

The off-response essentially marks the offset of sound energy (Hari et al., 1987). However, under certain circumstances, which probably include the present case, it may also signify the cessation of more specific sound features. This hypothesis is corroborated by our finding that the latency of the off-response gradually decreased with increasing pitch salience (Fig. 8b). The fact that the latency variability was substantially smaller than for the POR, may indicate that reliable identification of the onset of temporal regularity (POR) involves more complex and time consuming processes than the identification of its offset (off-response).

In the BB conditions, the off-responses were significantly larger (34%) and faster (7 ms) than in the NB conditions. If the effects were basically mediated by generators responding to the unspecific offset of sound energy, they would reflect the compressive properties of the auditory periphery. In particular, the disruption of a BB signal, exhibiting a relatively higher net firing rate (c.f. Fig. 1b), would result in a larger and faster off-response. If, on the other hand, the effects were essentially mediated by pitch-sensitive generators, they would indicate that the disrupted BB condition had a relatively higher pitch salience (c.f. Fig. 2). In the current experiment, both effects probably contributed to the observed bandwidth-effect.

## 5.2. Hemispheric asymmetry

There were no hemispheric differences with regard to either the latency or the amplitude of the AEF components that we analysed. This result is consistent with Gutschalk et al. (2004), who observed hemispheric differences neither

for the activation of the regularity-insensitive posterior source (N100m and SF) nor for the regularity-sensitive anterior source (N150m and SPR). This might at first seem surprising since many studies suggest a greater involvement of the right hemisphere in pitch-related tasks (Robin et al., 1990; Zatorre and Samson, 1991). In a recent review article (Zatorre et al., 2002), it was pointed out that asymmetry appears with higher level tasks, like the tracking of pitch contours, and in the more distributed neural networks associated with higher-level tasks. Patterson et al. (2002) performed an fMRI study on the representation of noise, RIS-pitch, and RIS-melody in the auditory pathway and examined the symmetry at each neural stage. RIS-pitch produced relatively symmetric activation in Heschl's gyrus and planum temporale. Pronounced asymmetries were largely limited to RIS melodies, and they emerged as the activation moved from Heschl's gyrus out onto planum polare and superior temporal gyrus, both of which are part of secondary auditory cortex. This suggests that there is no clear hemispheric specialization for the processing of pitch in the N100m generators with a centre of activity on planum temporale (Gutschalk et al., 2002, 2004; Lütkenhöner and Steinsträter, 1998), the generators of the SF on planum temporale and lateral Heschl's gyrus (Gutschalk et al., 2002, 2004), the POR generator in medio-lateral Heschl's gyrus (Krumbholz et al., 2003), and the generators of the off-response in the supratemporal plane next to the generators of the N100m and SF (Hari et al., 1987).

Schneider et al. (2002) examined the extent of asymmetries in auditory cortex and the question of whether they are consistent across subjects. They compared the neural representation of pure tones in non-musicians, musical amateurs, and professionals. The amplitudes of the middle latency components (N19m-P30m complex, occurring between 19 and 30 ms post stimulus onset) were found to be substantially larger in the right hemisphere, at least in musically trained subjects. These functional asymmetries were reported to coincide with morphological differences in the generating structures. No corresponding asymmetries were established for the later N100m component, generated by a more distributed network. This suggests that even primary auditory areas may be lateralized, either as a sign of high musical aptitude or as a result of long-term musical training. In a follow-up study, Schneider et al. (2005) continued to search for subject specific factors. They reviewed earlier reports that individuals largely differ in their tendency to hear out the pitch of the missing fundamental, or the overtone spectrum of harmonic complex tones (Smooenburg, 1970; Houtsma and Fleuren, 1991).<sup>1</sup> Subjects who tended to hear out the overtone spectrum showed greater P50m dipole moments, which was associated with a larger grey matter volume, in right Heschl's gyrus. The reverse was true for subjects who tended to hear

<sup>1</sup> Please note that this does not apply to RIS, which has only one pitch corresponding to *d*.

out the missing fundamental. The source of the P50m was located in the lateral portion of Heschl's gyrus, close to the source of the POR (Krumbholz et al., 2003). The data in our study are not suitable for comparison with those in Schneider et al.'s (2005) study, because we did not assess our subject's ability to hear out the pitch of the missing fundamental. Moreover, our experiment was not designed to measure the P50m. We can conclude, however, that our failure to find hemispheric differences is what would be expected from the N100m results of Schneider et al. (2002).

### 5.3. Noise evoked neural activity: effect of spectral bandwidth

Increasing the bandwidth reduced the latency of the transient N100m (Fig. 6b), but did not produce a consistent amplitude increase (Fig. 6a). This somewhat surprising result led us to perform an analysis of the individual data. The AEF recordings consisted of five runs per hemisphere, each of which comprised the three  $n_{it}$ -conditions of the RIS, so there were 15 independent samples of the N100m response available for each bandwidth and hemisphere. The calculation of the N100m-dipole moments from the 15 samples was based on the source specifications (location and direction of the dipole) that had been derived from the averaged response. The results of paired  $t$ -tests on the bandwidth dependence of the N100m amplitude are illustrated in Fig. 10. The  $t$ -values are indicated on the ordinate. The dashed horizontal lines mark the critical  $t$ -values for significant effects in either direction (negative: higher amplitude for BB condition; positive: vice versa). The tested hemispheres (right: grey, left: black) were ordered with regard to the magnitude of the  $t$ -value. Of the 22 comparisons performed (subjects S1–S12; S2 excluded because of an insufficient GOF), seven are significant at  $p < 0.05$ , one at  $p < 0.01$ , one at  $p < 0.001$ , and two at  $p < 0.0001$ . Thus, bandwidth definitely had an effect on the N100m amplitude. However, the direction of the effect varied: In six hemispheres, the BB condition, and in five hemispheres,

the NB condition produced the larger dipole moment. Also the two cases with the most pronounced effect showed opposite directions: In the left hemisphere of subject S1, the ratio of the mean dipole moments for the BB and the NB condition was 1.28, whereas in the right hemisphere of subject S3 this ratio was 0.49. The analysis of the individual data revealed significant effects in 11 of 22 hemispheres, which is far beyond the statistical probability that single tests become significant because of an erroneous rejection of the null hypothesis.

This unsystematic behavior becomes more comprehensible if the source structure of the N100m is taken into account. The deflection is generated by several sources in auditory cortex (Loveless et al., 1996; Näätänen and Picton, 1987; Sams et al., 1993). A variation of critical stimulus parameters, such as spectral bandwidth, inevitably changes the relative contributions of these sources. Depending on the signs, amplitudes, and latencies of the respective dipole maxima, the net amplitude of the resultant deflection could either increase or decrease (Lütkenhöner et al., 2003, 2006). As the relative contributions of the N100m sources vary with the neuro-anatomy of the auditory cortex, different subjects (and hemispheres) may show inconsistent response patterns (c.f. Fig. 10). This fact, together with methodological differences, might explain why Soeta et al. (2005) obtained results that appear contradictory not only with the presents study, but also with our previous work (Lütkenhöner et al., 2006; Seither-Preisler et al., 2003). Using pure tones and noises with a maximum bandwidth of one octave, they found a decreasing N100m peak amplitude with increasing bandwidth, in all subjects. In an fMRI study, Hawley et al. (2005) found that an increase in the width of a noise band from third octave to two octaves enhanced the neural activity in the cochlear nucleus and the superior olivary complex of the brainstem, and in the inferior colliculus of the midbrain. Hall et al. (2001) performed an fMRI study with pure tones and harmonic complex tones of varying bandwidth, and observed that bandwidth-related activity increased in the medial and lateral parts of the superior temporal gyrus of auditory

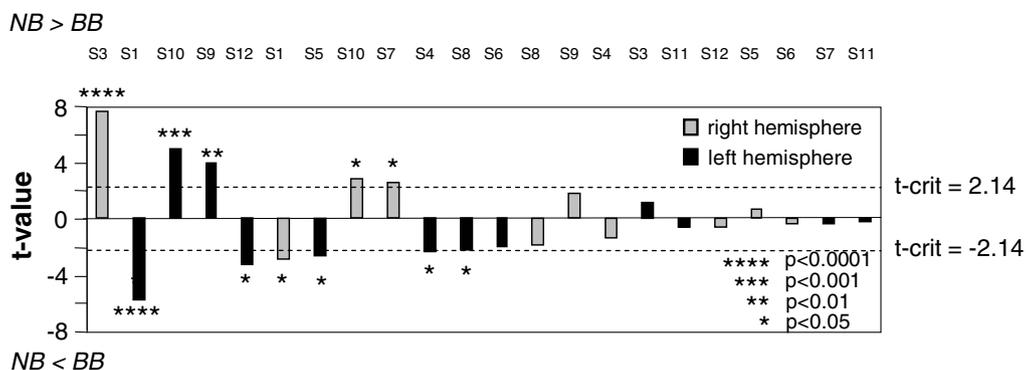


Fig. 10. Effect of spectral bandwidth on the N100m amplitudes in single hemispheres (grey: right, black: left; S1–S12: subjects, S2 excluded). The  $t$ -values obtained in paired  $t$ -tests are indicated on the ordinate. The dashed lines mark the critical  $t$ -values for significant effects in favor of the NB condition (positive values) and the BB condition (negative values). Results are ordered with regard to the magnitude of bandwidth-related differences. Eleven of the 22 performed comparisons are significant, indicating that the noise-bandwidth had an effect on the N100m amplitude.

cortex. In a MEG study with comparable stimuli, we obtained consistent results for the N100m (Seither-Preisler et al., 2003). However, harmonic complex tones provoke a clear pitch at their fundamental frequency. Our simulations with such stimuli (Fig. 2) show that an increase in bandwidth enhances both the level of stochastic and regular neural activity. Consequently, the increase in activity with bandwidth observed by Hall et al. (2001) and Seither-Preisler et al. (2003) may have been caused by a reduction in peripheral compression, as well as an increase in pitch salience. The present results appear to support the latter interpretation, at least with regard to the N100m: we were not able to reproduce a consistent bandwidth-related amplitude increase for noisy stimuli.

While the bandwidth effect on the N100m was highly subject-dependent and averaged out when calculating a grand average over all subjects, the effect on the noise-induced sustained field was much more consistent. The dipole moment of  $SF_{\text{noise}}$  was significantly higher in the BB condition (Fig. 6a), thus corroborating the model prediction of reduced peripheral compression for BB stimulation (Fig. 1b). Besides that, the dipole moment arising from noise stimulation was smaller than that arising from RIS-stimulation, provided that the RIS segment had a strong pitch (Fig. 9). This finding agrees with the generator structure proposed by Gutschalk et al. (2002, 2004) and suggests that, in the case of the  $SF_{\text{noise}}$ , only the intensity-source was active, while in the case of the  $SF_{\text{RIS}}$  the intensity and the pitch-sources were simultaneously active, resulting in higher overall response amplitudes.

## Acknowledgements

This study was supported by the Austrian Academy of Sciences as APART project 524 and by the Alexander von Humboldt Foundation (Institutspartnerschaft). R.P. was supported by the UK Medical Research Council (G9901257; G0500221). We especially thank Dr. Enriquez Lopez-Poveda for his instructive suggestions.

## References

- Bilsen, F., 1966. Repetition pitch: monaural interaction of a sound with the repetition of the same, but phase shifted, sound. *Acustica* 17, 295–300.
- Braak, H., 1978. On magnopyramidal temporal fields in the human brain-probable morphological counterparts of Wernicke's sensory speech region. *Anat. Embryol.* 132, 141–169.
- Cariani, P., 1999. Temporal coding of periodicity pitch in the auditory system: an overview. *Neural Plasticity* 6, 147–172.
- Galaburda, A., Sanides, F., 1980. Cytoarchitectonic organization of the human auditory cortex. *J. Comp. Neurol.*, 597–610.
- Gutschalk, A., Patterson, R., 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.
- Gutschalk, A., Patterson, R.D., Scherg, M., Uppenkamp, S., Rupp, A., 2004. Temporal dynamics of pitch in human auditory cortex. *Neuroimage* 22, 755–766.
- Hall, D.A., Haggard, M.P., Summerfield, A.Q., Akeroyd, M.A., Palmer, A.R., Bowtell, R.W., 2001. Functional magnetic resonance imaging measurements of sound-level encoding in the absence of background scanner noise. *J. Acoust. Soc. Am.* 109, 1559–1570.
- Hari, R., 1990. The neuromagnetic method in the study of the human auditory cortex. In: Grandori, F., Hoke, M., Romani, G.L. (Eds.), *Auditory Evoked Magnetic Fields and Electric Potentials*, Adv. Audiol., vol. 6. Karger, Basel, pp. 222–282.
- Hari, R., Pelizzone, M., Makela, J.P., Hallstrom, J., Leinonen, L., Lounasmaa, O.V., 1987. Neuromagnetic responses of the human auditory cortex to on- and offsets of noise bursts. *Audiology* 26, 31–43.
- Hawley, M.L., Melcher, J.R., Fullerton, B.C., 2005. Effects of sound bandwidth on fMRI activation in human auditory brainstem nuclei. *Hear. Res.* 204, 101–110.
- Houtsma, A.J.M., Fleuren, J.F.M., 1991. Analytic and synthetic pitch of two-tone complexes. *J. Acoust. Soc. Am.* 90 (3), 1674–1676.
- Krumbholz, K., Patterson, R., 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.
- Lopez-Poveda, E.A., Meddis, R., 2001. A human nonlinear cochlear filterbank. *J. Acoust. Soc. Am.* 110, 3107–3118.
- Loveless, N., Levanen, S., Jousmaki, V., Sams, M., Hari, R., 1996. Temporal integration in auditory sensory memory: neuromagnetic evidence. *Electroencephalogr. Clin. Neurophysiol.* 100, 220–228.
- 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., Seither-Preisler, A., 2003. Studies of tonotopy based on wave N100 of the auditory evoked field are problematic. *NeuroImage* 19, 935–949.
- Lütkenhöner, B., Seither-Preisler, A., Seither, S., 2006. Piano tones evoke stronger magnetic fields than pure tones or noise, both in musicians and non-musicians. *NeuroImage* 30, 927–937.
- Lütkenhöner, B., Krumbholz, K., Lammertmann, C., Seither-Preisler, A., Steinsträter, O., Patterson, R.D., 2003b. Localization of primary auditory cortex in humans by magnetoencephalography. *NeuroImage* 18, 58–66.
- Martin, B.A., Boothroyd, A., 1999. Cortical auditory event related potentials in response to periodic and aperiodic stimuli with the same spectral envelope. *Ear Hear.* 20, 33–44.
- Martin, B.A., Boothroyd, A., 2000. Cortical auditory evoked potentials in response to changes of spectrum and amplitude. *J. Acoust. Soc. Am.* 107, 2155–2161.
- Meddis, R., 1986. Simulation of mechanical to neural transduction in the auditory receptor. *J. Acoust. Soc. Am.* 79, 702–711.
- Meddis, R., 1988. Simulation of auditory-neural transduction: further studies. *J. Acoust. Soc. Am.* 83, 1056–1063.
- Meddis, R., O'Mard, L., 1997. A unitary model of pitch perception. *J. Acoust. Soc. Am.* 102, 1811–1820.
- Meddis, R., O'Mard, L.P., Lopez-Poveda, E.A., 2001. A computational algorithm for computing nonlinear auditory frequency selectivity. *J. Acoust. Soc. Am.* 109, 2852–2861.
- Moore, B.C., Glasberg, B.R., 1996. A revision of Zwicker's loudness model. *Acta Acust.* 82, 335–345.
- Moore, B.C., Glasberg, B.R., Baer, T., 1997. A model for the prediction of threshold, loudness, and partial loudness. *J. Audio. Eng. Soc.* 45, 224–240.
- Näättänen, 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.
- Ostroff, J.M., Martin, B.A., Boothroyd, A., 1998. Cortical evoked response to acoustic change within a syllable. *Ear Hear.* 19, 290–297.

- Patterson, R.D., Irino, T., 1998. Modeling temporal asymmetry in the auditory system. *J. Acoust. Soc. Am.* 104, 2967–2979.
- Patterson, R.D., Allerhand, M.H., 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.
- Patterson, R.D., Robinson, K., Holdsworth, J., 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.
- Pressnitzer, D., Patterson, R.D., Krumbholz, K., 2001. The lower limit of melodic pitch. *J. Acoust. Soc. Am.* 109, 2074–2084.
- Rivier, F., Clarke, S., 1997. Cytochrome oxidase, acetylcholinesterase, and NADPH-diphorase staining in human supratemporal and insular cortex: evidence for multiple auditory areas. *NeuroImage* 6, 288–304.
- Robin, D.A., Tranel, D., Damasio, H., 1990. Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain Lang.* 39, 539–555.
- Rupp, A., Uppenkamp, S., Bailes, J., Gutschalk, A., Patterson, R., 2003. Time constants in temporal pitch extraction: a comparison of psychophysical and neuromagnetic data. In: Pressnitzer, D., de Cheveigne, A., Mc Adams, S., Collet, L. (Eds.), *Auditory Signal Processing: Physiology, Psychoacoustics, and Models*, Proceedings of the 13th International Symposium on Hearing. Springer, New York, Dourdan, pp. 119–125.
- Sams, M., Hari, R., Rif, J., Knuutila, J., 1993. The human auditory sensory memory trace persists about 10 s: neuromagnetic evidence. *J. Cogn. Neurosci.* 5, 363–370.
- Schneider, P., Scherg, M., Dosch, H.G., Specht, H.J., Gutschalk, A., Rupp, A., 2002. Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat. Neurosci.* 5, 688–694.
- Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H.J., Dosch, H.G., Bleeck, S., Stippich, C., Rupp, A., 2005. Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nat. Neurosci.* 8, 1241–1247.
- Seither-Preisler, A., Krumbholz, K., Lütkenhöner, B., 2003. Sensitivity of the neuromagnetic N100m deflection to spectral bandwidth: a function of the auditory periphery? *Audiol. NeuroOtol.* 8, 322–337.
- Seither-Preisler, A., Krumbholz, K., Patterson, R., Seither, S., Lütkenhöner, B., 2004. Interaction between the neuromagnetic responses to sound energy onset and pitch onset suggests common generators. *Eur. J. Neurosci.* 19, 3073–3080.
- Seither-Preisler, A., Patterson, R., Krumbholz, K., Seither, S., Lütkenhöner, B., 2006. Evidence of pitch processing in the N100m component of the auditory evoked field. *Hear. Res.* 213, 88–98.
- Shamma, S.A., Chadwick, R.S., Wilbur, W.J., Morrish, K.A., Rinzel, J., 1986. A biophysical model of cochlear processing: Intensity dependence of pure tone responses. *J. Acoust. Soc. Am.* 80, 133–145.
- Shofner, W.P., Selas, G., 2002. Pitch strength and Stevens's power law. *Percept. Psychophys.* 64, 437–450.
- Smooenburg, G.F., 1970. Pitch perception of two-frequency stimuli. *J. Acoust. Soc. Am.* 48, 924–942.
- Soeta, Y., Nakagawa, S., Tonoike, M., 2005. Auditory evoked magnetic fields in relation to bandwidth variations of bandpass noise. *Hear. Res.* 202, 47–54.
- Sumner, C., Lopez-Poveda, E.A., O'Mard, L.P., Meddis, R., 2002. A revised model of the inner-hair cell and auditory-nerve complex. *J. Acoust. Soc. Am.* 111, 2178–2188.
- Terhardt, E., Stoll, G., Seewann, M., 1982. Pitch of complex signals according to virtual pitch theory: tests, examples and predictions. *J. Acoust. Soc. Am.* 71, 671–678.
- 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.
- Yates, G.K., Winter, I.M., Robertson, D., 1990. Basilar membrane nonlinearity determines auditory nerve rate-intensity functions and cochlear dynamic range. *Hear. Res.* 45, 203–220.
- Yost, W.A., 1996. Pitch of iterated rippled noise. *J. Acoust. Soc. Am.* 100, 511–518.
- Yost, W.A., Patterson, R., Sheft, S., 1996. A time domain description for the pitch strength of iterated rippled noise. *J. Acoust. Soc. Am.* 99, 1066–1078.
- Zatorre, R.J., Samson, S., 1991. Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain* 114, 2403–2417.
- Zatorre, R.J., Belin, P., Penhune, V.B., 2002. Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci.* 6, 37–46.
- Zwicker, E., Fastl, H., 1990. *Psychoacoustics – Facts and Models*. Springer, New York.