MEG Auditory Research

Alexander Gutschalk

Abstract This chapter reviews auditory research performed with magnetoencephalography (MEG) in normal listeners, with an emphasis on the auditory cortex. The first section provides an overview of basic characteristics of auditory evoked fields and their classification. The second section reviews the relationship between a selection of basic auditory features—including lateralization, periodicity, and spectral content—and auditory evoked fields generated in auditory cortex. The final section highlights recent MEG research in the field of auditory scene analysis, focusing specifically on auditory stream segregation, selective attention, and informational masking.

Keywords Auditory cortex · Auditory evoked fields · Selective adaptation · Pitch · Sound lateralization · Vowel · Auditory scene analysis · Stream segregation · Selective attention · Informational masking · Perceptual awareness

1 Introduction

Acoustic signals unfold on a multitude of timescales, ranging from the sub-millisecond processes supporting sound localization to the multi-second intervals necessary for music perception and speech comprehension. Sounds impinging on the ear are transformed into a frequency-specific neural code in the cochlea. This neural code in the auditory nerve has a sub-millisecond temporal precision and can phase-lock to periodic sound waves up to 5,000 Hz (Young and Sachs 1979). Frequency specificity is maintained in the ascending auditory pathway up to the auditory cortex, with an orderly mapping of frequency that is called tonotopy.

A. Gutschalk (✉)
Department of Neurology, University of Heidelberg,
Im Neuenheimer Feld 400, 69120 Heidelberg, Germany
e-mail: Alexander.Gutschalk@med.uni-heidelberg.de

While the ability to phase lock to the acoustic stimulus degrades along the ascending auditory pathway, some aspects of coding in human primary auditory cortex still maintain a millisecond precision and phase-locking capability up to at least 100 Hz (Brugge et al. 2009).

MEG is an excellent tool for studying the human auditory system for several reasons. First and foremost, MEG's temporal resolution matches the resolution with which the brain responds to sound. Second, owing to the situation of auditory cortex on the superior temporal plane, with dipole sources oriented primarily tangential to the head surface, MEG is particularly sensitive to activity generated there and can straight-forwardly discriminate activity arising from the left and right hemispheres. Third MEG acquisition is silent, a clear advantage when compared with modern fMRI acquisition sequences. The first auditory evoked response in MEG was published in the late 1970s (Reite et al. 1978). Using dipole source analysis, other early studies clearly demonstrated that these auditory-evoked fields were generated in the auditory cortex (Hari et al. 1980), and demonstrated tonotopy in the human auditory cortex (Romani et al. 1982). Today, more than 1,000 published studies of the auditory system have used MEG.

This review summarizes aspects of basic auditory neuroscience using MEG in healthy adult listeners. The chapter focuses on activity in the auditory cortex, with less focus on activity in other brain areas. The chapter starts in Sect. 2 with a classification of the different aspects of activity evoked by auditory stimuli as seen by MEG. Section 3 reviews the relationship between specific acoustic features and the MEG response, while Sect. 4 focuses on the perception of more complex auditory scenes. The selection of studies reviewed here is strongly biased towards studies using MEG because of the scope of this book; studies using EEG, a method very related to MEG, as well as intracranial EEG and fMRI studies, are mentioned only occasionally.

2 Classification of Auditory Evoked MEG Activity

The classification of auditory responses used in this chapter is primarily one that is based on the anatomical site of their generation, and as such is a view from source space. Three sites are dissociated: brainstem, auditory cortex, and multi-modal areas beyond auditory cortex. Most auditory studies using MEG have focused on auditory cortex, and therefore activity generated there comprises the largest part of this section. Traditionally, auditory evoked (magnetic) fields (AEF) have been subdivided into three latency ranges, in accordance with the classification of auditory evoked potentials (AEP) in EEG (Picton et al. 1974). In this chapter, the division of auditory evoked fields into early (up to 8 ms), middle (15–50 ms), and long-latency (>50 ms) ranges is introduced alongside the generator-based view. Still other types of activity are not covered by the latency classification, such as steady-state responses and induced activity (i.e., activity that is not precisely phase locked to

stimulus presentation). Each of these classifications has its own limitations, but some basic knowledge of how they have been used is helpful before discussing research that addresses questions of auditory neuroscience more specifically.

2.1 Brainstem

Occurring in the first 8 ms post stimulus onset, the early auditory-evoked field (EAEF) is also referred to as the auditory brainstem response (ABR). The ABR typically comprises five subsequent peaks, known as waves I–V. These components are small relative to either the ongoing MEG or later auditory evoked components, and therefore require large numbers of trials (thousands) in order to achieve an adequate signal to noise ratio. The typical stimuli used to evoke the ABR are clicks presented with inter-stimulus intervals (ISI) in the range of 50–100 ms. Waves I–V of the ABR have prominent spectral power in the range from 700 to 1,200 Hz. High sampling rates are therefore required to record the ABR and the low-pass filter should not be set below 1,000 Hz (better still 1,500 Hz). Highpass filters up to 200 Hz are typically used to suppress the low-frequency components of the later cortical responses that overlap the ABR because of the short ISI. There exist only a few published studies that have used MEG to study the ABR (Lütkenhöner et al. 2000; Parkkonen et al. 2009). These studies show that waves I–V can be recorded in MEG and that the estimated generators are consistent with their EEG counterparts (Scherg and von Cramon 1985). In brief, waves I and II are thought to be generated in the auditory nerve just beyond the cochlea, while wave V, with a latency of 5–6 ms post stimulus, is generated below the inferior colliculus, the obligatory auditory-midbrain nucleus, and probably reflects neuronal input to this structure.

2.2 Auditory Cortex

Both the middle- and long-latency AEF (MAEF and LAEF, respectively) are primarily generated in the auditory cortex (Fig. 1), and their separation at 50 ms is arbitrary. Historically, the MAEF peaks have been denoted with letters (e.g., N_am, P_am, N_bm, and P_bm), and the LAEF peaks with numbers (P_1m, N_1m, and P_2m). Alternatively, these peaks are labeled with their prototypical peak latency. In this nomenclature, the MAEF peaks are N19m, P30m, N40m and P50m; the LAEF peaks are known as P50m, N100m, and P200m. Denoting these peaks as negative (N) or positive (P) was originally in reference to the scalp vertex in EEG, but can be easily adopted with reference to the surface of the auditory cortex in MEG. The P50m has been considered both middle- (P_bm) and long-latency (P_1m), indicating one of the limitations of the latency-based nomenclature. Nevertheless, the dissociation of MAEF and LAEF is often useful and therefore these peaks will be introduced in separate paragraphs below.

2.2.1 Middle-Latency Auditory Evoked Fields

Most of the spectral energy of the early MAEF lies in the (lower) gamma band around 30–60 Hz, with a maximum around 40 Hz. For recording of the MAEF, the low-pass filter cutoff should therefore not be set below 100 Hz. A high-pass filter with cutoff frequencies in the range of 10–30 Hz is often used to suppress overlapping LAEF components (Fig. 2), because the typical ISI to record the MAEF is around 100–200 ms, and thus shorter than the LAEF. The most prominent peak of the MAEF is the P30m (Pelizzone 1987; Mäkelä et al. 1994; Pantev 1995). The preceding N19m is smaller, but has been consistently localized in Heschl's gyrus and close to the generator of the P30m (Hashimoto et al. 1995; Gutschalk et al. 1999; Rupp et al. 2002b; Parkkonen et al. 2009). It has been suggested that the N19m and the P30m share the same macroscopic generator in medial Heschl's gyrus, whereas the P50m is generated more lateral along Heschl's gyrus (Scherg et al. 1989; Yvert et al. 2001), a view that is supported by depth electrode recordings in patients with epilepsy (Liegeois-Chauvel et al. 1991; Liegeois-Chauvel et al. 1994).

With reference to microscopic anatomy, the sources of the N19m and P30m are in the auditory core area (Galaburda and Sanides 1980), most likely in the primary auditory cortex field A1. A less-likely alternative is that the N19m and the P30m are generated in the more medial core field CM (Hackett et al. 2001). The more lateral localization of the P50m would better match with a generator in the lateral core field R, but this is more speculative, and it is likely that other fields additionally contribute to the P_1m peak measured in MEG (Yvert et al. 2001). Laminar recordings of click-evoked activity in monkey A1 show a peak N8 that is generated in deep cortical layers (4 and 5), and a subsequent P24 which is generated predominantly in layer 3 (Steinschneider et al. 1992). One hypothesis is that human N19m is also generated by thalamocortical input into the granular layer 4.

2.2.2 Long-Latency Auditory Evoked Fields

Traditionally, the earliest peak of the LAEF is the P_1m, which has already been mentioned in the context of the MAEF. The earliest latency of the P_1m in response to pure-tone stimuli is typically in the range of 50 ms—hence P50m—(Pantev et al. 1996b). There is at least a second subcomponent of the P1 m with a peak latency around 70 ms (Yvert et al. 2001), and for clicks-train stimuli P_1m latencies around 60–80 ms are typically observed (Gutschalk et al. 2004a). One reason for P_1m variability is that the peak, especially when it is later than 50 ms, overlaps with the onset of the larger N_1m, which may reduce the P_1m amplitude and latency (Königs and Gutschalk 2012).

By far the most prominent peak of the AEF is the N_1m, which comprises a number of subcomponents (Näätänen and Picton 1987) whose specific features are reviewed in more detail below. Optimal recording of the N_1m requires an ISI of 500 ms or longer. The spectral content of the N_1m lies primarily in the theta band and the lower alpha band (approximately 3–10 Hz), such that low-pass filters
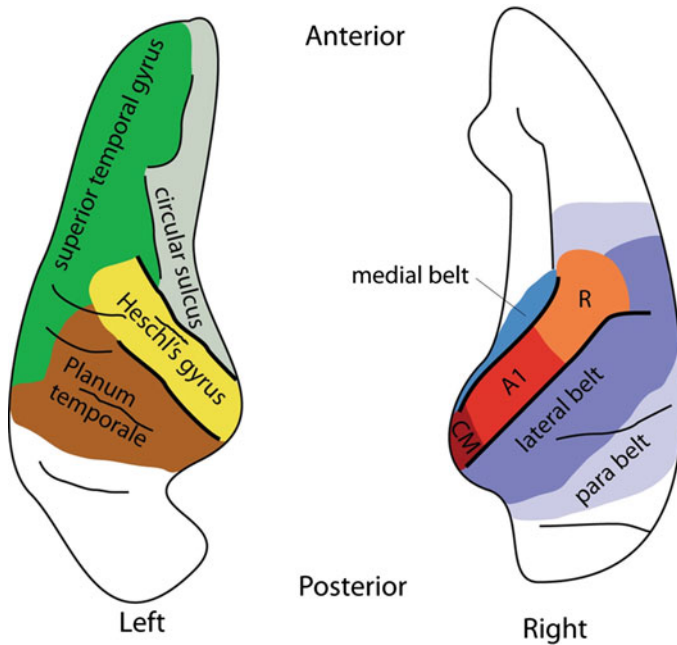
Fig. 1 Schematic of the human auditory cortex. The view is from the *top* on the superior temporal plane, which is buried inside the Sylvian fissure in the intact brain. The macroscopic anatomy is labeled on the *left*: Most lateral is the superior temporal gyrus (STG), most of whose surface extends to the lateral surface not seen in this view. Heschl's gyrus extends from postero-medial to anterior-lateral, where it meets the STG. The border between Heschl's gyrus and the STG is not sharply defined, and in some brains it appears as if the anterior STG is the continuation of Heschl's gyrus. Many subjects have more than one Heschl's gyrus, especially in the lateral part. The Planum temporale starts posterior from Heschl's gyrus and extends up to the temporo-parietal junction. There is no sharp border between the Planum temporale and the STG. A simplified schematic of histological auditory cortex fields is provided on the right: The core area (also primary auditory cortex, koniocortex, or Brodman Area 41) roughly coincides with the borders of Heschl's gyrus. Most anatomists subdivide the core region in at least two to three subregions. The nomenclature used here is adopted from the nomenclature used in the monkey (Hackett et al. 2001). The most medial field CM is not always considered a core field. The field A1 is often considered "primary auditory cortex" sensu stricto, but cannot be easily separated from the more lateral field R based on histology. These two fields have opponent tonotopic organizations (Formisano et al. 2003); in A1, high frequencies are localized postero-medially and low frequencies more antero-laterally, and vice versa in R, so that both fields share a common low-frequency border. An alternative nomenclature for the core fields is, from medial to lateral, Te1.1, Te1.0, and Te1.2 (Morosan et al. 2001). The lateral belt is located in the Planum temporale and extends to the lateral surface of the STG (Braak 1978). It can also be subdivided in at least two to three subfields oriented parallel to the core region, but there is only little information available from human anatomy (Rivier and Clarke 1997). Alternative names for areas that overlap with the lateral-belt definition are parakoniocortex, Brodman Area 42, or Te2 (Morosan et al. 2005). The anterior belt field is located in the circular sulcus, anterior from Heschl's gyrus. This area is also referred to as Prokoniokortex (Galaburda and Sanides 1980). The belt cortex is probably surrounded by the putative parabelt, which includes but may not be limited to Brodman Area 22 or Field Te3 (Morosan et al. 2005); these fields extend far into the lateral STG not seen on the view used here. Note that the different nomenclatures don't map on each other easily and that there is considerable inter-individual variability
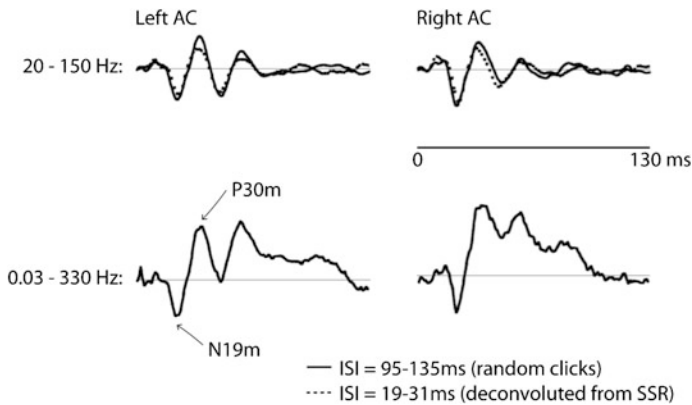
Fig. 2 Middle-latency auditory-evoked fields (MAEF). The data shown are source waveforms based on dipoles in medial Heschl's gyrus—supposedly in A1—averaged across six listeners. The stimuli were clicks presented with a randomized ISI in the range 95–135 ms (for the data plotted in solid lines). A comparison of two filter settings is shown for the upper (20–150 Hz) and lower (0.03–330 Hz) traces. As can be seen, the peaks N19m and P30m are clearly observed with both setting. The subsequent peaks N41m and P50m are elevated by a slower positivity; the latter is not clearly definable and because of the fast repetition rate might comprise a mixture of slower components. This positive shift has been removed by high-pass filtering in the upper traces. The dotted lines show the waveforms deconvolved from a steady-state response (SSR) with seven different rates (19–31 ms) in the same listeners. Note the high similarity between the early peaks and the traditionally obtained MAEF (Gutschalk et al. 1999)

down to 20 Hz cutoff frequency can usually be applied without any appreciable effect on component morphology (Fig. 3). High-pass filters are typically chosen in the range of 0.3–3 Hz, depending on the low-frequency noise level and whether later, slower components are also of interest.

The best-studied subcomponent of the N_1m, termed the N100m, peaks at a latency around 100 ms and is generated on the superior temporal plane (Hari et al. 1980). Based on co-registration with anatomical MRI, both Heschl's gyrus (Eulitz et al. 1995) and the planum temporale, just posterior to Heschl's gyrus (Lütkenhöner and Steinstrater 1998), are thought to be generators of this subcomponent. One important feature of the N100m is the large ISI range—up to 10 s—below which it will not reach its maximal amplitude (Hari et al. 1982; Pantev et al. 1993; Sams et al. 1993b; McEvoy et al. 1997). This effect is diminished for other N_1m subcomponents, which peak at slightly longer latencies (130–150 ms). One subcomponent was localized to the superior temporal gyrus (STG) (Lü et al. 1992), and might be identical to a radial N150 component described in EEG (Scherg and Von Cramon 1986). Another N_1m subcomponent has been consistently observed about 1 cm anterior to the main N100m peak and with a latency around 130 ms (Sams et al. 1993b; Loveless et al. 1996; McEvoy et al. 1997; Gutschalk et al. 1998). Note that the latencies of these N_1m subcomponent are not fixed but vary considerably with the onset and fine-structure of the stimuli used.
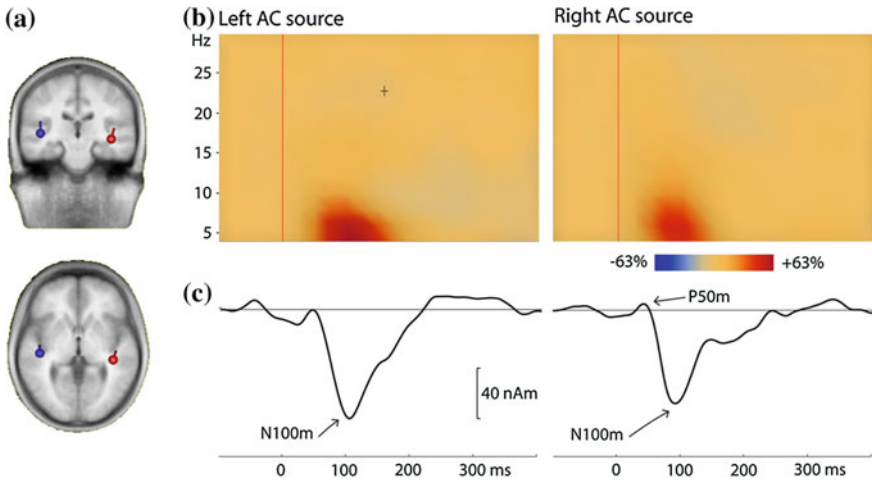
Fig. 3 Long-latency auditory-evoked fields (LAEF) from a single listener. The stimuli were 100-ms long pure tones with frequencies in the range 500–3,000 Hz, presented with a fixed ISI of 800 ms; frequency was randomly changed after 10 s. **a** Two dipoles were fitted to the averaged data in the time range 80–100 ms. **b** Time-frequency plots for the time range −100–400 ms and the frequency range 4–30 Hz. The plot shows the enhancement of power in comparison to the baseline in the time interval 100 ms before tone onset. Most of the signal power is in the theta band. **c** The averaged evoked response is shown for the same time range as the time-frequency analysis (low-pass filter = 30 Hz, no high-pass filter). The most prominent component is the N100m. These source waveforms as well as the time-frequency plots are based on the dipoles shown in (**a**)

The latency of the subsequent P_2m is around 150–250 ms (P200m). Sometimes the P_2m has been studied together with the N_1m by using a peak-to-peak measure. The few studies that specifically studied the P_2m found that the generator of the response is typically located anterior to the N100m (Hari et al. 1987; Pantev et al. 1996a).

For tones longer than about 50 ms, the P_2m is followed by a negative wave—the so-called sustained field—whose duration is directly linked to the stimulus duration. To obtain sustained fields, high-pass filters below 0.5 Hz or direct-coupled recordings should be used. The sustained field can be fitted by a dipole source that is typically located anterior to the N100m in auditory cortex (Hari et al. 1980; Hari et al. 1987; Pantev et al. 1994; Pantev et al. 1996a). Based on parametrical variation of sound features such as temporal regularity (see Sect. 3.4) or sound intensity, at least two subcomponents of the sustained field can be separated in lateral Heschl's gyrus and the planum temporale, similar to the N_1m subcomponents (Gutschalk et al. 2002). With respect to microscopic anatomy, the sources of the N_1m and the sustained field subcomponents are probably distributed across core and belt fields (Fig. 1).

Importantly, components of the N_1m are not only evoked at sound onset from silence, but by all kinds of changes within an ongoing sound (Mäkelä et al. 1988; Sams et al. 1993a). Finally, sounds that are played for a second or longer will also evoke an offset response. This offset response comprises mainly peaks N_1m and

P_2m, whose amplitude varies with sound duration like the onset peaks vary with the silent ISI (Hari et al. 1987; Pantev et al. 1996a).

2.2.3 Selective Adaptation and the Mismatch Negativity

As was briefly noted in the previous paragraph, the N_1m amplitude is determined in part by the ISI between the serial tones that are used to evoke the response (Hari et al. 1982; Imada et al. 1997). This observation is based on simple paradigms, where the same sound is repeated once or continuously. The response to each tone is reduced or adapted by the previous tone(s) of the sequence, and more so when the ISI is short. When two different tones are alternated instead, the adaptation of the N_1 depends additionally on how different these tones are from each other, as has been shown by several EEG studies (Butler 1968; Näätänen et al. 1988): when pure tones are used, the adaptation is strong when the frequencies of the two tones are near to each other; much less adaptation is observed when the tones are an octave or more apart. This phenomenon is referred to as *selective* or *stimulus-specific* adaptation. Selective adaptation is not limited to the N_1m, and has more recently been demonstrated for the P_1m (Gutschalk et al. 2005).

Another classical auditory stimulus paradigm that uses two tones dissociated by their tone frequency (or other features) is the auditory oddball paradigm (Näätänen et al. 1978). In contrast to the paradigm used to evaluate selective adaptation, the two tones are not simply alternated, but are presented at different probabilities. The more frequent tone is referred to as the standard, whereas the rare tone is referred to as the deviant. The ISI between subsequent tones is typically chosen around 300 ms, where the N_1m is almost completely suppressed. In this setting, a prominent negative response with peak latency around 130–200 ms is evoked by the rare deviants, but not by the frequent standard tones. This negative wave, called the mismatch negativity (MMN), is separated from other response components by subtracting the response to standards from the response to deviants. Many studies have examined the MMN and cannot be reviewed here in detail; extensive reviews on this component are already available (Garrido et al. 2009; May and Tiitinen 2010; Näätänen et al. 2011). Briefly, the MMN is not only evoked by differences in tone frequency, but by any sound difference between standard and deviant that is above the listener's threshold. Originally, the MMN was considered to be a component that is distinct from the other LAEF components reviewed in the previous section. However, this view has recently been challenged: a number of studies suggest that the MMN is identical to the anterior N_1m subcomponent, which is reduced in response to the standards but not in response to the deviants due to selective adaptation (May et al. 1999; Jääskeläinen et al. 2004; May and Tiitinen 2010). This view is supported by microelectrode studies in monkey, which suggest that—at least in A1—there is no evidence of an additional evoked component in the context of deviants presented in an oddball paradigm (Fishman and Steinschneider 2012). The associated debate of whether the MMN reflects (bottom-up) selective adaptation (May and Tiitinen 2010), or

(top-down) predictive coding (Garrido et al. 2009; Wacongne et al. 2012) is ongoing.

Finally, the MMN itself is not a single component with a stable topography, but comprises at least two subcomponents in the auditory cortex (Imada et al. 1993; Kretzschmar and Gutschalk 2010). Moreover, it has been suggested that the MMN receives contributions from generators in the frontal cortex (Schönwiesner et al. 2007). Second, a subsequent slow negativity that persists for 600 ms is additionally evoked by the oddball paradigm, which is also generated in the more anterior aspect of the auditory cortex along with the generator of the classical MMN (Kretzschmar and Gutschalk 2010).

2.2.4 Auditory Steady-State Responses

The auditory cortex is able to time-lock to periodic stimuli, a phenomenon that has been studied in particular at rates around 40 Hz (Romani et al. 1982; Mäkelä and Hari 1987) (Fig. 4). A periodic brain response that is imposed by a periodic stimulus is referred to as steady- state response (SSR) in EEG and MEG research. Steady-state responses require an evoked component whose inherent spectral power overlaps with the rate of the periodic repetition. As a result, the spectral representation of an SSR is a narrow band at the frequency of the periodic stimulus and sometimes its harmonics. Accordingly, a relationship between the 40-Hz SSR and the early MAEF peaks, whose spectral maximum is close to 40 Hz, was suggested early on (Galambos et al. 1981; Hari et al. 1989), and steady-state responses in the range of 30–50 Hz can be explained by assuming an identical response convolved with the periodic pulse train used as the stimulus. Conversely, when the underlying response is deconvolved on the basis of this assumption (Gutschalk et al. 1999), it shows high similarity with the early MAEF peaks recorded with a transient stimulus paradigm (Fig. 2). The main source of the 40-Hz SSR is in the medial half of Heschl's gyrus, and thus most likely in the primary area A1 (Fig. 1). This has been demonstrated by source analysis of MEG data (Pantev et al. 1996b; Gutschalk et al. 1999; Brookes et al. 2007), and was confirmed by intracranial recordings (Brugge et al. 2009) and fMRI (Steinmann and Gutschalk 2011). Note that other aspects of the 40-Hz SSR are not readily explained by ongoing, non-refractory MAEF activity. For example, the 40-Hz SSR shows a buildup of activity over about 250 ms before it reaches its constant amplitude (Ross et al. 2002), and this process starts over when, for example, a short sound in another frequency band is presented in parallel (Ross et al. 2005b). Potentially, these effects are related to secondary, more lateral generators of the 40-Hz SSR along Heschl's gyrus (Gutschalk et al. 1999) up to the superior temporal gyrus (Nourski et al. 2013).

Steady-state responses are not limited to the 40-Hz range: Higher frequency SSRs are observed in relationship to the ABR, known as the frequency following response, but this application has so far been limited to EEG research. In the lower frequency range, it has been demonstrated that SSRs power decreases with increasing modulation rate between 1.5 and 30 Hz (Wang et al. 2012); at the single
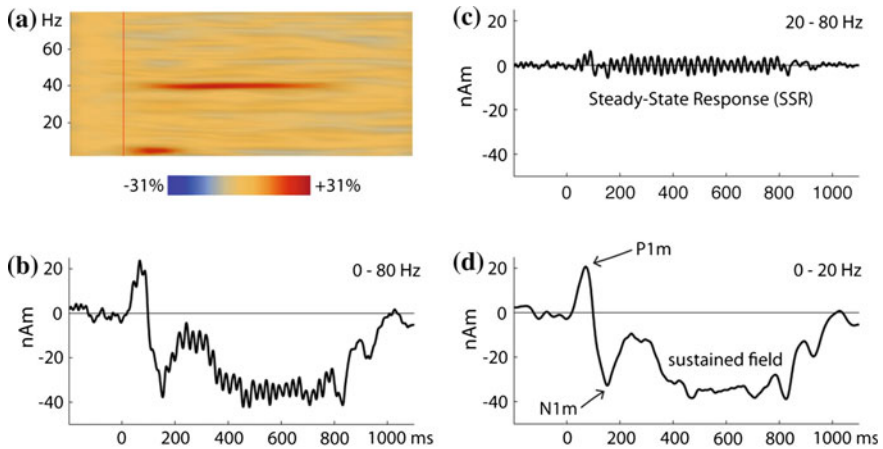
Fig. 4 Auditory 40-Hz steady-state response (SSR) from a single listener. The stimuli were 800-ms long trains of short tone pulses (500 Hz or 1,000 Hz) presented at a rate of 40 Hz. The waveforms are estimated for a dipolar source in the left auditory cortex; highly similar responses were observed on the right (not shown). **a** Time-frequency plots for the time range −200–1,100 ms and the frequency range 1–80 Hz. The plot shows the enhancement of power in comparison to the baseline in the time interval 200 ms before the train onset. The SSR is seen as narrow activity band at 40 Hz, which persists for the whole stimulus duration. The onset response is reflected by a transient increase of power in the theta band. **b** Source waveform for the averaged evoked response filtered from 0–80 Hz. In this setting, the LAEF and the SSR are mixed. Because of the broad frequency separation between these components demonstrated in A, they are separated well with different filter settings. **c** SSR in the frequency range 20–80 Hz, otherwise identical to B. **d** LAEF in the frequency range 0–20 Hz. Note the strong sustained field that is not captured by the time-frequency analysis

subject level, a reliable SSR was generally obtained at 1.5, 3.5, and 31.5 Hz, but only variably at 7.5 and 15.5 Hz stimulation rate. The apparently latency was in the range of 100–150 ms, and there was only a weak dependence on the bandwidth of the stimulus carrier. For an SSR at 4 Hz, it was independently demonstrated that the SSR is stronger for stimuli with a non-sinusoidal amplitude modulation and a more rapid sound onset (Prendergast et al. 2010).

2.2.5 Auditory Non-phase-locked Activity

Separating auditory evoked fields from the background activity by response averaging is based on the assumption that there is little or no jitter between subsequent trials. Stronger jitter may blur the shape of the evoked response in the lower frequency (theta) range. In the higher frequency (gamma) range, jitter may easily exceed the phase duration of a single cycle, such that the variable phase relationship between stimulus and response may results in a cancelation of the response by the averaging procedure. Similar response cancelation by averaging occurs for rhythmic activity that appears in a circumscribed time window but not

tightly locked to the auditory stimulus. Techniques other than response averaging are required to evaluate such non-phase-looked activity. One possibility is to perform time-frequency analysis on a single-trial level and remove phase information before summation across trials. The increase in response power is typically plotted relative to a pre-stimulus baseline (Figs. 3 and 4). This technique is equally sensitive for phase-locked and non-phase locked activity.

Traditionally, gamma activity in the auditory cortex has been evaluated in a narrow frequency band around 40 Hz (Pantev 1995). More recently, activity in the auditory cortex has been observed in a wide frequency range of 70–250 Hz: this high-gamma activity in human auditory cortex has been clearly demonstrated in intracranial recordings on the superior temporal gyrus (Crone et al. 2001; Edwards et al. 2005; Dykstra et al. 2011) as well as in medial Heschl's gyrus (Brugge et al. 2009). It has been suggested that high-gamma activity covaries more closely with spiking activity than with evoked potentials in the lower spectral range (Steinschneider et al. 2008; Ray and Maunsell 2011). Measuring gamma activity in the auditory cortex with MEG is more difficult than in the visual system (Kahlbrock et al. 2012; Millman et al. 2013). However, some recent MEG studies raise hope that high-gamma activity can indeed be evaluated non-invasively based on MEG recordings (Todorovic et al. 2011; Sedley et al. 2012).

2.3 Beyond the Auditory Cortex

While activity in the auditory cortex is modulated by active listening, as discussed in more detail in Sect. 4, all response components reviewed so far are readily recorded in a passive mode, where the subject is not attending to the auditory stimulation and may even be involved into reading a book, watching a silent movie, or another task unrelated to the auditory stimulation. Once a task is added that is directly related to the auditory stimulation, however, additional activity can be elicited, the generators of which are supposedly located in multimodal areas beyond the auditory cortex. The most-frequently-studied response elicited during auditory tasks is the P3 or P300. Sources of the P3 have been studied with depth electrodes in patients suffering from epilepsy (Halgren et al. 1998), and in combined EEG-fMRI studies (Linden 2005), suggesting, amongst others, generators in parietal, prefrontal, and cingulate cortex. So far, only a few MEG studies have explored the generators of the P_3m, suggesting mostly sources in the temporal and frontal lobes (Rogers et al. 1991; Anurova et al. 2005; Halgren et al. 2011). It remains to be determined, whether P3 generators in other sites are also accessible to MEG. Cortical activity related to auditory cognition beyond the auditory cortex is certainly not limited to the P3, but an extensive review of this topic is beyond the scope of this chapter. The near future will likely bring a wealth of new contributions on the functional relationship between the auditory cortex and areas in the frontal and parietal lobe for auditory cognition.

3 Stimulus Specificity of Auditory MEG Activity

This section reviews a selection of basic sound features and how they are reflected in MEG activity originating in the auditory cortex. Only a brief introduction to the background and psychophysics is provided along with each paragraph, and the reader is referred to the available textbooks on psychological acoustics (Moore 2012), phonetics (Stevens 2000), or auditory physiology (Schnupp et al. 2011) for more details and references to the original publications.

3.1 Temporal Resolution and Integration

Temporal coding of sound is differently reflected in the MAEF and LAEF. The early MAEF peaks are very robust to fast stimulus repetition: When two pulses are repeated at ISIs between 1–14 ms (Rupp et al. 2000), a clear response to the second pulse is observed at ISIs $>=$ 4 ms, and the response is nearly completely recovered at ISIs $>=$ 14 ms. The continuous time-locking capability of the MAEF is also demonstrated by the 40-Hz SSR (Gutschalk et al. 1999; Brugge et al. 2009), which shows phase-locking to inter-click intervals of less than 20 ms.

A classical psychoacoustic paradigm to test temporal resolution is gap detection, where a short interruption in an ongoing sound is used as the stimulus. For example, listeners are able to detect interruptions of few milliseconds duration in a continuous broadband noise. When this stimulus is applied in MEG, gaps as short as 3 ms are sufficient to evoke a significant MAEF response (Rupp et al. 2002a), which is in accordance with psychoacoustic thresholds. Moreover, the higher perceptual thresholds observed at the beginning of a noise burst (5 or 20 ms after onset) are paralleled by a lack of MAEF (Rupp et al. 2004).

The subsequent P_1m and N_1m are distinctly different with regard to their suppression at short ISI: when periodic click trains are interrupted by omission of one or more clicks, the onset response after the interruption does not show a significant P_1m when the interruption is 12 and 24 ms. At gap durations of 48 ms, the P_1m is partly recovered, and it has regained almost completely at gaps of 196 ms (Gutschalk et al. 2004b). The time interval required for complete recovery is even longer for the N_1m: Some recovery, especially of the anterior N_1m generator, is observed between 70–150 ms in a two-tone paradigm (Loveless et al. 1996). With ongoing stimulation, the N_1m is reliably observed at ISIs of 300 ms and more (Carver et al. 2002), but some reduction of the response is observed up to 5–10 s (see Sect. 2.2.2). Note that the N_1m can also be evoked by all sorts of transients and transitions in ongoing sound, and not only by sound onset. For example, short gaps of 6 ms in an ongoing noise produce not only a P30m, but also a prominent N_1m (Rupp et al. 2002a). This should not be mistaken as evidence that the N_1m shows similarly fine and fast time-locking as the P30m, but rather reflects the perceptual salience of the transient gap. In contrast to the N19m-P30m, the

N_1m may also reflect auditory events integrated over longer time intervals. Early studies suggested that the N_1m integrates over a time interval of approximately 30–50 ms (Joutsiniemi et al. 1989), because the response amplitude increases with the tone duration for intervals up to this length. More recent studies indicate that temporal integration at the level of the N_1m is not captured by a fixed time window and depends on parameters such as the onset dynamics (Biermann and Heil 2000) and temporal structure of the eliciting stimulus (Krumbholz et al. 2003).

3.2 Stimulus Lateralization

Spatial hearing in the horizontal plane is based on two main cues: one is the difference of sound intensity between the ears caused by the head shadow, the interaural level difference (ILD). The other is the timing difference between the ears, or interaural time difference (ITD). For humans, ITD is predominantly used for lower frequencies, whereas ILD is more important for higher frequencies. The relationship between perceived lateralization and the exact physical parameters is variable, depending on the shape and size of the head and ears. To produce spatial hearing perception, arrays of speakers grouped in some distance around the listener in an anechoic room are the gold standard. In MEG, insert earphones are typically used, in which case one relies on direct manipulation of ITD and ILD. Note, however, that this sound delivery produces somewhat non-ecological percepts of sound sources inside of the head. More exact perceptual lateralization with earphones can be achieved with head-related transfer functions, for which the exact physical parameters are measured with microphones placed at the position of the ears. The simplest method of sound lateralization with earphones is monaural presentation, which is again not an ecological stimulus for normal hearing subjects, but can be viewed as an extreme variant of ILD. Moreover, monaural presentation is easy to implement and has a long tradition in experimental psychology and audiology.

Important processing steps of binaural lateralization cues occur early in the brainstem, and are not readily accessible by MEG. Many MEG studies of sound lateralization have instead focused on its effect on the inter-hemispheric balance between the left and right auditory cortex. It was established early on that the N_1m evoked by monaural sounds is around 15–30 % larger for contralateral compared to isopilateral stimulation, and that the latency of the N_1m is 7–12 ms shorter for contralateral stimulation (Reite et al. 1981; Pantev et al. 1986; Mäkelä et al. 1993). For the P_1m, similar amplitude but smaller latency differences in the range of 1–5 ms were reported (Mäkelä et al. 1994). A stronger modulation of response amplitude in the range of 50 % for contra- in comparison to ipsilateral ear stimulation has been observed for the P30m at the sensor level (planar gradiometers), although the effect was smaller in dipole source waveforms (Mäkelä et al. 1994). However, an EEG source analysis study of the N19-P30 found only an amplitude lateralization in the

range of 6 % and no latency difference (Scherg and Von Cramon, 1986). Currently, little additional data is available to resolve this discrepancy.

ITDs around the maximal physiological range (700 μs) produce lateralization of N_1m-peak amplitudes that can be almost as strong as with monaural presentation (McEvoy et al. 1993; Gutschalk et al. 2012). Moreover, earlier N_1m latencies are observed in the auditory cortex contralateral to the perceptual lateralization (McEvoy et al. 1993). In contrast, no significant effect of ITD is observed for the P30m (McEvoy et al. 1994). Recent MEG studies on the coding of ITD in the auditory cortex support a model with a population rate code for opponent left and right channels, in accordance with earlier work in cat (Stecker et al. 2005), by demonstrating that selective adaptation of the N_1m depended more strongly on whether the adapter and probe were in the same hemifield than on the actual difference in azimuth (Salminen et al. 2009).

So far, the review of contralateral representation in the auditory cortex is simplified, because the balance of activity between the left and right auditory cortex is not symmetric for left- and right-ear stimulation. An amplitude bias towards the right hemisphere has been observed first for the N_1m (Mäkelä et al. 1993), but is even more prominent for the 40-Hz SSR and the sustained field (Ross et al. 2005a): lateralization by ear modulates these responses more strongly in the right AC, and as a result the hemispheric bias is strongly lateralized towards the right auditory cortex for left-ear stimulation and almost counterbalanced for right-ear stimulation (Ross et al. 2005a; Gutschalk et al. 2012). This lateralization bias is not limited to monaural presentation. For example, a combination of ILD and ITD cues, or the use of head-related transfer functions, produces stronger effects on N_1m lateralization than either cue alone (Palomaki et al. 2005), but most prominently in the right auditory cortex. Potentially, this right-hemisphere bias is related to a dominant role of the right hemisphere for spatial processing (Kaiser et al. 2000; Spierer et al. 2009). On the other hand, the bias towards the right may be limited to situations where stimuli are presented in quiet, whereas a lateralization bias towards the left has been observed when sounds are presented under perceptual competition (Okamoto et al. 2007a; Elhilali et al. 2009; Königs and Gutschalk 2012). Finally, the interpretation of hemispheric balance is complicated by anatomical asymmetry in the auditory cortex: stronger cortical folding in the left hemisphere produces stronger signal cancelation in left auditory cortex. The cancelation reduces the MEG signal over the left auditory cortex and biases the MEG response towards larger right-hemisphere responses when in fact equally strong generators can be assumed in both sides (Shaw et al. 2013).

3.3 Sound Frequency

The spectral content of sound is decomposed during sensory transformation in the cochlea, and the resulting tonotopic representation is maintained throughout the ascending auditory pathway, including the auditory cortex. The first demonstration

of a tonotopic map in human auditory cortex made use of MEG, applying dipole source analysis to 32-Hz SSRs evoked by amplitude-modulated pure tones (Romani et al. 1982). This study revealed that the source of the SSR is more medial for higher, and more lateral for lower tone frequencies. The direction of tonotopy, as well as the mapping of dipole locations on structural MRI (Pantev et al. 1996b), is in accordance with a generator of the 40-Hz SSR in the primary auditory cortex field A1. Tonotopy has also been studied for other response components. Studies of the N_1m (Pantev et al. 1988; Pantev et al. 1996b) and the sustained field (Pantev et al. 1994) revealed similar high-low frequency gradients from medial to lateral cortex, as was demonstrated for the SSR. However, it is likely that current source localization techniques are insufficient for modeling synchronous activity in multiple tonotopic fields of the auditory cortex. While the 40-Hz SSR is probably generated in an area focal enough to reflect only one tonotopic gradient, LAEF components are more likely generated in multiple regions of the auditory cortex.

Another reflection of stimulus frequency is by the peak latency of the AEF: because of the propagation delay in the cochlea, AEF latencies are shorter for higher compared to lower stimulus frequencies (Scherg et al. 1989; Roberts and Poeppel 1996). Chirp stimuli (frequency glides from low to high) have been designed to compensate for the propagation delay of the cochlea (Dau et al. 2000). The N19m-P30m evoked by such a chirp is larger than the response evoked by a click or a reversed chirp, because the chirp synchronizes the activity in high and low frequency channels (Dau et al. 2000; Rupp et al. 2002b).

Finally, MEG allows for studying the interaction between stimuli, depending on their frequency separation. One approach that has already been mentioned (Sect. 2.2.3), frequency-selective adaptation, reveals the frequency specificity of cortical processing by reduced adaptation between serial tones when the adapter and probe tones are different in frequency. Another involves tagging simultaneously-presented tones with different amplitude-modulation rates (John et al. 1998). Applying this technique to record the SSR at multiple amplitude-modulation rates around 40-Hz revealed a reduction of amplitude that is more broadly tuned than would have been predicted based on cochlea tuning (Ross et al. 2003). This interaction between simultaneous tones may persist for alternating tones presented at fast repetition rates (20–40 Hz): the alternation of two different tones produces a smaller SSR when the tones are separated by more than a critical band compared to the repetition of identical tone bursts (Gutschalk et al. 2009). Note that the latter finding is opposite to selective adaptation of the P_1m and N_1m, where stronger responses are observed for larger frequency separation between alternating tones. A potential source of the SSR reduction is lateral inhibition. However, a study that explored evidence of lateral inhibition in the auditory cortex found evidence for it only at the level of the N_1m, but not for the SSR (Pantev et al. 2004).

3.4 Pitch and Sound Regularity

Pitch perception is associated with periodic sounds, such as those typically produced by the human voice or musical instruments. In music, pitch is the basic perception required to form melodies. While pure tones evoke a unique pitch percept directly corresponding to their sound frequency, the situation is more complex for everyday periodic sounds in our environment. Briefly, two neural mechanisms supporting pitch perception have been proposed: temporal models that are based on phase-locked neural discharges, primarily in the auditory nerve and spectral-based models relying on distinct loci of maximal displacement of the basilar-membrane. While temporal models assume that pitch is extracted purely in the temporal domain, spectral models estimate pitch based on regular spacing of basilar-membrane maxima from a periodic stimulus' harmonic structure. Many present-day models rely on both spectral and temporal sound features.

One approach to study pitch specificity is to compare regular, periodic sounds with irregular, non-periodic sounds that are otherwise matched in their spectral and temporal envelope. For example, regular click trains are associated with a salient pitch; this pitch can be reduced when the interval between successive clicks is jittered, to the degree that the pitch perception is even completely suppressed (Gutschalk et al. 2002): Regular click trains evoke a much more prominent sustained field than irregular click trains, and source analysis shows that the sustained field evoked by irregular click trains is best explained by dipoles in the planum temporale. Assuming that the components of the sustained field evoked by irregular click trains are also evoked by regular click trains, the pitch-specific component of the sustained field can be separated by calculating the difference between the responses evoked by regular and irregular click trains. This pitch-specific difference response is best explained by dipoles in lateral Heschl's gyrus. In addition to the anatomical separation, these two sources reveal a functional double dissociation: Manipulation of sound intensity predominantly modulates sustained activity in the more posterior source in planum temporale. Conversely, manipulation of click-train regularity predominantly modulates activity in the more anterior source in Heschl's gyrus (Fig. 5).

Another stimulus used to study pitch is so-called iterated rippled noise (Yost et al. 1996); here, a noise is repeatedly copied to itself with a fixed time delay, which equals the inverse of the fundamental frequency (f_0). At the transition from a matched noise to an iterated rippled noise stimulus, a prominent N_1m-like response is evoked, whose peak latency is longer for lower f_0 (Krumbholz et al. 2003); this response has been referred to as the pitch-onset response (POR). The same transient response is evoked at the transition from irregular to regular click trains (Gutschalk et al. 2004a), at the onset of a binaural (Huggins) pitch (Chait et al. 2006), or at the transition between different types of IRN (Ritter et al. 2005). The source of the pitch-onset N_1m is also located in lateral Heschl's gyrus, whereas the sound-onset N_1m observed for irregular click trains or noise maps to the planum temporale. This dissociation is similar to the source configuration of the sustained field, mentioned
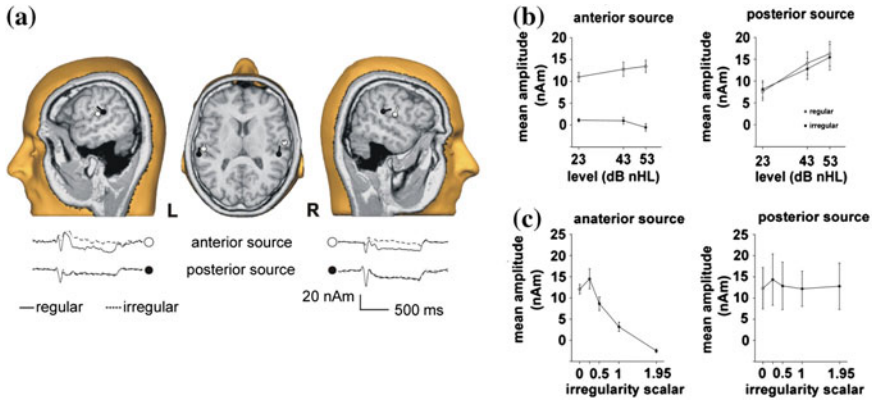
Fig. 5 a Influence of click-train regularity—and supposedly pitch salience—on the sustained field and N_1m (exemplary listener). The 1000-ms long click trains were either regular (inter-click interval 5 ms) or irregular (inter-click interval 2.5–7.5 ms); only the regular click trains produce a salient periodicity pitch. One set of dipoles was fitted to the sustained field evoked by irregular click trains (*black*, in Planum temporale). The other set of dipoles was fitted to the difference between the sustained fields evoked by regular minus irregular click trains (*white*, in lateral Heschl's gyrus), supposedly representing pitch- or regularity-specific activity. As can be seen in the source waveforms, the N_1m and sustained field imaged by the anterior source are only observed for regular click trains, whereas the N_1m and sustained field in the posterior source are identical for regular and irregular click trains. **b** Effect of click train intensity on the sustained-field strength in the anterior (*left*) and posterior (*right*) sources (mean ± standard error, N = 12). Intensity only affects activity in the posterior source significantly. **c** Effect of click train regularity (ISI range = 5 ms ± 5 ms * irregularity scalar) on the sustained-field strength in the anterior (*left*) and posterior (*right*) sources (mean ± standard error, N = 11). Regularity only affects activity in the anterior source significantly. Panels A and B reproduced with permission from Elsevier (Gutschalk et al. 2002); panel C represents unpublished data obtained in the same listeners

earlier. Moreover, spatio-temporal dipole modeling allows for separating the pitch-onset and sound-onset components of the N_1m in situations where the periodic sound is presented out of silence (Gutschalk et al. 2004a). Both, the pitch-onset N_1m as well as the sustained pitch response reflect the stimulus history. The amplitude of the pitch-onset N_1m increases with the directly preceding ISI; the sustained field varies depending on the ratio of regular and irregular stimuli occurring in a stimulus sequence on a time scale of seconds to minutes (Gutschalk et al. 2007b).

Specificity for pitch in lateral Heschl's gyrus had also been suggested based on fMRI (Patterson et al. 2002), but this has recently been questioned because it was shown that the fMRI signal evoked by iterated rippled noise is dominated by the presence of temporal fluctuations that are unrelated to pitch (Barker et al. 2012). Note, that these fluctuations evoke ongoing activity in the theta-band in MEG, whereas the N_1m and sustained field components evoked by periodicity are similar for click trains and iterated rippled noise (Steinmann and Gutschalk 2012).

As a final note, it should be mentioned that the interpretation of these regularity-specific responses in terms of pitch perception might be too exclusive.

A number of studies suggest that these responses could also be related to a more general processing of stimulus regularity: a prominent N_1m is, for example, evoked at the transition from random tones (duration = 15, 30 or 60 ms) to a constant tone, whereas a much weaker response was observed when the transition was from constant to random (Chait et al. 2007). With respect to the sustained field, it was demonstrated that the periodic repetition of frozen noise evokes stronger sustained fields than random white noise down to repetition rates of 5 Hz (Keceli et al. 2012), which is well below the lower limit where musical pitch is typically observed (Pressnitzer et al. 2001).

3.5 Vowels and Other Speech Sounds

Vowels are one of the basic elements of speech, and their classification for speech is determined by formants, which are basically peaks in certain parts of the spectrum. The spectral shape of the human voice in general, and thus also of vowels, is formed by the upper vocal tract. MEG studies demonstrated that the N_1m evoked by vowel onset cannot be explained by a linear superposition of their frequency content (Diesch and Luce 2000). It has been suggested instead that the source localization and latency of the N_1m represent abstract phonological features such as place of articulation (Obleser et al. 2004).

As mentioned in Sect. 3.4, the human voice is a prototype of a periodic sound source, due to the periodic pulsations of the vocal folds during voiced speech. Speech periodicity may be disturbed, for example in whispering, or in hoarse, pathological speech, and in this case the sustained field is reduced (Yrttiaho et al. 2009). However, the sustained field does not only reflect the vowels' periodicity, but is also enhanced by spectral formant features that determine the phonological vowel quality: This was first shown with the comparison of pure tones and sine vowels (Eulitz et al. 1995). Using damped sine pulses, the periodicity pitch and the vowels formant structure can be separately violated, producing sounds that have periodicity pitch and/or vowel quality or neither. This way, the sustained field components evoked by pitch, formant structure, and the control sound can be separately evaluated. The source-analysis results showed that the sustained field evoked by the periodicity pitch and the one evoked by the formant structure are co-located in lateral Heschl's gyrus, whereas the residual sustained field was located more posterior (Gutschalk and Uppenkamp 2011). This result raises the possibility that lateral Heschl's gyrus plays a general role in speech sound extraction, or is alternatively related to a more general mechanism of regularity extraction (see Sect. 3.4). This question is of considerable interest, because fMRI studies typically do not find enhanced activity in auditory cortex for speech in contrast to non-speech sounds; for example, the same vowel and non-vowels stimuli evaluated in fMRI evoke enhanced activity only in the superior temporal sulcus (Uppenkamp et al. 2006). This discrepancy between MEG and fMRI can probably be explained by the

finding that sustained fields in MEG have only a weak (Gutschalk et al. 2010) or no (Steinmann and Gutschalk 2012) correlate at all in BOLD fMRI.

Vowels are only one category of speech-specific (phonetic) elements. Topographical differences between N_1m responses have also been found for different consonants, which depended not only on the physical sound's structure but also on its intelligibility (Obleser et al. 2006). In summary, findings accumulated with MEG and other techniques indicate that the transformation of sound into basic speech-specific (phonological) categories starts in the auditory cortex on the superior temporal plane, and it remains to be determined how much of this process is already completed there.

4 Auditory Scene Analysis

Most of the studies reviewed so far explored the processing of sounds emanating sequentially from a single source. This is not the most frequent constellation in ecological environments, where multiple sounds sources are often active interleaved or at once. The title of the seminal monograph "auditory scene analysis" (Bregman 1990) provides the heading for research that explores how the brain separates multiple sound sources. The subsequent sounds emanating from one source, for example the speech from one person, or the melody played on a musical instrument, are herein referred to as auditory streams. Auditory streams are of similar importance for auditory cognitive neurosciences as the conception of objects for the visual neurosciences.

4.1 Auditory Stream Segregation

One of the basic and most commonly used paradigms to study auditory scene analysis is the stream-segregation or streaming paradigm. In the simplest version of this paradigm, two pure tones A and B are alternated ($ABAB...$) at a rate of around 5–10 Hz with the frequency separation Δf. When Δf is small (up to a few semitones), the sequence is heard as a stream of alternating tones, a trill (Miller and Heise 1950). The streaming phenomenon is observed at larger Δf: here, A and B tones are perceived as two separate streams, each with its own beat and rhythm. This can be well demonstrated with the $ABA\_$ triplet paradigm (Van Noorden 1975), where the underscore stands for a pause whose duration is equal to the tones. When the triplets are heard as one stream, they are associated with a characteristic galloping rhythm. In contrast, two isochronous streams are perceived in the case of streaming. When $ABA\_$ tone triplets are presented in MEG, the response strength of B tones depend on the Δf (Gutschalk et al. 2005): the P_1m is strongly suppressed by the preceding A tone when the tones are close in frequency. For $\Delta f = 4$–6 semitones, there is less adaptation (or suppression) caused by the A tones, and the P_1 m evoked at $\Delta f = 12$

semitones is almost the size of the P_1m evoked by B tones in the absence of any A tones (Fig. 6). This effect is similar to the selective-adaptation phenomenon discussed in Sect. 2.2.3 for the N_1m. In fact, selective adaptation of the N_1m was also observed, but for the fast repetition rates typically used for streaming, the N_1m remains relatively small overall. Importantly, selective adaptation of the response in auditory cortex was correlated with the listeners rating of how easy it was for them to hold to the two-stream perception, suggesting that the selective adaption observed in MEG is linked to neurophysiological processes important for streaming perception. Similar results were obtained by other investigators (Snyder et al. 2006; Chakalov et al. 2012).

Selective adaptation of the P_1m in streaming contexts is not limited to situations where Δf is the segregation cue. Selective release of P_1m adaptation has also been observed when streaming was based on periodicity pitch, using stimuli that were prepared such that they did not provide spectral cues that can be resolved by frequency analysis in the cochlea (Gutschalk et al. 2007a). Finally, selective release of P_1m adaptation was observed with streaming based on lateralization by ITD and was stronger for conditions where streaming was more frequently observed (Carl and Gutschalk 2013). In both cases, for streaming based on pitch and based on ITD, the sources of selective adaption are located in the same area around Heschl's gyrus including core as well as belt areas of the auditory cortex (Schadwinkel and Gutschalk 2010). It therefore appears that the separation of sound sources based on different segregation cues converges at the level of the auditory cortex, potentially providing a general mechanism for sound source separation.

A more direct way to study the relationship between neurophysiology and perception is based on perceptual bistability. The relationship between for example Δf and streaming perception is not deterministic; the same sequence can alternatively be perceived as one or two streams, especially in the intermediate Δf range (Van Noorden 1975), and the perception may flip back and forth between the two perceptual organizations. When listeners indicate the reversal towards one stream with one key, and the reversal towards two streams with another key, the MEG activity evoked by an ongoing sequence with fixed Δf can be averaged with respect to the perception. The results show that the response evoked by the B tones is stronger in intervals where listeners heard two streams compared to intervals where they heard one stream (Gutschalk et al. 2005). This result is similar to the growth of the P_1m evoked by B tones with larger Δf, albeit the effect size in the bistability experiment was smaller than in the Δf experiment.

4.2 Auditory Selective Attention

Two separate streams of tones are also presented in another classical paradigm, but with a different focus: the ISI between subsequent tones is randomized and one stream is presented to the left and another one to the right ear. Within each stream
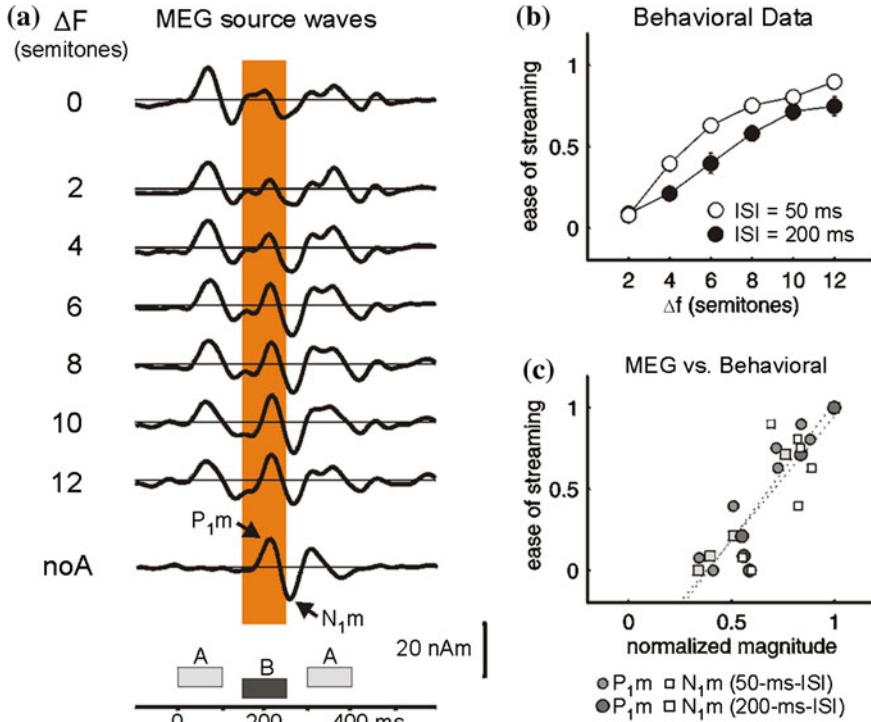
Fig. 6 Relationship between streaming perception and frequency selective adaptation of the P_1m and N_1m (modified from Gutschalk et al. 2005). **a** Auditory cortex source waveforms of the response evoked by sequences of repetitive *ABA_* triplets (average across listeners; n = 14). The frequency (Δf) difference between *A* and *B* is indicated on the *left*. The P_1m and N_1m evoked by *B* tones are strongly suppressed for $\Delta f = 0$ and 2 semitones, which were not perceived as two streams. There is a marked release of this adaptation for $\Delta f = 4$ semitones and beyond, which can be perceived as one or two streams. At $\Delta f = 10$ semitones, the amplitude of the response is almost the same size as the response evoked by *B* tones without any interfering *A* tones. **b** Ease of streaming for the sequences used in panel a and similar sequences with a longer ISI (n = 13). Listeners tried to hear two streams and indicated after the end of the sequence how easy it was to hear two streams on a continuous scale between 0 and 1 (0 = impossible, 1 = very easy). **c** Scatter plot of the average, normalized MEG amplitudes (P_1m and N_1m) versus the average ease of streaming. The correlation was r = 0.91 (p < 0.0001) for the P_1m and r = 0.83 (p < 0.001) for the N_1m

there are standards and deviants, like in the oddball paradigm introduced in Sect. 2.2.3, and the listeners task is to monitor the occurrence of deviants in only one of the two streams. This paradigm has not been used to study if one or two streams are perceived—the latter was rather implicitly assumed by the setup—but to evaluate how selectively listening to one of the streams modulates the auditory evoked activity. An early EEG study demonstrated that the N_1 is prominently

larger for the tones (standards as well as deviants) in the ear that the listener attended to (Hillyard et al. 1973). Later on, it was demonstrated in MEG that the attentional enhancement of vertex-negative responses originates in the auditory cortex (Rif et al. 1991; Woldorff et al. 1993). One of these studies (Rif et al. 1991) used a setup where the two streams were not separated by ear, but only by their frequency (1,000 vs. 3,000 Hz). The enhancement of surface-negative activity in the auditory cortex was observed in the time interval of the N_1m, or alternatively in the latency range of the P_2m when a longer ISI was used (Rif et al. 1991). There has been some discussion of whether the enhanced negative response evoked by attended streams reflects enhancement of the N_1m or a separate response component called the processing negativity (Näätänen 1982) or the late negative difference wave (Hansen and Hillyard 1980). In any case, there is no doubt that auditory cortex activity in the N_1m latency range can be enhanced by selectively listening to one stream in certain stimulus configurations.

It is less well settled whether attention also modulates response components that are associated with earlier processing stages, such as the P_1m and the 40-Hz SSR. In the P_1m interval, one study found that the response in this interval was more negative with attention, supposedly reflecting the early onset of N_1m enhancement (Rif et al. 1991). Two other studies found an enhanced positive response in the time interval 20–50 ms (Woldorff et al. 1993; Poghosyan and Ioannides 2008), potentially reflecting enhancement of processes related to the P_1m. A few reports also suggest that the 40 Hz SSR is modulated by intra-modal auditory versus visual attention (Ross et al. 2004; Saupe et al. 2009). However, the effect size of attentional amplitude enhancement for the 40-Hz SSR is generally small, and it has been pointed out that the effect is much stronger for the N_1m and the sustained field (Okamoto et al. 2011). One intracranial study suggests that the 20-Hz SSR is modulated when one of two concurrent amplitude-modulated tones is selectively attended (Bidet-Caulet et al. 2007). A recent dichotic MEG study found that the 40-Hz SSR in right auditory cortex was reduced for attended targets in the ipsilateral, right ear, and non-significantly enhanced for attended targets in the contralateral, left ear (Weisz et al. 2012). In summary, these studies suggest that the 40 Hz SSR in primary auditory cortex can be modulated by attention in certain contexts, but that the effect size of the attentional modulation is small in comparison to the response amplitude, as well as compared to the modulation observed at later processing stages.

Response enhancement by selective attention is not limited to simple tone stimuli, but can also be observed for more complex sounds, for example when two competing speakers are played to the left and right ear, and the listeners are instructed to report the information from one ear only. This classical dichotic paradigm (Cherry 1953), typically cited in the context of the cocktail party phenomenon, was recently adapted for MEG with an elegant analysis method: instead of averaging from tone onset, Ding and Simon extracted the envelope of each speaker and deconvolved the time course of activity in the auditory cortex using crosscorrelation between the signal envelope and the MEG time series (Ding and Simon, 2012b). The results revealed a response similar to the classical evoked

response with peaks P1 m and N_1m. Moreover, when the listeners selectively listened to one of the speakers, the associated N_1m like response was prominently enhanced. This effect is not limited to the dichotic paradigm, but was also observed when two speakers, for example a male and a female, were presented to both ears without spatial separation, and the listeners were instructed to selectively listen to one of the speakers (Ding and Simon 2012a).

One model for the selective response enhancement observed for attended streams is a simple gain model, which assumes that the response to the attended signal is enhanced. However, the response modulation in the auditory cortex by attention may be more selective. For example, it has been shown that selectively attending to a spatial cue modulates activity in more posterior areas of the auditory cortex, whereas attention to phonetic content predominantly modulates activity in more anterior areas of the auditory cortex (Ahveninen et al. 2006). It has also been suggested that attention towards a tone sharpens the spectral tuning in auditory cortex: When a pure tone is presented in a notch-filtered noise, the attentional enhancement is larger for narrow than for broader notches (Okamoto et al. 2007b), and no response enhancement is observed for tones presented without a concurrent masker (Ahveninen et al. 2011). The authors of the study suggested that this is because the notched noise adapts the broadly tuned activity evoked by pure tones in the absence of attention, but not the sharpened, more focal activation when the tone is attended to.

Directing attention involves a number of areas outside the auditory cortex, such as the frontal eye fields and the temporo-parietal junction (Larson and Lee 2012), as well as more dorsal parietal areas (Sieroka et al. 2003). The exact role of each of these areas is still being explored, and is not reviewed here in detail.

4.3 Auditory Perceptual Awareness

The streaming and attention paradigms reviewed above are typically designed such that the presence of each stream is easily noted, even though smaller details or changes of the target stream may sometimes be missed because of interference from the competing streams. Thus, listeners are typically able to deploy their attention towards a specific stream without major efforts. The situation may be different when more complex soundscapes are used, where multiple streams compete for the listeners processing capacity, such that the listener is not aware of each stream's presence at a time. This phenomenon is known as informational masking (Durlach et al. 2003). In contrast to energetic masking, where two sounds that overlap in their spectrum compete for sensory transformation in the cochlea, informational masking is thought to originate in the central nervous system. To avoid additional energetic masking, a spectral separation between target and masker (the protected region) is typically used. Accordingly, once a stream has been detected in the presence of an informational masker, the perception of the stream is salient, because the target tones are clearly above the sensory threshold.
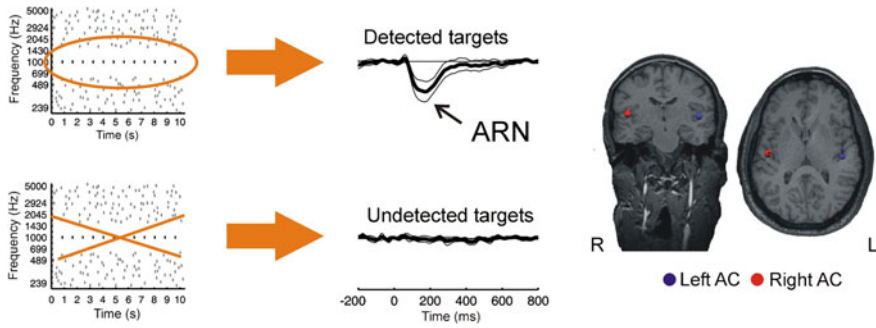
Fig. 7 Relationship between auditory perceptual awareness versus informational masking and MEG activity in the auditory cortex. Streams of target tones are presented for 10.4 s with a stimulus-onset asynchrony of 800 ms and in the context of a random multi-tone masker. The target-tone frequency was randomly chosen for each 10 s sequence (range 489–2924 Hz). Listeners indicated with a mouse button when they detected a regular target stream. Considering that at least two tones were heard before each button press, about half of the target tones were heard and the other half was masked. When the response time locked to target tones was averaged, no significant evoked response was observed for undetected targets (*lower trace*). In contrast, detected targets evoked a prominent negative response in auditory cortex in the time interval 50–250 ms, the awareness related negativity (ARN). Example stimuli are available along with the original, open access online publication (Gutschalk et al. 2008)

An informational masking stimulus that has been adapted for MEG research is illustrated in Fig. 7 (Gutschalk et al. 2008): the target is a regular tone stream, with fixed frequency and ISI. The masker comprises multiple tones, which are arranged in several frequency bands and whose ISI is independently randomized. This type of masker is called a multi tone masker; the randomization of the masker onsets was introduced for application in MEG, to cancel out responses that are phase locked to masker tones, and be able to evaluate selectively the neural response evoked by targets. Because the target frequency varied in subsequent trials, listeners cannot simply monitor a fixed frequency region, but need to listen (search) for the regular target stream. Listeners were instructed to press a mouse button whenever they heard out the regular target stream, and these behavioral responses were used to dissociate epochs where the listeners were aware of the target stream, and those where they were not aware of the target's presence. MEG revealed a prominent negative response in the auditory cortex in the latency range 50–250 ms, with apeak latency around 120–200 ms after tone onset. No late negativity was evoked by target tones in epochs where listeners were not aware of their presence.

In contrast, the 40 Hz SSR (Gutschalk et al. 2008) and the P_1m (Königs and Gutschalk 2012) were evoked by detected and undetected target tones alike. Moreover, the results from an fMRI and MEG study show that stronger activity for detected compared to undetected targets is observed in medial Heschl's gyrus, and thus most probably in the primary auditory cortex (Wiegand and Gutschalk 2012). These results suggest that there is a coexistence of two types of neural activity in the (primary)

auditory cortex: one type (40 Hz SSR) is more closely related to the physical stimulus and the other type (ARN) reflects the perception rather than the sound input.

The source location of the ARN was not statistically different from the N_1m evoked passively when the targets were presented in silence in one study (Gutschalk et al. 2008) and only about 5 mm apart in another study (Königs and Gutschalk 2012). Moreover, the hemispheric balance of both, the ARN and the N_1m, is modulated to similar amounts by sound lateralization (Königs and Gutschalk 2012). It is therefore possible, that the generators of the ARN and N_1m are—at least in part—identical. As has been noted in the previous sections, the N_1m is an automatic response and shows little or no modulation by attention in situations where tones are presented without competing auditory stimuli e.g. (Ahveninen et al. 2011). In contrast, the ARN is not evoked at all when attention is distracted to a different task, e.g. in a dichotic paradigm (Gutschalk et al. 2008). Another study that applied informational masking in MEG found that the SSR evoked by a 4 Hz target stream was enhanced when listeners detected frequency deviants within that stream, but not when they detected a temporal elongation of tones within the multi-tone masker (Elhilali et al. 2009).

While a clear attentional modulation of the N_1m is already observed, for example, when one of two interleaved streams is selectively attended (Rif et al. 1991) or when an attended tone is presented within a simultaneous noise masker (Okamoto et al. 2007b), the N_1m is still evoked automatically by the unattended stream in these cases, and the listener is typically aware of the unattended stream's presence. One explanation for these different observations could be that processes reflected by the N_1m/ARN are only modulated by attention under sensory competition (Desimone and Duncan 1995; Lavie 2006), and that at high levels of sensory competition the reduction of these neural processes is so prominent that they are insufficient for perceptual awareness. The latter case would then produce informational masking. At this point, we don't know if informational masking can already be overcome by bottom-up activity in the auditory cortex, or if the deployment of attentional resources directed by the frontal lobe is additionally required. The relative role of modality specific sensory cortex on the one hand, and activity in prefrontal areas for perceptual awareness, on the other hand, is still diversely discussed across sensory modalities (Dehaene & Changeux 2011; Meyer 2011), and remains an important topic for future research.

References

Ahveninen J, Hämäläinen M, Jääskeläinen IP, Ahlfors SP, Huang S, Lin FH, Raij T, Sams M, Vasios CE, Belliveau JW (2011) Attention—driven auditory cortex short—term plasticity helps segregate relevant sounds from noise. Proc Nat Acad Sci USA 108:4182–4187

Ahveninen J, Jääskeläinen IP, Raij T, Bonmassar G, Devore S, Hämäläinen M, Levanen S, Lin FH, Sams M, Shinn-Cunningham BG, Witzel T, Belliveau JW (2006) Task-modulated "what" and "where" pathways in human auditory cortex. Proc Nat Acad Sci USA 103:14608–14613

Anurova I, Artchakov D, Korvenoja A, Ilmoniemi RJ, Aronen HJ, Carlson S (2005) Cortical generators of slow evoked responses elicited by spatial and nonspatial auditory working memory tasks. Clin Neurophysiol 116:1644–1654

Barker D, Plack CJ, Hall DA (2012) Reexamining the evidence for a pitch-sensitive region: a human fMRI study using iterated ripple noise. Cereb Cortex 22:745–753

Bidet-Caulet A, Fischer C, Besle J, Aguera PE, Giard MH, Bertrand O (2007) Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. J Neurosci 27:9252–9261

Biermann S, Heil P (2000) Parallels between timing of onset responses of single neurons in cat and of evoked magnetic fields in human auditory cortex. J Neurophysiol 84:2426–2439

Braak H (1978) The pigment architecture of the human temporal lobe. Anat embryol (Berlin) 154:213–240

Bregman AS (1990) Auditory scene analysis. MIT Press, Cambridge

Brookes MJ, Stevenson CM, Barnes GR, Hillebrand A, Simpson MI, Francis ST, Morris PG (2007) Beamformer reconstruction of correlated sources using a modified source model. Neuroimage 34:1454–1465

Brugge JF, Nourski KV, Oya H, Reale RA, Kawasaki H, Steinschneider M, Howard MA 3rd (2009) Coding of repetitive transients by auditory cortex on Heschl's gyrus. J Neurophysiol 102:2358–2374

Butler RA (1968) Effect of changes in stimulus frequency and intensity on habituation of the human vertex potential. J Acoust Soc Am 44:945–950

Carl D, Gutschalk A (2013) Role of pattern, regularity, and silent intervals in auditory stream segregation based on inter-aural time differences. Exp Brain Res 224:557–570

Carver FW, Fuchs A, Jantzen KJ, Kelso JA (2002) Spatiotemporal analysis of the neuromagnetic response to rhythmic auditory stimulation: rate dependence and transient to steady-state transition. Clin Neurophysiol 113:1921–1931

Chait M, Poeppel D, de Cheveigne A, Simon JZ (2007) Processing asymmetry of transitions between order and disorder in human auditory cortex. J Neurosci 27:5207–5214

Chait M, Poeppel D, Simon JZ (2006) Neural response correlates of detection of monaurally and binaurally created pitches in humans. Cereb Cortex 16:835–848

Chakalov I, Draganova R, Wollbrink A, Preissl H, Pantev C (2012) Modulations of neural activity in auditory streaming caused by spectral and temporal alternation in subsequent stimuli: a magneto encephalographic study. BMC Neuroscience 13:72

Cherry C (1953) Some experiments on the recognition of speech, with one and two ears. J Acoust Soc Am 25:975–981

Crone NE, Boatman D, Gordon B, Hao L (2001) Induced electrocorticographic gamma activity during auditory perception. Clin Neurophysiol 112:565–582

Dau T, Wegner O, Mellert V, Kollmeier B (2000) Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion. J Acoust Soc Am 107:1530–1540

Dehaene S, Changeux JP (2011) Experimental and theoretical approaches to conscious processing. Neuron 70:200–227

Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. Annu Rev Neurosci 18:193–222

Diesch E, Luce T (2000) Topographic and temporal indices of vowel spectral envelope extraction in the human auditory cortex. J Cogn Neurosci 12:878–893

Ding N, Simon JZ (2012a) Emergence of neural encoding of auditory objects while listening to competing speakers. Proc Nat Acad Sci USA 109:11854–11859

Ding N, Simon JZ (2012b) Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. J Neurophysiol 107:78–89

Durlach NI, Mason CR, Kidd G Jr, Arbogast TL, Colburn HS, Shinn-Cunningham BG (2003) Note on informational masking. J Acoust Soc Am 113:2984–2987

Dykstra AR, Halgren E, Thesen T, Carlson CE, Doyle W, Madsen JR, Eskandar EN, Cash SS (2011) Widespread brain areas engaged during a classical auditory streaming task revealed by intracranial EEG. Front Hum Neurosci 5:74

Edwards E, Soltani M, Deouell LY, Berger MS, Knight RT (2005) High gamma activity in response to deviant auditory stimuli recorded directly from human cortex. J Neurophysiol 94:4269–4280

Elhilali M, Xiang J, Shamma SA, Simon JZ (2009) Interaction between attention and bottom-up saliency mediates the representation of foreground and background in an auditory scene. PLoS Biol 7:e1000129

Eulitz C, Diesch E, Pantev C, Hampson S, Elbert T (1995) Magnetic and electric brain activity evoked by the processing of tone and vowel stimuli. J Neurosci 15:2748–2755

Fishman YI, Steinschneider M (2012) Searching for the mismatch negativity in primary auditory cortex of the awake monkey: deviance detection or stimulus specific adaptation? J Neurosci 32:15747–15758

Formisano E, Kim DS, Di Salle F, van de Moortele PF, Ugurbil K, Goebel R (2003) Mirror-symmetric tonotopic maps in human primary auditory cortex. Neuron 40:859–869

Galaburda A, Sanides F (1980) Cytoarchitectonic organization of the human auditory cortex. J Compar Neurol 190:597–610

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

Garrido MI, Kilner JM, Stephan KE, Friston KJ (2009) The mismatch negativity: a review of underlying mechanisms. Clin Neurophysiol 120:453–463

Gutschalk A, Brandt T, Bartsch A, Jansen C (2012) Comparison of auditory deficits associated with neglect and auditory cortex lesions. Neuropsychologia 50:926–938

Gutschalk A, Hämäläinen MS, Melcher JR (2010) BOLD responses in human auditory cortex are more closely related to transient MEG responses than to sustained ones. J Neurophysiol 103:2015–2026

Gutschalk A, Mase R, Roth R, Ille N, Rupp A, Hähnel S, Picton TW, Scherg M (1999) Deconvolution of 40 Hz steady-state fields reveals two overlapping source activities of the human auditory cortex. Clin Neurophysiol 110:856–868

Gutschalk A, Micheyl C, Melcher JR, Rupp A, Scherg M, Oxenham AJ (2005) Neuromagnetic correlates of streaming in human auditory cortex. J Neurosci 25:5382–5388

Gutschalk A, Micheyl C, Oxenham AJ (2008) Neural correlates of auditory perceptual awareness under informational masking. PLoS Biol 6:e138

Gutschalk A, Oldermann K, Rupp A (2009) Rate perception and the auditory 40-Hz steady-state fields evoked by two-tone sequences. Hear Res 257:83–92

Gutschalk A, Oxenham AJ, Micheyl C, Wilson EC, Melcher JR (2007a) Human cortical activity during streaming without spectral cues suggests a general neural substrate for auditory stream segregation. J Neurosci 27:13074–13081

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

Gutschalk A, Patterson RD, Scherg M, Uppenkamp S, Rupp A (2004a) Temporal dynamics of pitch in human auditory cortex. Neuroimage 22:755–766

Gutschalk A, Patterson RD, Scherg M, Uppenkamp S, Rupp A (2007b) The effect of temporal context on the sustained pitch response in human auditory cortex. Cereb Cortex 17:552–561

Gutschalk A, Patterson RD, Uppenkamp S, Scherg M, Rupp A (2004b) Recovery and refractoriness of auditory evoked fields after gaps in click trains. Eur J Neurosci 20:3141–3147

Gutschalk A, Scherg M, Picton TW, Mase R, Roth R, Ille N, Klenk A, Hähnel S (1998) Multiple source components of middle and late latency auditory evoked fields. In: Kakigi R, Hashimoto I (eds) Recent advances in human neurophysiology. Elsevier, Amsterdam, pp 270–278

Gutschalk A, Uppenkamp S (2011) Sustained responses for pitch and vowels map to similar sites in human auditory cortex. Neuroimage 56:1578–1587

Hackett TA, Preuss TM, Kaas JH (2001) Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. J Compar Neurol 441:197–222

Halgren E, Marinkovic K, Chauvel P (1998) Generators of the late cognitive potentials in auditory and visual oddball tasks. Electroencephalogr Clin Neurophysiol 106:156–164

Halgren E, Sherfey J, Irimia A, Dale AM, Marinkovic K (2011) Sequential temporo-fronto-temporal activation during monitoring of the auditory environment for temporal patterns. Hum Brain Mapp 32:1260–1276

Hansen JC, Hillyard SA (1980) Endogenous brain potentials associated with selective auditory attention. Electroencephalogr Clin Neurophysiol 49:277–290

Hari R, Aittoniemi K, Jarvinen ML, Katila T, Varpula T (1980) Auditory evoked transient and sustained magnetic fields of the human brain. Localization of neural generators. Exp Brain Res 40:237–240

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

Hari R, Kaila K, Katila T, Tuomisto T, Varpula T (1982) Interstimulus interval dependence of the auditory vertex response and its magnetic counterpart: implications for their neural generation. Electroencephalogr Clin Neurophysiol 54:561–569

Hari R, Pelizzone M, Mäkelä JP, Hallstrom J, Leinonen L, Lounasmaa OV (1987) Neuromagnetic responses of the human auditory cortex to on- and offsets of noise bursts. Audiology 26:31–43

Hashimoto I, Mashiko T, Yoshikawa K, Mizuta T, Imada T, Hayashi M (1995) Neuromagnetic measurements of the human primary auditory response. Electroencephalogr Clin Neurophysiol 96:348–356

Hillyard SA, Hink RF, Schwent VL, Picton TW (1973) Electrical signs of selective attention in the human brain. Science 182:177–180

Imada T, Hari R, Loveless N, McEvoy L, Sams M (1993) Determinants of the auditory mismatch response. Electroencephalogr Clin Neurophysiol 87:144–153

Imada T, Watanabe M, Mashiko T, Kawakatsu M, Kotani M (1997) The silent period between sounds has a stronger effect than the interstimulus interval on auditory evoked magnetic fields. Electroencephalogr Clin Neurophysiol 102:37–45

Jääskeläinen IP, Ahveninen J, Bonmassar G, Dale AM, Ilmoniemi RJ, Levanen S, Lin FH, May P, Melcher J, Stufflebeam S, Tiitinen H, Belliveau JW (2004) Human posterior auditory cortex gates novel sounds to consciousness. Proc Nat Acad Sci USA 101:6809–6814

John MS, Lins OG, Boucher BL, Picton TW (1998) Multiple auditory steady-state responses (MASTER): stimulus and recording parameters. Audiology 37:59–82

Joutsiniemi SL, Hari R, Vilkman V (1989) Cerebral magnetic responses to noise bursts and pauses of different durations. Audiology 28:325–333

Kahlbrock N, Butz M, May ES, Schnitzler A (2012) Sustained gamma band synchronization in early visual areas reflects the level of selective attention. Neuroimage 59:673–681

Kaiser J, Lutzenberger W, Preissl H, Ackermann H, Birbaumer N (2000) Right-hemisphere dominance for the processing of sound-source lateralization. J Neurosci 20:6631–6639

Keceli S, Inui K, Okamoto H, Otsuru N, Kakigi R (2012) Auditory sustained field responses to periodic noise. BMC Neuroscience 13:7

Königs L, Gutschalk A (2012) Functional lateralization in auditory cortex under informational masking and in silence. Eur J Neurosci 36:3283–3290

Kretzschmar B, Gutschalk A (2010) A sustained deviance response evoked by the auditory oddball paradigm. Clin Neurophysiol 121:524–532

Krumbholz K, Patterson RD, Seither-Preisler A, Lammertmann C, Lütkenhoner B (2003) Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. Cereb Cortex 13:765–772

Larson E, Lee AK (2012) The cortical dynamics underlying effective switching of auditory spatial attention. Neuroimage 64:365–370

Lavie N (2006) The role of perceptual load in visual awareness. Brain Res 1080:91–100

Liegeois-Chauvel C, Musolino A, Badier JM, Marquis P, Chauvel P (1994) Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. Electroencephalogr Clin Neurophysiol 92:204–214

Liegeois-Chauvel C, Musolino A, Chauvel P (1991) Localization of the primary auditory area in man. Brain 114(Pt 1A):139–151

Linden DE (2005) The p300: where in the brain is it produced and what does it tell us? Neuroscientist 11:563–576

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

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

Lütkenhöner B, Lammertmann C, Ross B, Pantev C (2000) Brain stem auditory evoked fields in response to clicks. NeuroReport 11:913–918

Lütkenhöner B, Steinstrater O (1998) High-precision neuromagnetic study of the functional organization of the human auditory cortex. Audiol Neurootology 3:191–213

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

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

Mäkelä JP, Hari R (1987) Evidence for cortical origin of the 40 Hz auditory evoked response in man. Electroencephalogr Clin Neurophysiol 66:539–546

Mäkelä JP, Hari R, Leinonen L (1988) Magnetic responses of the human auditory cortex to noise/square wave transitions. Electroencephalogr Clin Neurophysiol 69:423–430

May P, Tiitinen H, Ilmoniemi RJ, Nyman G, Taylor JG, Näätänen R (1999) Frequency change detection in human auditory cortex. J Comput Neurosci 6:99–120

May PJ, Tiitinen H (2010) Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. Psychophysiology 47:66–122

McEvoy L, Hari R, Imada T, Sams M (1993) Human auditory cortical mechanisms of sound lateralization: II. Interaural time differences at sound onset. Hear Res 67:98–109

McEvoy L, Levanen S, Loveless N (1997) Temporal characteristics of auditory sensory memory: neuromagnetic evidence. Psychophysiology 34:308–316

McEvoy L, Mäkelä JP, Hämäläinen M, Hari R (1994) Effect of interaural time differences on middle-latency and late auditory evoked magnetic fields. Hear Res 78:249–257

Meyer K (2011) Primary sensory cortices, top-down projections and conscious experience. Prog Neurobiol 94:408–417

Miller GA, Heise GA (1950) The trill threshold. J Acoust Soc Am 22:637–638

Millman RE, Prendergast G, Hymers M, Green GG (2013) Representations of the temporal envelope of sounds in human auditory cortex: Can the results from invasive intracortical "depth" electrode recordings be replicated using non-invasive MEG "virtual electrodes"? Neuroimage 64:185–196

Moore BCJ (2012) An introduction to the psychology of hearing. Emerald, Bingley

Morosan P, Rademacher J, Schleicher A, Amunts K, Schormann T, Zilles K (2001) Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. Neuroimage 13:684–701

Morosan P, Schleicher A, Amunts K, Zilles K (2005) Multimodal architectonic mapping of human superior temporal gyrus. Anat embryol (Berlin) 210:401–406

Näätänen R (1982) Processing negativity: an evoked-potential reflection of selective attention. Psychol Bull 92:605–640

Näätänen R, Gaillard AW, Mantysalo S (1978) Early selective-attention effect on evoked potential reinterpreted. Acta Psychologica (Amsterdam) 42:313–329

Näätänen R, Kujala T, Winkler I (2011) Auditory processing that leads to conscious perception: a unique window to central auditory processing opened by the mismatch negativity and related responses. Psychophysiology 48:4–22

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

Näätänen R, Sams M, Alho K, Paavilainen P, Reinikainen K, Sokolov EN (1988) Frequency and location specificity of the human vertex N1 wave. Electroencephalogr Clin Neurophysiol 69:523–531

Nourski KV, Brugge JF, Reale RA, Kovach CK, Oya H, Kawasaki H, Jenison RL, Howard MA 3rd (2013) Coding of repetitive transients by auditory cortex on posterolateral superior temporal gyrus in humans: an intracranial electrophysiology study. J Neurophysiol 109:1283–1295

Obleser J, Lahiri A, Eulitz C (2004) Magnetic brain response mirrors extraction of phonological features from spoken vowels. J Cogn Neurosci 16:31–39

Obleser J, Scott SK, Eulitz C (2006) Now you hear it, now you don't: transient traces of consonants and their nonspeech analogues in the human brain. Cereb Cortex 16:1069–1076

Okamoto H, Stracke H, Bermudez P, Pantev C (2011) Sound processing hierarchy within human auditory cortex. J Cogn Neurosci 23:1855–1863

Okamoto H, Stracke H, Ross B, Kakigi R, Pantev C (2007a) Left hemispheric dominance during auditory processing in noisy environment. BMC Biol 5:52

Okamoto H, Stracke H, Wolters CH, Schmael F, Pantev C (2007b) Attention improves population-level frequency tuning in human auditory cortex. J Neurosci 27:10383–10390

Palomaki KJ, Tiitinen H, Mäkinen V, May PJ, Alku P (2005) Spatial processing in human auditory cortex: the effects of 3D, ITD, and ILD stimulation techniques. Cogn Brain Res 24:364–379

Pantev C (1995) Evoked and induced gamma-band activity of the human cortex. Brain Topogr 7:321–330

Pantev C, Elbert T, Makeig S, Hampson S, Eulitz C, Hoke M (1993) Relationship of transient and steady-state auditory evoked fields. Electroencephalogr Clin Neurophysiol 88:389–396

Pantev C, Eulitz C, Elbert T, Hoke M (1994) The auditory evoked sustained field: origin and frequency dependence. Electroencephalogr Clin Neurophysiol 90:82–90

Pantev C, Eulitz C, Hampson S, Ross B, Roberts LE (1996a) The auditory evoked "off" response: sources and comparison with the "on" and the "sustained" responses. Ear Hear 17:255–265

Pantev C, Hoke M, Lehnertz K, Lütkenhöner B, Anogianakis G, Wittkowski W (1988) Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. Electroencephalogr Clin Neurophysiol 69:160–170

Pantev C, Lütkenhöner B, Hoke M, Lehnertz K (1986) Comparison between simultaneously recorded auditory-evoked magnetic fields and potentials elicited by ipsilateral, contralateral and binaural tone burst stimulation. Audiology 25:54–61

Pantev C, Okamoto H, Ross B, Stoll W, Ciurlia-Guy E, Kakigi R, Kubo T (2004) Lateral inhibition and habituation of the human auditory cortex. Eur J Neurosci 19:2337–2344

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

Parkkonen L, Fujiki N, Mäkelä JP (2009) Sources of auditory brainstem responses revisited: contribution by magnetoencephalography. Hum Brain Mapp 30:1772–1782

Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD (2002) The processing of temporal pitch and melody information in auditory cortex. Neuron 36:767–776

Pelizzone M, Hari R, Mäkelä JP, Huttunen J, Ahlfors S, Hämäläinen M (1987) Cortical origin of middle-latency auditory evoked responses inman. Neurosci. Lett. 82:303–307

Picton TW, Hillyard SA, Krausz HI, Galambos R (1974) Human auditory evoked potentials. I. Evaluation of components. Electroencephalogr Clin Neurophysiol 36:179–190

Poghosyan V, Ioannides AA (2008) Attention modulates earliest responses in the primary auditory and visual cortices. Neuron 58:802–813

Prendergast G, Johnson SR, Green GG (2010) Temporal dynamics of sinusoidal and non-sinusoidal amplitude modulation. Eur J Neurosci 32:1599–1607

Pressnitzer D, Patterson RD, Krumbholz K (2001) The lower limit of melodic pitch. J Acoust Soc Am 109:2074–2084

Ray S, Maunsell JH (2011) Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. PLoS Biol 9:e1000610

Reite M, Edrich J, Zimmermann JT, Zimmerman JE (1978) Human magnetic auditory evoked fields. Electroenceph Clin Neurophysiol 45:114–117

Reite M, Zimmerman JT, Zimmerman JE (1981) Magnetic auditory evoked fields: interhemispheric asymmetry. Electroencephalogr Clin Neurophysiol 51:388–392

Rif J, Hari R, Hämäläinen MS, Sams M (1991) Auditory attention affects two different areas in the human supratemporal cortex. Electroencephalogr Clin Neurophysiol 79:464–472

Ritter S, Dosch HG, Specht HJ, Rupp A (2005) Neuromagnetic responses reflect the temporal pitch change of regular interval sounds. Neuroimage 27:533–543

Rivier F, Clarke S (1997) Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: evidence for multiple auditory areas. Neuroimage 6:288–304

Roberts TP, Poeppel D (1996) Latency of auditory evoked M100 as a function of tone frequency. NeuroReport 7:1138–1140

Rogers RL, Baumann SB, Papanicolaou AC, Bourbon TW, Alagarsamy S, Eisenberg HM (1991) Localization of the P3 sources using magnetoencephalography and magnetic resonance imaging. Electroencephalogr Clin Neurophysiol 79:308–321

Romani GL, Williamson SJ, Kaufman L (1982) Tonotopic organization of the human auditory cortex. Science 216:1339–1340

Ross B, Draganova R, Picton TW, Pantev C (2003) Frequency specificity of 40-Hz auditory steady-state responses. Hear Res 186:57–68

Ross B, Herdman AT, Pantev C (2005a) Right hemispheric laterality of human 40 Hz auditory steady-state responses. Cereb Cortex 15:2029–2039

Ross B, Herdman AT, Pantev C (2005b) Stimulus induced desynchronization of human auditory 40-Hz steady-state responses. J Neurophysiol 94:4082–4093

Ross B, Picton TW, Herdman AT, Pantev C (2004) The effect of attention on the auditory steady-state response. Neurol Clin Neurophysiol 2004:22

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

Rupp A, Gutschalk A, Hack S, Scherg M (2002a) Temporal resolution of the human primary auditory cortex in gap detection. NeuroReport 13:2203–2207

Rupp A, Gutschalk A, Uppenkamp S, Scherg M (2004) Middle latency auditory-evoked fields reflect psychoacoustic gap detection thresholds in human listeners. J Neurophysiol 92:2239–2247

Rupp A, Hack S, Gutschalk A, Schneider P, Picton TW, Stippich C, Scherg M (2000) Fast temporal interactions in human auditory cortex. NeuroReport 11:3731–3736

Rupp A, Uppenkamp S, Gutschalk A, Beucker R, Patterson RD, Dau T, Scherg M (2002b) The representation of peripheral neural activity in the middle-latency evoked field of primary auditory cortex in humans(1). Hear Res 174:19–31

Salminen NH, May PJ, Alku P, Tiitinen H (2009) A population rate code of auditory space in the human cortex. PLoS ONE 4:e7600

Sams M, Hämäläinen M, Hari R, McEvoy L (1993a) Human auditory cortical mechanisms of sound lateralization: I. Interaural time differences within sound. Hear Res 67:89–97

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

Saupe K, Schröger E, Andersen SK, Müller MM (2009) Neural mechanisms of intermodal sustained selective attention with concurrently presented auditory and visual stimuli. Front Hum Neurosci 3:58

Schadwinkel S, Gutschalk A (2010) Activity associated with stream segregation in human auditory cortex is similar for spatial and pitch cues. Cereb Cortex 20:2863–2873

Scherg M, Hari R, Hämäläinen MS (1989) Frequency-specific sources of the auditory N19-P30-P50 response detected by a multiple source analysis of evoked magnetic fields and potentials. In: Williamson SJ, Hoke M, Sroink G, Kotani M (eds) Advances in biomagnetism. Plenum Press, New York

Scherg M, von Cramon D (1985) A new interpretation of the generators of BAEP waves I-V: results of a spatio-temporal dipole model. Electroencephalogr Clin Neurophysiol 62:290–299

Scherg M, Von Cramon D (1986) Evoked dipole source potentials of the human auditory cortex. Electroencephalogr Clin Neurophysiol 65:344–360

Schnupp JW, Nelken I, King AJ (2011) Auditory neuroscience: making sense of sound. MIT Press, Cambridge, MA

Schönwiesner M, Novitski N, Pakarinen S, Carlson S, Tervaniemi M, Näätänen R (2007) Heschl's gyrus, posterior superior temporal gyrus, and mid-ventrolateral prefrontal cortex have different roles in the detection of acoustic changes. J Neurophysiol 97:2075–2082

Sedley W, Teki S, Kumar S, Overath T, Barnes GR, Griffiths TD (2012) Gamma band pitch responses in human auditory cortex measured with magnetoencephalography. Neuroimage 59:1904–1911

Shaw ME, Hämäläinen MS, Gutschalk A (2013) How anatomical asymmetry of human auditory cortex can lead to a rightward bias in auditory evoked fields. Neuroimage 74:22–29

Sieroka N, Dosch HG, Specht HJ, Rupp A (2003) Additional neuromagnetic source activity outside the auditory cortex in duration discrimination correlates with behavioural ability. Neuroimage 20:1697–1703

Snyder JS, Alain C, Picton TW (2006) Effects of attention on neuroelectric correlates of auditory stream segregation. J Cogn Neurosci 18:1–13

Spierer L, Bellmann-Thiran A, Maeder P, Murray MM, Clarke S (2009) Hemispheric competence for auditory spatial representation. Brain 132:1953–1966

Stecker GC, Harrington IA, Middlebrooks JC (2005) Location coding by opponent neural populations in the auditory cortex. PLoS Biol 3:e78

Steinmann I, Gutschalk A (2011) Potential fMRI correlates of 40-Hz phase locking in primary auditory cortex, thalamus and midbrain. Neuroimage 54:495–504

Steinmann I, Gutschalk A (2012) Sustained BOLD and theta activity in auditory cortex are related to slow stimulus fluctuations rather than to pitch. J Neurophysiol 107:3458–3467

Steinschneider M, Fishman YI, Arezzo JC (2008) Spectrotemporal Analysis of Evoked and Induced Electroencephalographic Responses in Primary Auditory Cortex (A1) of the Awake Monkey. Cereb Cortex 18:610–625

Steinschneider M, Tenke CE, Schroeder CE, Javitt DC, Simpson GV, Arezzo JC, Vaughan HG Jr (1992) Cellular generators of the cortical auditory evoked potential initial component. Electroencephalogr Clin Neurophysiol 84:196–200

Stevens KN (2000) Acoustic phonetics. MIT Press, Cambridge

Todorovic A, van Ede F, Maris E, de Lange FP (2011) Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG study. J Neurosci 31:9118–9123

Uppenkamp S, Johnsrude IS, Norris D, Marslen-Wilson W, Patterson RD (2006) Locating the initial stages of speech-sound processing in human temporal cortex. Neuroimage 31:1284–1296

Van Noorden LPAS (1975) Temporal coherence in the perception of tone sequences. University of Technology, Eindhoven

Wacongne C, Changeux JP, Dehaene S (2012) A neuronal model of predictive coding accounting for the mismatch negativity. J Neurosci 32:3665–3678

Wang Y, Ding N, Ahmar N, Xiang J, Poeppel D, Simon JZ (2012) Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: MEG evidence. J Neurophysiol 107:2033–2041

Weisz N, Lecaignard F, Müller N, Bertrand O (2012) The modulatory influence of a predictive cue on the auditory steady-state response. Hum Brain Mapp 33:1417–1430

Wiegand K, Gutschalk A (2012) Correlates of perceptual awareness in human primary auditory cortex revealed by an informational masking experiment. Neuroimage 61:62–69

Woldorff MG, Gallen CC, Hampson SA, Hillyard SA, Pantev C, Sobel D, Bloom FE (1993) Modulation of early sensory processing in human auditory cortex during auditory selective attention. Proc Nat Acad Sci USA 90:8722–8726

Yost WA, Patterson R, Sheft S (1996) A time domain description for the pitch strength of iterated rippled noise. J Acoust Soc Am 99:1066–1078

Young ED, Sachs MB (1979) Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. J Acoust Soc Am 66:1381–1403

Yrttiaho S, Alku P, May PJ, Tiitinen H (2009) Representation of the vocal roughness of aperiodic speech sounds in the auditory cortex. J Acoust Soc Am 125:3177–3185

Yvert B, Crouzeix A, Bertrand O, Seither-Preisler A, Pantev C (2001) Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. Cereb Cortex 11:411–423