Sound Processing Hierarchy within Human Auditory Cortex

Hidehiko Okamoto[1], Henning Stracke[1], Patrick Bermudez[1,2,3], and Christo Pantev[1]

Abstract

■ Both attention and masking sounds can alter auditory neural processes and affect auditory signal perception. In the present study, we investigated the complex effects of auditory-focused attention and the signal-to-noise ratio of sound stimuli on three different auditory evoked field components (auditory steady-state response, N1m, and sustained field) by means of magnetoencephalography. The results indicate that the auditory steady-state response originating in primary auditory cortex reflects the signal-to-noise ratio of physical sound inputs (bottom–up process) rather than the listener's attentional state (top–down process), whereas the sustained field, originating in nonprimary auditory cortex, reflects the attentional state rather than the signal-to-noise ratio. The N1m was substantially influenced by both bottom–up and top–down neural processes. The differential sensitivity of the components to bottom–up and top–down neural processes, contingent on their level in the processing pathway, suggests a stream from bottom–up driven sensory neural processing to top–down driven auditory perception within human auditory cortex. ■

INTRODUCTION

Conscious sensory perception not only depends on external signal inputs but also on the attentional state of the recipients. Attentionally driven neural activity changes have frequently been investigated in the visual modality (Rotermund, Taylor, Ernst, Kreiter, & Pawelzik, 2009; Stokes, Thompson, Nobre, & Duncan, 2009; Bar et al., 2006; Fenske, Aminoff, Gronau, & Bar, 2006; Hopfinger, Buonocore, & Mangun, 2000; Mangun, 1995). However, numerous studies have demonstrated that auditory-focused attention can increase neural activity as well (e.g., Fritz, Elhilali, David, & Shamma, 2007a, 2007b). For instance, attentional effects in the human auditory modality have been measured with functional magnetic resonance imaging (fMRI), which has very high spatial but rather low temporal resolution (Paltoglou, Sumner, & Hall, 2009; Salmi, Rinne, Koistinen, Salonen, & Alho, 2009; Woods et al., 2009; Rahne et al., 2008; Rinne et al., 2007, 2008; Johnson & Zatorre, 2005, 2006). In order to overcome this limitation, methods such as electroencephalography and magnetoencephalography (MEG), which benefit from millisecond-scale temporal resolution, can be used to measure attentional effects on human auditory evoked responses with different latencies. For instance, the auditory evoked N1 response (and its magnetic counterpart, the N1m), which exhibits a typical latency of around 0.1 sec after sound onset (Näätänen & Picton, 1987), is enhanced by attention focused on the auditory input (Woldorff et al.,

1993; Picton & Hillyard, 1974). Ross, Picton, Herdman, and Pantev (2004) demonstrated that auditory attention can also amplify the sustained field (SF) response, which essentially is a stimulus-locked DC-shift of the magnetic field amplitude evolving subsequently to the transient evoked responses (Pantev, Eulitz, Elbert, & Hoke, 1994). However, the potential effect of attention on the auditory steady-state response (ASSR), which is typically elicited by a sequence of clicks (Hari, Hamalainen, & Joutsiniemi, 1989; Galambos, Makeig, & Talmachoff, 1981), Gaussian tone pulses (Pantev, Roberts, Elbert, Ross, & Wienbruch, 1996), or an amplitude-modulated tone (Engelien, Schulz, Ross, Arolt, & Pantev, 2000; Rees, Green, & Kay, 1986), is still contentiously debated. Some studies have observed an effect of auditory attention on ASSR amplitude (Müller, Schlee, Hartmann, Lorenz, & Weisz, 2009; Bidet-Caulet et al., 2007; Skosnik, Krishnan, & O'Donnell, 2007; Ross et al., 2004), whereas others have not (Linden, Picton, Hamel, & Campbell, 1987).

There is already considerable experimental evidence regarding the generator sites of the different components of the auditory evoked response in human auditory cortex. The N1 and N1m generators are thought to be located in lateral aspects of Heschl's gyrus and the temporal plane (Eggermont & Ponton, 2002; Pantev et al., 1995). In contrast, the ASSR seems to originate more medially in primary auditory cortex (Pantev et al., 1996; Makela & Hari, 1987), and the SF in the auditory belt region (Gutschalk, Patterson, Rupp, Uppenkamp, & Scherg, 2002; Pantev et al., 1994). Although both the sources of the N1m and the SF can be explained by a single equivalent current dipole and both components originate in nonprimary auditory cortex, the

---

[1]University of Muenster, Germany, [2]McGill University, Canada, [3]BRAMS Laboratory, Canada

source of the SF is indeed spatially distinct from the N1m source, demonstrating that the underlying neuronal populations are at least partially different (Mackert et al., 1999; Eulitz, Diesch, Pantev, Hampson, & Elbert, 1995). Given these different cortical source sites, auditory attention might differentially affect the generators of ASSR, N1m, and SF. Recently, an fMRI experiment by Petkov et al. (2004) demonstrated that the primary, medial part of human auditory cortex is a stimulus-driven area that is always and strictly activated by a sound, regardless of the listener's state of attention. Even unattended sounds activate medial auditory cortex in a way very similar to attended ones. In contrast, the activation of nonprimary, lateral auditory cortex strongly depends on the state of attention, regardless of the type of sound input. These results support the hypothesis that there are differential effects of attention on primary and nonprimary auditory cortices; focused attention as a top–down process seems to affect mainly neural responses in nonprimary auditory cortical areas, whereas the physical features of the sound (bottom–up) mainly affect those in primary auditory cortex. Combinations of these top–down and bottom–up driven neural processes could lead to improved auditory performance compared to that which might be obtained by either process alone. However, the time courses of bottom–up and top–down effects on neural activity in auditory cortex remain elusive.

Following the results summarized above, the goal of the present study was to employ MEG to investigate the effects of attention and masking sounds on three auditory evoked components representing population-level neural activity in primary or nonprimary auditory cortex. The test stimulus (TS), a 40-Hz amplitude-modulated tone, was presented simultaneously with a masking white noise (WN), which varied parametrically in sound level during attentive versus distracted listening conditions (Figure 1). The utilization of the 40-Hz amplitude-modulated TS enabled us to investigate ASSR, N1m, and SF simultaneously (Engelien et al., 2000).

METHODS

Subjects

Sixteen healthy subjects (8 women, age range = 23–30 years, mean age = 26.2 years), without any history of psychiatric or neurological disorders, participated in the present study. All subjects were right-handed (assessed via Edinburgh Handedness Inventory; Oldfield, 1971) and had normal hearing as confirmed by clinical pure tone audiometry. All participants gave written informed consent for participation in the study in accordance with procedures approved by the Ethics Committee of the Medical Faculty, University of Muenster.

Stimuli and Experimental Design

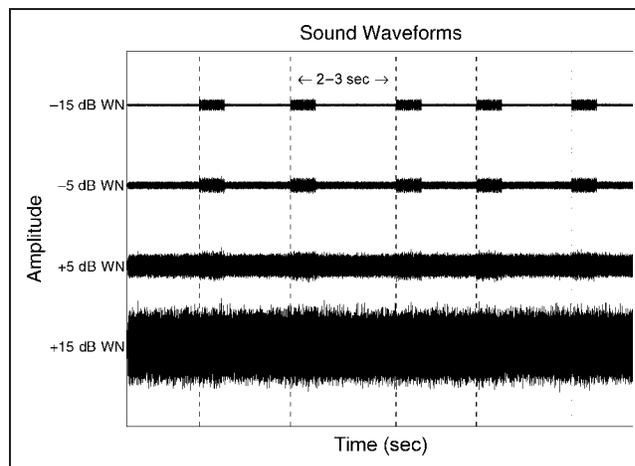We presented a binaural test sound signal (TS) simultaneously with 8000 Hz low-pass filtered WN (sampling rate =

Figure 1. Sound waveforms of noises overlapping the test stimulus (TS) of 0.7 sec duration are displayed from highest (top, −15 dB) to lowest (bottom, +15 dB) signal-to-noise ratio. The TS waveform is not clearly visible in the combined waveform in the bottom row due to the much smaller intensity of the TS compared to WN. The vertical dashed (1st to 4th) and dotted (5th) lines represent the onset of the standard and deviant (containing frequency shift) TS, respectively.

48,000 Hz). The TS, with a carrier frequency of 1000 Hz and a duration of 0.7 sec, was amplitude-modulated with a modulation frequency of 40 Hz and a modulation depth of 100%. The WN power was either −15 dB, −5 dB, +5 dB, or +15 dB relative to the TS power (cf. Figure 1) and had a duration of 60 sec (0.01 sec rise and fall times) and therefore was present throughout the presentations of the TS. The sound onset asynchrony between subsequent TS presentations was pseudorandomized between 2 and 3 sec, resulting in 24 TS being presented during each WN signal. The TS deviated from the standard TS in 50% of trials for the behavioral pretest session and 10% of trials for the MEG session. In these "deviant" trials, the TS contained a carrier frequency change starting either 0.175, 0.35, or 0.525 sec after TS-onset. Responses elicited by these deviant TS were excluded from MEG data analysis.

All sound stimuli were prepared as digital sound files and were delivered with the Presentation software package (Neurobehavioral Systems, Albany, CA). All sounds were delivered through plastic tubes and silicon earpieces were individually fitted to each subject's ear. Hearing thresholds for the TS were determined for each ear before both behavioral pilot studying and MEG measurements. The TS was presented binaurally at an intensity level of 35 dB above individual sensation level. WNs were also presented binaurally. In advance of the MEG session, we performed a behavioral pretest in the magnetically shielded room in order to balance the deviant detection difficulty levels between WN conditions. Thirty TS (15 standard TS and 15 deviant TS) were presented during each WN condition to 16 subjects, who also participated in the subsequent MEG measurement, which was performed on a different day. In order to investigate effects of attention, we presented six different stimulation blocks per subject during

the MEG measurement. At the beginning of three blocks constituting the attentive listening condition, subjects were instructed to press a response button as quickly as possible with their left or right index finger (8 subjects each) when they perceived an upward shift in frequency within a TS (deviant TS: 10%). Based on the behavioral pretest, frequency shifts were set to either 10 Hz (in case of the −15 dB WN condition), 11 Hz (−5 dB WN), 14 Hz (+5 dB WN), or 40 Hz (+15 dB WN). During the remaining three blocks, constituting the distracted listening condition, subjects performed a visual target detection task in order to prevent them from paying attention to the sound stimuli and to keep them in a stable alert state. Visual stimuli consisted of one to nine crosses, which could appear simultaneously in nine fixed positions on the screen (3 rows × 3 columns). Participants were instructed to fix their eyes on the cross located at the center of the screen (always visible during the MEG measurement) and to press a response button as quickly as possible when they detected four neighboring crosses arranged into a small square. The visual stimulation was totally independent from the auditory stimulation. The visual task was solely intended to draw participants' attention away from the auditory stimuli. Further details regarding the visual task are described in a previous article (Stracke, Okamoto, & Pantev, 2009). Although, the bottom–up auditory and visual inputs were identical between attentive and distracted listening conditions, the focus of attention was either on the auditory inputs (attentive listening) or the visual inputs (distracted listening). The initial condition (attentive or distracted listening) was pseudorandomized between subjects (resulting in 8 subjects for each), and attentive and distracted listening blocks alternated within subjects. During each attention condition, 144 standard TS trials were presented for each WN condition.

Data Acquisition and Analysis

Auditory evoked fields (AEFs) were measured with a helmet-shaped 275-channel whole-head neurogradiometer (VSM Med-Tech Ltd., Coquitlam, BC, Canada) in a silent magnetically shielded room. The magnetic field signals were digitally sampled at a rate of 600 Hz. Epochs of magnetic field data elicited by the standard TS, including 0.3 sec prestimulus and 0.8 sec poststimulus intervals, were averaged selectively for each WN condition after rejection of artifact-contaminated epochs containing field changes larger than 3 pT. The origins of locations and orientations of the equivalent dipolar sources of the different evoked response components were determined in a Cartesian coordinate system with an origin at the midpoint of the medial–lateral axis (y-axis) connecting the preauricular points of both ears. The posterior–anterior axis (x-axis) ran between the nasion and the origin, and the inferior–superior axis (z-axis) ran through the origin perpendicularly to the x–y-plane.

For the analysis of the ASSR, the grand-averaged magnetic field signals across all WN conditions were initially band-pass filtered between 32 and 48 Hz. Following this, a 40-Hz sine wave was fitted to each magnetic waveform within the time range from 0.4 to 0.7 sec in order to increase the signal-to-noise ratio of the evoked responses prior to the dipole-fit procedure (Ross, Herdman, & Pantev, 2005). The locations and orientations of fixed single equivalent current dipoles corresponding to the maximal global field power, measured as root-mean-square across all sensors, were then approximated above right and left hemispheres for each subject individually. The resulting source locations and orientations were fixed, and the source strengths were approximated for the 40-Hz fitted magnetic waveforms in each WN condition. Then, the maximal source strengths for each condition and hemisphere were calculated.

For the analysis of the N1m responses, the grand-averaged magnetic field signals were first 30-Hz low-pass filtered and the baseline was corrected relative to a 0.3-sec prestimulus interval. The locations and orientations of fixed single equivalent current dipoles corresponding to the N1m responses were individually approximated to the averaged magnetic field distribution of all sensors by using a 0.01-sec time window centered at the time point of maximal global field power around 0.15 sec after stimulus onset. The estimated source for each subject in each hemisphere was fixed in its location and orientation, and the source strengths were calculated for all time points for each WN condition. Thereafter, the maximal source strengths were calculated for each condition and hemisphere.

In order to analyze the auditory evoked SF, the grand-averaged magnetic field responses were first 5-Hz low-pass filtered and the baseline was corrected relative to a 0.3-sec prestimulus interval. The fixed source locations and orientations were approximated between 0.4 and 0.7 sec for the grand-averaged MEG waveforms of all conditions. The estimated source for each subject in each hemisphere was fixed in its location and orientation and the average source strengths between 0.4 and 0.7 sec were calculated for each WN condition and used for further analysis.

The source strengths of ASSR, N1m, and SF elicited by the TS for each WN condition were normalized with respect to the mean ASSR, N1m, and SF source strengths across all WN conditions for each subject and for each hemisphere individually. These normalized ASSR, N1m, and SF source strengths were evaluated via a repeated measures analysis of variance (ANOVA) using three factors (attention: attentive and distracted; noise level: −15 dB, −5 dB, +5 dB, and +15 dB; AEF component: ASSR, N1m, and SF). Additionally, the normalized source strengths of each component (ASSR, N1m, and SF) were evaluated separately via repeated measures ANOVAs using attention and noise level as factors.

RESULTS

The means and standard deviations of the error rates obtained during the behavioral pretests were 15.4 ± 9.4% (10 Hz frequency shift in the −15 dB WN condition),

15.9 ± 10.8% (11 Hz frequency shift in the −5 dB WN condition), 15.1 ± 8.8% (14 Hz frequency shift in the +5 dB WN condition), and 15.1 ± 8.6% (40 Hz frequency shift in the +15 dB WN condition). The means and standard deviations of the hit rates obtained during the MEG session were 66.1 ± 27.6% (−15 dB WN condition), 70.4 ± 27.6% (−5 dB WN condition), 67.7 ± 29.9% (+5 dB WN condition), and 70.7 ± 28.0% (+15 dB WN condition). There was no significant difference between the different WN conditions. Clearly identifiable auditory evoked fields were obtained from all subjects in the MEG measurements. The mean and standard deviation of the trial numbers averaged for each condition was 136.9 ± 7.5 (range 114–144). Figure 2 shows the sensor waveforms, contour maps, and calculated equivalent current dipole locations of one representative subject overlaid onto the individual MRI brain reconstruction. Clear dipolar patterns are visible over the right hemisphere. The source estimation goodness-of-fit means and standard deviations were 90.7 ± 2.2% for the ASSR, 96.5 ± 1.8% for the N1m, and 93.8 ± 2.3% for the SF, further justifying the use of single equivalent dipoles in each hemisphere for the analysis of the present data. The group-averaged source locations of the N1m, ASSR, and SF components are shown in Figure 3, which demonstrates clearly that the centers of the estimated source locations were significantly different between components. The ASSR estimated sources were located more medially compared to the N1m response, and the SF was characterized by more anterior, medial, and inferior estimated source locations compared to the N1m response, indicating that these three evoked components were at least partially generated by distinct neural populations. These results are in line with previous studies demonstrating that the ASSR

is more medially located compared to the N1m response (Engelien et al., 2000; Pantev et al., 1996), and that the SF has a more anterior, more medial, and more inferior source than the N1m response (Eulitz et al., 1995; Pantev et al., 1994).

The 30-Hz low-pass filtered N1m source strength waveforms grand-averaged across all subjects for the time range of −0.3 to 0.3 sec, as well as the 5-Hz low-pass filtered SF source strength waveforms grand-averaged across all subjects for the time range of 0.4 to 0.8 sec, are displayed in Figure 4. The clear N1m response after TS-onset and the stable SF between 0.4 and 0.7 sec are clearly discernable. The N1m responses in the two loudest WN conditions (+15 and +5 dB WN) had longer durations and smaller peak amplitudes compared to the two softer WN conditions (−15 and −5 dB WN). As can be seen in Figure 4, the grand-averaged SF source waveforms were clearly amplified during attentive listening. However, the differences in SF source strengths between different WN conditions during attentive listening were not as pronounced as with the N1m responses.

The repeated measures ANOVA applied to the normalized source strengths resulted in significant main effects for attention [$F(1, 31) = 220, p < .001$] and noise level [$F(3, 93) = 250, p < .001$]. There were also significant Attention × Noise level [$F(3, 93) = 6.1, p < .001$], Attention × AEF component [$F(2, 62) = 79, p < .001$], and Noise level × AEF component [$F(6, 90) = 50, p < .001$] interactions. The significant interactions between attention and AEF component and between noise level and AEF component indicated that the effects of attention and noise significantly differed between the auditory evoked components (ASSR, N1m, and SF). The means of the normalized ASSR, N1m,

Figure 2. The top row represents individual auditory evoked magnetic fields grand-averaged across all conditions. The bottom row represents isocontour maps of the magnetic fields and estimated equivalent current dipoles corresponding to three evoked components (ASSR, N1m, and SF from left to right) overlaid onto the individual MRI. The magnetic waveforms are 32–48 Hz band-pass filtered (ASSR), 30-Hz low-pass filtered (N1m), or 5-Hz low-pass filtered (SF). Black lines represent inward flow of magnetic fields to the brain, black lines represent outward flow. The contour maps show clear dipolar patterns above right auditory cortex. Gray circles and bars represent the estimated dipole locations and orientations.
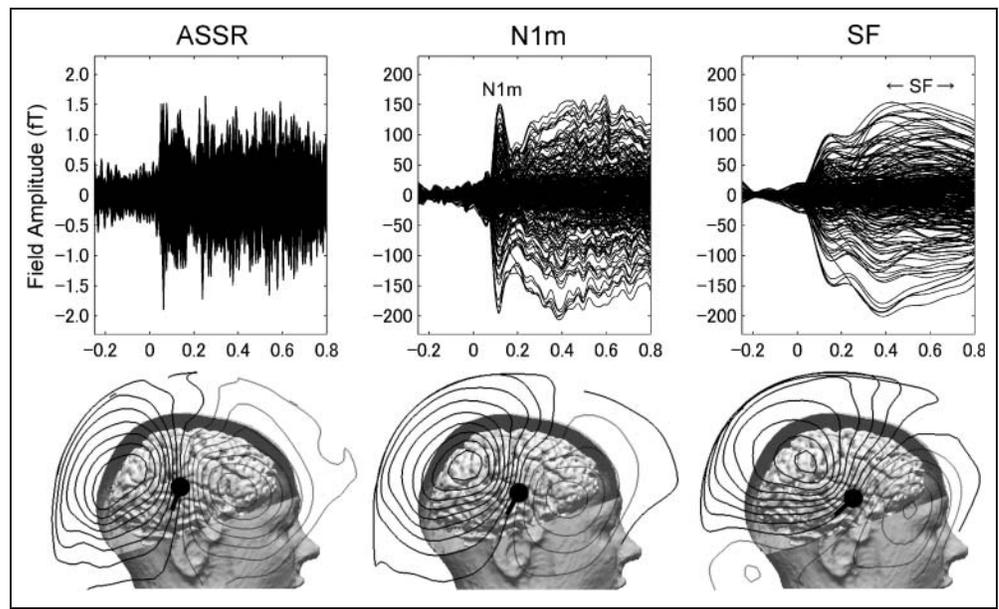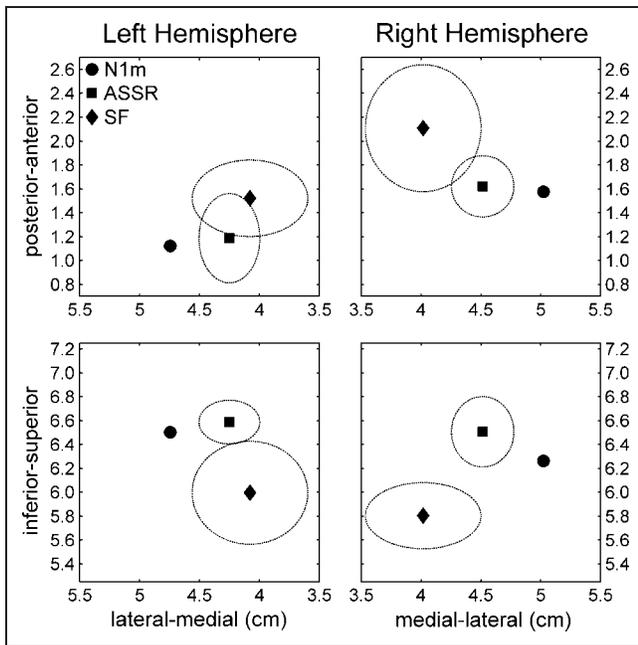
Figure 3. Mean localizations of the N1m (circle symbol), ASSR (square symbol), and SF sources [diamond symbols in the y–x-plane (medial–lateral vs. posterior–anterior directions) and the y–z plane (medial–lateral vs. inferior–superior directions)]. The ellipses around ASSR and SF locations denote the 95% confidence limits of the differences from N1m locations.

Figure 5. The group means of the normalized auditory steady-state response (ASSR) source strengths with error bars denoting the 95% confidence limits for the group means. Filled and open circles denote the responses during attentive and distracted listening.

and SF source strengths for each WN condition are presented in Figures 5, 6, and 7, respectively. The normalized ASSR source strengths strongly depended on the sound level of the simultaneously presented WN. Comparatively, attention had much less effect on the normalized ASSR source strengths (Figure 5). This differs from the case of the normalized SF source strengths, which were hardly in-

fluenced by the simultaneously presented WN, although they could be doubled by attention as compared to distracted listening (Figure 7). The normalized N1m response was influenced by both the WN and attention (Figure 6), with a softer WN and focused attention resulting in comparatively larger N1m responses.

The repeated measures ANOVAs calculated for each evoked component separately resulted in significant main effects of attention and noise level for all components [attention: ASSR, $F(1, 31) = 9.6, p < .005$; N1m, $F(1, 31) = 48$,

Figure 4. The graph displays grand-averaged source waveforms of the N1m (−0.3 to 0.3 sec) and the SF (0.4 to 0.8 sec) for all noise conditions during attentive and distracted listening conditions.

Figure 6. The group means of the normalized N1m source strengths with error bars denoting the 95% confidence limits for the group means. Filled and open squares denote the responses during attentive and distracted listening.

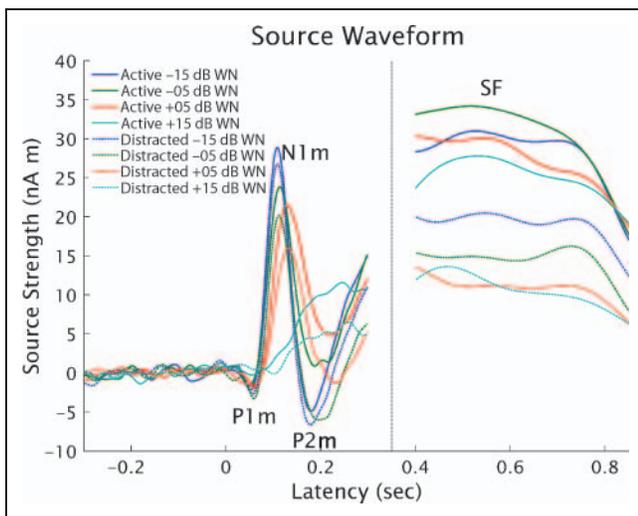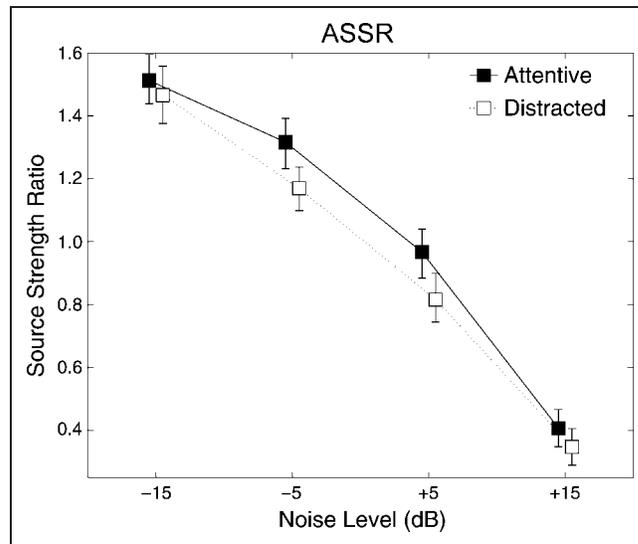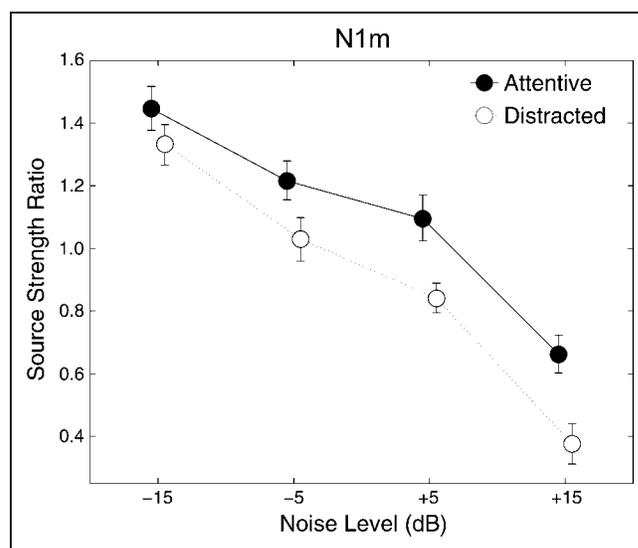**Figure 7.** The group means of the normalized sustained field (SF) source strengths with error bars denoting the 95% confidence limits for the group means. Filled and open diamonds denote the responses during attentive and distracted listening.

$p < .001$; SF, $F(1, 31) = 270, p < .001$; noise level: ASSR, $F(3, 93) = 210, p < .001$; N1m, $F(3, 93) = 150, p < .001$; SF, $F(1, 31) = 14, p < .001$]. Moreover, there were significant interactions between attention and noise level in case of N1m and SF [N1m: $F(3, 93) = 3.9, p < .02$; SF: $F(3, 93) = 4.4, p < .007$]. The results indicated that the attentional gain effect was significant for all AEF components including the ASSR, which was characterized by the smallest attentional effect among the three AEF components. The effects of the WNs were also significant for all AEF components.

In order to examine whether the slow readiness field (Deecke, Scheid, & Kornhuber, 1969) had a significant effect on the SF during attentive listening, the normalized SF source strengths in the hemispheres ipsilateral and contralateral to the finger with which the button press was made were evaluated via planned comparison. There were no significant effect of the side of button press [(Attentive_SF_Contralateral − Distracted_SF_Contralateral) − (Attentive_SF_Ipsilateral − Distracted_SF_Ipsilateral): $F(1, 15) = 0.046, p = .834$].

DISCUSSION

The present study confirmed that the auditory evoked fields elicited by the TS depended on both the signal-to-noise ratio (i.e., −15 dB, −5 dB, +5 dB, or +15 dB) of the external sounds and the internal attentional state (attentive vs. distracted listening) of the subjects. Results demonstrated that the external and internal factors differentially impacted the ASSR, N1m, and SF components of the auditory evoked responses, which are known to be generated at different cortical sites. N1m sources have a nonprimary

auditory cortex origin (Pantev et al., 1995; Liegeois-Chauvel, Musolino, Badier, Marquis, & Chauvel, 1994), whereas the ASSR emerges mainly in primary auditory cortex (Engelien et al., 2000; Pantev et al., 1996; Makela & Hari, 1987). The site of origin of the SF is still actively debated and has been posited to be either the supratemporal region (Pantev et al., 1994) or to consist of separate sources adjacent to primary auditory cortex (Gutschalk et al., 2002). As shown in Figure 5, ASSR source strengths were strongly affected by the external signal-to-noise ratio of the sound stimuli but to a much lesser degree by the attentional state. In contrast, SF source strengths strongly depended on the attentional state, whereas the external signal-to-noise ratio was less relevant (Figure 7). N1m source strengths were modulated by both external and internal factors concurrently (Figure 6); the N1m was larger during attentive listening and soft WN conditions as compared to distracted listening and loud WN conditions. These results suggest that bottom–up inputs strongly influenced activity in primary auditory cortex, whereas the top–down neural processes strongly affected nonprimary auditory cortex.

Sensory stimulation (auditory and visual inputs) was identical between the attentive and distracted listening conditions. Moreover, the task difficulty was similar between the different WN conditions, hence, vigilance and attention levels were also comparable. By virtue of these experimental design details, we were able to isolate and examine the effects of both bottom–up and top–down neural influences on the cortical generators of N1m, ASSR, and SF at the same time.

Simultaneously presented masking sounds have already been shown to affect both the ASSR (Galambos & Makeig, 1992) and N1m responses (Okamoto, Stracke, Ross, Kakigi, & Pantev, 2007; Morita et al., 2006; Hari & Makela, 1988) to an auditory test signal. The present study did indeed demonstrate that relatively loud masking noises caused remarkable ASSR and N1m source strength decrements (Figures 5 and 6) but also that the masking sounds could significantly decrease the SF source strength (Figure 7), even though the effect was much smaller compared to the ASSR and N1m.

Moreover, we found that all three auditory evoked responses (ASSR, N1m, and SF) were significantly larger in the active listening condition than in the distracted listening condition. By presenting a test sound simultaneously with different band-eliminated noise maskers during attentive versus distracted listening conditions, previous studies (Okamoto, Stracke, Zwitserlood, Roberts, & Pantev, 2009; Stracke et al., 2009; Okamoto, Stracke, Wolters, Schmael, & Pantev, 2007) demonstrated that auditory-focused attention can enhance the neural processing of task-relevant sounds and, at the same time, can suppress task-irrelevant neural activity. Therefore, the significant ASSR, N1m, and SF source strength differences between the attentive and distracted listening conditions in the present study might reflect either an attentional gain effect on the task-relevant neural activity corresponding to the test stimulus, or an

attentional inhibitory effect on the task-irrelevant neural activity corresponding to the noise, or most likely, both. Notably, herein, we succeeded in measuring for the first time the simultaneous effects of attention and signal-to-noise ratio of external sounds on three different auditory evoked components (ASSR, N1m, and SF), each characterized by different latencies and source locations.

Compared to the N1m and SF responses, the ASSR source strengths showed significant yet smaller differences between attentive and distracted listening conditions (Figure 5). However, the effect of attention on primary auditory cortex is not yet fully understood. Several studies (Müller et al., 2009; Paltoglou et al., 2009; Poghosyan & Ioannides, 2008; Bidet-Caulet et al., 2007; Skosnik et al., 2007; Ross et al., 2004) demonstrated attentional effects on human primary auditory cortex, which was enhanced when elicited by task-relevant auditory signals compared to task-irrelevant auditory signals, whereas other studies (Petkov et al., 2004; Linden et al., 1987) did not find significant differences between attentive and distracted listening conditions in human primary auditory cortex. Reasons for this discrepancy likely include task requirement differences. In the present study, during active listening, we used a frequency change detection task which obliged participants to continuously pay attention to the TS, and we used an active visual task during distracted listening. Consequently, attention was not focused on the 40-Hz modulation frequency, but rather on the carrier frequency. However, we still found significantly larger ASSR source strengths in the active compared to the distracted listening condition. The visual task during distracted listening might have resulted in a stronger attentional contrast compared to the studies of Petkov et al. (2004) and Linden et al. (1987), leading to the significant attentional effect in primary auditory cortex observed in our results. However, we must emphasize that this significant attentional gain effect was much smaller compared to that observed on the later evoked components (N1m and SF). Furthermore, it is noteworthy that the neural activity in primary auditory cortex seemed to be strongly driven by the signal-to-noise ratio of auditory inputs (Figure 5). It was not difficult for the subjects to consciously perceive the TS even in the lowest signal-to-noise condition (+15 dB WN). Nonetheless, the results showed that the ASSR was almost suppressed in the +15 dB WN condition. Thus, the ASSR obtained by means of MEG would not mainly represent active cognitive neural activity, but rather sensory neural activity elicited by external sound inputs (note how there is no 40-Hz rhythm visible in Figure 1 due to small signal-to-noise ratio).

The normalized SF source strengths were significantly affected by attention. The SF source strengths were significantly larger during attentive compared to distracted listening and were less dependent on the signal-to-noise ratio of external sound input (Figure 7). Few studies have investigated the effects of attention on SF source strengths. Picton, Woods, and Proulx (1978) showed that the SF was enhanced when participants paid attention to the duration of test sounds, whereas the SF did not change when they paid attention to intensity or pitch of the test signals. In case of the intensity or pitch discrimination task, participants did not need to continuously pay attention to the sound signals because the pitch and intensity could be identified right at the beginning of the sound stimuli. Therefore, the SF might be enhanced only when attention is continuously focused on the auditory signals. In the present study, the participants would have continuously focused their attention on the auditory signals because the task was to detect a temporal gap, which might have occurred at either 0.175, 0.35, or 0.525 sec after TS-onset. This continuous attention during the whole sound signal presentation probably resulted in the significant SF increments. Considering the high goodness-of-fit values for the dipole solutions, and given that there was no effect of hemisphere for the button press finger side, the SF in our data does not represent cortical activity in the motor area, but rather neural activity within auditory cortex.

The SF responses might reflect the subjects' awareness of the test sounds, as suggested by Gutschalk, Micheyl, and Oxenham (2008). By measuring auditory evoked fields during informational masking, these authors demonstrated that a long latency response (0.05–0.25 sec after sound onset) was discernable only when the subjects were aware of the target auditory signals, whereas both detected and undetected targets elicited equally robust ASSR. These results are congruent with our findings in showing that the SF strongly represents top–down attentional neural activity, whereas the ASSR strongly reflects the signal-to-noise ratios of bottom–up sound inputs. Considering the latency of the SF (0.4–0.7 sec), our findings might result from feedback transmitted via the top–down pathway within auditory cortex (Nahum, Nelken, & Ahissar, 2008; Eggermont, 2001; Wallace, Kitzes, & Jones, 1991). In the present study, even though the TS onsets were not visible in the sound waveform of the +15 dB WN condition (cf. Figure 1), the stimuli were easily detected by the auditory system. After the reception of the TS-related signals by nonprimary auditory cortex, attention might have increased the TS-related neural activity with attentive listening via the top–down attentional system in order to successfully detect the frequency shift within the TS, as was demanded of subjects by the behavioral task. This would result in the amplification of the SF responses elicited by the standard TS, which is indeed what we observed.

In sum, ASSR, N1m, and SF responses were differently influenced by auditory-focused attention and external sounds. This finding indicates that the effects of top–down and bottom–up neural inputs differed at different cortical sites, even though the task in the present study was not specifically related to one of these auditory evoked components. The ASSR, originating in primary auditory cortex, strongly reflects the signal-to-noise ratios of bottom–up sound inputs and is relatively weakly influenced by the top–down attentional state. In contrast, the SF is largely dependent on the subject's attentional state and is much

less dependent on the signal-to-noise ratios of sound inputs. The N1m is considerably influenced by both the signal-to-noise ratios of bottom–up sound inputs and top–down attention processes. Hence, our results demonstrate a hierarchy from sensory neural processing in primary auditory cortex to active perception within nonprimary human auditory cortical structures. Taken together, these findings would demonstrate neural encoding stages ranging from sensory neural inputs to active perception *within* human auditory cortex, and a means for the objective measurement of subjective sound perception.

REFERENCES

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., et al. (2006). Top–down facilitation of visual recognition. *Proceedings of the National Academy of Sciences, U.S.A., 103,* 449–454.

Bidet-Caulet, A., Fischer, C., Besle, J., Aguera, P. E., Giard, M. H., & Bertrand, O. (2007). Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. *Journal of Neuroscience, 27,* 9252–9261.

Deecke, L., Scheid, P., & Kornhuber, H. H. (1969). Distribution of readiness potential, pre-motion positivity, and motor potential of the human cerebral cortex preceding voluntary finger movements. *Experimental Brain Research, 7,* 158–168.

Eggermont, J. J. (2001). Between sound and perception: Reviewing the search for a neural code. *Hearing Research, 157,* 1–42.

Eggermont, J. J., & Ponton, C. W. (2002). The neurophysiology of auditory perception: From single units to evoked potentials. *Audiology & Neuro-otology, 7,* 71–99.

Engelien, A., Schulz, M., Ross, B., Arolt, V., & Pantev, C. (2000). A combined functional in vivo measure for primary and secondary auditory cortices. *Hearing Research, 148,* 153–160.

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. *Journal of Neuroscience, 15,* 2748–2755.

Fenske, M. J., Aminoff, E., Gronau, N., & Bar, M. (2006). Top–down facilitation of visual object recognition: Object-based and context-based contributions. *Progress in Brain Research, 155,* 3–21.

Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007a). Auditory attention-focusing the searchlight on sound. *Current Opinion in Neurobiology, 17,* 437–455.

Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007b). Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? *Hearing Research, 229,* 186–203.

Galambos, R., & Makeig, S. (1992). Physiological studies of central masking in man. I: The effects of noise on the 40-Hz steady-state response. *Journal of the Acoustical Society of America, 92,* 2683–2690.

Galambos, R., Makeig, S., & Talmachoff, P. J. (1981). A 40-Hz auditory potential recorded from the human scalp. *Proceedings of the National Academy of Sciences, U.S.A., 78,* 2643–2647.

Gutschalk, A., Micheyl, C., & Oxenham, A. J. (2008). Neural correlates of auditory perceptual awareness under informational masking. *PLoS Biology, 6,* 1156–1165.

Gutschalk, A., Patterson, R. D., Rupp, A., Uppenkamp, S., & Scherg, M. (2002). Sustained magnetic fields reveal separate sites for sound level and temporal regularity in human auditory cortex. *Neuroimage, 15,* 207–216.

Hari, R., Hamalainen, M., & Joutsiniemi, S. L. (1989). Neuromagnetic steady-state responses to auditory stimuli. *Journal of the Acoustical Society of America, 86,* 1033–1039.

Hari, R., & Makela, J. P. (1988). Modification of neuromagnetic responses of the human auditory cortex by masking sounds. *Experimental Brain Research, 71,* 87–92.

Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top–down attentional control. *Nature Neuroscience, 3,* 284–291.

Johnson, J. A., & Zatorre, R. J. (2005). Attention to simultaneous unrelated auditory and visual events: Behavioral and neural correlates. *Cerebral Cortex, 15,* 1609–1620.

Johnson, J. A., & Zatorre, R. J. (2006). Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. *Neuroimage, 31,* 1673–1681.

Liegeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., & Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: Evaluation and topography of the middle latency components. *Electroencephalography and Clinical Neurophysiology, 92,* 204–214.

Linden, R. D., Picton, T. W., Hamel, G., & Campbell, K. B. (1987). Human auditory steady-state evoked potentials during selective attention. *Electroencephalography and Clinical Neurophysiology, 66,* 145–159.

Mackert, B. M., Wubbeler, G., Burghoff, M., Marx, P., Trahms, L., & Curio, G. (1999). Non-invasive long-term recordings of cortical "direct current" (DC-) activity in humans using magnetoencephalography. *Neuroscience Letters, 273,* 159–162.

Makela, J. P., & Hari, R. (1987). Evidence for cortical origin of the 40 Hz auditory evoked response in man. *Electroencephalography and Clinical Neurophysiology, 66,* 539–546.

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology, 32,* 4–18.

Morita, T., Fujiki, N., Nagamine, T., Hiraumi, H., Naito, Y., Shibasaki, H., et al. (2006). Effects of continuous masking noise on tone-evoked magnetic fields in humans. *Brain Research, 1087,* 151–158.

Müller, N., Schlee, W., Hartmann, T., Lorenz, I., & Weisz, N. (2009). Top–down modulation of the auditory steady-state response in a task-switch paradigm. *Frontiers in Human Neuroscience, 3,* 1.

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.

Nahum, M., Nelken, I., & Ahissar, M. (2008). Low-level information and high-level perception: The case of speech in noise. *PLoS Biology, 6,* 978–991.

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

Okamoto, H., Stracke, H., Wolters, C. H., Schmael, F., & Pantev, C. (2007). Attention improves population-level frequency tuning in human auditory cortex. *Journal of Neuroscience, 27,* 10383–10390.

Okamoto, H., Stracke, H., Zwitserlood, P., Roberts, L. E., & Pantev, C. (2009). Frequency-specific modulation of population-level frequency tuning in human auditory cortex. *BMC Neuroscience, 10,* 1.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9,* 97–113.

Paltoglou, A. E., Sumner, C. J., & Hall, D. A. (2009). Examining the role of frequency specificity in the enhancement and suppression of human cortical activity by auditory selective attention. *Hearing Research, 257,* 106–118.

Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., et al. (1995). Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *Electroencephalography and Clinical Neurophysiology, 94,* 26–40.

Pantev, C., Eulitz, C., Elbert, T., & Hoke, M. (1994). The auditory evoked sustained field: Origin and frequency dependence. *Electroencephalography and Clinical Neurophysiology, 90,* 82–90.

Pantev, C., Roberts, L. E., Elbert, T., Ross, B., & Wienbruch, C. (1996). Tonotopic organization of the sources of human auditory steady-state responses. *Hearing Research, 101,* 62–74.

Petkov, C. I., Kang, X., Alho, K., Bertrand, O., Yund, E. W., & Woods, D. L. (2004). Attentional modulation of human auditory cortex. *Nature Neuroscience, 7,* 658–663.

Picton, T. W., & Hillyard, S. A. (1974). Human auditory evoked potentials: II. Effects of attention. *Electroencephalography and Clinical Neurophysiology, 36,* 191–199.

Picton, T. W., Woods, D. L., & Proulx, G. B. (1978). Human auditory sustained potentials. II. Stimulus relationships. *Electroencephalography and Clinical Neurophysiology, 45,* 198–210.

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

Rahne, T., Deike, S., Selezneva, E., Brosch, M., Konig, R., Scheich, H., et al. (2008). A multilevel and cross-modal approach towards neuronal mechanisms of auditory streaming. *Brain Research, 1220,* 118–131.

Rees, A., Green, G. G., & Kay, R. H. (1986). Steady-state evoked responses to sinusoidally amplitude-modulated sounds recorded in man. *Hearing Research, 23,* 123–133.

Rinne, T., Balk, M. H., Koistinen, S., Autti, T., Alho, K., & Sams, M. (2008). Auditory selective attention modulates activation of human inferior colliculus. *Journal of Neurophysiology, 100,* 3323–3327.

Rinne, T., Stecker, G. C., Kang, X., Yund, E. W., Herron, T. J., & Woods, D. L. (2007). Attention modulates sound processing in human auditory cortex but not the inferior colliculus. *NeuroReport, 18,* 1311–1314.

Ross, B., Herdman, A. T., & Pantev, C. (2005). Right hemispheric laterality of human 40 Hz auditory steady-state responses. *Cerebral Cortex, 15,* 2029–2039.

Ross, B., Picton, T. W., Herdman, A. T., & Pantev, C. (2004). The effect of attention on the auditory steady-state response. *Neurology & Clinical Neurophysiology, 2004,* 22.

Rotermund, D., Taylor, K., Ernst, U. A., Kreiter, A. K., & Pawelzik, K. R. (2009). Attention improves object representation in visual cortical field potentials. *Journal of Neuroscience, 29,* 10120–10130.

Salmi, J., Rinne, T., Koistinen, S., Salonen, O., & Alho, K. (2009). Brain networks of bottom–up triggered and top–down controlled shifting of auditory attention. *Brain Research, 1286,* 155–164.

Skosnik, P. D., Krishnan, G. P., & O'Donnell, B. F. (2007). The effect of selective attention on the gamma-band auditory steady-state response. *Neuroscience Letters, 420,* 223–228.

Stokes, M., Thompson, R., Nobre, A. C., & Duncan, J. (2009). Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proceedings of the National Academy of Sciences, U.S.A., 106,* 19569–19574.

Stracke, H., Okamoto, H., & Pantev, C. (2009). Interhemispheric support during demanding auditory signal-in-noise processing. *Cerebral Cortex, 19,* 1440–1447.

Wallace, M. N., Kitzes, L. M., & Jones, E. G. (1991). Intrinsic inter- and intralaminar connections and their relationship to the tonotopic map in cat primary auditory cortex. *Experimental Brain Research, 86,* 527–544.

Woldorff, M. G., Gallen, C., Hampson, S. W., Hillyard, S. A., Pantev, C., Sobel, D., et al. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences, U.S.A., 18,* 8722–8726.

Woods, D. L., Stecker, G. C., Rinne, T., Herron, T. J., Cate, A. D., Yund, E. W., et al. (2009). Functional maps of human auditory cortex: Effects of acoustic features and attention. *PLoS ONE, 4,* e5183.