JARO
Journal of the Association for Research in Otolaryngology

Neural Coding of Sound Intensity and Loudness in the Human Auditory System

Markus Röhl[1], and Stefan Uppenkamp[1,2]

[1]*Medizinische Physik, Fakultät V, Universität Oldenburg, 26111 Oldenburg, Germany*
[2]*Forschungszentrum Neurosensorik, Universität Oldenburg, 26111 Oldenburg, Germany*

ABSTRACT

Inter-individual differences in loudness sensation of 45 young normal-hearing participants were employed to investigate how and at what stage of the auditory pathway perceived loudness, the perceptual correlate of sound intensity, is transformed into neural activation. Loudness sensation was assessed by categorical loudness scaling, a psychoacoustical scaling procedure, whereas neural activation in the auditory cortex, inferior colliculi, and medial geniculate bodies was investigated with functional magnetic resonance imaging (fMRI). We observed an almost linear increase of perceived loudness and percent signal change from baseline (PSC) in all examined stages of the upper auditory pathway. Across individuals, the slope of the underlying growth function for perceived loudness was significantly correlated with the slope of the growth function for the PSC in the auditory cortex, but not in subcortical structures. In conclusion, the fMRI correlate of neural activity in the auditory cortex as measured by the blood oxygen level-dependent effect appears to be more a linear reflection of subjective loudness sensation rather than a display of physical sound pressure level, as measured using a sound-level meter.

Keywords: fMRI, categorical loudness scaling, auditory cortex

Correspondence to: Stefan Uppenkamp · Medizinische Physik, Fakultät V · Universität Oldenburg · 26111 Oldenburg, Germany. Telephone: +49-441-7983955; fax: +49-441-7983902; email: stefan. uppenkamp@uni-oldenburg.de

INTRODUCTION

According to a classical definition, the term loudness describes the "magnitude of an auditory sensation" (Fletcher and Munson 1933). It is the perceptual correlate of sound intensity but also depends on a number of other acoustical variables, such as frequency, spectral bandwidth, stimulus duration, temporal fluctuations, or monaural vs. binaural stimulus presentation (Fletcher and Munson 1933; Zhang and Zeng 1997; Grimm et al. 2002; Verhey and Uhlemann 2008; Epstein and Florentine 2009). Also non-auditory factors like context effects and personality traits like anxiety can affect loudness (Stephens 1970; Algom and Marks 1990; Gabriel et al. 1997; Menzel et al. 2008). Loudness plays, in particular, an important role in hearing impairment, associated with sound distortion that accompanies age as presbyacusis for many of us. While the transformation of mechanical vibrations into the neuronal code at the level of the cochlea is well understood, less is known about the transformation of the neural code into perceptual measures in the upper auditory pathway, e.g., loudness perception. The aim of this study is, therefore, to provide a detailed analysis of the interrelation of physical sound intensity, perceived loudness, and the corresponding responses in functional MRI as an indirect measure of neural activity in the human central auditory system based on the blood oxygen level-dependent (BOLD) effect.

Several neuroimaging studies on the neural coding of sound intensity in the human auditory system have commonly shown that neural activation increases as a function of sound intensity in auditory areas (Hegerl et al. 1994; Jäncke et al. 1998; Mohr et al. 1999; Hall et al. 2001, Bilecen et al. 2002; Brechmann et al. 2002;

Gutschalk et al. 2002; Hart et al. 2002, 2003; Lasota et al. 2003; Mulert et al. 2005; Sigalovsky and Melcher 2006; Langers et al. 2007; Ernst et al. 2008; Röhl and Uppenkamp 2010; Röhl et al. 2011). There is still some dispute, however, over the precise location of intensity coding, e.g., whether it is more related to primary areas like the Heschl's gyrus as described by Hart et al. (2002) or more related to secondary auditory areas like the Planum Temporale as described by Gutschalk et al. (2002). It is also not finally resolved whether the activation magnitude (e.g., the BOLD-signal intensity in functional magnetic resonance imaging (fMRI) studies) grows linearly (Hall et al. 2001; Langers et al. 2007) or nonlinearly (Hart et al. 2002) as a function of sound pressure level.

There are various reasons that can explain apparent discrepancies across studies, like differences between the applied neuroimaging techniques, differences in the data analysis strategies (Kriegeskorte et al. 2009), the spectral characteristics and dynamic range of acoustic stimuli, the limited number of participants per study (Genovese et al. 1997; Thirion et al. 2007), their age (Hegerl et al. 1994), their attention to acoustic stimuli (Jäncke et al. 1999), their musical experience and musical ability (Schneider et al. 2005), their personality traits (Juckel et al. 1995; Röhl and Uppenkamp 2010), their emotional states (Quirk et al. 1997), or their neurotransmission (Juckel et al. 2008). The effect of these factors needs to be considered to reveal the neuronal underpinning of loudness sensation. At the same time, the respective interrelations of the corresponding neural activation with sound intensity on one hand and with the perceptual measure of loudness on the other hand need to be disentangled, to allow for a reliable identification of that stage in the auditory pathway where the transformation into the perceptual correlate of intensity is completed.

One very effective approach to allow for a disentanglement of loudness and intensity was pursued by Langers et al. (2007). They compared neural activation of a group of normal-hearing and hearing-impaired listeners that differed in their intensity–loudness relationship. They observed that the growth of activation with sound intensity did significantly differ between both subject groups, but not as a function of loudness level. From these findings, they concluded that fMRI activation at the level of the auditory cortex (AC) is more closely related to loudness rather than to sound intensity. However, the reported findings in that study might also be related to other factors, e.g., functional neuroplasticity in hearing-impaired subjects (Bilecen et al. 2000) or age-related differences in neural activation (Hegerl et al. 1994), as the mean age of both investigated groups differed by about three decades. The study of Langers et al. gives also no indication whether or not loudness coding is already completed at lower stages of the auditory pathway, e.g., inferior colliculi (IC). Inter-

individual differences in loudness sensations might also be explained by differences in the compression at the level of the basilar membrane only.

To avoid possible confounding effects, we examined a very homogenous group of normal-hearing listeners with auditory fMRI. We employed the variability in loudness sensation across listeners for a broadband pink-noise stimulus to disentangle the effects of sound intensity and perceived loudness on the fMRI correlate of neural activation. Inter-individual differences in loudness sensation, e.g., what is the sound pressure level that makes a participant to judge a stimulus as "loud," were measured by categorical loudness scaling (Heller 1985).

METHODS

The full experiment for each listener consisted of two separate appointments. In both sessions, hearing thresholds were determined and a categorical loudness scaling procedure was performed. The first session took place in the sound booth. It was also used for standard audiometry. The second session took place in the MRI scanner room. During this session, neuroimaging was performed before the psychoacoustic experiment. The main purpose of the experiment was to capture inter-individual differences in loudness perception to be compared with sound-induced neural activation in different stages of the auditory pathway.

Participants

Forty-five participants were recruited through advertisements placed on the notice board at the University of Oldenburg. Inclusion criteria were the following: male, age between 18 and 30 years, and normal hearing (i.e., audiometric thresholds better than 20 dB HL for all frequencies except 8 kHz, where up to 35 dB HL were accepted). Six subjects had hearing thresholds higher than 20 dB HL. Exclusion criteria were psychiatric or neurological disorders and any contraindication for MRI. All volunteers gave written informed consent to the study, which was approved by the ethics committee of the University of Oldenburg.

Acoustic stimuli

Continuous, unmodulated, and interaurally uncorrelated pink noise with a duration of 5 s was used as an acoustic stimulus throughout this study, to avoid any confounding effects due to additional temporal and spectral features. For the fMRI recording, all stimuli were played binaurally via an MR-compatible, dynamic headphone system (MR confon GmbH, Magdeburg, Germany) at a sampling rate of 44.1 kHz. The helium pump of the MRI scanner was switched off during the acquisition of the BOLD images.

The headphones were calibrated with a fiber-optic microphone (Sennheiser GmbH & Co. KG, Wedemark, Germany) and a custom-made, MR-compatible acoustic coupler that conforms to the IEC 60318-3:1998 standard. This coupler was manufactured by the mechanical workshop of our university. The accuracy of calibration of the MRI sound delivery system was within 2–3 dB. For the psychoacoustical experiments in the sound booth, the stimuli were presented via headphones Sennheiser HDA 200. They were calibrated with a condenser microphone (model 4134, Brüel & Kjær GmbH) and a coupler (artificial ear model 4153, Brüel & Kjær GmbH).

Determination of hearing thresholds

The thresholds of hearing were determined by a three-alternative forced choice procedure. Each of the three intervals was marked acoustically by the corresponding number softly spoken by a female voice. In the MRI scanner room, the participants gave their responses verbally, in the silent booth by pressing a button on a keyboard.

Categorical loudness scaling

Categorical loudness scaling is a psychoacoustic measurement procedure to capture individual subjective loudness perception (Pascoe 1978; Heller 1985). During the procedure, subjects gave their rating on a response scale with 11 response alternatives. The response scale included five named loudness categories, "very soft–soft–medium–loud–very loud"; four numbered intermediate response alternatives; and two named limiting categories, "inaudible" and "too loud." These categories were transformed into numbers from 0 categorical units (cu) to 50 cu in steps of 5 cu. The procedure is based on the reasonable assumption of an indigenous metric for each listener (Heller 1985), with a linear relationship between the perceived loudness and this scale, i.e., equal spaces between the different categories along the whole dynamic range. Acoustic stimuli were presented in random order, where the level of two successive presentations always differed by at least 10 dB and less than 50 dB. The lowest presentation level was 0 dB SPL; the highest presentation level was 3 dB above the individual uncomfortable loudness level for a 1-kHz sinusoid. In most cases, this maximum presentation level for the pink noise was about 100 dB SPL. The range of levels was covered equally spaced at a minimum step size of 2 dB. In the sound booth, subjects gave their ratings by a simple mouse click on a response scale displayed on a computer screen. In the MRI environment, a paper version of this scale was attached to the MRI bore by a sticky tape, at a distance of 20 cm in front of the subject's eyes. The subjects gave their ratings verbally, which were monitored via a microphone from the control room by

the investigator. In order to make the overall acoustic situation during the psychoacoustic measurement in the MRI scanner as similar as possible to the real fMRI recording, the timing of the former was adopted to the latter. Therefore, each pink-noise stimulus was presented upon completion of an imaging sequence (21 EPI slices lasting for 2.7 s) which itself produced a lot of stray noise caused by gradient coil switching. One presentation cycle lasted for 15 s which included 7 s of silence for the subjects to tell aloud their rating before the next cycle began. This was different from the procedure inside the sound booth where the response by the mouse triggered the next presentation with a delay of 500 ms. One experimental session took about 50 presentation cycles which lasted about 13 min in the scanner and about 5 min in the sound booth.

Functional magnetic resonance imaging

The measurement was performed on a MRI scanner SIEMENS Sonata 1.5 T, equipped with a standard single channel head coil. Twenty-one transversal slices of 3.9 mm thickness angled away from the eyes, centered at the posterior commissure, were acquired covering the superior temporal lobes, including the primary auditory cortex. Functional MRI using echo planar imaging (EPI) sequences (time of echo 63 ms; volume acquisition time 2.7 s; flip angle 90°; matrix size 64×64; field of view 192× 192 mm^2, interleaved acquisition mode, 1/8 overlap) was performed in a sparse imaging paradigm (Edmister et al. 1999; Hall et al. 1999) at a time of repeat (TR) of 7.7 s. Images were acquired at the end of a 5-s stimulus interval, so that the presentation of the auditory stimuli was not masked by the scanner noise.

The employed stimulus paradigm of clustered volume acquisition and sparse imaging at this comparatively short TR of 7.7 s is a compromise between efficiency in time and appropriate separation of scanner noise and stimulus. It is expected that the BOLD response has come to some sort of saturation after the continuous 5-s stimulus. Activation in response to the noise caused by the EPI scans will build up and decay, on a time scale within 8 or 10 s (cf. Hall et al. 1999). It might therefore overlap a little with the target bit of the response. In previous studies, it was demonstrated that the current choice of timing parameters is nevertheless still successful for a reasonable separation of both components of the response (Ernst et al. 2008, 2010). Therefore, to make sufficient use of scanner time with many repetitions per condition and subject, the tight timing was kept for the current study.

The pink-noise stimulus was presented at sound pressure levels from 20 dB SPL in increments of 10 dB up to 80 or 90 dB SPL in most cases, depending on the individual loudness rating of "too loud" which was avoided for all listeners. Five of the participants tolerated

a maximum level of 100 dB SPL. One experimental session was split into four runs interrupted by approximately two minutes of rest. Acoustic stimuli were presented in random order, where the level of two successive presentations always differed by at least 10 dB and less than 50 dB to avoid artifacts caused by subject's head movement due to sudden large leaps. Each level was presented 12 times during a run, including the silence condition which was the baseline condition of the fMRI experiment.

A simple detection task was employed to ensure that the participants were attending to the acoustic stimuli. Deviants in this task differed from standard stimuli in a way that the sound intensity was not kept constant but after 2 s decreased by 10 dB for a short period of 330 ms. The average number of deviants was five; the maximum number was 10. They were randomly distributed over each run. Participants had to count the number of deviants within each run. From pilot experiments, it was known that the performance in this deviant detection task was always close to 100% correct and therefore too good to serve as a measure that can differentiate between attentional states. To assess subject's attention, we used an additional rating scheme consisting of five response alternatives: "close to sleeping–less attentive–attentive–very attentive–wide awake." These categories were transformed into numbers from one to five. The participants were asked after each of the four runs for the number of counted deviants and their attention rating. A T1-weighted structural image was also acquired after the fMRI session to obtain individual anatomical landmarks.

Data analysis

Data analysis was performed with SPM5 (Wellcome Department of Cognitive Neurology). A set of standard preprocessing steps including realignment, normalization to the Montreal Neurological Institute (MNI) standard coordinate space (Brett et al. 2002b), interpolation to $2 \times 2 \times 2$ mm^3, and smoothing with 6 mm FWHM Gaussian filter were carried out. FMRI time series were high-pass-filtered with a cutoff frequency of 1/128 s. A first-order autoregressive model was used to handle the physiological and non-physiological low-frequency noise characteristics of fMRI time series. The fMRI data were analyzed with a general linear model (GLM) in which the number of regressors corresponded to the number of different stimulation levels used, including the silence condition. Since the fraction of deviants with respect to the number of scans per run was very small, between 4% and 5%, deviants were handled as normal sounds in the data analysis. Contrast images were calculated for the difference between the sound and the silence conditions for each presentation level (20, 30, 40... dB SPL) and each

participant. Neural activation in the auditory cortex was determined using one-sample t tests with a threshold for significance of $p < 0.05$, corrected for multiple comparisons with the false discovery rate method (FDR). Based on these first-level images, a region-of-interest (ROI) analysis was performed to extract neural activation as a function of sound intensity by two parameters, the volume of activation and the mean percent signal change. The volume of activation was calculated as the sum of significant (suprathreshold) voxels from the activation pattern within rectangular boxes surrounding the left and right auditory cortex. The MNI coordinates of the ROI surrounding the AC were $|x| > 20$ mm, $|y| < 50$ mm, and -20 mm $< z < 50$ mm. These boundaries were defined with respect to the observed variability of sound-induced activation across listeners. The main motivation was a rough separation of activation in the left AC, right AC, and brainstem structures. This comparatively generous definition of the ROI ensures that no functional activation in auditory areas is missed, even if it does not necessarily overlap with textbook anatomical landmarks. The percent signal change was calculated by determining the difference between the mean BOLD-signal intensity of the preprocessed EPI images of consecutive sound and silence conditions within an auditory mask. Due to the restriction of consecutive presentations, this procedure acts like a high-pass filter with a mean cutoff frequency at 1/131 Hz. The auditory masks consisted of those voxels for which any sound condition showed a significant activation against silence. These masks were created for each participant. Neural activation in the inferior colliculi and the medial geniculate bodies (MGB) was investigated in a similar way but using a small volume correction procedure instead of the false discovery rate method to adjust the p value. The threshold for significance was set to $p < 0.01$ as in Sigalovsky and Melcher (2006). The small volume was defined by spheres of 5 mm radius as in Griffiths et al. (2001). The center of the spheres was determined by visual inspection of the anatomical images and functional activation, respectively. For the IC, the mean center was found to be at $x = \pm 4$ mm, $y = -36$ mm, and $z = -12$ mm. For the MGB, the mean center was found to be at $x = \pm 12$ mm, $y = -26$ mm, and $z = -8$ mm. These coordinates correspond well to Griffiths et al. (2001) and von Kriegstein et al. (2008).

Hypothesis

As the loudness has been previously associated with the magnitude of activation (Langers et al. 2007) and not with the volume of activation (Hart et al. 2003), fMRI activation is analyzed in the present study with respect to the percentage change of the BOLD signal relative to the silence condition. We hypothesize that

in cortical auditory areas, the variability of the PSC for a fixed sound pressure level between subjects is directly linked to differences in their loudness sensation. If this hypothesis would not hold, the variance would probably be caused by other factors, e.g., differences in the hemodynamic efficiency across participants (Logothetis and Wandell 2004). We assume that the coding of loudness is not completed before AC, since also non-auditory factors like context effects and personality traits like anxiety can act on perceived loudness (Stephens 1970; Algom and Marks 1990; Gabriel et al. 1997; Menzel et al. 2008). Therefore, we hypothesize that inter-individual differences in loudness sensation cannot explain a significant amount of the variance in the sound-induced activation for subcortical stages of the auditory pathway such as MGB or IC.

RESULTS

The results are described in five steps. First, psychoacoustical findings from categorical loudness scaling in the two examined hearing environments, sound booth and MRI, are presented. Then, the general response characteristics in relation to sound intensity are analyzed. The third part of this section focuses on the question if and at what stage of the upper auditory pathway perceived loudness is represented in the brain activation as measured with fMRI. In the fourth step, the role of participants' attention on sound-induced activation is taken into consideration. In the last step, the quality of the PSC measurements in subcortical and cortical structures has been assessed to verify the major conclusions of this study.

Categorical loudness scaling

Mean hearing thresholds (with standard deviations across listeners) for the 5-s-long pink-noise stimulus in the silent booth were at 7.7±3.8 dB SPL; in the MRI scanner, they were at 15.4±2.6 dB SPL. The difference was significant ($p<0.001$). A typical result of loudness scaling for one participant and a fit of the data with a polynomial function of third degree, e.g., the loudness curve, is presented in Figure 1. This fit is later used to transform sound pressure levels to individual categorical loudness judgments. It can be seen for this listener that the relation between categorical loudness units and sound level is almost linear. A similar result was found for most participants in this study regarding both scalings, in the MRI scanner room (coefficient of determination $R^2=0.93\pm0.02$) and in the silent booth ($R^2=0.90\pm0.04$). Linear regression of the scaling results in the silent booth yielded an offset of -3.6 ± 1.8 dB SPL and a slope of 4.6±0.5 cu/10 dB SPL.
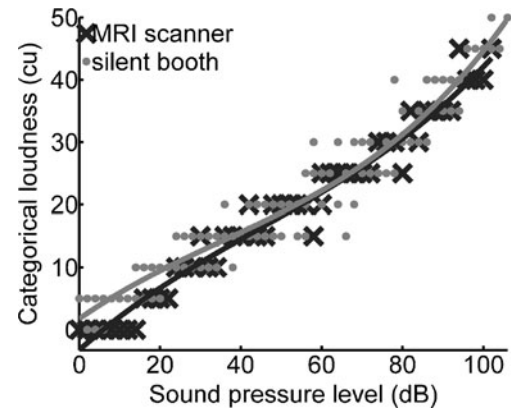

FIG. 1. Typical results of loudness scaling for two different hearing environments for one participant and its polynomial approximation of third degree. During the psychophysical measurements inside the MRI Scanner room, a 2.7-s gradient noise burst (TR=15 s) preceded every 5-s pink-noise stimulus.

Scaling results in the MRI scanning room yielded an offset of -4.4 ± 1.6 dB SPL and a slope of 4.7±0.5 cu/10 dB SPL (group mean averages). The difference in the offset of 0.8 dB SPL was small but significant ($p<0.01$). Using the polynomial relationship for each participant, the three categorical loudness ratings, "soft" (15 cu), "medium" (25 cu), and "loud" (35 cu), corresponded to the sound pressure levels of 46±5, 68±6, and 85±6 dB SPL inside the MRI scanner. The dynamic range that was associated with a specific loudness rating increased with categorical loudness from 21 dB for "soft" to 29 dB for "loud."

The use of higher-order polynomials for a fit of the loudness curve (as obtained in the MRI scanner) did significantly improve the coefficient of determination as compared to the linear regression, both for a second-order polynomial ($R^2=0.95\pm0.02$) and for a third-order polynomial ($R^2=0.96\pm0.02$). However, the nonlinear portion was very small with respect to the presentation levels applied during the fMRI session (up to the individual rating of "very loud"). The mean correlation between presentation level and perceived (categorical) loudness was almost one ($r=0.99\pm0.01$).

General response characteristics of the auditory system in relation to sound intensity

Sound-induced activation could be detected in the auditory cortex for at least one sound level for all of the 45 participants, in the IC for 43, and in the MGB for 42. The probability map of the auditory masks covering those voxels for which any sound condition showed a significant activation against silence and that are used to calculate the PSC are depicted in Figure 2A, B. Their mean volume summed over both hemispheres was 22.1±17.3 cm³ in the AC, 0.6±0.3 cm³ in the IC, and 0.4±0.3 cm³ in the MGB. Activation in the AC was observed
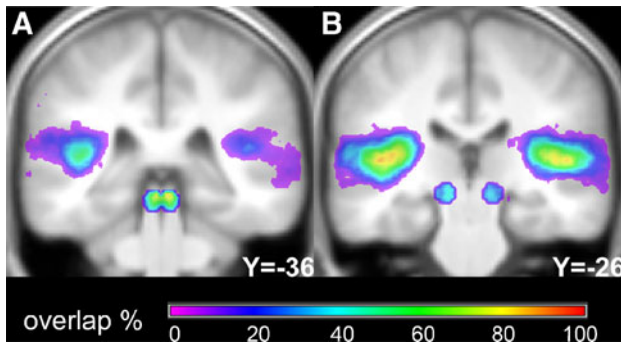
FIG. 2.　A, B Probability map of the auditory masks to derive the percent signal change. An individual auditory mask consisted of those voxels that showed a significant ($p<0.05$, FDR) sound-induced activation for any sound pressure level.

across the supratemporal plane in HG and adjacent areas, e.g., Planum Temporale. The smallest sound pressure level for which a significant activation could be observed was 20 dB SPL both for the AC (three subjects), the IC (five subjects), and the MGB (three subjects). In the AC and the IC, more than half of the participants showed a significant activation for presentation levels from above 40 dB SPL. The MGB were activated at 40 dB SPL in a third of the participants only.

Figure 3A depicts the mean volume of activation for the whole group of participants including their respective standard deviations. The volume of activation increased exponentially with level in all investigated stages of the auditory pathway. Therefore, the decadic logarithm of the volume of activation, as illustrated in Figure 3A, grows almost linearly for the AC ($R^2=0.92$), the MGB ($R^2=0.93$), and IC ($R^2=0.87$). The same holds for the mean percent signal change from baseline, as depicted in Figure 3B, which grows almost linearly for the AC ($R^2=0.89$), as well as for the MGB ($R^2=0.87$) and IC ($R^2=0.86$). The correlation between the mean percent signal change and the mean volume of activation (averaged over the range from 20 to 80 dB SPL) was not significant in the AC ($r=0.20$, $p=0.20$) and in the IC ($r=0.15$, $p=0.35$). In the

MGB, for which neural activation was generally weaker as compared to AC and IC, there was a significant correlation ($r=0.53$, $p=3.3e{-}4$).

FMRI activation of the auditory system: does it reflect loudness or rather sound intensity?

Inter-individual differences of fMRI activation of the investigated stages of the auditory pathway, as illustrated by the error bars in Figure 3B for the activation magnitude, are now compared with inter-individual differences in loudness sensation to answer the question, whether and at what stage in the auditory pathway the transformation of sound intensity into loudness comes into play. Since both perceived loudness and neural activation, as measured by the PSC, were almost linearly related to sound pressure level, we calculated the corresponding growth rate for each participant and each examined stage. This allowed for an effective comparison between neural activation and perceived loudness. The correlation across subjects between the slopes of the PSC and the loudness function was significant for the AC ($r=0.44$, $p=0.002$), but not for the IC ($r=-0.16$, $p=0.30$) and MGB ($r=0.06$, $p=0.71$). The difference of correlation coefficients between cortex and IC ($p=0.004$, using Fisher's Z) was significant. The difference of correlation coefficients between cortex and MGB missed significance slightly ($p=0.06$). A stronger relation between loudness and percent signal change was observed when data analysis was restricted to larger BOLD-signal changes, as measured at a sound pressure level of 80 dB SPL. This was the highest presentation level within this study that was played to all listeners. The PSC showed a statistically more satisfying correlation to the individual perceived categorical loudness for the AC ($r=0.60$, $p=1.09e{-}5$). The IC ($r=0.02$, $p=0.92$) and MGB ($r=0.19$, $p=0.24$) showed no such significant correlation. The three panels of Figure 4 illustrate this relation within the auditory masks for the three examined stages of the upper auditory pathway and a sound pressure level of 80 dB SPL. In summary,
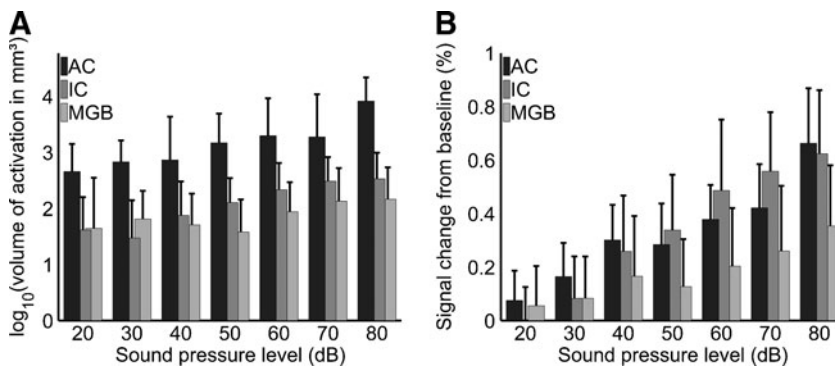
FIG. 3.　General response characteristics of fMRI parameters derived from statistical parametric maps for 45 participants. A The group's mean volume of activation within the predefined regions of interest as a function of sound pressure level on a logarithmic scale ($p<0.05$, FDR). B Group's mean percent signal change relative to baseline within individual auditory masks as a function of sound pressure level.
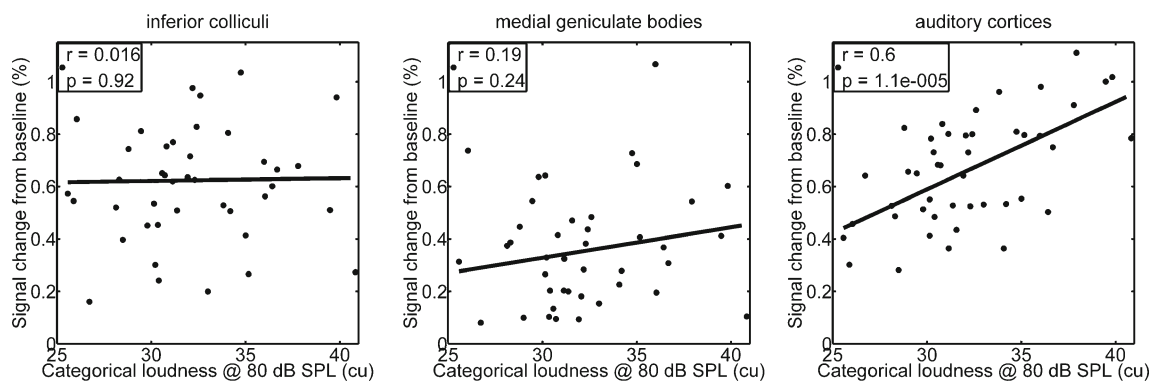
FIG. 4. Percent signal change relative to baseline plotted as a function of perceived categorical loudness for a fixed presentation level of 80 dB SPL.

fMRI activation in the AC is more related to perceived loudness than to sound intensity.

Participants' attention

Data on attention were collected in 44 out of the 45 participants. The participants' performance in the deviant detection task was very high. On average, 21 deviants were presented during the whole fMRI session. Twenty-eight out of 44 subjects were correct or mis-counted by just one deviant in the total experiment. The mean error rate for all listeners was smaller than one. Subjects' self-ratings of alertness, A, decreased slightly with scanning time from run #1 ($A=3.7\pm0.8$), to run #2 ($A=3.0\pm0.9$), and run #3 ($A=2.9\pm0.9$), but partly recovered in the last of the four runs ($A=3.4\pm0.8$). In general, subjects' self-rating of alertness was between "attentive" and "very attentive" over the whole experiment ($A=3.3\pm0.9$) in line with the good performance in the deviant detection task.

Assessment of the quality of the PSC measurements in IC and HG

Cardiac gating, i.e., synchronizing EPI acquisition with a physiological signal reflecting the cardiac cycle, like a pulse oximeter signal or electrocardiogram, has previously been used to improve the BOLD-signal quality, especially from brainstem structures (e.g., Guimaraes et al. 1998; Griffiths et al. 2001). Since a cardiac-gating procedure was not performed during this study for the detection of neural activation in subcortical areas, we employed an additional quality assessment of the IC data to verify the main findings of our study. The assessment was based on the signal-to-noise ratio (SNR) of the BOLD signal achieved within specific ROIs for each participant. The regions of interests used to assess the quality of the PSC measurements had a volume of 3.79 cm³ for HG and 1.2 cm³ for the IC, respectively. The HG ROI was taken from the MarsBaR AAL ROI package (Brett et al. 2002a) and was covered by 58±23%

by the auditory masks; the IC ROI was identical with the definition of the spheres that were used for the small volume correction as described in "Data analysis" section. The SNR, expressed on a logarithmic scale in dB, was calculated from the square of the ratio of the mean and standard deviation of the BOLD signal intensities within the ROIs. The IC data showed large differences between subjects with respect to the signal-to-noise ratio as illustrated in Figure 5. The mean SNR of the BOLD signal in IC was 13.7±7.8 dB, while in the HG it was 19.3±1.3 dB. The SNR showed a weak significant correlation with the volume of the auditory masks in HG ($r=-0.33$, $p=0.029$) and a pronounced positive correlation at the level of IC ($r=0.63$, $p=2.54e-6$). A threshold for the SNR of at least 10 dB was now introduced in a second step of the analysis of the IC data, in order to distinguish noisy from "more reliable" individual data. Following this ad hoc criterion, 33 of 45 participants were allocated to the "more reliable" data set.

Based on this group of 33 participants, the correlation analysis has been repeated. The mean SNR of the BOLD signal in HG in this subset was 19.6± 1.3 dB, while in the IC it was 18.1±2.0 dB. The quantitative results of the correlation analysis between
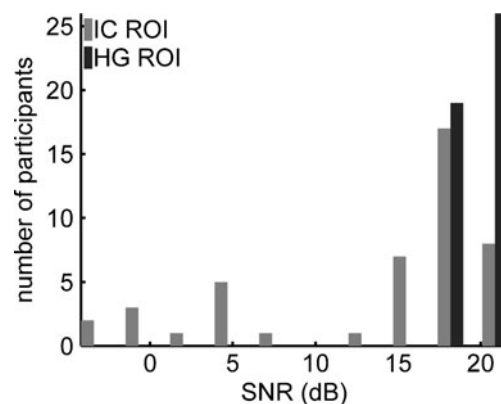
FIG. 5. Distribution of the signal-to-noise ratio within subcortical and cortical regions of interest employed for the quality assessment of the PSC measurements in IC and HG.

neural activation and perceived loudness as presented in the section "FMRI activation of the auditory system: Does it reflect loudness or rather sound intensity?" could be confirmed in this subset. In subcortical brain regions, the slopes of the loudness and the PSC as a function of sound level showed no correlation when compared across subjects. This was true for the IC ($r=0.026$, $p=0.89$) as well as for the MGB ($r=0.17$, $p=0.37$). In the AC, the positive and significant correlation reached a more statistically satisfying level ($r=0.53$, $p=0.0017$). Similar results were observed for the particular stimulus levels that were associated with the individual loudness judgment "loud."

DISCUSSION

The present study investigated the relation between sound pressure level, categorical loudness, and parameters derived from auditory fMRI activation maps using 5-s bursts of pink noise as an acoustic stimulus. The main findings of the current study are: (1) The growth of categorical loudness as a function of sound pressure level is almost linear for a 5-s unmodulated pink-noise stimulus. Scaling results obtained in the silent booth and in the MRI scanner room are very similar, apart from a difference of approximately 10 dB in hearing thresholds. (2) In the IC, MGB, and auditory cortex, the volume of activation increases exponentially, and the percent signal change from baseline (i.e., silence condition) increases almost linearly as a function of sound pressure level. (3) The BOLD response in the auditory cortex appears to be a linear reflection of subjective loudness sensation rather than a display of measured sound pressure level. (4) A similar relationship between loudness and activation was not manifested at lower stages of the central auditory pathway.

Psychoacoustics in different hearing environments

Categorical loudness as a function of sound pressure level showed an almost linear growth, which was very similar for the silent booth and the MRI environment. There was no significant difference in the slope of the loudness curves. There was a significant, although small difference in their offset: A stimulus presented in the MRI scanner required an increase of sound pressure level by 0.8 dB to achieve the same categorical loudness. Although this difference in the offset was statistically significant—and at a first glance—would be in line with the plausible assumption that participants adjust their loudness ratings to the noisier environment inside the MRI scanner, this difference was still smaller than the accuracy of the calibration of the MRI sound delivery system. Our results therefore

provide further evidence for the usefulness of the sparse temporal sampling paradigm in auditory fMRI from a psychoacoustician's perspective (Edmister et al. 1999; Hall et al. 1999). This finding suggests that the results on loudness coding from this study can largely be generalized, since the sound booth (very quiet) and the fMRI sound environment (comparatively noisy) may be considered as two opposite poles of the daily sound environment with respect to noise.

General response characteristics of the human auditory system in relation to sound intensity

In the present study, we found strong evidence that both investigated fMRI parameters, the volume of activation and the percent signal change from baseline, show a monotonic growth with sound pressure level. This is in line with previous studies on sound intensity (Hegerl et al. 1994; Jäncke et al. 1998; Mohr et al. 1999; Hall et al. 2001; Bilecen et al. 2002; Brechmann et al. 2002; Gutschalk et al. 2002; Hart et al. 2002, 2003; Lasota et al. 2003; Mulert et al. 2005; Sigalovsky and Melcher 2006; Langers et al. 2007; Ernst et al. 2008). Probably due to the comparatively large number of participants in our study and the large examined range of levels and corresponding perceived loudness, it was possible to specify that the volume of activation increases exponentially and the percent signal change from baseline almost linearly with sound pressure level in all examined stages of the auditory pathway. The latter is in line with the study by Langers et al. (2007). They described the relation between activation magnitude and sound intensity as predominantly linear for normal-hearing listeners. However, it is not in line with the study by Hart et al. (2002). They had reported a significant quadratic component. Possible reasons for this difference might be caused by the effect of additional factors as listed in the introduction of this paper, e.g., specific features related to the particular type of stimuli, like spectral content or modulation.

Sound-induced activation could reliably be detected along all investigated stages of the central auditory pathway for the great majority of participants. The detection of neural activation in brainstem or midbrain areas could probably be systematically improved by applying cardiac-gating procedures with prospective or retrospective methods (Glover et al. 2000; Griffiths et al. 2001; Sigalovsky and Melcher 2006; Harvey et al. 2008). However, as the results suggest, this is not generally required, at least not for the inferior colliculus, if the quality of the PSC measurement is being assessed properly. This was also demonstrated by other studies using the same 1.5-T MRI scanner (Ernst et al. 2008) or another 1.5-T model (Langers et al. 2005). The repetition time of 7.7 s was shorter as compared to other fMRI

studies, e.g., 10 s as in Langers et al. (2007). A silent fMRI sequence with a repetition time of 20 s as proposed by Yang et al. (2000) had been tested as an alternative paradigm during a pilot experiment. With this procedure, a baseline scan would usually need to be acquired after each active condition for statistical comparison. However, given a maximum comfortable scanning time of about one hour for one participant, the efficiency of the "silent procedure" appeared to be less suitable, since sound-induced activation was by far lower.

Transformation of sound intensity into perceived loudness

In this study, we demonstrated a significant correlation of PSC in the auditory cortex with the perceived loudness for a fixed presentation level. Since a similar relationship could not be detected for the PSC in the IC or MGB in our data, we suggest the interpretation that the perceptual measure of loudness is only completed at cortical level. This should, however, not be interpreted as indication that IC in brainstem and MGB in the thalamus would not be involved in loudness coding. Much of the preprocessing relevant for the transformation of sound intensity into the perceptual measure of loudness is related to stimulus compression, which is a consequence of cochlear nonlinearity and therefore should already be present in the ascending auditory pathway. It is possible that a subset of IC neurons might represent loudness in one way or the other. Possible coding strategies might be the overall neural activation for a subset of loudness-sensitive neurons, or even a loudness tuning of specific sub-populations with a comparatively small dynamic range, each of these sets representing a limited loudness range only. If loudness was already represented at the level of IC, it was still not visible in our fMRI activation data, possibly due to the limited resolution of a 1.5-T scanner or the signal variability related to the cardiac cycle. Therefore, the straightforward interpretation of our results would be that a full representation of perceived loudness is only completed at the level of the auditory cortex.

Our findings about loudness coding in AC are largely in line with the results of Langers et al. (2007). One additional result is that the AC is the first stage of the central auditory pathway for which a representation of loudness coding can be observed in the fMRI activation map, whereas at the level of the IC, the BOLD signal as a correlate of neural activation might be more a reflection of physical sound intensity. Our findings are largely free of confounds caused by differences in the age of the participants or by neuroplasticity as a result of hearing impairment, since all our participants were from a very homogenous group of healthy young normal-hearing listeners. However, the comparatively weak activation observed in the MGB in this study leaves the question open whether loudness coding can in part be related to neural processes in the thalamus. This needs to be employed in more detail in future studies as there are many efferent connections between AC and MGB suggesting a possible top-down regulation of loudness-associated neural activity, e.g., during loudness adaptation. In addition, the variance in the PSC data that could not be explained by individual differences in loudness sensation may be a possible source of further information about neural signal processing in the auditory pathway. This should be analyzed in more detail in future studies.

Relation between percent signal change and volume of activation

In contrast to the relation between PSC and loudness, there was no initial hypothesis about a possible relation between the volume of activation and inter-individual differences in loudness sensation in this study. Hart et al. (2003) had demonstrated that neural activity in the primary auditory cortex as measured by the sound-level-dependent growth in the activated volume cannot be considered as a simple reflection of the perceptual representation of loudness. Previous fMRI and EEG studies also indicated that the volume of activation is related rather to a different, non-auditory factor. Juckel et al. (1995) showed with EEG that the intensity dependence of the mean current source density of auditory-evoked potentials in the PAC is positively correlated to the personality trait "novelty seeking." Mulert et al. (2005) showed in a corresponding fMRI experiment that changes in the mean current source density are related to changes in the extent of fMRI activation within a region of interest covering the primary auditory cortex. They found no correlation between the mean current source density and the activation magnitude (PSC). Therefore, the volume of activation was expected to be related to the personality trait "novelty seeking." This was indeed demonstrated in a previous study (Röhl and Uppenkamp 2010). In our group of participants, we observed no significant correlation between the slope of the loudness curve and the personality trait "novelty seeking" ($r=-0.22$, $p=0.14$). According to the findings of Juckel et al. (1995) and Mulert et al. (2005) and to the findings of our previous study (Röhl and Uppenkamp 2010), a significant correlation between the mean volume of activation and the mean PSC could not been observed.

However, since both fMRI parameters, activation volume and PSC, are associated with a stronger neural activity and, therefore, should be related to each other, one may ask for a reason for this lack of correlation. We assume that the lack of correlation is

related to the exact choice of the acoustic stimulus in this study, pink noise. The strong increase of the volume of activation with sound intensity probably reflects the uniform excitation of the basilar membrane across all frequency channels by the broadband pink-noise stimulus. Assuming that this uniform pattern of activation is maintained up to the auditory cortex, the critical value for the percent signal change, for which the difference between a certain sound condition and the baseline becomes significant, should be similar for all of the frequency channels—and corresponding voxels. Once the BOLD-signal change is above the threshold for significance, nearly all auditory cortex voxels in the statistical parametric map will suddenly be classified as "activated" by the GLM. Therefore, the exponential increase of the activated volume with sound pressure level, as observed in the present study, essentially reflects a step function which is smoothed by the inter-individual differences in loudness perception, by the spatial low-pass filter in the preprocessing of the data and by some statistical noise. This fundamental difference between the growth functions of the volume of activation and the percent signal change (almost a linear function) is the most likely reason for the lack of a significant correlation between both parameters in the AC and in the IC. However, this may only hold for stimuli that cause an almost uniform excitation of the basilar membrane. Other stimuli may result in more similar growth functions for the volume of activation and the percent signal change, with a stronger correlation between both parameters. In those cases, it might be more difficult to employ auditory fMRI to track subjective loudness with the focus on the PSC. It might be also more difficult to employ auditory fMRI to track the personality trait impulsivity or "novelty seeking" as suggested by Juckel et al. (1995) or by our previous study with the focus on the volume of activation (Röhl and Uppenkamp 2010). However, the list of possible non-auditory factors as given in the introduction of this paper is very long, underlining that even with fMRI, such endeavors require great care to be effective.

The normal range of individual differences in loudness sensation

In our study, a range of sound pressure levels of almost 30 dB was rated as similarly loud across the normal-hearing listeners. This range of individual differences in loudness sensation within the examined subject group is by far larger than would be expected for the results from other controlled psychoacoustic experiments, like, e.g., the detection of a tone in a masking noise. None of the variables we registered in our study (e.g., hearing thresholds, age, personality traits, musical ability, and experience) seemed to be able to explain a significant amount of the observed

variance in loudness sensation (data not shown here). On the other hand, it is very unlikely that this variance is only a result of methodological issues of categorical loudness scaling, since these individual differences were shown to be linked to individual differences in neural activation of the AC, as reflected by the BOLD signal in auditory fMRI. Therefore, the origin of these individual differences in loudness judgments deserves further attention.

REFERENCES

ALGOM D, MARKS LE (1990) Range and regression, loudness scales, and loudness processing: toward a context-bound psychophysics. J Exp Psychol Hum Percept Perform 16:706–727

BILECEN D, SEIFRITZ E, RADU EW, SCHMID N, WETZEL S, PROBST R, SCHEFFLER K (2000) Cortical reorganization after acute unilateral hearing loss traced by fMRI. Neurology 54:765–767

BILECEN D, SEIFRITZ E, SCHEFFLER K, HENNING J, SCHULTE AC (2002) Amplitopicity of the human auditory cortex: an fMRI study. NeuroImage 17:710–718

BRECHMANN A, BAUMGART F, SCHEICH H (2002) Sound-level-dependent representation of frequency modulations in human auditory cortex: a low-noise fMRI study. J Neurophysiol 87:423–433

BRETT M, ANTON JL, VALABREGUE R, POLINE JB (2002A) Region of interest analysis using an SPM toolbox. In: 8th International Conference on Functional Mapping of the Human Brain (Sendai, Japan). Neuroimage 16(2): abstract 497

BRETT M, JOHNSRUDE IS, OWEN AM (2002B) The problem of functional localization in the human brain. Nature Reviews Neurosci 3:243–249

EDMISTER WB, TALAVAGE TM, LEDDEN PJ, WEISSKOFF RM (1999) Improved auditory cortex imaging using clustered volume acquisitions. Hum Brain Mapp 7:89–97

EPSTEIN M, FLORENTINE M (2009) Binaural loudness summation for speech and tones presented via earphones and loudspeakers. Ear and Hearing 30:234–237

ERNST SMA, VERHEY JL, UPPENKAMP S (2008) Spatial dissociation of changes of level and signal-to-noise ratio in auditory cortex for tones in noise. NeuroImage 43:321–328

ERNST SMA, UPPENKAMP S, VERHEY JL (2010) Cortical representation of release from auditory masking. NeuroImage 49:835–842

FLETCHER H, MUNSON WM (1933) Loudness, its definition, measurement and calculation. J Acoust Soc Am 5:82–108

GABRIEL B, KOLLMEIER B, MELLERT V (1997) Influence of individual listener, measurement room and choice of test-tone levels on the shape of equal-loudness level contours. Acta Acustica united with Acustica 83:670–683

GENOVESE CR, NOLL DC, EDDY WF (1997) Estimating test–retest reliability in functional MR imaging. 1: Statistical methodology. Magn Reson Med 38:497–507

GLOVER GH, LI TQ, RESS D (2000) Image-based method for retrospective correction of physiological motion effects in fMRI: RETRO-ICOR. Magn Reson Med 44:162–167

Griffiths TD, Uppenkamp S, Johnsrude I, Josephs O, Patterson RD (2001) Encoding of the temporal regularity of sound in the human brainstem. Nat Neurosci 4:633–637

Grimm G, Hohmann V, Verhey JL (2002) Loudness of fluctuating sounds. Acta Acustica united with Acustica 88:359–368

Guimaraes AR, Melcher JR, Baker JR, Ledden P, Rosen BR, Kiang NYS, Fullerton BC, Weisskoff RM (1998) Imaging subcortical auditory activity in humans. Hum Brain Mapp 6:33–41

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

Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW (1999) "Sparse" temporal sampling in auditory fMRI. Hum Brain Mapp 7:213–223

Hall DA, Haggard MP, Summerfield AQ, Akeroyd MA, Palmer AR, Bowtell RW (2001) Functional magnetic resonance imaging measurements of sound-level encoding in the absence of background scanner noise. J Acoust Soc Am 109:1559–1570

Hart HC, Palmer AR, Hall DA (2002) Heschl's gyrus is more sensitive to tone level than non-primary auditory cortex. Hear Res 171:177–190

Hart HC, Hall DA, Palmer AR (2003) The sound-level-dependent growth in the extent of fMRI activation in Heschl's gyrus is different for low- and high-frequency tones. Hear Res 179:104–112

Harvey AK, Pattinson KT, Brooks JCW, Mayhew SD, Jenkinson M, Wise RG (2008) Brainstem functional magnetic resonance imaging: disentangling signal from physiological noise. J Magn Reson Imaging 28:1337–1344

Hegerl U, Gallinat J, Mrowinski D (1994) Intensity dependence of auditory evoked dipole source activity. Int J Psychophysiol 17:1–13

Heller O (1985) Hörfeldaudiometrie mit dem Verfahren der Kategorienunterteilung. Psychologische Beiträge 27:478–493

Jäncke L, Shah NJ, Posse S, Grosse-Ryuken M, Müller-Gärtner HW (1998) Intensity coding of auditory stimuli: an fMRI study. Neuropsychologia 36:875–883

Jäncke L, Mirzazade S, Shah NJ (1999) Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. Neurosci Lett 266:125–128

Juckel G, Schmidt LG, Rommelspacher H, Hegerl U (1995) The tridimensional personality questionnaire and the intensity dependence of auditory evoked dipole source activity. Biol Psychiatry 37:311–317

Juckel G, Hegerl U, Giegling I, Mavrogiorgou P, Wutzler A, Schuhmacher C, Uhl I, Brüne M, Mulert C, Pogarell O, Rujescu D (2008) Association of 5-HT1B receptor polymorphisms with the loudness dependence of auditory evoked potentials in a community-based sample of healthy volunteers. Am J Med Genet Part B, Neuropsychiatric Genetics 147B:454–458

Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci 12:535–540

Langers DRM, van Dijk P, Backes WH (2005) Lateralization, connectivity and plasticity in the human central auditory system. NeuroImage 28:490–499

Langers DRM, van Dijk P, Schoemaker ES, Backes WH (2007) fMRI activation in relation to sound intensity and loudness. NeuroImage 35:709–718

Lasota KJ, Ulmer JL, Firszt JB, Biswal BB, Daniels DL, Prost RW (2003) Intensity-dependent activation of the primary auditory cortex in functional magnetic resonance imaging. J Comput Assist Tomogr 27:213–218

Logothetis NK, Wandell BA (2004) Interpreting the BOLD signal. Ann Rev Physiol 66:735–769

Menzel D, Fastl H, Graf R, Hellbrück J (2008) Influence of vehicle color on loudness judgments. J Acoust Soc Am 123:2477–2479

Mohr CM, King WM, Freeman AJ, Briggs RW, Leonard CM (1999) Influence of speech stimuli intensity on the activation of auditory cortex investigated with functional magnetic resonance imaging. J Acoust Soc Am 105:2738–2745

Mulert C, Jäger L, Propp S, Karch S, Störmann S, Pogarell O, Möller HJ, Juckel G, Hegerl U (2005) Sound level dependence of the primary auditory cortex: simultaneous measurement with 61-channel EEG and fMRI. NeuroImage 28:49–58

Pascoe DP (1978) An approach to hearing aid selection. Hearing Instruments 29:12–16

Quirk GJ, Armony JL, LeDoux JE (1997) Fear conditioning enhances different temporal components of tone-evoked spike trains in auditory cortex and lateral amygdala. Neuron 19:613–624

Röhl M, Uppenkamp S (2010) An auditory fMRI correlate of impulsivity. Psychiatry Res Neuroim 181:145–150

Röhl M, Kollmeier B, Uppenkamp S (2011) Spectral loudness summation takes place in the primary auditory cortex. Hum Brain Mapp 32:1483–1496

Schneider P, Sluming V, Roberts N, Scherg M, Goebel R, Specht HJ, Dosch HG, Bleeck S, Stippich C, Rupp A (2005) Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. Nat Neurosci 8:1241–1247

Sigalovsky IS, Melcher JR (2006) Effects of sound level on fMRI activation in human brainstem, thalamic and cortical centers. Hear Res 215:67–76

Stephens SD (1970) Personality and the slope of loudness function. Q J Exp Psychol 22:9–13

Thirion B, Pinel P, Meriaux S, Roche A, Dehaene S, Poline JB (2007) Analysis of a large fMRI cohort: statistical and methodological issues for group analyses. NeuroImage 35:105–120

Verhey JL, Uhlemann M (2008) Spectral loudness summation for sequences of short noise bursts. J Acoust Soc Am 123:925–934

von Kriegstein K, Patterson RD, Griffiths TD (2008) Task-dependent modulation of medial geniculate body is behaviorally relevant for speech recognition. Curr Biol 18:1855–1859

Yang Y, Engelien A, Engelien W, Xu S, Stern E, Silbersweig DA (2000) A silent event-related functional MRI technique for brain activation studies without interference of scanner acoustic noise. Magn Reson Med 43:185–190

Zhang CY, Zeng FG (1997) Loudness of dynamic stimuli in acoustic and electric hearing. J Acoust Soc Am 102:2925–2934