Auditory cortex activation by infrasonic and low-frequency sound of equalized individual loudness

Robert Kühler, Johannes Hensel, Christian Koch
Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, Germany.

Martin Bauer, Tillmann Sander-Thömmes
Physikalisch-Technische Bundesanstalt (PTB), Berlin, Germany.

Summary

Categorical loudness scaling as standardized in ISO 16832:2006 provides an easy and fast procedure to determine the loudness over the whole dynamic range of the auditory system. Within this method the loudness is rated by the subject on a scale with defined categories, such as "soft", "moderate", "loud", etc. In clinical audiology this method is widely used to determine the amount of recruitment in hearing impaired listeners. This method is extended to infrasonic ($f < 20$ Hz) and low frequencies ($f \geq 20$ Hz) in order to investigate the loudness perception for such frequencies. 30 otologically normal subjects performed the categorical loudness scaling procedure for monaural stimulation with pure tones between 8 and 125 Hz. Stimuli were presented by means of a newly designed insert-earphone sound source for infrasonic frequencies. The loudness functions showed a significant decrease of the dynamic range towards lower frequencies. To investigate the hearing mechanism at infrasonic frequencies more extensively, brain responses were measured by means of Magnetoencephalography (MEG). The variation of the M100 brain response stimulated by pure tones with decreasing frequencies from 250 Hz down to 8 Hz for 16 otologically normal subjects is investigated. The stimuli were ramped sine tones of equal individually perceived loudness determined before by the categorical loudness scaling method described above. Brain responses were recorded by a Yokogawa MEG system and averaged to obtain the M100 response. The positions of the underlying magnetic field generators were estimated offline, using a dipole fit routine. Stable responses were measured for stimulation frequencies between 20 and 250 Hz. At 8 and 12 Hz the results are difficult to interpret.

PACS no. 43.66.Cb, 43.28.Dm, 43.64.Ri

1. Introduction

Contrary to popular belief, hearing of infrasound ($f < 16$ Hz) is possible and may eventually cause distress which differs from person to person [1]. However, the mechanisms of this auditory perception by human beings are not yet fully understood. Within the scope of the European research project "EARS" [2], these mechanisms are to be investigated in detail by means of objective procedures of medical engineering, i.e. magnetic resonance imaging (MRI) and magnetoencephalography (MEG). We tackle the problem of the acoustic stimulation under sensitive magnetic conditions posed by both MEG and MRI with a newly developed insert earphone sound source that is able to generate acoustic stimulation with both a high sound pressure level at infrasonic frequencies and low harmonic distortion [3]. Loudness is, next to the pitch perception, one of the auditory key perceptions. For lower frequencies the pitch perception and the tonal character of pure tones becomes weaker, and both cease around 20 Hz. Below this frequency, tones are perceived as discontinuous. On the other hand the loudness perception remains active. The goal of this study was to investigate the change in loudness perception towards very low frequencies and the validation of objective auditory evoked brain responses to pure tones with decreasing frequency. The paper describes the determination of equal-loudness-level contours by means of categorical loudness scaling and the measurement of brain responses by MEG for infrasonic and low-frequencies pure tones of equalized loudness.

2. Loudness perception

Loudness as an attribute of the auditory sensation corresponds among other to the physical sound intensity and is one of the key perceptions of sound in everyday life and in case of unwanted noise plays a crucial role for annoyance. To investigate the loudness perception for low and infrasonic frequencies we extended the method of categorical loudness scaling standardized in ISO 16832 [4] to frequencies below 125 Hz.

2.1. Method

The stimuli were pure tones with frequencies of 8, 12, 16, 20, 32, 40, 63 and 125 Hz with an equal duration of 1600 ms. All signals had $\cos^2$ on- and offset ramps with a duration of 6 times the oscillation period for $f \geq 40$ Hz and 4 times the period for $f < 40$ Hz. Loudness perception was measured monaurally (right ear) at each frequency. All subjects received a written instruction into the loudness scaling experiment. Their task was to rate the loudness of the given tone stimulus using 11 response alternatives with defined categories ranging from "not heard", "soft", "medium", to "loud" and "extremely loud" with intermediate steps. The resulting loudness in categorical units (CU) from 0 to 50 represents the subjectively perceived loudness expressed in the mentioned 11 response alternatives (see figure 1). The whole experiment was self-controlled by the subjects via a personal computer with monitor and mouse. The loudness scaling was based on the adaptive procedure described in [5] and in agreement with ISO 16832 [4]. This procedure consisted of two phases, which was not noticeable for the subjects. First the auditory dynamic range of the subject was roughly estimated and in the second phase more data for the "medium" loudness levels were collected. The first phase in our study started with a stimulus at 80 phon according to the ISO 226 [6] equal-loudness-level contours (ELC) for $f \geq 20$ Hz, and according to the ELCs proposed by [7] for $f < 20$ Hz. If a response between "not heard" and "extremely loud" was given, the stimulus level was increased until the response "extremely loud" was given or the maximum stimulus level was reached. Beforehand, the maximum levels for the frequencies 125, 63, 40 and 32 Hz had been set to

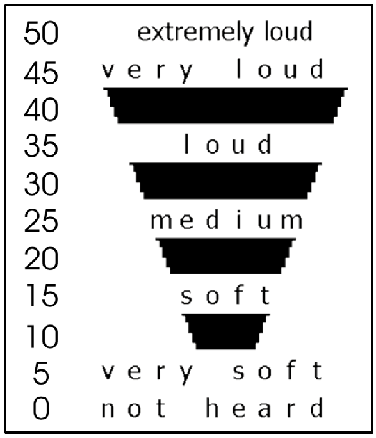

Figure 1: Category scale with 11 response alternatives including intermediate steps used by the subject to rate the loudness. The numbers on the left side indicate the categorical units (CU) which were not visible to the subjects. (Modified from [5, 8]).

80 phon according to [6] plus additional 20 dB. For all frequencies below 32 Hz the maximum level had been set fixed to 130 dB SPL to prevent the listener from harmful sound exposure. Then the level was decreased until the sound was inaudible and increased again until it become audible again. The adaptive level step sizes ranged from 5 to 15 dB in 5 dB steps. In the second phase the remaining categorical loudness levels were estimated by linear interpolation and presented in random order. This process of "estimating and presenting" was performed two times in this study. The separately measured signals at distinct frequencies were presented in random order except for a training tone of 60 Hz and the first stimulus of 125 Hz. All subjects performed the loudness scaling at all eight frequencies twice consecutively with a minimum pause of an hour.

Subjects: 30 test subjects (10 female and 20 male) aged between 18 and 30 years (mean age: 24.8 years) participated in the loudness measurements. All subjects were otologically normal (as assessed by means of the ISO 389-9 [9] questionnaire). No hearing thresholds by standard audiometry were determined. 19 subjects reported experience in hearing experiments.

Sound presentation: The experiment itself was controlled by a MATLAB-based software framework, controlling a 24 bit DA converter (RME Fireface UC). All stimuli were digitally generated online with 96 kHz sampling rate and presented monaurally by the insert earphone source described in [3]. The high sampling rate (with respect to the applied test stimuli) was chosen because this setting provided the lowest possible corner frequency for the low-frequency roll-off of the RME soundcard.

Sound calibration: Sound pressure level (SPL) calibrations were carried out using a low-frequency pressure-field ½'' microphone (Brüel & Kjær 4193 + UC0211, with its vent exposed to the sound field) in a cylindrical cavity with an equivalent ear canal volume of 1.3 cm^3 to which the sound tube from the insert earphone source was coupled via an Etymotic ER3-14A ear tip.

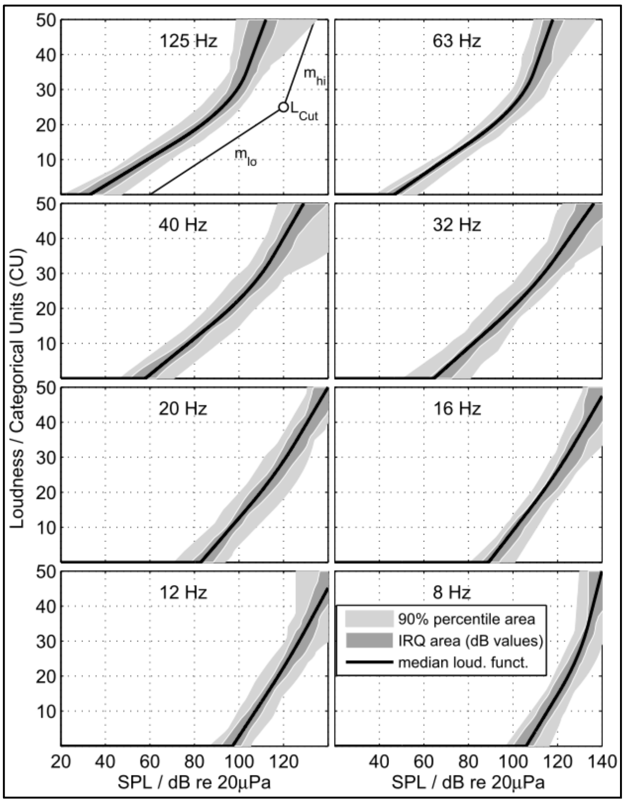

Figure 2. Median loudness functions across 30 ears (solid black line) obtained with the categorical loudness scaling procedure, interquartile range (IRQ, dark gray area), and the area between the 5% and 95% percentile (lighter gray area), for eight pure tones.

2.2. Results

After completion of a measurement cycle, a model function was fitted to the data that consisted of two linear parts with independent slope values (m_{lo}, m_{hi}) and an intersection point at L_{Cut} = 25 CU. The transition area between the two independent linear parts was smoothed using a Bezier fit [5]. Figure 2 shows the loudness functions formed by the median values of the estimated function parameters (m_{lo}, m_{hi}, L_{Cut}) over all individual loudness function parameters. Additionally shown are the interquartile range (IRQ) and the 90% percentile area of the CU vs. SPL of all individual loudness functions. The shape of the loudness function and the percentile range for the 125 Hz pure tone is comparable to the loudness function given in Appendix A of ISO 16832 [4]. No loudness functions for lower frequencies are provided in ISO 16832. For pure tones from 125 down to 32 Hz, all functions show a steeper slope for higher SPLs (m_{hi}) than for lower SPLs (m_{lo}). Below 32 Hz the median loudness function slopes can be described by a single linear function with one slope value. The increasing slope values of the loudness functions to lower frequencies indicate a significantly reduction of the dynamic range of the auditory system. This plays a critical role in the assessment of noise annoyance due to low

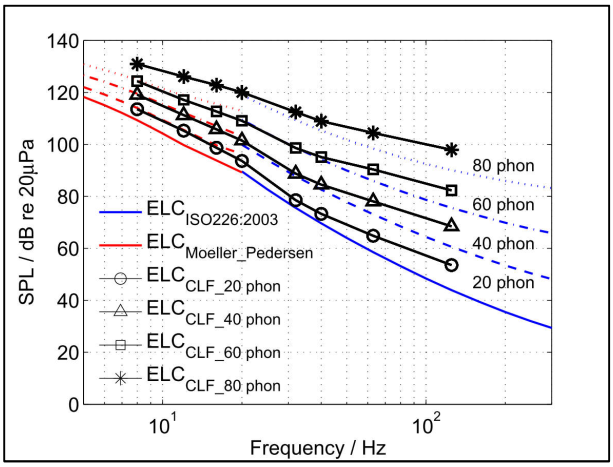

Figure 3. Equal-loudness-level contours (ELC) for 20, 40, 60 and 80 phon derived from the median categorical loudness functions (CLF) compared to ELCs from ISO226 [6] and Moeller and Pedersen [7].

frequency sound. To relate the loudness in CU to established noise emission limits expressed in phons or dBA, ELCs were determined and compared to standardized ELCs in ISO 226 [6].

To derive ELCs in phon from the loudness data in CU, reference loudness data for 1 kHz from literature were used. Therefore the loudness in CU of a third-octave band noise centered at 1 kHz was determined for levels of 20, 40, 60 and 80 dB SPL using the median loudness function from [10], resulting in CU values that correspond to 20, 40, 60, and 80 phons. The desired levels in dB SPL were then derived from the median loudness functions (figure 2) for the respective CU values. Figure 3 shows the resulting ELCs. The solid black lines (with symbols) indicate the ELCs derived from the median categorical loudness functions (ELC_{CLF}) determined in this study, and the blue lines show those given in ISO 226 [6]. Red lines indicate non-standardized ELCs for $f \leq 20$ Hz derived from a literature review by Moeller and Pedersen [7]. The ELC_{CLF} curves match only the 60 phon-ELC_{ISO226} curve at 20, 32, and 40 Hz. The differences between the ELC_{CLF} and ELC_{ISO226} for 20, 40 and 80 phon for f ≥ 20 Hz are smaller than 5 dB except the ELC_{CLF} at 63 and 125 Hz with differences ranging from 6 - 8 dB. For infrasonic frequencies a difference of about 6 dB compared to [7] is independently seen for all frequencies and ELCs. The dynamic range from 20 to 80 phon which is 45 dB at 125 Hz is reduced to 18 dB at 8 Hz.

3. Brain responses

In the auditory frequency range between 20 Hz and 20 kHz, brain responses are well known and the underlying processes are studied by many groups. Especially the brain response called M100, which occurs around 100 ms after an auditory stimulus onset and whose source is located in the auditory cortex, is of particular importance [11]. Insight into the processes of these brain responses is given by neuroimaging, such as electro-encephalographic (EEG) or MEG mapping. In order to comply with the very fast dynamics of brain processes we used MEG to investigate the M100 brain response to stimuli from audible sound frequencies of 250 Hz down to infrasound of 8 Hz. In particular, we investigated the variation of the M100 response and their corresponding generators with respect to decreasing stimulus frequencies.

3.1. Method

A commercial 128 channel first order gradiometer helmet-shaped MEG-system (Yokogawa) was used to record the magnetic components of auditory brain responses. It provides a gradiometer baseline of 50 mm and a pick-op coil radius of 25 mm. The subject is positioned recumbently. The average sensor noise inside the magnetically shielded room was estimated to be 7 fT/√Hz. The recorded data were sampled at 1 kHz and an analog low-pass filter set to 500 Hz and a high-pass filter set to 0.3 Hz were applied before digitization. Prior to the measurement 5 marker coils and an ultrasound device (3DSpace, Zebris) were used to determine the position of the subjects head relative to the helmet MEG system by means of calculating the transformation between the coil positions measured magnetically and the positions measured with the ultrasound device. This lead to an uncertainty of the head position of less than 5 mm.

Subjects: 16 test subjects (8 female and 8 male) aged between 20 and 30 years (mean age: 23.7 years) participated in the MEG measurements. All subjects were otologically normal and had performed the loudness measurement described above.

Stimuli paradigm: The auditory stimuli were $\cos^2$-ramped pure sine tones, which were fed into the right ear and their parameters are summarized in table I. The interstimulus interval and the order of the stimuli were varied randomly. A total measurement time of 45 minutes allowed around 75 averages per stimulus. As seen in table I individual sound pressure levels were applied resulting in an equally perceived loudness of all stimuli for every subject.

Sound calibration: Calibrations were carried out according to the loudness measurements, see section 2.1. During the listening tests, the SPL was monitored by a fiber-optical microphone (MO 2000, Sennheiser, Germany), which does not cause any magnetic interference and can therefore be fitted close to the ear even in the MEG environment.

Analysis of data: The recorded data was averaged to obtain event related fields (ERF) corresponding to the delivered auditory frequency. The signal time intervals corresponding to the onset ramps of the stimuli were omitted and the M100 responses were compared with respect to latency, amplitude and shape. Forward modeling was assumed by a spherical volume conductor with a homogeneous conductivity and two electric current dipoles representing the two auditory cortices. Source reconstruction were performed by means of a non-linear dipole fit technique [12] provided by Field Trip [13] to estimate the source position for every stimulus-evoked brain response.

Table I. Parameter of the auditory stimuli for MEG measurements. L_{25} stands for the level in dB SPL according the individual categorical loudness function at 25 CU.

Freq. (Hz)	250	125	63	40	20	12	8
Length (s)	1	1	1	1.5	2	2.5	3
Ramps (ms)	20	32	60	100	150	250	375
Level (dB_{SPL})	80	L_{25}	L_{25}	L_{25}	L_{25}	L_{25}	L_{25}

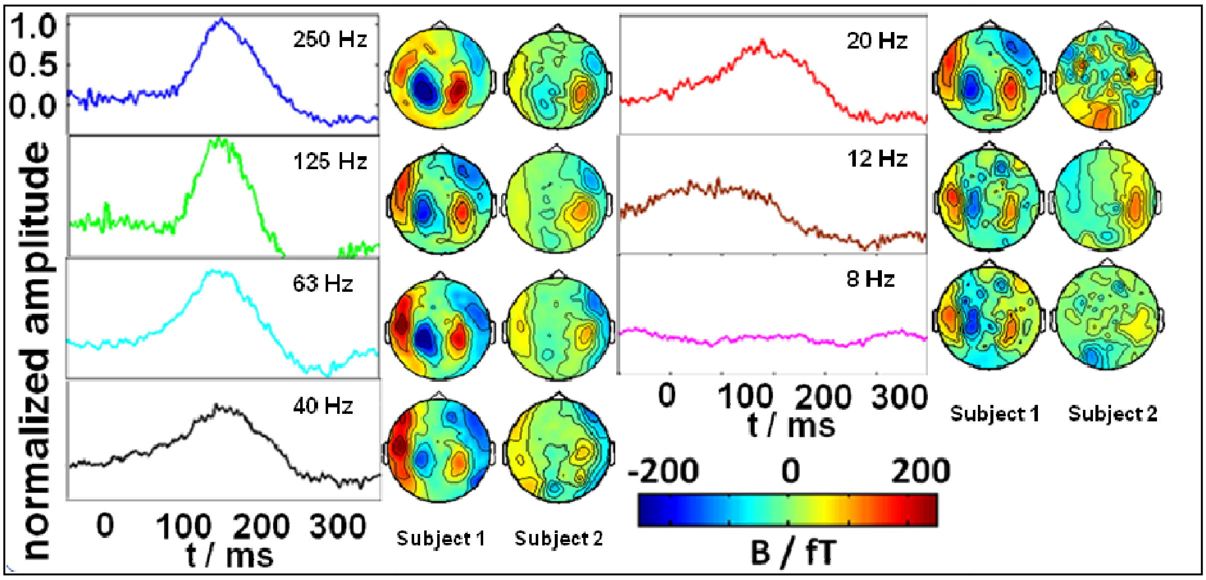

Figure 4. Time signals of the magnetic field strength and field maps of representative subjects. The time signals show a stable latency of 120 ms down to 20 Hz. The amplitude also remains stable within 20% in the frequency range between 20 and 250 Hz. The two subjects and there field maps are representatives of two groups of subjects, for one group a dipolar pattern is observed down to 8 Hz stimulation (Subject 1) and for the other group the very low frequency stimulation leads to irregular field maps instead of a dipolar pattern in the field map (Subject 2).

3.2. Results

All subjects reported perception of the stimuli down to 8 Hz. The average brain responses to tones with a frequency between 250 and 20 Hz showed nearly the same amplitude of around 250 fT. The response amplitude to the very low frequencies of 12 and 8 Hz reduced to around 100 and 50 fT, respectively. After subtracting the duration of the onset ramp, a stable M100-latency of 120 ms is observed. The recorded field maps suggest two subgroups of subjects. The first group has a stable dipolar response pattern down to 8 Hz; the other group shows only irregular field distributions instead of a dipolar pattern below 40 Hz. Most of the dipolar patterns observed in the response to the 8 and 12 Hz stimuli are different due to the orientation and the appearance of *two* dipolar patterns inside the spherical volume conductor compared to the current dipolar response pattern observed for the stimuli between 20 and 250 Hz. The described behavior of the responses is shown in figure 4; the time signals of the channel with the strongest amplitude in the evoked response of a single subject is shown on the left, the associated field map at the peak of the time signals for both subgroups of subjects in the middle. All active dipoles are estimated in the area of the auditory cortices (figure 5). The confidence interval of the dipoles representing activity relating to the 8, 12 and 20 Hz tones is wider than that of the other dipoles at higher stimulus frequencies, because the decreasing signal-to-noise ratio (SNR) to very low frequencies limits the accuracy of the electric current dipole model significantly (Subject 2 in figure 4).

4. Conclusion

We derived ELCs for infrasonic and low-frequency pure tones from categorical loudness functions measured with loudness scaling by means of categories according to ISO 16832 [4]. The resulting ELCs for 20, 40, 60, and 80 phon are comparable to standardized ELCs from ISO226 [6] and curves from literature [7] except an overall offset about 3 - 8 dB depending on frequency. An explanation for the differences may be the use of a reference categorical loudness function for a third-octave band noise centered at 1 kHz from literature [10], instead of an own measured loudness function at 1 kHz to derive the corresponding CU values for 20, 40, 60, and 80 phon respectively dB SPL. Due to the poor sound presentation abilities of the used (low-frequency) insert-earphone sound source of higher frequencies no reference measurement was possible at that time. Also the monaural presentations in our study can be a possibly explanation for at least 2 - 3 dB difference to the binaural data from [6, 7]. Despite the differences it was shown that the dynamic range of the human auditory system decreases dramatically with decreasing frequency. This plays a critical role in the assessment of noise annoyance due to infrasonic frequencies. In MEG measurements we used individual loudness functions to present the auditory stimuli with equal perceived loudness except for 250 Hz (see table I). All subjects reported a hearing sensation down to 8 Hz. MEG measurements showed a stimulus-evoked activation of the auditory cortex with a stable latency for frequencies from 250 down to 20 Hz. The responses to 8 and 12 Hz differ in amplitude of the M100 and the resulting dipolar patterns compared to the results at higher frequencies. Although with subjectively equal perceived stimuli loudness, the M100 response nearly disappears for 8 Hz. Two subgroups of subjects have been identified by means of the field distribution maps. For one group a dipolar pattern is observed down to 8 Hz, the other group shows irregular field distributions instead of a dipolar pattern to lower frequencies. This raises the question of whether the perception of infrasound by the human ear is still processed in the auditory cortex alone or whether other cognitive processes (e.g. somatosensory) are involved. Future measurements in both MEG and functional MRI will help to answer this question.

Acknowledgement

Financial support from the European Metrology Research Programme (EMRP, programme Health, grant "HLT01 EARS") is gratefully acknowledged. The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union.

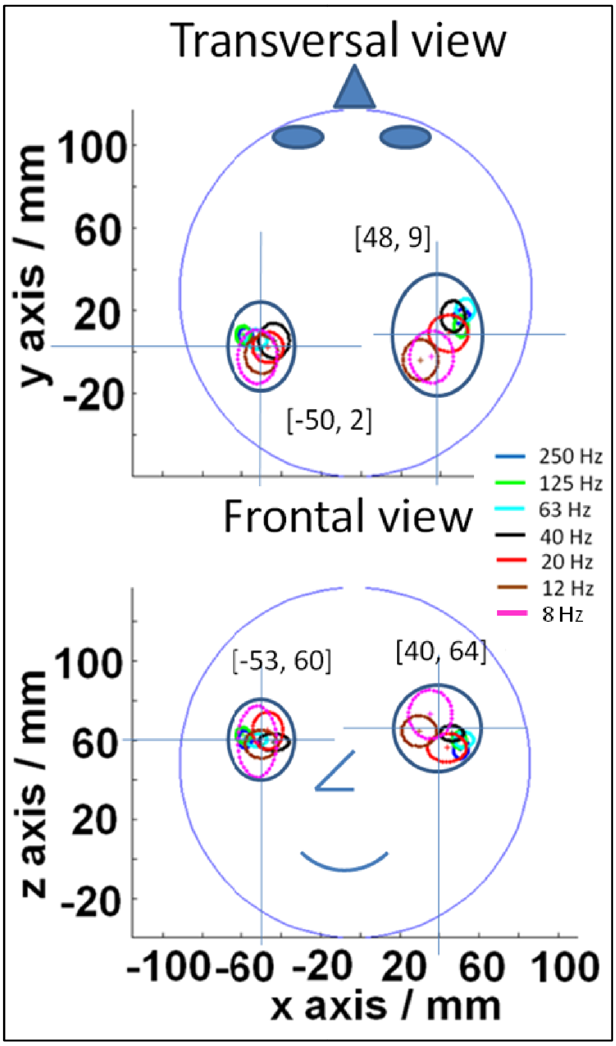

Figure 5. Estimation of the dipole location in a transversal and frontal slice for the group of subjects. The large blue circle represents the extent of the normalized volume conductor model. The small circles are the 95% confidence intervals of the dipole localizations using a non linear dipole fit routine. The colours represent the stimulus frequency and correspond to the colours of the time signals in figure 4. All activations are consistent with the auditory cortices.

References

[1] H. G. Leventhall: Low Frequency Noise and Annoyance. Noise & Health 6 (2004) 59-72.

[2] EARS Project: Metrology for a universal ear simulator and the perception of non-audible sound. European Metrology Research Programme (EMRP), http://www.ears-project.eu/emrp/ears.html (2015-02-18)

[3] R. Kuehler, T. Fedtke, J. Hensel: Infrasonic and low frequency insert earphone hearing threshold. J. Acoust. Soc. Am. (2015) EL (in print).

[4] ISO 16832: Acoustics – Loudness scaling by means of categories (ISO 16832:2006).

[5] T. Brand, V. Hohmann: An adaptive procedure for categorical loudness scaling. J. Acoust. Soc. Am. 112 (2002) 1597-1604.

[6] ISO 226: Acoustics – Normal equal-loudness-level contours (ISO 226:2003).

[7] H. Moeller, C. S. Pedersen: Hearing at Low and Infrasonic Frequencies. Noise & Health 6 (2004) 37-57

[8] HörTech gGmbH: Operating manual Categorical Loudness Scaling. Rev. 1.3.a, issue 21.06.2005. Oldenburg Measurement Applications.

[9] ISO 389: Acoustics – Reference zero for the calibration of audiometric equipment -- Part 9: Preferred test conditions for the determination of reference hearing threshold levels (ISO 389-9:2009)

[10] W. Heeren, V. Hohmann, J. E. Appell, J. L. Verhey: Relation between loudness in categorical units and loudness in phons and sones. J. Acoust. Soc. Am. 133 (2013) EL314-EL319.

[11] T. P. Roberts, P. Ferrari, S. M. Stufflebeam, D. Poeppel: Latency of the auditory evoked neuromagnetic field components: stimulus dependence and insights toward perception. J. clin. neurophysiol, 17 (2000), 114-129.

[12] B. Lütkenhöner: Möglichkeiten und Grenzen der neuromagnetischen Quellenanalyse. Habilitationsschrift (1992).

[13] R. Oostenveld, P. Fries, E. Maris, J.-M. Schoffelen: FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. Computational Intelligence and Neuroscience, vol. 2011, Article ID 156869, 9 pages (2011).