9-12 AUGUST 2015
inter.noise 2015
SAN FRANCISCO, CALIFORNIA USA

Experiments on the loudness-transfer of headphone-based virtual acoustics

Florian Völk[1]

AG Technische Akustik, MMK, Technische Universität München
Arcisstraße 21, 80333 München, Germany

WindAcoustics UG (haftungsbeschränkt)
Mühlbachstraße 1, 86949 Windach, Germany

Bio-Inspired Information Processing, Institute of Medical Engineering,
Technische Universität München, Boltzmannstraße 11, 85748 Garching, Germany

Hugo Fastl[2]

AG Technische Akustik, MMK, Technische Universität München
Arcisstraße 21, 80333 München, Germany

Binaural synthesis is a virtual-acoustics technology based on the convolution of sound signals with impulse responses modeling the propagation paths between sources and listener. The convolution products are typically presented by headphones. The frequency-dependent correction level necessary for a binaural-synthesis system to elicit the reference-scene loudness is referred to as loudness-transfer function. An ideal binaural-synthesis system provides frequency-independent loudness-transfer functions for every listener. The frequency dependence of a binaural-synthesis system's inter-individually averaged loudness-transfer function has been shown to depend on the hardware, the implementation, and the degree of individualization. In this contribution, perceptually acquired loudness-transfer functions of binaural-synthesis systems are discussed from an auditory-adapted perspective with regard to listening experiments, especially sound-quality judgments. The results provide quantitative estimates of the accuracy of sound-quality judgments and noise ratings achievable with different headphone-based binaural-synthesis implementations.

[1]email: voelk@tum.de, voelk@windacoustics.com
[2]email: fastl@mmk.ei.tum.de

1 INTRODUCTION

Binaural synthesis (BS) aims at acoustically synthesizing or reproducing a reference scene by means of headphones (Wightman and Kistler 1989a,b, Møller 1992). This goal would be reached if the physical signals detected by the eardrums, the so-called ear signals, were identical in the reference and synthesis situations (Møller 1992, Völk 2011, 2013). Validating this goal, however, requires to identify and measure the ear signals. Both, ear-signal identification and measurement are subject to an ongoing scientific discussion and are likely not possible in a strict sense (Stinson 1985, Neely and Gorga 1998, Schmidt and Hudde 2009). Additionally, setting up a fully three-dimensional, highly accurate dynamic system that takes into account listener movements requires considerable effort (Møller 1992, Völk 2013). For that reason, most implementations are restricted to some extent regarding resolution, dynamics, the simulated degrees of freedom, as well as the degree of individualization (e. g. Wightman and Kistler 1989a, Völk 2013). These restrictions prevent the validation by physical comparison against the reference, as the synthesized signals will intentionally deviate from those of the reference scene.

As a physical validation is hard to impossible, BS systems are frequently being validated perceptually, by means of listening experiments. Early studies focused predominantly on the achievable directional-localization accuracy, revealing that even static systems and dynamic systems restricted to head rotations can show remarkable accuracy, and that individualization further improves that accuracy (e. g. Wenzel et al. 1993, Bronkhorst 1995, Møller et al. 1996, Middlebrooks 1999). A second criterion being studied was the naturalness of the hearing sensations. Silzle (2002a,b) reported BS implemented based on artificial-head recordings to sound less natural than a human-head based implementation. Usher and Martens (2007) found a non-individualized system to be rated more natural than an individualized implementation. This at first glance unexpected result may be understood taking into account that the most natural situation not necessarily equals the most authentic one. Therefore, naturalness is not suited as a quality criterion for BS aiming at reproducing a specific reference scene. For that reason, a perceptual BS quality criterion more selective than the accuracy of the absolute localization is desirable, as also stated by Martens et al. (2010).

In search of more selective criteria for validating the quality of a binaurally-synthesized real scenario, distance perception (Hartmann and Wittenberg 1996, Völk et al. 2008, Völk 2009), directional resolution (e. g. Völk et al. 2012), and sound color (Völk et al. 2011, Völk and Fastl 2011, 2013, Völk 2013) have been studied. As a result, we proposed to address directional resolution by measuring the minimum-audible angle, and to address sound color by perceptually acquiring the system-under-test's loudness transfer function (LTF) with regard to the reference scene (Völk 2013). The LTF is defined as the frequency-dependent correction level required at the BS input to elicit, with narrow-band stimuli, equal loudness of reference and synthesis. A perfect BS would show a frequency-independent LTF of 0 dB.

In this contribution, the implications of typical shortcomings of BS systems with different individualization degrees on sound-quality evaluation results acquired with BS are discussed. Deviations between synthesized and reference scene are quantified based on the respective LTFs both, physically and from a psychoacoustic perspective. As the exemplary reference scene, a single loudspeaker box located in front of the subjects in rooms with different acoustical conditions was chosen.

2 SETUP AND PROCEDURE

According to Møller (1992) and Völk (2010, 2012, 2013), the preferable approach to BS is individual miniature-microphone recording of the necessary impulse responses at the entrances to the blocked auditory canals. As individual recording and equalization-filter design is time consuming, completely or partially non-individualized procedures are employed sometimes, especially when using BS in applications, as for example sound-quality evaluations.

In detail, non-individual in contrast to individual recording means that the impulse responses describing the transfer paths from the sound sources to the ears were measured on a subject different from the listener. For individual respectively non-individual equalization, the listener's own respectively other headphone-transfer functions are inverted to form the equalization target. A filter useful in implementations is then obtained by reducing amplification requirements of the filter target not provided by the equipment, for example by high- and low-pass filtering and regularization (Kirkeby and Nelson 1999, Norcross et al. 2004). A target for equalization that is to a certain degree suited for different subjects can be acquired by averaging the magnitude-transfer functions of a certain number of subjects. The so-called average-magnitude equalization filter is then obtained by regularization and combination with a realistic linear phase response. Using such a phase response, it is also possible to equalize only the magnitude-transfer functions individually or non-individually. Different combinations of the aforementioned components recording and equalization were included in this study.

BS with adaptive signal processing adjusting the impulse responses to the current listening situation is referred to as dynamic in contrast to static binaural synthesis. Both, static and dynamic systems are addressed, where the dynamic system was restricted to rotational head movements in that a single set of 360 pairs of impulse-responses was used, measured for a full rotation of the subject in 1° steps. Translational head movements did only cause ear signal adaption if the direction of the source relative to the head was affected.

In order to address the effect of different room-acoustic conditions, the experiments were conducted in two different laboratory rooms. Figure 1 shows the corresponding reverberation times, measured at the listening position with the reference-scene loudspeaker box at its reference-scene position as the sound source.

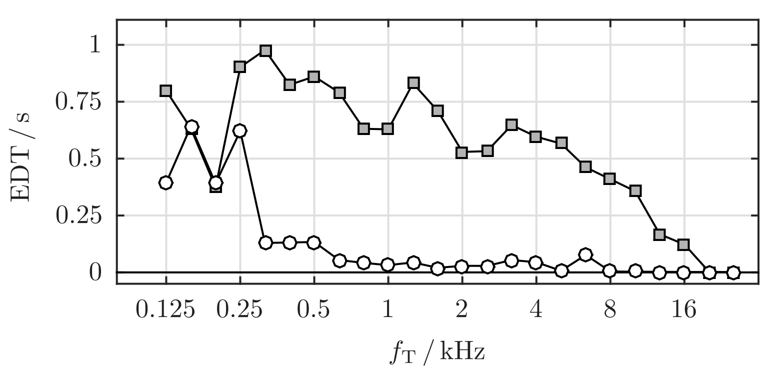

Fig. 1 – Third-octave band early-decay times of the rooms where the listening experiments took place: Laboratory 1 (open circles) and laboratory 2 (filled squares).

Laboratory 1 (open circles) is designed to resemble a highly damped living room and therefore provides a low reverberation time, especially at high frequencies. Laboratory 2 (filled squares) is a small laboratory with an average reverberation time of about 600 ms at low and mid frequencies, which decays towards higher frequencies. These room-acoustical conditions were included to address possibly occurring differences between the loudness-transfer functions due to room acoustics.

In every room, a chair was positioned in front of a loudspeaker box (Klein + Hummel O 98) at 1.5 m distance. The complete setup was centered around the room midpoint, with neither the chair nor the loudspeaker directly at the midpoint. In each condition, the impulse responses for the BS were measured either individually or non-individually. In order to allow for comfortably and accurately carrying out the loudness comparisons, the reference situation to be simulated by BS was defined as a subject listening to a loudspeaker through non-operational headphones. That way, the subjects did not have to remove the headphones when comparing reference and synthesis. As a consequence, the impulse responses describing the transfer paths from the loudspeaker to the eardrums were measured through the headphones.

For the measurements as well as the listening experiments, an optical position control based on the principle of a pinhole camera was used to adjust the chair so that the midpoint of the seated subject's inter-aural axis lied horizontally and vertically with an accuracy of ± 3 cm on the loudspeaker radiation axis. During the measurements, the chair was rotated around the midpoint of the inter-aural axis in 6° steps, the rotational center and step size being assured by a head tracking system (Polhemus 3 Space FasTrack) with an accuracy of ± 1 cm and $\pm 0.5°$ respectively. Measurements with head nodding or tilting deviations from the horizontal plane of more than $\pm 0.75°$ were repeated until a smaller deviation was reached. After the measurements, the horizontal grid resolution was increased to 1° by cubic-spline interpolation over the time shifted impulse responses (cf. Christensen et al. 1999).

The measurements were carried out using exponential sine sweeps (5 s duration, frequency range from 10 Hz to 22050 Hz; cf. Farina 2000, Müller and Massarani 2001, Völk et al. 2009). Measurements, data analysis, and signal processing were done at double-precision word length with 44.1 kHz sampling rate using modules of the WindAcoustics Suite (2015). For all recordings, Sennheiser KE 4-211-2 electret microphones embedded in modified foam ear plugs were inserted 2 to 4 mm in the auditory canals, completely blocking them.

Two different headphone models have been included in this study: Stax λ pro NEW and Sennheiser HD 800. The respectively used model is indicated along with the results. As only a single specimen per model was used, the results may not be representative for the model. However, the differences between the results will support the necessity to select headphones for binaural playback carefully (Møller et al. 1995a,b, Völk 2010, 2012). No head fixation was applied and the subjects were allowed to turn their heads. In order to reduce visual influences, the experiments were carried out in complete darkness.

3 METHOD AND STIMULI

The experimental method used was loudness adjustment with Békésy-Tracking according to Fastl and Zwicker (2007, cf. also von Békésy 1947, Zwicker and Feldtkeller 1955). The subjects had to listen in turn to the real and the corresponding binaurally synthesized loudspeaker, *without* taking off the headphones. It was their task to continuously adjust the level of the

BS so that equal loudness was elicited by the synthesis and the real loudspeaker. For that purpose, tone impulses were presented alternately by the BS and the loudspeaker. After each pair, the frequency was changed automatically. The level $L_{\mathrm{ls_{in}}}$ at the loudspeaker input remained constant while the level $L_{\mathrm{bs_{in}}}$ of the BS was either increased or decreased after each pair. The subject had to change the direction of level variation using a button and was asked to change the direction every time a loudness difference within a pair occurred. This procedure resulted in a frequency-dependent zigzag-pattern, alternating around the equal-loudness contour. In detail, two tone impulses with 0.4 s duration and 5 ms Gaussian gating, separated by 0.1 s pause were compared. The second tone was presented by the BS and adjusted to the first (reference), presented at the same frequency by the loudspeaker. The loudspeaker was calibrated with broadband pink noise to a level of 58 dB SPL. The BS level varied with 1.5 dB step size, starting 10 dB above the level eliciting approximately the same loudness as the loudspeaker. Two successive pairs were separated by 0.4 s pause.

Each experimental run was divided in two parts, one with increasing frequency (equally spaced on the critical-band rate scale ac. to Fastl and Zwicker 2007), starting from 1.3 kHz upwards to 20 kHz, and one decreasing from 1.7 kHz downwards to 20 Hz, at 0.05 Bark step size. To reduce methodical artifacts, 20 steps of the results in the overlapping region, at the start of each section, were not included in computing the combined result. The individual result was calculated by interpolating the mean values of every two neighboring turning points. Thereby, frequency dependent average and maximum deviations between two runs of about ±2 dB and ±4 dB were achieved (validated with three experienced subjects).

4 RESULTS

In this section, perceptively acquired loudness-transfer functions $L_{\mathrm{bs_{in}}} - L_{\mathrm{ls_{in}}}$ of different BS implementations are shown by means of the inter-individual medians and inter-quartile ranges of at least eight experienced, normal-hearing subjects per condition. Along with the results, calculated loudness and sharpness deviations that would be introduced by an intensity frequency dependence comparable to the loudness-transfer function are given for typical noise spectra at different playback levels. These instrumental predictions of two psychoacoustic magnitudes indicate important factors of the impact of the respective playback system on sound-quality evaluations carried out with the corresponding BS implementation.

For the loudness and sharpness calculations, noise signals of 10 s duration and overall sound-pressure levels L_{p} between 0 dB and 120 dB SPL were spectrally weighted with the inter-individual medians of the loudness-transfer functions and then analyzed according to DIN 45 631/A1 (2010) and DIN 45 692 (2009) using WindAcoustics Suite (2015), resulting in the test-signal loudness N_{T} and sharpness S_{T}. As a reference, loudness N_{R} and sharpness S_{R} were also calculated for the unweighted noise signals. On that basis, the loudness deviation

$$N_{\mathrm{dev}} = (N_{\mathrm{T}}/N_{\mathrm{R}} - 1) \cdot 100\% \tag{1}$$

and the sharpness deviation

$$S_{\mathrm{dev}} = (S_{\mathrm{T}}/S_{\mathrm{R}} - 1) \cdot 100\% \tag{2}$$

were calculated. In order to reduce the influence of the statistical properties of the noises on the results, the calculation was carried out five times per signal. The medians over the repetitions are shown and discussed. Three different noise spectra were included: broadband white and pink noise as well as critical-band wide narrow-band noises of 25 different center frequencies, equally spaced on the critical-band rate scale. Critical bandwidth and critical-band rate were calculated according to Völk (2015).

4.1 Non-Individual Dynamic Synthesis, Average-Magnitude Equalization

Figure 2 shows the results for a non-individual dynamic BS system, simulating a loudspeaker box in the slightly reverberant laboratory 1 (cf. figure 1). The upper panel a) represents the loudness-transfer function, panel b) the calculatd loudness (solid) and sharpness (dashed) deviations introduced by the BS of broadband white (dark) and pink noises (light), at different levels. Panel c) shows the corresponding loudness deviations for critical-band wide narrow-band noises at different levels and center-frequencies.

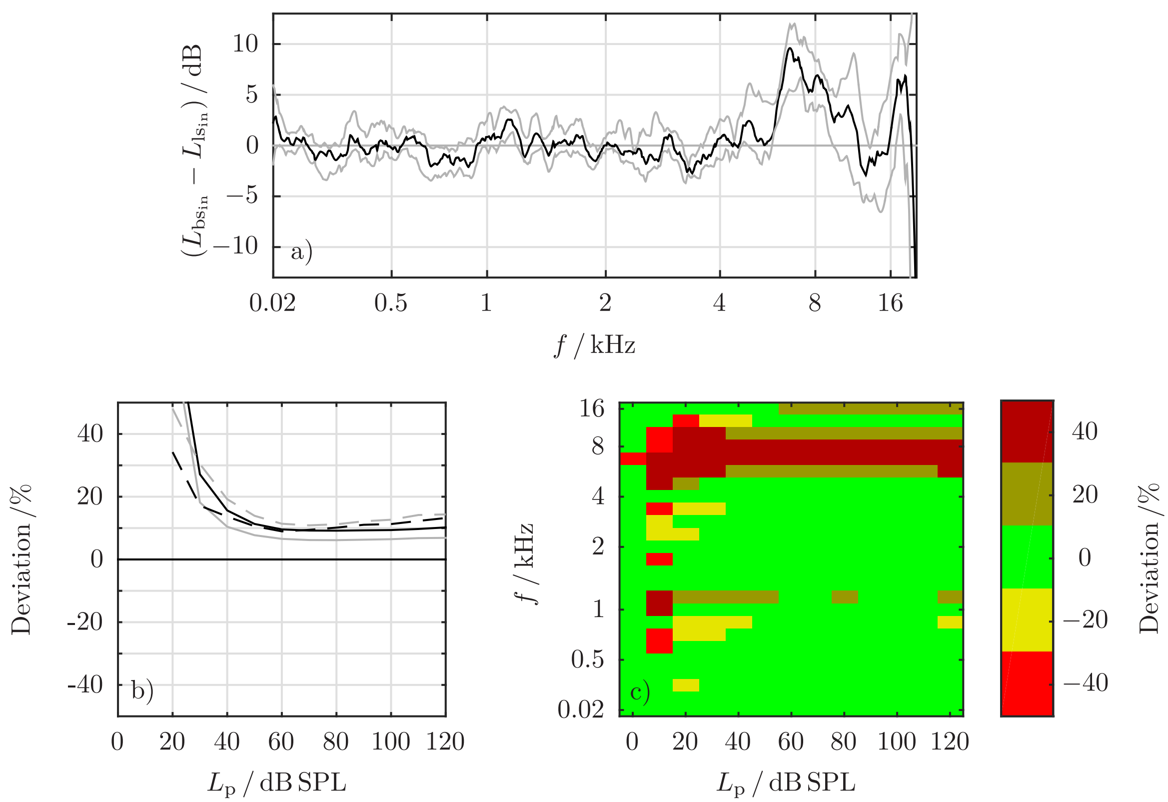

Fig. 2 – a) Inter-individual median (black) and inter-quartile range (gray) of the loudness-transfer function of non-individual dynamic binaural synthesis of a frontally located loudspeaker box in the slightly reverberant laboratory 1. Synthesis with Stax λ pro NEW headphones and average-magnitude equalization. b) Corresponding instrumentally-predicted loudness (solid) and sharpness (dashed) deviations for broadband white (dark) and pink noise (light). c) Corresponding instrumentally-predicted loudness deviations for critical-band-wide narrow-band noise.

The loudness-transfer function in panel a) shows a pronounced peak in the area around some 7 kHz, where the BS requires on average more than 5 dB more input level than the loudspeaker to sound equally loud. This peak increases the predicted loudness and sharpness of broadband sounds at supra-threshold levels by about 10%, as shown in panel b). At low levels close to threshold in quiet, the peak elicits a sensation at lower levels than the reference system, causing the steep increase in loudness and sharpness deviation towards low sound-pressure levels. The loudness deviation for narrow-band noises in panel c) reveals the effect of the peak by a more or less level-independent loudness deviation of about 40% in the frequency range somewhat below 8 kHz.

The peak of the loudness transfer function in figure 2 was traced back at least partly to the suitability of the headphone specimen for BS (Völk and Fastl 2011, Völk 2012, 2013). While using better-suited headphones reduces the extent of the artifact of the loudness-transfer function, as shown in panel a) of figure 3, some artifacts remain.

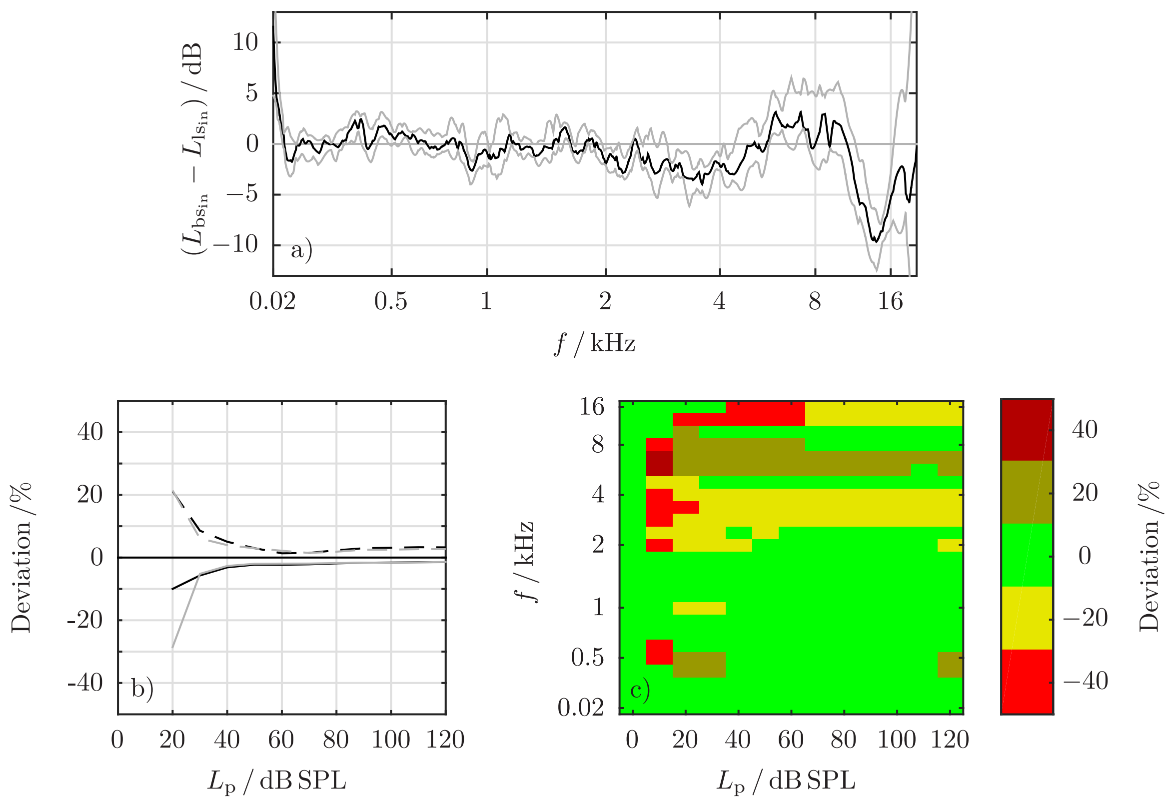

Fig. 3 – Results as described in the caption of figure 2 but for non-individual dynamic binaural synthesis of the reverberant laboratory 2 with Sennheiser HD 800 headphones and average-magnitude equalization.

However, the artifacts cause deviations of the predicted loudness and sharpness for broadband sounds less than about ±5%, at levels above 45 dB, as shown in panel b) of figure 3. For narrow-band noises, the loudness deviations exceed ±20% at some mid and high frequencies.

4.2 Individual Static Synthesis, Average-Magnitude Equalization

Figure 4 shows the experimental results and corresponding instrumental predictions for *individualized* (instead of non-individualized as above) static (instead of dynamic as above)

BS of the reverberant laboratory 2, combined with average-magnitude equalization. The high-frequency loudness-transfer-function artifact of figures 2 and 3 is less pronounced (panel a), to an extent reducing the predicted loudness deviations for broadband sounds to virtually zero and the sharpness deviations to values below about 4% for levels above some 40 dB.

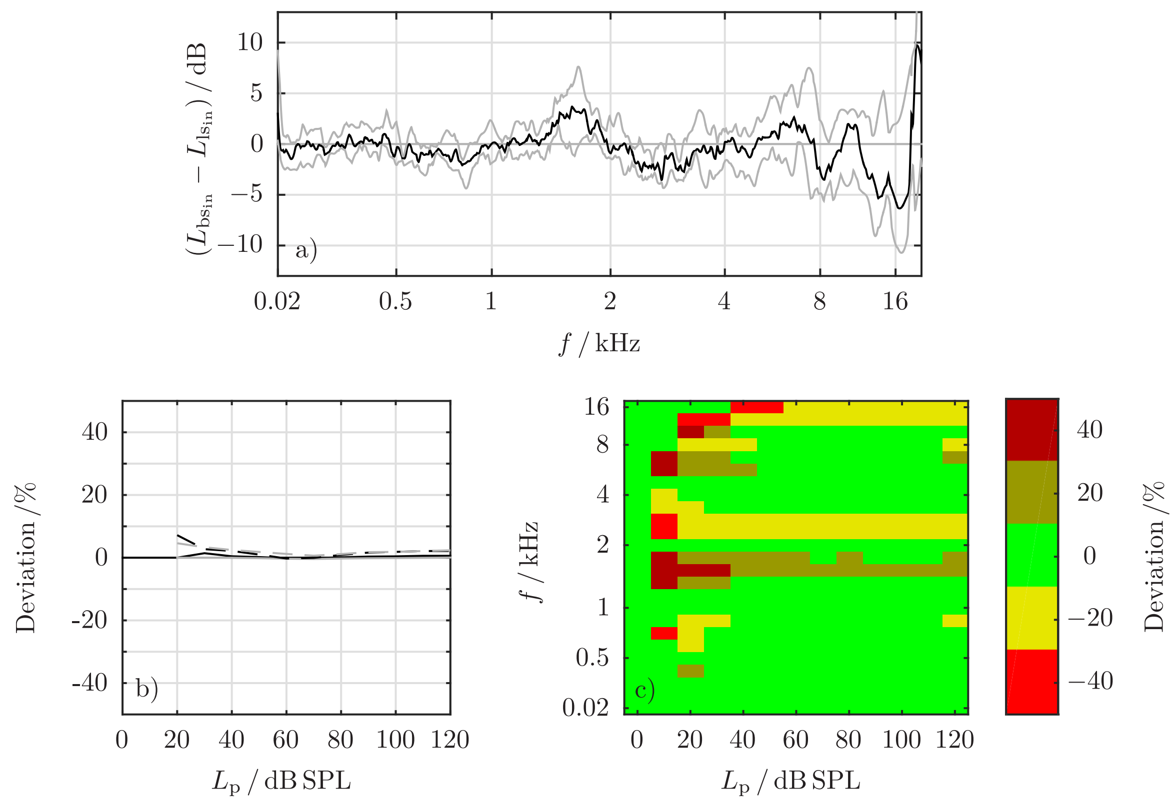

Fig. 4 – Results as described in the caption of figure 2 but for individual static binaural synthesis of the reverberant laboratory 2 with Sennheiser HD 800 headphones and average-magnitude equalization.

The predicted loudness deviations are also reduced, compared to non-individualized BS, for narrow-band noises, especially at high frequencies. However, in the frequency range between about 1.3 kHz and 1.9 kHz, the loudness-transfer function in panel a) shows a peak, which is clearly reflected in panel c), that is in the loudness deviation of narrow-band signals. The artifact's magnitude lies in the range of 20%, its frequency range almost exactly at the beginning of the tracking procedure (starting upwards at 1.3 kHz and downwards at 1.7 kHz).

4.3 Individual Static Synthesis, Individual-Magnitude Equalization

With individual-magnitude equalization, individualized BS provides the results shown in figure 5. The loudness-transfer function in panel a) proceeds, in the frequency range below about 12 kHz, on average frequency independently within the accuracy of the measurement procedure (cf. section 3). The high-frequency artifact around 16 kHz is reduced in level and bandwidth compared to the less individualized BS configurations shown above.

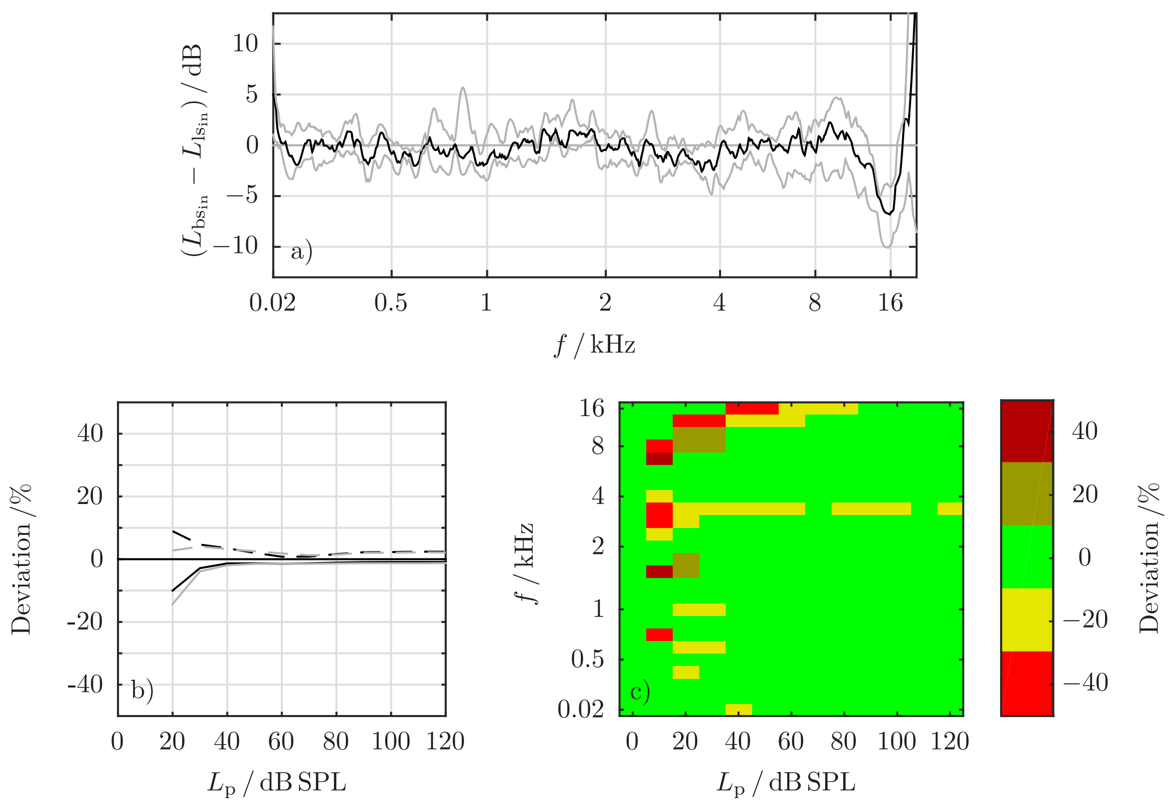

Fig. 5 – Results as described in the caption of figure 2 but for individual static binaural synthesis of the reverberant laboratory 2 with Sennheiser HD 800 headphones and individual-magnitude equalization.

The instrumentally predicted loudness and sharpness deviations introduced by the loudness-transfer function in panel a) are, for levels above 40 dB that is above threshold effects, in good approximation frequency independent, for broadband (panel b) and narrow-band noises (panel c). Near threshold in quiet, the at least partly due to the procedure ragged nature of the loudness transfer function causes rather pronounced relative differences, as the reference loudness may still be zero (inaudible), while a peak in the loudness-transfer function causes audibility in the test condition.

5 DISCUSSION AND CONCLUSIONS

In this contribution, perceptually acquired loudness-transfer functions of binaural-synthesis systems implemented with blocked-auditory-canal measurement are analyzed from an auditory-adapted perspective and with regard to conducting listening experiments, especially sound-quality judgments. A loudness-transfer function is the frequency-dependent correction level necessary for an audio-transmission system to elicit the reference-scene loudness. Consequently, an ideal binaural-synthesis system would provide frequency-independent loudness-transfer functions for every listener. In order to estimate loudness and sharpness deviations from the reference scene introduced by binaural-synthesis, instrumental analyses were conducted.

The results support the necessity of selecting headphones for binaural synthesis with blocked-auditory-canal recording, for example using the criterion proposed by Møller et al. (1995b) and extended by Völk (2012). The non-individualized system with the rather unsuited headphone specimen studied here in combination with average-magnitude equalization causes deviations of the predicted loudness and sharpness from the reference scene exceeding 10% even for broadband sounds (cf. figure 2, panel b). The instrumentally predicted loudness deviations for narrow-band sounds (panel c) are frequency dependent. At low and medium center frequencies, little to no deviations occur, whereas the deviations exceed ±40% at center frequencies above 5 kHz. While the system is suited to address low-frequency narrow-band stimuli, other sound-quality studies conducted using this system, as for example addressing the sound color or loudness of broadband stimuli, will be corrupted by the playback system.

With the selected headphone specimen, non-individualized binaural synthesis with average magnitude equalization provides at levels above 40 dB loudness and sharpness deviations for broadband noises below ±5% (figure 3, panel b). At lower levels, threshold effects in combination with the evaluation method used here may compromise the results. Apart from these artifacts, also the loudness of narrow-band stimuli centered at frequencies below 2 kHz is reproduced well (panel c). At higher frequencies, deviations exceeding ±20% occur, which are due to a high-frequency artifact in the loudness-transfer function (panel a). This artifact and the corresponding loudness deviation is reduced with individualized synthesis and average-magnitude equalization (figure 4). Further optimization is possible by switching to individual-magnitude equalization (figure 5). The deviation of the loudness of narrow-band noises centered between 1 kHz and 2 kHz may be attributed to the tracking procedure (cf. section 4), but may also be due to the missing individualization of the equalization, combined with individual recording and static synthesis. The effect of the dip in the loudness-transfer function remaining with individual-magnitude equalization around 16 kHz (figure 5) is lower than in the other conditions, and decreases further with increasing sound-pressure level.

Concluding, the suitability of different binaural-synthesis implementations for sound-quality evaluations depends on the stimuli to be studied. For low-frequency narrow-band sounds (centered below some 3 kHz), non-individualized synthesis and equalization can lead to correct loudness-transfer, even without specifically selected headphones. For ensuring loudness and sharpness errors below ±5% with broadband stimuli, at least the selection of appropriate headphones is necessary. If sound color or high frequency narrow-band effects are to be addressed, the individualization of impulse responses and equalization is required.

REFERENCES

Bronkhorst A. W.: Localization of real and virtual sound sources. J. Acoust. Soc. Am. **98**, 2542–2553 (1995)

Christensen F., H. Møller, P. Minnaar, J. Plogsties, S. K. Olesen: Interpolating between Head-Related Transfer Functions Measured with Low Directional Resolution. In *107th AES Convention* (1999) (Preprint 5047)

DIN 45 631/A1: *Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum – Verfahren nach E. Zwicker – Änderung 1: Berechnung der Lautheit zeitvarianter Geräusche; mit CD-ROM (Calculation of loudness level and loudness from the sound spectrum – Zwicker method – Amendment 1: Calculation of the loudness of time-variant sound; with CD-ROM).* (Deutsche Norm, Beuth Verlag, Berlin, 2010)

DIN 45 692: *Messtechnische Simulation der Hörempfindung Schärfe (Measurement technique for the simulation of the auditory sensation of sharpness).* (Deutsche Norm, Beuth Verlag, Berlin, 2009)

Farina A.: Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique. In *108th AES Convention* (2000) (Preprint 5093)

Fastl H., E. Zwicker: *Psychoacoustics – Facts and Models.* 3rd Edition (Springer, Berlin, Heidelberg, 2007)

Hartmann W. M., A. Wittenberg: On the externalization of sound images. J. Acoust. Soc. Am. **99**, 3678–3688 (1996)

Kirkeby O., P. A. Nelson: Digital Filter Design for Inversion Problems in Sound Reproduction. J. Audio Eng. Soc. **47**, 583–595 (1999)

Martens W. L., A. Guru, D. Lee: Effects of individualised headphone response equalization on front/back hemifield discrimination for virtual sources displayed on the horizontal plane. In *20th Int. Congress on Acoustics (ICA)* (2010)

Middlebrooks J. C.: Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency. J. Acoust. Soc. Am. **106**, 1493–1510 (1999)

Møller H.: Fundamentals of Binaural Technology. Applied Acoustics **36**, 171–218 (1992)

Møller H., D. Hammershøi, C. B. Jensen, M. F. Sørensen: Transfer Characteristics of Headphones Measured on Human Ears. J. Audio Eng. Soc. **43**, 203–217 (1995a)

Møller H., C. B. Jensen, D. Hammershøi, M. F. Sørensen: Design Criteria for Headphones. J. Audio Eng. Soc. **43**, 218–232 (1995b)

Møller H., M. F. Sørensen, C. B. Jensen, D. Hammershøi: Binaural Technique: Do We Need Individual Recordings? J. Audio Eng. Soc. **44**, 451–469 (1996)

Müller S., P. Massarani: Transfer-Function Measurement with Sweeps. J. Audio Eng. Soc. **49**, 443–471 (2001)

Neely S. T., M. P. Gorga: Comparison between intensity and pressure as measures of sound level in the ear canal. J. Acoust. Soc. Am. **104**, 2925–2934 (1998)

Norcross S. G., G. A. Soulodre, M. C. Lavoie: Subjective Investigations of Inverse Filtering. J. Audio Eng. Soc. **52**, 1003–1027 (2004)

Schmidt S., H. Hudde: Accuracy of acoustic ear canal impedances: Finite element simulation of measurement methods using a coupling tube. J. Acoust. Soc. Am. **125**, 3819–3827 (2009)

Silzle A.: Auswahl und Tuning von HRTFs (Selection and Tuning of HRTFs). In *Fortschritte der Akustik, DAGA '02*, 734–735 (Dt. Gesell. für Akustik e. V., Oldenburg, 2002a)

Silzle A.: Selection and Tuning of HRTFs. In *112th AES Convention* (2002b) (Convention Paper 5595)

Stinson M. R.: The spatial distribution of sound pressure within scaled replicas of the human ear canal. J. Acoust. Soc. Am. **78**, 1596–1602 (1985)

Usher J., W. L. Martens: Perceived Naturalness of Speech Sounds Presented Using Personalized Versus Non-Personalized HRTFs. In *13th Int. Conference on Auditory Display (ICAD)* (2007)

Völk F.: Externalization in data-based Binaural Synthesis: Effects of Impulse Response Length. In *Int. Conference on Acoustics NAG/DAGA 2009*, 1075–1078 (Dt. Gesell. für Akustik e. V., Berlin, 2009)

Völk F.: Messtechnische Verifizierung eines datenbasierten binauralen Synthesesystems (Metrological verification of a data-based binaural-synthesis system). In *Fortschritte der Akustik, DAGA 2010*, 1049–1050 (Dt. Gesell. für Akustik e. V., Berlin, 2010)

Völk F.: System Theory of Binaural Synthesis. In *131st AES Convention* (2011) (Convention Paper 8568)

Völk F.: Headphone Selection for Binaural Synthesis with Blocked Auditory Canal Recording. In *132nd AES Convention* (2012) (Convention Paper 8677)

Völk F.: *Interrelations of Virtual Acoustics and Hearing Research by the Example of Binaural Synthesis*, PhD thesis, Technische Universität München (2013)

Völk F.: Updated analytical expressions for critical bandwidth and critical-band rate. In *Fortschritte der Akustik, DAGA 2015* (Dt. Gesell. für Akustik e. V., Berlin, 2015)

Völk F., H. Fastl: Locating the Missing 6 dB by Loudness Calibration of Binaural Synthesis. In *131st AES Convention* (2011) (Convention Paper 8488)

Völk F., H. Fastl: Physical correlates of loudness transfer functions in binaural synthesis. In *21st Int. Congress on Acoustics (ICA)* (2013)

Völk F., A. Gottschalk, H. Fastl: Evaluation of Binaural Synthesis by Minimum Audible Angles. In *Fortschritte der Akustik, DAGA 2012*, 321–322 (Dt. Gesell. für Akustik e. V., Berlin, 2012)

Völk F., F. Heinemann, H. Fastl: Externalization in binaural synthesis: effects of recording environment and measurement procedure. In *Acoustics '08*, 6419–6424 (2008)

Völk F., A. Dunstmair, T. Riesenweber, H. Fastl: Bedingungen für gleichlaute Schalle aus Kopfhörern und Lautsprechern (Requirements for equally-loud sounds from headphones and loudspeakers). In *Fortschritte der Akustik, DAGA 2011*, 825–826 (Dt. Gesell. für Akustik e. V., Berlin, 2011)

Völk F., M. Straubinger, L. Roalter, H. Fastl: Measurement of Head Related Impulse Responses for Psychoacoustic Research. In *Int. Conference on Acoustics NAG/DAGA 2009*, 164–167 (Dt. Gesell. für Akustik e. V., Berlin, 2009)

von Békésy G.: A new audiometer. Acta Oto-laryngol. **35**, 411–422 (1947)

Wenzel E. M., M. Arruda, D. J. Kistler, F. L. Wightman: Localization using nonindividualized head-related transfer functions. J. Acoust. Soc. Am. **94**, 111–123 (1993)

Wightman F. L., D. J. Kistler: Headphone simulation of free-field listening. I: Stimulus synthesis. J. Acoust. Soc. Am. **85**, 858–867 (1989a)

Wightman F. L., D. J. Kistler: Headphone simulation of free-field listening. II: Psychophysical validation. J. Acoust. Soc. Am. **85**, 868–878 (1989b)

WindAcoustics Suite: *Modular software package for auditory-adapted signal and system analysis, real-time audio-signal processing, audio recording, and sound measurements.* (WindAcoustics UG [haftungsbeschränkt], www.windacoustics.com, 2015)

Zwicker E., R. Feldtkeller: Über die Lautstärke von gleichförmigen Geräuschen (On the loudness of uniform sounds). Acustica **5**, 303–316 (1955)