

Control of earphone produced binaural signals

Hammershøi, Dorte; Hoffmann, Pablo F.

Published in:
Acustica United with Acta Acustica

Publication date:
2011

Document Version
Early version, also known as pre-print

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Hammershøi, D., & Hoffmann, P. F. (2011). Control of earphone produced binaural signals. *Acustica United with Acta Acustica*, 97(Supplement 1), S 135.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Control of earphone produced binaural signals

Dorte Hammershøj, Pablo F. Hoffmann
Section of Acoustics, Aalborg University, Aalborg, Denmark.

Summary

While most people keep a high attention to the significance of the binaural recording method, whether it is e.g. individual or non-individual (as e.g. artificial head recording), many pay less attention to the type of earphone used to reproduce the binaural signals, and to the accurate control of the ideal 1:1 reproduction of the signals at the eardrum. This paper identifies and discusses two special cases of earphone reproduction. Further work on the analysis and quantification of calibration errors is planned.

PACS no. 43.60.Dn, 43.66.Pn

1. Introduction

Various methods exist, which aim at an authentic reproduction of a listening event, including the authentic experience of space, direction and distance to the sound source(s), and any reverberance, echoes or similar qualities of the original acoustic space.

One group of methods is based on the delivery of the original eardrum signals of the listener for the listener during playback (e.g. [1, 2, 3, 4, 5]). Binaural signals that represent the user eardrum signals are either recorded in the ears of the listener or synthesized by signal processing to render eardrum equivalent signals. Such binaural signals would ideally carry all the original or desired spatial information for an authentic reproduction of the listening event. The aim of any system for reproduction of binaural signals is therefore to deliver the signals undistorted to the eardrum of the listener.

The concept of the binaural technique has been understood for a while, and mastered to near perfection in various laboratory setups. The best results have been demonstrated for recordings, which are made in the ears of each individual listener, and reproduced through earphones, for which the characteristics have been measured on each individual listener, and the reproduction calibrated using carefully designed personal electronic equalization filters (e.g. [6]).

For most scenarios in the real world, it will typically be impossible to use individual recordings for each individual listener. Many studies have therefore investigated the significance of using non-individual recordings, trying to quantify the deterioration in various ways (e.g. [7]).

The most general compromise is to use an artificial head for sound recording or for measurement of the head-related transfer functions, which are required for binaural synthesis (e.g. [8]). Various artificial heads have been developed with the purpose of determining the best compromise for generic recording of binaural signals [9, 10]. It is generally understood that non-individual recordings or synthesis will never provide the same performance, as if the recording or synthesis is based on the listener's own ear acoustics.

Certain types of perceptual confusions predominate only when the listener is instructed to keep his/her head still. A great deal of attention has been given to e.g. front-back confusions (e.g. [11, 12]), which fortunately seem to occur only in the static listening situation. With appropriate tracking of the listener it is possible to facilitate a dynamic binaural synthesis, which generally facilitates a more authentic listening experience, and almost eliminate front-back confusions [13].

Binaural signals are typically reproduced by the use of circumaural or supraaural headphones or earphones of similar type as used for HiFi (e.g. [14]) or audiological testing (e.g. [15]). The primary advantage of the earphones is the effective channel separation, which ensures that the left ear is only reached by the left ear signal and vice versa for the right ear. This is not obtained when the signals are reproduced by loudspeakers, because the sound from the left ear will also reach the right ear. The crosstalk is an integral part of the traditional stereo technique, but is an undesired component in the reproduction of binaural signals.

It is possible to electronically counter-compensate the cross-talk from the loudspeaker opposite to the receiving ear (e.g. [16, 17, 18, 19]), but the setup is always quite sensitive to the physical surroundings, and the exact position of loudspeakers and listener (e.g. [20, 21]).

In recent years the use of earphones has increased with the technological advances in the field of portable music players, smartphones etc. The most common earphones are small insert earphones, which are either placed in the conchae in the outer part of the human ear, or inserted into the ear canal, where the earphone will typically fill out the outer part of the canal. These types of earphones were not popular only 5-10 years back, but they are now more commonly used, and also provide a much improved physical fit, better sound quality and appear more robust in use.

The use of insert earphones for reproduction of binaural signals, and possibly even for the development of general hear-through assistive hearing devices ([22, 23, 24]), calls for a review of the theory of earphone calibration. Two special cases, using i) blocked-entrance recording and semi-open-entrance calibration are discussed, or ii) blocked-entrance recording and eardrum equivalent calibration, and examples of relevant transfer functions given. The theory applies also for the measurement and application of head-related transfer functions in binaural synthesis.

2. Theory

Møller [2] presented the relevant theory for the analysis of the sound transmission from the binaural recording to the reproduction over earphones. It was shown that recordings could be made at any point in the ear canal (or outer ear) *from which* the further transmission to the eardrum would not depend on direction of incidence.

In [25] the directional dependence from various points in the ear canal to the eardrum was analyzed. It was concluded that recordings could be made at the very entrance of the ear canal, even when blocked. It was also argued that it was favorable to record at the blocked entrance, where the impact of individual idiosyncrasies was small compared to e.g. recording at the entrance of the open ear canal, or at the eardrum. This method is used widely both by academia and industry, and the examples here included is therefore based on non-individual blocked-entrance recording (or measurement of HRTFs).

In the ideal case, where the listener's head can also be used for recording (or measurement of HRTFs), one simply has to design a headphone equalization filter, which has the inverse characteristics of the transfer function of the earphones, when measured in the physical point in the ear canal, where the recording is made. This simplification presumes that the acoustical coupling of the earphones and ear canal is similar to the coupling of the ear canal to the free field. It was shown in [14] and [15] to be the case for several earphones.

In the non-ideal case, where the recording is made on one head, the calibration of the headphone on another, and the listener is neither of the two, then one

has to analyze more accurately, which error terms the various compromises induce. Using the methods presented in [2], it is possible to quantify the resulting errors and discuss this in view of the given situation.

In the following, P will represent frequency domain representations of sound pressure signals or transfer functions, and the subscripts ED will refer to eardrum signals or transfer functions, OE to open-entrance signals or transfer functions, sOE to semi-open entrance recordings or transfer functions, and BE to blocked entrance signals or transfer functions. Subscripts FA will refer to the free-air situation (either during listening in the real life situation, during recording or during measurement of HRTFs) and EP will refer to the listening situation, when the earphone is positioned in the listener's ear (either during measurement of the transfer characteristics or during listening).

The subscripts *rec*, *cal*, *data*, and *list* are used to indicate, whether a given recording or transfer function stem from the recording head, the head used for the calibration, from data in literature or standards, or from the listener. An asterisk '*' indicate recordings or transfer functions, which are estimated.

In the present paper, two cases of in-ear reproduction situations are identified and discussed generally. Detailed analyses of the various error terms will follow in later communication.

2.1. Recording at semi-open entrance, calibrating for blocked-entrance

As argued above, when binaural reproduction is considered for many listeners, blocked-entrance recordings are desirable because they ameliorate individual idiosyncrasies. However, some situations may require binaural recordings on human ears with the additional constraint that the recording system should not disrupt normal hearing. In this recording situation, one typically places miniature microphones at the entrance to the open ear canal. Even though the microphones are small, their dimensions are still not small enough so that their effect on the ear canals' sound field can be ignored. This recording situation is represented by $P_{FA,sOE,rec}$.

A processing scheme has earlier been proposed to transform open-entrance recordings $P_{FA,sOE,rec}$ into blocked-entrance versions $P_{FA,BE,rec}^*$ [24]. The idea is based on the assumption that the ratio of $P_{FA,BE,rec}$ and $P_{FA,sOE,rec}$ represents the most individual part of the sound transmission, which however is independent of direction and can be determined for a given well-defined free-air situation before the recording start. The ratio can subsequently be used in the design target for filters that compensate for the directionally independent but distinct features of the individual from which the recordings are made. The processing scheme is expressed in the following equa-

tion, where FA' denotes the given free-air sound field used:

$$P_{FA, BE, rec}^* = \frac{P_{FA', BE, rec}}{P_{FA', sOE, rec}} \cdot P_{FA, sOE, rec} \quad (1)$$

It was shown in [14] that a range of traditional headphones present a load to the ear which compare to that of a free field. Such a headphone can in principle substitute the free-air sound field for the determination of the blocked vs. semi-open entrance ratio, if a well-defined free-air sound field is hard to establish. The alternative processing scheme can be described in the following equation:

$$P_{FA, BE, rec}^* = \frac{P_{EP, BE, rec}}{P_{EP, sOE, rec}} \cdot P_{FA, sOE, rec} \quad (2)$$

Fig. 1 shows two examples of the ratio $P_{EP, sOE, rec}/P_{EP, BE, rec}$ from two different listeners. The shape of these ratios are comparable to those measured in [25] up to approximately 6–7 kHz. In these examples, the magnitudes of the peak and notch observed in the range 2–4 kHz are less pronounced than those reported in [25], and this is attributed to the effect of the microphones' size on $P_{EP, sOE, cal}$.

Now, that we have a scheme for transforming semi-open entrance recordings into blocked-entrance recordings, one needs to equalize the reproduction channel for the particular use of insert earphones in the same way as any other blocked-entrance recording made for in-ear delivery.

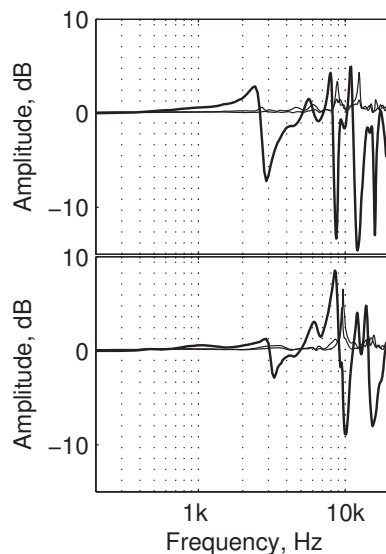


Figure 1. Open-to-blocked ratios. Top and bottom panels show data from two different listeners respectively. Thick lines indicate $P_{EP, sOE, rec}/P_{EP, BE, rec}$ ratios computed from the mean of 5 repeated measures for each transfer function. Thin lines show the standard deviations.

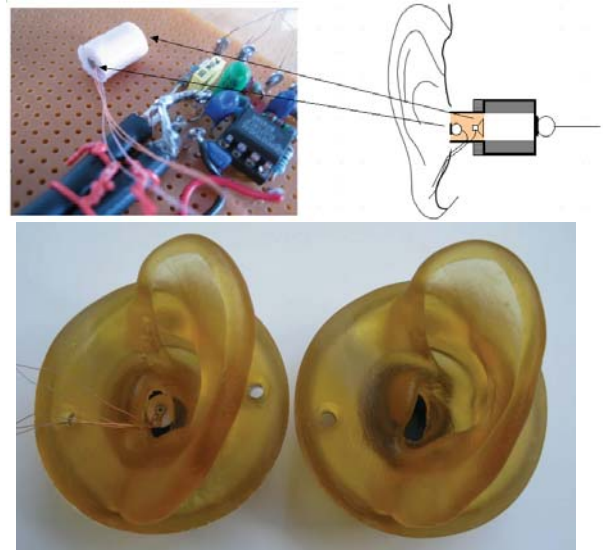


Figure 2. **Top row:** Illustration of the hear-through device prototype and schematic of its placement in the ear coupler used for measurements. **Bottom row:** Pinnae simulators with one sample having the hear-through device flush with the entrance to the ear canal.

2.2. Recording at blocked-entrance, calibrating at eardrum

Among the applications that demand the reproduction of binaural signals over insert earphones are those using hear-through assistive devices, e.g. auditory augmented environments [23]. A hear-through device should allow for simultaneous capturing and reproduction of binaural signals, with the additional characteristic of being able to superimpose artificial signals onto the real-life recordings. We have conducted a set of preliminary measures on a prototype in order to evaluate the degree of acoustic transparency one can achieve with off-the-shelf materials.

Preliminary measurements were conducted in an anechoic chamber using an artificial head. The top row of Fig. 2 shows an illustration of the hear-through device prototype and an schematic indicating how the device is placed in the extensions of the ear coupler used for measurements (IEC 60711 ear simulator G.R.A.S. Type RA0045). The components of the hear-through device were basically a miniature electret microphone (Knowles electronics FG23629), a preamplifier, a miniature loudspeaker (Knowles Electronics FK-23451-000), and the necessary custom-made circuit for the power supply (9V DC). The microphone and speaker were embedded in a cylinder made of silicon material and shaped to perfectly fit the ear canal extension used in the artificial head (further description below). The output from the microphone amplifier is directly connected to the input of the amplifier of the miniature loudspeaker. This allowed to control the input as well as the output gains separately.

For the simulation of the external ear two rubber pinna replica were employed during all measurements. Ear canal extensions were coupled to both pinnae so that the outer surface of the extension was flush with the entrance to the ear canal. The hear-through device was mounted on one of the extensions. This is illustrated in Fig. 2. The use of two pinnae and ear canal extensions instead of only one was primarily for convenience since placing the hear-through device in and out for each measurement condition proved to be too time consuming and prone to increase the inherent error in the measurement system. Ear canal extensions were coupled to the IEC 60711 ear simulator (G.R.A.S. Type RA0045) via a retaining collar. With this setup both ear replica could be easily assembled and disassembled depending on the condition to be measured.

Fig. 3 shows results for sound sources located in front of the artificial head and 90 degrees to the left. For frequencies below 2.5 kHz the response of the hear-through device exhibited a roll-off until reaching a floor level around 1 kHz of about -18 to -20 dB relative to the response of the natural condition without wearing the hear-through device. This pattern may be attributed to the possible leaking of low frequencies in spite of the efforts done to perfectly fit the embedding of the hear-through device to the ear coupler. For frequencies above approximately 2.5 kHz the hear-through device showed responses with shapes comparable to those observed for the natural condition, with a steeper high-frequency roll off for frequencies above 12 kHz. Between 2.5 and 9 kHz the frequency response of the hear-through device matched the response of the natural condition within ± 5 dB, with a best case scenario of ± 3 dB within 2.8 and 8 kHz.

3. Discussion

For the case discussed in Section 2.1, a blocked-entrance recording was rendered through post-processing of the original semi-open-entrance recording. For the correct 1:1 transmission to be obtained, the next step is to provide an earphone calibration, where the inverse of the headphone transfer function is determined for the same physical point, i.e. at the blocked entrance. This cannot be measured or estimated directly for insert earphones, and has no physical equivalent. The alternative is to measure in-situ at eardrum for the listener, which however is highly unfavorable. The only relevant alternative is therefore to measure the transfer characteristics on a standardized ear simulator, as the IEC 60711.

In Section 2.1, the recordings were deliberately "transferred" to blocked entrance by Eq. (1) or (2), and this gives two options for the calibration, 1) to

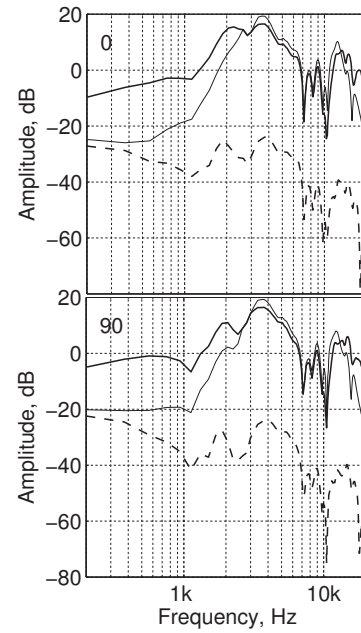


Figure 3. Hear-through device prototype measurement for the frontal direction (top panel) and the left direction (bottom panel). Thick and thin lines indicate frequency responses for the condition without wearing the device and for the condition with the hear-through device on. Dashed lines indicate frequency responses for the condition with the hear-through device off.

"transfer" the blocked-entrance recording to ear simulator "eardrum" (similar method used in [27]):

$$P_{FA,ED',cal}^* = \frac{P_{FA,ED',data}}{P_{FA,BE,data}} \cdot P_{FA,BE,rec}^* \quad (3)$$

where ED' symbolize the "eardrum" of the ear simulator, or 2) to transfer the earphone characteristics $P_{EP,ED',cal}$ to blocked entrance:

$$\frac{P_{EP,BE',cal}^*}{E_{EP}} = \frac{P_{FA,BE',data}}{P_{FA,ED',data}} \cdot \frac{P_{EP,ED',cal}^*}{E_{EP}} \quad (4)$$

where BE' represent the blocked entrance of the ear simulator. Option 2) will have no physical equivalent for verification, but is still a legitimate option which will facilitate cancellation of the directionally independent term, which is involved during calibration of the earphone ($P_{FA,ED',cal}/P_{FA,BE',cal}$), if the same ear simulator is used both for calibration and in the determination of data (from e.g. standards).

In Section 2.2. it was demonstrated that the hear-through device gives an almost transparent transmission for frequencies above 2.5 kHz, and thus do not call for a revised design target, as the case described above. The primary challenge in this context appear to be the control of low frequencies, where the significance of individual idiosyncracies can be ignored.

4. Conclusion

Two scenarios have been presented for which the significance of individual outer ear acoustics plays a role, and for which the earphone calibration is not trivial. Further work on the analysis and quantification of the various transmission elements and the errors induced during calibration remains.

Acknowledgement

This work has in part been supported by the EU FP6 project "Multimodal Interfaces for Capturing and Transfer of Skill" (known by the acronym SKILLS; IST-FP6 #035005, <http://www.skills-ip.eu>).

References

- [1] J. Blauert: Spatial hearing. The psychophysics of human sound localization. The MIT Press, USA-Cambridge MA, 1996.
- [2] H. Møller: Fundamentals of binaural technology. *Applied Acoustics* **36**(3/4) (1992) 171–218.
- [3] F. Wightman, D. Kistler: Measurement and validation of HRTFs for use in hearing research. *Acta Acustica united with Acustica* **91**(3) (2005) 429–439.
- [4] D. Hammershøi, H. Møller: Methods for binaural recording and reproduction. Review article in *Acta Acustica united with Acustica* **88**(3) (2002) 303–311.
- [5] D. Hammershøi, H. Møller: Binaural technique - Basic methods for recording, synthesis and reproduction. - In: *Communication Acoustics*. J. Blauert (ed.). Springer-Verlag, 2005.
- [6] H. Møller, M. F. Sørensen, C. B. Jensen, D. Hammershøi: Binaural technique: Do we need individual recordings? *J. Audio Eng. Soc.* **44**(6) (1996) 451–469.
- [7] H. Møller, C. B. Jensen, D. Hammershøi, M. F. Sørensen: Using a typical human subject for binaural recording. *Proceedings of 100th Audio Eng. Soc. Conv.*, Copenhagen, May 11–14, 1996, preprint 4157, pp. 1–18. Abstract in: *J. Audio Eng. Soc.* **44**(7/8) (1996) 632.
- [8] H. Møller, M. F. Sørensen, D. Hammershøi, C. B. Jensen: Head-related transfer functions of human subjects. *J. Audio Eng. Soc.* **43** (1995) 300–321.
- [9] H. Møller, D. Hammershøi, C. B. Jensen, M. F. Sørensen: Evaluation of artificial heads in listening tests. *J. Audio Eng. Soc.* **47**(3) (1999) 83–100.
- [10] P. Minnaar, S. K. Olesen, F. Christensen, H. Møller: Localization with binaural recordings from artificial and human heads. *J. Audio Eng. Soc.* **49** (2001) 323–336.
- [11] P. Zahorik, P. Bangayan, V. Sundareswaran, K. Wang, C. Tam: Perceptual recalibration in human sound localization: Learning to remediate front-back reversals. *J. Acoust. Soc. Am.* **120**(1) (2006) 343–359.
- [12] P. X. Zhang, W. M. Hartmann: On the ability of human listeners to distinguish between front and back. *Hearing research*. **11**(1–2) (2009) 30–46.
- [13] J. Sandvad: Dynamic aspects of auditory virtual environments", *Proceedings of 100th Audio Eng. Soc. Conv.*, Copenhagen, May 11–14, 1996, preprint 4226, pp. 1–14. Abstract in: *J. Audio Eng. Soc.* **44**(7/8) (1996) 644.
- [14] H. Møller, D. Hammershøi, C. B. Jensen, M. F. Sørensen: Transfer characteristics of headphones measured on human ears. *J. Audio Eng. Soc.* **43** (1995) 203–217.
- [15] D. G. Ciric, D. Hammershøi: Coupling of earphones to human ears and to standard coupler. *J. Acoust. Soc. Am.* **120** (2006) 2096–2107.
- [16] M. R. Schroeder, B. S. Atal: Computer simulation of sound transmission in rooms. *IEEE Intl. Conv. Rec.* **11** (1963) 150A–155.
- [17] B. S. Atal, M. R. Schroeder: Apparent sound source translator. U.S. Patent 3,236,949, Feb. 1966.
- [18] D. H. Cooper, J. L. Bauck: Prospects for transaural recording. *J. Audio Eng. Soc.* **37** (1989) 3A–19.
- [19] H. Møller: Reproduction of artificial-head recordings through loudspeakers. *J. Audio Eng. Soc.* **37** (1989) 30–33.
- [20] D. B. Ward, G. W. Elko: Optimum loudspeaker spacing for robust crosstalk cancellation. *Proceedings of the 1998 International Conference on Acoustics, Speech and Signal Processing*. **6** i–lxiii.
- [21] Y. Lacouture Parodi: A systematic study of binaural reproduction systems through loudspeakers: A multiple stereo-dipole approach. Ph.D. dissertation, Aalborg University, 2010.
- [22] D. S. Brungart, A. J. Kordik, C. S. Eades, B. D. Simpson: The effect of microphone placement on localization accuracy with electronic pass-through earplugs. In *Applications of Signal Processing to Audio and Acoustics*, 2003 IEEE Workshop, New Paltz, New York, USA, (2003) 149–152.
- [23] A. Härmä, J. Jakka, M. Tikander, M. Karjalainen, T. Lokki, J. Hiipakka, G. Lorho: Augmented reality audio for mobile and wearable appliances. *J. Audio Eng. Soc.* **52** (2004) 618–639.
- [24] D. Hammershøi, P. Hoffmann, S. K. Olesen, P. Rubak: Capturing blocked-entrance binaural signals from open-entrance recordings. Presented at *Acoustics' 08: the 155th Acoust. Soc. Am. Meeting*, 5th Forum Acusticum, 9th Congr. Acoustique, June 29 to July 4, 2008, Paris, France. Abstract in: *J. Acoust. Soc. Am.* **123** (2008) 3246 (paper 2pNSa11). Abstract in: *Acta Acustica united with Acustica* **94** (2008) S278.
- [25] D. Hammershøi, H. Møller: Sound transmission to and within the human ear canal. *J. Acoust. Soc. Am.* **100** (1996) 408–427.
- [26] D. S. Brungart, B. W. Hobbs, J. T. Hamil: A comparison of acoustic and psychoacoustic measurements of pass-through hearing protection devices. In *Applications of Signal Processing to Audio and Acoustics*, 2007 IEEE Workshop, New Paltz, New York, USA, (2007) 70–73.
- [27] D. Hammershøi, H. Møller: Determination of noise immission from sound sources close to the ears. *Acta Acustica united with Acustica* **94**(1) (2008) 114–129.