INTERNATIONAL SYMPOSIUM on MUSICAL and ROOM ACOUSTICS

ISMRA 2016

September 11-13, 2016

La Plata, Buenos Aires, Argentina

Evaluation of Concert Halls / Opera Houses: Paper ISMRA2016-30

Accurate reproduction of non-individual binaural recordings without head tracking through individual headphone equalization

David Griesinger[a]

[a] David Griesinger Acoustics USA dgriesinger@verizon.net

Abstract

Acoustic research of all kinds desperately needs methods that allow accurate and instant comparisons of the sound in different seats and different venues. Binaural recording could provide such a reference if it can be made accurate enough. Early work by Schroeder and others showed that recording at the eardrums of a listener, and playing back with individual eardrum equalization, captures sound accurately without head tracking. But their methods were cumbersome. We show that the major problems with binaural reproduction do not lie with individual pinna functions, but with the highly individual resonances in the concha and ear canals. The eardrum response of different individuals with the same headphones varies by +-8dB in the vital frequency range of 500Hz to 6000Hz. This talk describes the physics responsible for these variations and how headphones affect them. We have developed a software application that can quickly and simply equalize these variations through an equal-loudness technique. Once equalized for an individual music of all types, especially binaural recordings, can be perceived as frontally localized and startlingly beautiful.

Keywords: Concert Halls, Binaural Recording, Headphones

INTERNATIONAL SYMPOSIUM on MUSICAL and ROOM ACOUSTICS
ISMRA 2016
September 11-13, 2016
La Plata, Buenos Aires, Argentina

Accurate reproduction of non-individual binaural recordings without head tracking through individual headphone equalization

1 Introduction

Timbre is the vital clue the ear and brain need to localize sounds of all types, but timbre, as perceived by the eardrum, depends dramatically on the pinna, concha, and ear canal resonances that concentrate sound pressure on that surface. For the author this pressure increase is 18dB at 3000Hz. But these resonances are highly individual, sufficiently so that they can be used as fingerprints [1]. But any change in the impedance at the entrance to the ear canal alters these resonances, and many headphone types simply eliminate them. Even if current measurement techniques for headphones did not ignore these resonances, the resonances are sufficiently different for different individuals that we believe a universally accurate equalization for headphones does not exist.

We have developed a software application that allows a user to accurately match the timbre of a headphone to that of a frontal loudspeaker, using the user's own eardrum as a microphone. The procedure is simple, painless, and inexpensive.

2 A brief history of headphone research

In September of 1940 Leo Beranek was given the job of directing a laboratory at Harvard charged with solving the severe communication problems aboard heavy bombers. He had to devise methods of measuring and standardizing the frequency response of headphones, and figuring out how to attach them comfortably to a pilot while minimizing noise intrusion. But part of the project was finding ways of testing the progress on live subjects, for which purpose a separate psychological laboratory was set up under Smitty Stephens. One of the first researchers there was found and hired by Beranek – J.C.R. Licklider. [2]

Together the group made substantial progress in both hardware and testing. The goal was the best possible communication of information. The testing was done by intelligibility tests of speech in the presence of recorded aircraft noise at full volume. The subjects were conscientious objectors, young men similar to soldiers and airmen. The project was thorough, goal oriented, and successful.

It is reasonable to assume that if a person could make a recording of the sound pressure at each eardrum, and then play it back in such a way that the sound pressure was precisely duplicated, that the original sound impression would be exactly reproduced. Manfred Schroeder

[3] demonstrated this technique using a dummy head by Mellert [4] that attempted eardrum equalization from a lateral source. Schroeder's playback system employed loudspeakers, mathematically derived crosstalk canceling filters, and steel probe tubes at the listener's eardrums. This reproduced the "eardrum" pressure of the dummy head precisely for each listener. The system was deemed successful at the time, although the choice of equalizing a dummy head from a lateral source is not optimum. Worse, Schroeder attempted to emulate an orchestra with two stereo speakers on stage, which is disastrous.

These experiments were continued at IRCAM in Paris. In this case the recording was from steel probe tubes at the subject's eardrums, and playback was with Schroeder's crosstalk cancellation method. Live musicians were used. Subjects needed to have their heads in a clamp both for the recording and the playback, or risk damage to the eardrums. But the IRCAM experiments are reported to have been very successful. The original sound field was reproduced exactly [5]. Both Schroeder and the IRCAM experiments avoided the problem of headphone equalization by not using headphones, but their system re-created the eardrum pressures of each subject individually.

Since then the goal of headphone research has been to find a standard procedure that results in optimum sound for any individual. An intuitively promising procedure was standardized as DIN 45-619. DIN 45-619 used an equal loudness procedure whereby a listener switched rapidly between a 1/3 octave noise band from one or more laterally placed loudspeakers, and the same noise band reproduced through headphones. An attenuator in the headphone circuit was adjusted until the two tones had equal loudness. The attenuator settings, when duplicated with an equalizer, are the needed corrections to equalize the headphones. As we mentioned, the choice of lateral sources is not optimal, and averaging many listeners to find an ideal equalization for all of them is unlikely to work for everyone. But if a frontal source was used to find an equalization for a particular individual, DIN 45-619 could work very well. It would accurately reproduce the timbre of the loudspeaker at the eardrum of the listener. Unfortunately the procedure is tedious and time-consuming.

DIN 45-617 was abandoned in a quest to find simpler measurement methods. Current standards favour measuring the sound pressure at the entrance to the ear canal. See references [6], [7], [8], and [9]. See also ITU-T Recommendation P.57 type 3.3, and IEC coupler 60711, which is the current standard coupler for both Kemar and the B&K HATS. The "ear canal" in this coupler is a straight cylinder 1cm long, just long enough to test an insert phone. None of these standards attempt to duplicate the impedance of a human ear canal entrance, the resonances of the canal, nor the resistive impedance of the eardrum. In our view they are fundamentally flawed.

The choice of measuring the pressure at the entrance to an ear canal and not at the eardrum was based on the assumption that if the sound pressure from a headphone at an ear canal entrance could be optimized and standardized, then the average listener would hear a natural timbre. We have found no headphone for which this assumption is correct.

In an attempt to minimize the effects of individual variations in the directional dependent component of HRTFs, the concept of equalizing headphones to emulate an omnidirectional sound field – so called "diffuse field equalization"- was developed. In our view this is also an error. Human hearing is most directionally acute in the forward direction, and almost all sources we are intensely interested in recording and playing back through headphones are frontal. It is the forward equalization that we need to get right.

Current equalization standards are chasing a will-o'-the-wisp. They purport to put scientific support behind headphone quality even though every individual will hear a different timbre. This is convenient for manufacturers, who are free to sell headphones on the basis of style and price. For the general public it does not matter if the timbre is correct or not, as the outer hair cells in the basilar membrane act as a continuous multi-band compressor. The ear adapts quickly minutes to even gross errors in frequency response. But this does not mean the sound at the eardrum has natural timbre or produces forward localization.

3 Variation of HRTFs with azimuth and elevation

Figure one shows Bill Gardner's MIT Kemar data [10] for the contralateral ear at zero degrees elevation and 0, 30, and 60 degrees azimuth. Figure two shows the same for zero degrees azimuth and 0, 30, and 60 degrees elevation. Notice that the variation with azimuth in figure one is largely confined to head-shadowing. This is why stereo loudspeaker reproduction works a well as it does. The high frequency localization notch is almost constant. In figure 2 it is the elevation notches above 6kHz that vary. The timbre of the middle frequencies do not change very much.

(In figure 1 The MIT data has been equalized so the frontal HRTF is frequency flat up to the deep vertical localization notch. This is NOT the eardrum pressure of a human, although this equalization can be useful for recording.)

Figure 1 shows that we need not vary the frequency notches above 6kH to reproduce azimuth. But considerable experience shows that to get a *forward* and *frontal* image we need to accurately reproduce the frequencies *below 6kHz at the eardrum*. The careful equalization of the MIT Kemar dummy makes this look easy. But the eardrum pressure of a human is grossly nonlinear and individual from 500Hz to 6kHz. The pinna, concha and ear canal form a horn, which has evolved to concentrate sound energy on the eardrum. The horn is resonant and varies strongly with individuals.

INTERNATIONAL SYMPOSIUM on MUSICAL and ROOM ACOUSTICS
ISMRA 2016
September 11-13, 2016
La Plata, Buenos Aires, Argentina

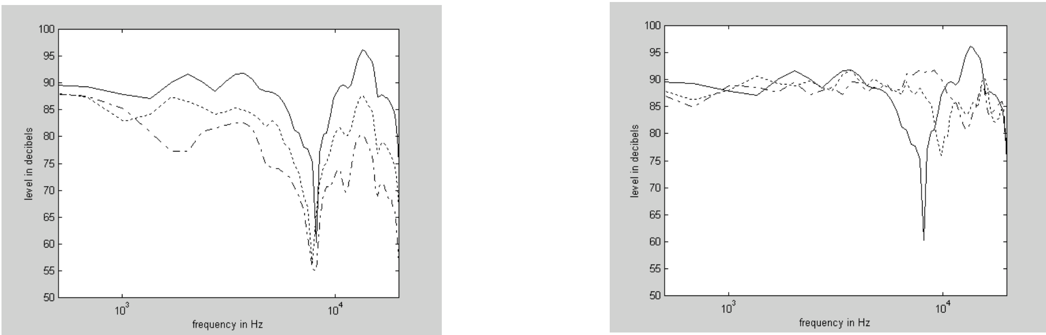

Figure 1: MIT Kemar data for the contralateral ear at zero degrees elevation and 0, 30, and 60 degrees azimuth, and data for zero degrees azimuth an 0, 30, and 60 degrees elevation.

4 Ear canal resonances

Like a trumpet, the concha, ear canal, and eardrum form a resonant instrument. For the author's ears two parametric filters are required to model the resonances, one at 3000Hz and one at 2700Hz. The pressure at the eardrum is boosted 18dB at 3000Hz. Like all trumpets, the frequencies and amplitudes of these resonances are altered when anything changes the impedance at the bell. All headphones the author has tested alter these resonances. Insert phones eliminate them. See figure 2.

Figure 2: Right: An enlargement of the author's right ear, with Ville Pulkki pointing to the opening of the horn. Left: a negative cast of the author's left ear canal that clearly shows the horn shape.

In [8] Møller discusses individual equalization of headphones, but he does not measure the eardrum pressure. He adjusts the headphone response by measuring at the entrance to the ear canal where blocked ear canal measurements of HRTFs were made. [The author finds his description of this process confusing. Was the headphone measurement made with the ear canal open or closed?] We assume the ear canal was open. Møller states in [9] "When aiming at knowledge about the actual sound pressure at the eardrum of a specific subject, no alternative to eardrum measurements exists. … Identical pressure divisions only exist — in principle — when the radiation impedance is undisturbed, which requires that no object is mounted close to the ear. Although we believe that most headphones do affect the radiation impedance, we have in another study, seen that the effect of many traditional headphones is not so severe that it significantly alters the pressure division." In [9] Møller refers to these headphones as "FEC" phones, and states that many headphones meet this requirement.

Our data from loudness matching shown in figure 4 shows that at least for all the headphones we tested the criterion is not met. As an additional check, we measured the three most open headphones available to the author at this time to see if the sound pressure near the entrance to the ear canal could predict the pressure at the eardrum. If the headphone impedance at the ear canal entrance was very close to the impedance of free air, this should be the case.

One of the author's dummy heads has accurate castings of his own ears all the way to the eardrum, the impedance of which is modelled with a long tube. The microphone output closely matches the sound pressure at the author's eardrums, both for a free-field and with headphones. We measured the response from a frontal loudspeaker to the eardrum of the author's dummy, and at the same time from a microphone at the mouth of the ear canal. We repeated the measurements with three headphones. The sound pressure at the mouth of the ear canal did not predict the pressure at the eardrum for any of them.

Figure 3 left: Blue: the pressure at the eardrum position from a free field frontal plane wave. Red: The pressure slightly inside the ear canal opening. The blue curve is typical for the author's ears. The measurement was not anechoic. Right: The same measurement from a Stax model 303 Classic electrostatic headphone. Similar but different results were obtained from AKG 701 and 501 headphones.

It can be seen from figure 3 that the difference between the red and blue curves with the headphone was not the same as that from the free field. We conclude that individually equalizing headphones from a measurement near the ear canal mouth is very unlikely to be effective, whether the ear canal is blocked or open. The equalizations found for our subjects found by equal loudness at the eardrum with the AKG 701 and the Sennheiser 600 shown in figure 4 support this conclusion. There is considerable variation between different individuals.

5 Reproducing a natural timbre

A frequency-flat frontal loudspeaker is the essential reference for timbre in the audio world. Such speakers are considered essential for audio production, and Toole has found [11] that loudspeakers with the most linear on and off axis response are preferred in blind listening tests both by experts and the general public. We believe the same is true for headphones. But to achieve a flat response for headphones the frequency response at the listener's eardrum must match that of a frequency linear frontal loudspeaker. If we do this carefully enough the listener will perceive standard recordings as frontal, and accurate.

Is it possible to make binaural recordings with a non-individual dummy head, and play them successfully to other individuals? Binaural recordings made from my head contain my individual elevation data at frequencies above 6kHz. These will not necessarily match those of another listener. But most foreground signals of interest are frontal, and the graphs in figure one show that we do not need to precisely reproduce the listener's individual elevation data above 6kHz to achieve plausible azimuth. But we need to reproduce accurate timbre if there is to be frontal localization.

6 Headphone equalization through loudness matching

We have developed a software application that facilitates individual equalization of headphones through a loudness matching procedure similar to DIN 45-617 and the ISO 266 2003 procedure for determining equal loudness curves. A subject sits in front of a frequency linear loudspeaker that produces signals that alternate once a second between tones or noise bands at a reference frequency, and tones or noise bands at a test frequency. We chose a reference frequency of 500Hz. The subject can select to use sine tones, noise bands, or filtered harmonic tones as test signals. Noise and harmonic tones give the best results.

The subject adjusts a 27 band 1/3 octave Q=5 graphic equalizer until the test signals match the loudness of the reference. The equalization, in dB, that results is their personal equal loudness data. They then put on the headphones and find their personal equal loudness data for that headphone. The difference between their headphone data and their loudspeaker data is the desired headphone equalization. This is loaded into the equalizer. We find it is additionally useful to have the subject balance the perceived left-right azimuth of the headphone tones. Not all ears have the same equal loudness curves, and neither do headphone drivers. People with some mild hearing loss in one ear find the balancing procedure useful (including the author.)

Once the equalization has been found the subject can listen to pink noise or music through their personal equalization. Almost everyone finds the image is frontal, and the timbre of pink noise accurate. Our binaural recordings can be startlingly real. The subject's equalization settings are written as a .txt file, along with their equal loudness data. The app also creates a .wav file of an impulse response of their equalization that can be convolved with music to equalize it for that pair of headphones.

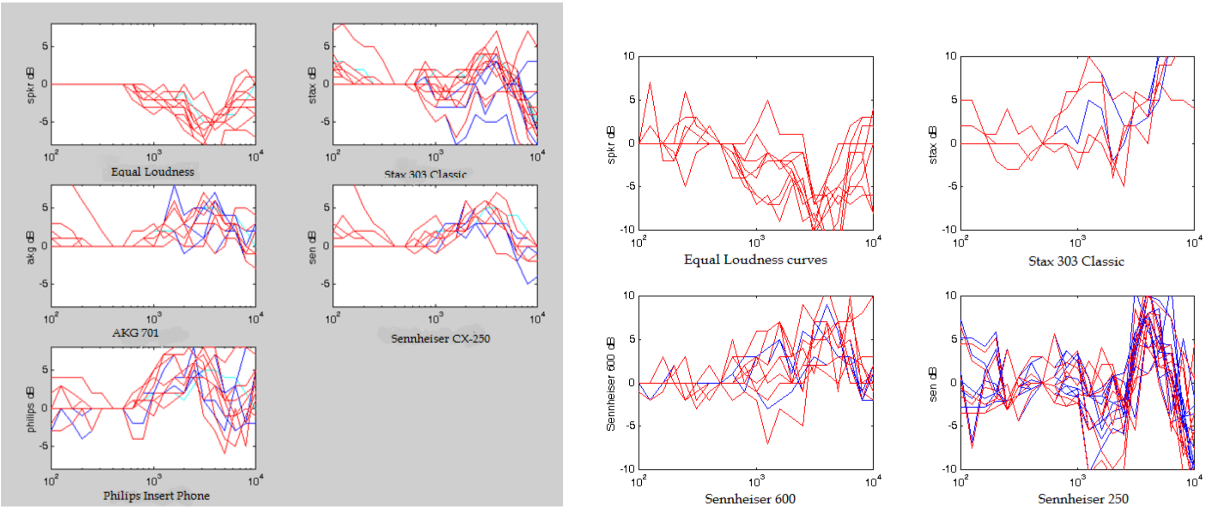

Figure 4 Left: Equal loudness data and the headphone correction needed for four headphones by students at Aalto University. Right: similar data from students at Rensselaer, but this data is the perceived headphone timbre – the inverse of the data from Aalto. Note the large variation in timbre for different individuals. The equal loudness data has the same sign in both data sets.

Our procedure requires a reference loudspeaker, particularly for frequencies above 250Hz. We find lower frequencies can be assumed to be equally loud. With the help of a calibrated smart-phone real-time analyser small single driver speakers can be inexpensively equalized with the app to be flat within one dB from 200Hz to 10kHz.

We have conducted experiments with the headphone app with the help of Ville Pulkki at Aalto University in Finland, and Jonas Braasch at Rensselaer University in the US. Students familiar with sound recording find the procedure easy and fast. Older or more naive subjects take more time to get facile, but they all can do it. The results have been uniformly good. Almost everyone achieves frontal localization. The perception of presence (or “proximity” see [14]) is clear, although the exact distance in both binaural and natural hearing depends more on vision or expectation than any acoustic cue.

7 Accuracy versus preference

Not every subject prefers the individual equalization they find with this method, although the great majority (especially the young students) do. We test the equalization by playing broadband pink noise through the calibrated speaker and having them listen to the same noise through the headphones. If the two are not perceived to have the same timbre we ask them to re-do some of the frequency bands. Eventually they find the timbres to be nearly the same. But they may not like the sound.

The large boost in the sound pressure at the eardrum that corresponds to the dip in the equal loudness curve at ~3kHz is audible, and some subjects may prefer a headphone equalization closer to equal-loudness. But it is not natural, and it does not result in frontal, out of head localization.

Sometimes a subject that is very familiar with a particular headphone is initially unwilling to accept the equalized phone, which seems midrange-heavy. But any doubt by a particular subject in our experimental result disappears when we play one of our binaural recordings, many of which are of great performances in great halls. There is almost always a sense of “being there” and it is quite difficult to get them to turn it off.

8 Conclusions

Our experiments with individual headphone equalization are on-going but sparse. The author lacks access to legions of eager students, and published references on individual headphone equalization *at the eardrum* are hard to come by. Eardrum equalization is widely believed to be either unnecessary or impractical.

But it is simple to prove that it works, and that conventional equalization does not. We believe that the data here, and experience with binaural examples, speak for themselves. Individual equalization of headphones through loudness matching results in accurate timbre and frontal localization without head tracking. I encourage sceptical readers to email me to try the app and hear some binaural examples. They will find that switching individual equalization in and out is very convincing.

With this in mind we offer a few personal conclusions from 30 years of experimentation with individual headphone equalization.

First, the search for a *universal* equalization for headphones has not produced a headphone design that reliably results in frontal, out of head localization. Such an equalization probably does not exist, although some headphone types may be less individually variable than others.

Second, we believe that if head tracking is required for frontal localization with a non-individually equalized headphone the timbre is sufficiently incorrect that there will be errors of judgement both for acoustic research and for balancing a recording. Adding head tracking to an incorrectly equalized headphone only makes the errors in judgement more convincing. It does not correct problems with timbre.

Third, we show in another preprint for this conference that the sound of an ensemble in a hall can be convincingly recreated from a single binaural measurement. A binaural impulse response at a particular seat can be manipulated to create at least six different azimuths. Convolving the results with six of Lokki’s anechoic recordings and listening with individually equalized headphones is accurate and convincing. The effects of different reflections can then be studied by modifying the impulse responses.

Fourth, we believe that the oversimplification of the sound sources used by Schroeder and others to interpret the acoustic measures enshrined in ISO 3382 has led to more confusion than success. We need to start over with more realistic sources and playback systems, paying particular attention to the effect proximity has on the perception of the hall. Lokki is making progress in this field, but binaural technology with individual headphone equalization is simpler, less expensive, and possibly more accurate. It deserves a prominent place in this research. More on this subject can be found in [12], [13], [14], and [15].

References

[1] A link to recent work by NEC on developing personal identification technology based on individual differences in ear canal resonances. *https://www.helpnetsecurity.com/2016/03/10/biometric-tech-uses-sound-to-distinguish-ear-cavity-shape/*

[2] Beranek, L.; Riding the Waves: A Life in Sound, Science, and Industry. *Macmillan*

[3] Schroeder, M., Gottlob, D., Siebrasse, K. Comparative study of European concert halls: correlation of subjective preference with geometric and acoustic parameters. JASA 56, number 4, 1195-1201 (1974)

[4] Mellert, V. Construction of a Dummy Head after New Measurements of Thresholds of Hearing. JASA 51 number 4, 1359-1361 (1971)

[5] Personal communication from Eckhard Kahle

[6] Theile, G.; On the standardization of the Frequency Response of High-Quality Studio Headphones. JASA (34) 12 1986 pp 956-969.

[7] Spikofski, G. The diffuse field transfer function of studio-quality headphones. EBU Review – Technical 229 (1988) pp 111-126.

[8] Hammershøi, D., Møller, M., Sound transmission to and within the human ear canal. JASA vol. 100, No. 1, July 1996

[9] Møller, M. Sørensen, M., Jensen, C., Hammershøi, D. Binaural Technique: Do We Need Individual Recordings J. Audio Eng Soc. Vol. 44 No. 6, (1966)

[10] Bill Gardner's data for Kemar HRTFs *www.media.**mit**.edu/~kdm/**hrtf**doc/**hrtf**doc.html*

[11] Toole, F. The Acoustics and Psychoacoustics of Loudspeakers and rooms." Routledge 2007

[12] Griesinger, D. Binaural hearing and headphones. A power point presentation originally written for a convention of the Verband des Deutches Tonmiesters. It is available on the author's web-site. Binaural Hearing, Ear Canals, and Headphones"

[13] Hiipakka, M. Estimating pressure at the eardrum for binaural reproduction. A thesis from Aalto University. http://lib.tkk.fi/Diss/2012/isbn9789526048659/isbn9789526048659.pdf

[14] Lokki, T., Pätynen, J., Kuusinen, A., Tervo, S. "Disentangling preference ratings of concert hall acoustics using subjective sensory profiles." *The Journal of the Acoustical Society of America, 132,* 3148-3161 (2012).

[15] Pätynen, J., Lokki, T. Evaluation of Concert Hall Auralization with Virtual Symphony Orchestra. *Building Acoustics, **18**, 349-366* (2011).