



Audio Engineering Society

Convention Paper 9620

Presented at the 141st Convention
2016 September 29–October 2 Los Angeles, USA

This Convention paper was selected based on a submitted abstract and 750-word precis that have been peer reviewed by at least two qualified anonymous reviewers. The complete manuscript was not peer reviewed. This convention paper has been reproduced from the author's advance manuscript without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. This paper is available in the AES E-Library, <http://www.aes.org/e-lib>. All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Accurate timbre and frontal localization without head tracking through individual eardrum equalization of headphones

David Griesinger¹

¹ David Griesinger Acoustics 221 Mt Auburn St. #540 Cambridge MA 02138, USA

Correspondence should be addressed to David Griesinger (dgriesinger@verizon.net)

ABSTRACT

The ear and brain perceive the vertical position of sounds by matching the timbre detected at the eardrum of a listener to timbre patterns built up by that individual over a long period of time. But the eardrum timbre depends dramatically on ear canal resonances between 1000Hz and 6000Hz that boost the pressure at the eardrum as much as 20dB. These resonances are highly individual, and are either eliminated or altered by headphones. In-head localization is the result. We have developed an app that uses an equal-loudness procedure to measure and restore the natural timbre. Accurate timbre and frontal localization are then perceived without head-tracking, and binaural recordings can be stunningly realistic.

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, they 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

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 researches 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.

Since then standards have been developed for testing headphones. The goal was to find a measurement method that would optimize the sound for any individual. One of the most promising was DIN 45-619, which attempted to duplicate the timbre at the eardrum of a listener of a lateral loudspeaker sound source. The choice of a lateral source was problematic, and the object of the standard – to find an average response that would work for anyone – is probably unobtainable. But at least the standard used eardrum measurement of real people with real ears.

DIN 45-617 was abandoned in favour of standards that measured the sound pressure at the entrance to the ear canal [3], [4], [5], [6]. 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 duplicate the impedance of a human ear canal entrance, nor the resistive impedance of the eardrum.

The choice of measuring the pressure at the entrance to an ear canal and not the eardrum was based on the assumption that if the sound pressure at the ear canal entrance could be optimized and standardized, then the average listener would hear a natural timbre.

For general use by the public it does not matter if the assumption is right or wrong. The outer hair cells in the basilar membrane act as a continuous multi-band compressor. The ear adapts within a few minutes to even gross errors in frequency response. But this does not mean the sound at the eardrum has a natural timbre.

3 Ear canal resonances

Figure one shows Bill Gardner’s MIT Kemar data [7] 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. The high frequency localization notch is almost constant. In figure 2 it is the elevation notches above 6kHz that vary. The middle frequencies do not change.

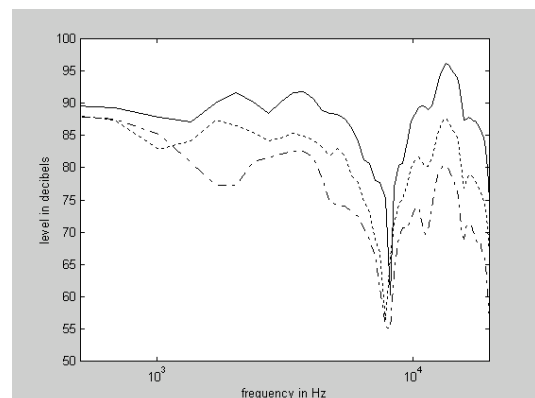


Figure 1: MIT Kemar data for the contralateral ear at zero degrees elevation and 0, 30, and 60 degrees azimuth.

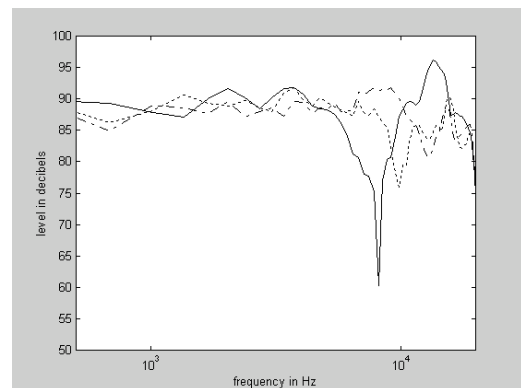


Figure 2: MIT Kemar data for zero degrees azimuth and 0, 30, and 60 degrees elevation

In both figures 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 3: An enlargement of the author's right ear, with Ville Pulkki pointing to the opening of the horn.



Figure 4: A negative cast of the author's left ear canal from the concha to the eardrum. The trumpet shape is clear. The length from the bell to the eardrum is over four centimetres.

We conclude that if our goal is to reproduce a convincing azimuth we need not vary the frequency notches above 6kHz. 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 from zero to 6kHz. The pinna, concha and ear canal form a horn, which has evolved to concentrate sound energy on the eardrum.

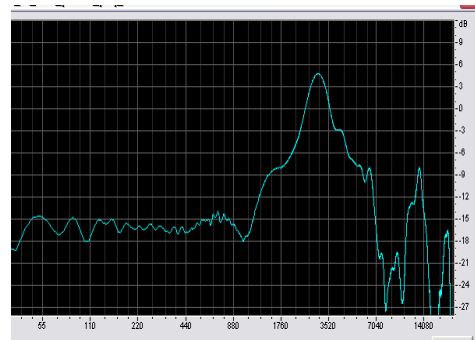


Figure 5: Pressure at the author's eardrum from a frequency linear frontal plane-wave. The boost is more than 18dB at 3000Hz.

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. 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 figures 3, 4, and 5.

In [6] Møller discusses individual equalization of headphones. He appears to do this by adjusting the headphone response so that it matches the free field response at the same point where a blocked ear canal pressure was taken. [The author finds the written description confusing.] We assume the ear canal was open. He states in [5] "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 [6] Møller refers to these headphones as “FEC” phones, and states that many headphones meet this requirement.

Our data from loudness matching shown in figures 11 and 12 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.



Figure 6: A probe microphone glued just inside the ear canal of the author’s model head. The model has castings of his ears all the way to the eardrum, and the eardrum impedance modelled with a resistance tube. The free-field frontal response at the eardrum closely matches his own head.

We measured the response from the eardrum of the dummy and from the probe from a frontal loudspeaker in the absence of a headphone, and then with the three headphones.



Figure 7: Blue: the pressure at the eardrum position from a frontal plane wave with an open ear canal. Red: The pressure slightly inside the ear canal opening. The blue curve is typical for the author’s ears. The measurement was not anechoic.

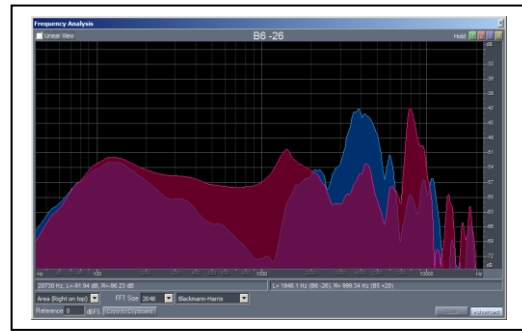


Figure 8: The same measurement from a Stax model 303 Classic electrostatic headphone.



Figure 9: The same measurement for an AKG 701 headphone.



Figure 10: The same measurement for an AKG 501 headphone.

It can be seen from figures 7, 8, 9, and 10 that the headphone difference between the red and blue curves was not the same as that from the free field. The Stax electrostatic phone was particularly different. We conclude that individually equalizing headphones from a measurement near the ear canal mouth is not effective. The equalization we find by equal loudness at the eardrum for multiple individuals with the AKG 701 and the Sennheiser 600 shown in figures 11 and 12 support this conclusion. They show considerable variation.

4 Reproducing a natural timbre

We believe there are no shortcuts. If we want to reproduce a sound through a headphone with the timbre of a frontal sound source we must reproduce the spectrum of the ear canal resonances at the eardrum.

This idea is not new. Early work by researchers such as Schroeder, Gottlob, and Siebrasse [8], and Mellert [9] attempted to find correlations with various acoustic measures with sound quality of binaural recordings. Their playback method, crosstalk cancellation calculated with probe microphones at the listener's eardrums, individually equalizes the recordings to the listener's eardrums. A head clamp was required.

Experiments at IRCAM recorded live sound from two probe microphones almost touching the eardrums of each subject. A head clamp was required, so head tracking was both impossible and unnecessary. They used Schroeder's crosstalk cancellation method to reproduce the recorded eardrum pressure. The listener's head was again put in a clamp, and the same steel probes were used to measure the response of two loudspeakers in front. The crosstalk and the frequency response was then mathematically adjusted, precisely re-creating the eardrum pressure of the recording. Live music was reproduced convincingly without head tracking [10].

There is nothing magical about Schroeder's crosstalk method. All that is necessary is that a pair of headphones is equalized *at the listener's*

eardrums to reproduce the timbre of a natural sound. The author has developed small probe microphones with soft silicon tips that sit comfortably on or next to the eardrums. He has recorded data and live concerts all over the world. When a pair of earphones are equalized with the same probes in the same place the sound pressure is reproduced exactly, and the result can be stunning.

Although both the recording and the playback are matched to my ears, the recordings play back remarkably well for other people if the headphones are individually equalized. Like Bill Gardner's MIT data from Kemar, the author equalizes the recordings such that a sweep from a calibrated frontal loudspeaker has a flat frequency response up to about 6kHz. The vertical localization notches above that frequency are left in place. This equalization turns my head and ears into a close analogue of a studio microphone, but with a very different directivity.

A frequency-flat frontal loudspeaker is the essential reference for timbre in the audio world. Such speakers are needed to accurately play standard audio recordings through loudspeakers.. Toole has found [11] that loudspeakers with the most linear on and off axis response are preferred in blind listening tests. 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.

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.

5 Headphone equalization through loudness matching

We have developed a software application that facilitates individual equalization of headphones through a loudness matching procedure similar to the one used to determine ISO equal loudness curves.

ISO 226:2003 specifies that plane waves of sine tones at different frequencies, alternating with a reference tone at 1000Hz are presented from a frontal loudspeaker to a subject. The subject adjusts the level of the tone under test until it is perceived equally loud as the reference. For most listeners the result is repeatable to ± 1 decibel.

We adapted the method to equalize headphones. 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. They all give similar 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 becomes their personal equal loudness data. They then put on the headphones and find their personal equal loudness data for that headphone. We find it is additionally useful to have the subject balance the perceived left-right azimuth of the headphone tones. Not all ears are the same, and neither are headphone drivers. People with some mild hearing loss in one ear also find the balancing procedure very useful (including the author.) The difference between their headphone data and their loudspeaker data is the desired headphone equalization. This is loaded into the equalizer.

The subject can then listen to pink noise or music of their choice through their personal equalization. Almost everyone finds the image is frontal, and the timbre of pink noise and recordings accurate. My 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 that pair of headphones.

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 single driver speakers can be inexpensively equalized with the app.

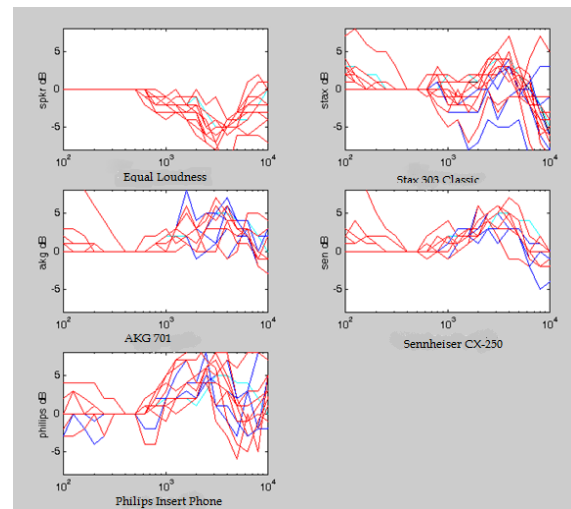


Figure 11: Data from students at Aalto University for their personal equal loudness and equalization for four different phone types.

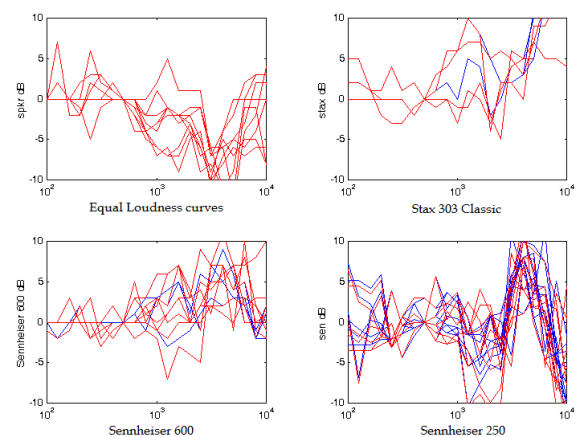


Figure 6: Data from students at Rensselaer University for equal loudness and three different phone types.

We have conducted experiments with the headphone app with the help of Ville Pulkki at Aalto University in Finland, and Jonas Braash 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 is clear, although distance perception is variable. We believe that the distance perceived for proximate sources in both binaural and natural hearing depends more on vision or expectation than any acoustic cue.

6 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 it.

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 might 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.

7 Conclusions

Our experiments with individual headphone equalization are on-going but sparse. The author lacks access to legions of eager students. Publishable conclusions on individual headphone equalization are hard to come by. It is widely believed to be either unnecessary or impractical.

But it is simple to prove that it works. We believe that the data here, and experience with a few binaural examples we provide on our web-page speak for themselves. We find that 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 nearly anything you play through individually equalized headphones sounds more natural.

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

1. 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.
2. We believe that if head motion is required for frontal localization with a non-individually equalized headphone the timbre is sufficiently incorrect that it will create 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.
3. The author had hoped to have a commercial cell-phone version of the app by the time of this conference, but this has been more difficult than expected. Enthusiasm from the people who have tried the app indicates that when a cell phone version is available how people use headphones will change. The app is a potentially disruptive technology.

4. We show in another preprint for this conference that with the help of individual equalization the sound of an ensemble on the stage of 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. Convolution of the results with Lokki's anechoic recordings and listening with individually equalized headphones is convincing. The effects of different reflections can then be studied.
5. 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 should play an important part.

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] L. Beranek, "Riding the Waves: A Life in Sound, Science, and Industry." Macmillan
- [3] Theile, G.; On the standardization of the Frequency Response of High-Quality Studio Headphones. *JASA* (34) 12 1986 pp 956-969.
- [4] G. Spikofski, "The diffuse field transfer function of studio-quality headphones." *EBU Review – Technical* 229 (1988) pp 111-126.
- [5] D. Hammershøi, H. Møller, "Sound transmission to and within the human ear canal." *JASA* vol. 100, No. 1, July 1996
- [6] M. Møller, M. Sørensen, C. Jensen, D. Hammershøi, "Binaural Technique: Do We Need Individual Recordings" *J. Audio Eng Soc.* Vol. 44 No. 6, (1966)
- [7] Bill Gardner's data on Kemar HRTFs www.media.mit.edu/~kdm/hrtfdoc/hrtfdoc.html
- [8] M. Schroeder, D. Gottlob, K. Siebrasse, "Comparative study of European concert halls: correlation of subjective preference with geometric and acoustic parameters." *JASA* 56, number 4, 1195-1201 (1974)
- [9] V. Mellert, "Construction of a Dummy Head after New Measurements of Thresholds of Hearing." *JASA* 51 number 4, 1359-1361 (1971)
- [10] Personal communication from Eckhard Kahle
- [11] F. Toole, "The Acoustics and Psychoacoustics of Loudspeakers and rooms." Routledge 2007
- [12] D. Griesinger, "Binaural hearing and headphones." A power point presentation originally written for a convention of the Verband des Deuteschen Tonmiesters. It is available on the author's web-site. [Binaural Hearing, Ear Canals, and Headphones](http://www.dgriesinger.com/BinauralHearing_EarCanals_andHeadphones)
- [13] M. Hiipakka, "Estimating pressure at the eardrum for binaural reproduction." A thesis from Aalto University. <http://lib.tkk.fi/Diss/2012/isbn9789526048659/isbn9789526048659.pdf>
- [14] T. Lokki, J. Pätynen, A. Kuusinen, S. Tervo, "Disentangling preference ratings of concert hall acoustics using subjective sensory profiles." *The Journal of the Acoustical Society of America*, 132, 3148-3161 (2012).
- [15] J. Pätynen, T. Lokki, "Evaluation of Concert Hall Auralization with Virtual Symphony Orchestra." *Building Acoustics*, 18, 349-366 (2011).