Audio Engineering Society

Convention Paper 10232

Presented at the 147th Convention
2019 October 16 – 19, New York

Personalized and Self-Adapting Headphone Equalization Using Near-Field Response

Adrian Celestinos[1], Elisabeth McMullin[1], Ritesh Banka[1], and Pascal Brunet[1]

[1] *Samsung Research America, DMS Audio, Valencia CA 91355*

Correspondence should be addressed to Adrian Celestinos (a.celestinos@samsung.com)

ABSTRACT

Variability in the acoustical coupling of headphones to human ears depends on a number of factors. Placement, size of user's head and ears, the headband and ear-pad material, are all major contributors to the sound quality delivered by the headphone to the user. By measuring the transfer function from the driver terminals to a miniature microphone set near the driver inside the cavity produced by the headphone and the ear, the degree of acoustical coupling and the fundamental frequency of the cavity volume was acquired. An individualized equalization based on these measurements was applied to every user. Listeners rated the personalized EQ significantly higher than a generic target response and slightly higher than the bypassed headphone.

1 Introduction

Variability in the coupling of headphones with a listener's head and ears is a common problem in headphone design. Variations in how headphones are coupled to a user's ears (e.g., how the headphones are worn on the user's ears, how tightly or loosely the headphones are placed against the the pinnae, anatomical differences, etc.) can negatively affect the output of the headphones. An improved approach that addresses this problem would be advantageous for headphone sound reproduction. In order to evaluate headphones as a means of the reproduction of binaural signals, Møller et al. [1] performed measurements of 14 headphones on 40 human subjects. Results showed that amplitude characteristics of the transfer functions were distinguished by smooth fluctuations at low frequencies, and many discrete resonances at high frequencies, when measured at the entrance of the blocked ear canal.

When over-ear headphones are worn, the acoustical system formed by the headphone and the ear works as a volume cavity below 4 kHz. Above these frequencies, standing waves build up inside the cavity, causing the pressure at the eardrum to be strongly dependent on the headphone placement [2, 3]. Masiero and Fels [2] also reported that the transfer function of headphones varies dramatically at high frequencies depending on the fit of the headphone. This means that even small displacements of the headphone after equalization will lead to irregularities in the resulting frequency response. Considering this, individual equalization should be used when possible. In another study, Paquier and Koehl [4] reported that the spectral modifications caused by headphone positioning variability led to audible differences.

In this study, measurements of the transfer function from the driver's electrical terminals to a point very close to the driver's diaphragm, inside the cavity produced by the headphone cup and the external ear, have been explored. This method is convenient to implement since modern headphones with noise cancellation already include microphones and DSP. This transfer function includes the ear-canal resonance and shows the degree of acoustical coupling between the headphone and the user's ears. By applying an individual correction, the intended sound quality can be maintained. This process can also serve as an automatic measurement of leakages, allowing for the correction of the output to a predetermined equalization target for over-ear headphones and potentially other type of headphones.

This paper addresses the problem of variability in the coupling of headphones to the user's ears, and the possible output corrections according to the individual characteristics of the user.

2 Methods

In this section, a series of experiments and methodologies to measure and equalize headphones will be presented.

2.1 Transfer Function Measurement

The measurement of the transfer function (TF) from the electrical terminals of the driver to a point inside the cavity provides the degree of coupling of the headphone to the individual characteristics of the user's ears. The impulse response to obtain the TF has been acquired using the logarithmic sweep method described in [5]. Additionally, the frequency response has been acquired using pink noise. The power spectral density (PSD) was computed using the so-called weighted overlapped segment averaging (WOSA), introduced by Welch [6].

The transfer-function measurements have been performed on a KEMAR head and torso simulator fitted with pinnae simulators, ear canal extensions, and IEC 60318-4 Ear Simulators. In Figure 1, two measurements of a circumaural closed-back headphone at KEMAR's ear simulator are shown. Both measurements have been performed with the same output voltage. The solid line is a measurement of the right ear using loose coupling between the headphone and KEMAR's head. On the dotted curve, the headphone

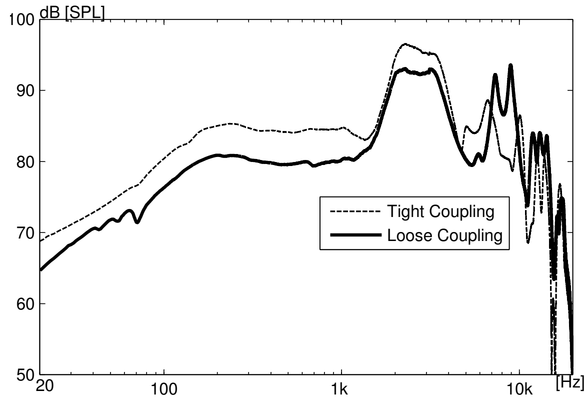

Fig. 1: Measurements of a circumaural closed-back headphone on KEMAR's ear simulator.

has been pushed firmly and kept towards the dummy head using tape. One can observe the differences in the output response having much more output when the headphone was kept pushed hard towards the dummy head producing a tight coupling between the earphone and the ear cavity.

2.1.1 Near-Field Microphone

A miniature MEMS (Micro-Electro-Mechanical System) Knowles SPM0687LR5H-1 microphone was mounted inside the cavity of a modified circumaural over-ear closed-back headphone. The microphone was

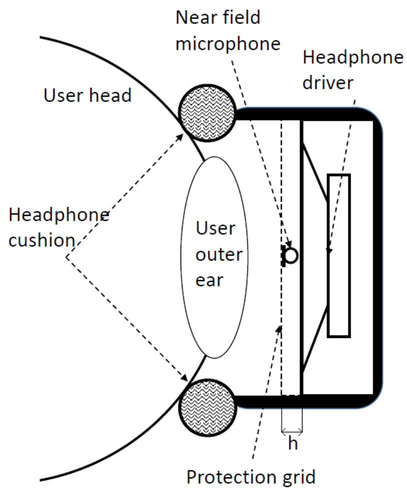

Fig. 2: Circumaural closed-back headphone on user's head including the near-field microphone.

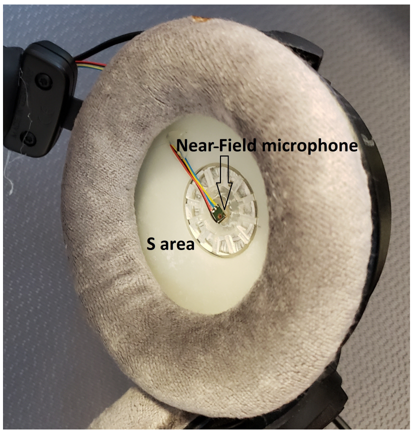

Fig. 3: Modified circumaural over-ear closed/open-back headphone with the MEMS microphone attached to the protection grid of the driver.

fixed on the grid at about $h = 5$ mm from the driver's diaphragm, as seen in Figures 2 and 3. In the same manner, a microphone was attached to a customized open-back headphone. In this paper, this measurement point is referred to as the near-field microphone.

By measuring the sound pressure level inside the cavity formed by the headphone, one can evaluate the degree of coupling. In Figure 4, the result of the measurement obtained at the ear simulator is compared with the sound pressure measured at the near-field microphone very close to the driver. The same type of measurement performed on the open-back headphone is presented in Figure 5. One can observe that both measurements (at the near-field microphone and at the ear simulator) are almost identical from 20 Hz to 800 Hz, and from 800 Hz to 1.8 kHz present similar tendencies but start to separate from each other. This behavior is detected in both, closed-back and open-back headphones.

The volume formed by the cavity of the circumaural headphone attached to the human ear is small compared with the wavelengths at frequencies below 1.5 kHz. Below these frequencies, the cavity will act as a pressure chamber where the pressure is the same at any point in the cavity. This can be observed in Figure 1 where the result of the measurement obtained at the artificial ear simulator compares with the sound pressure measured at the near-field microphone very close to the driver. It can be observed that both measurements are similar from 20 Hz to 800 Hz. From 800 Hz to 1.8 kHz, the measurements are still similar but begin to separate.

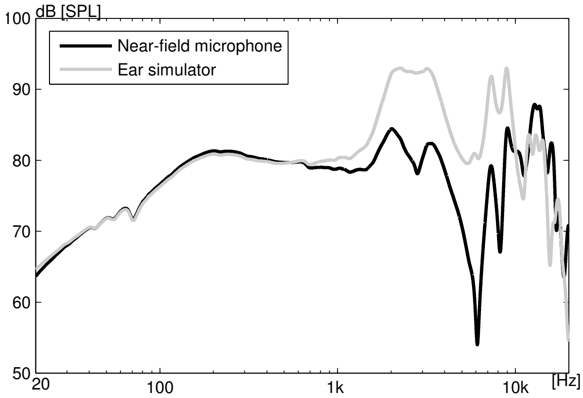

Fig. 4: Measurements of a circumaural closed-back headphone on KEMAR.

An important aspect of measuring the response of the headphone inside the cavity with a near-field microphone is that the fundamental or resonance frequency of the enclosure formed by the earphone, the outer ear, and the ear canal can be detected. This can be seen in Figure 4 around 2 kHz, expressed as a peak in the frequency response measured by the near-field microphone. In the same figure on the black line, the 2nd harmonic of the fundamental resonance frequency can be found within the range from 2.7 kHz to 5 kHz. In some cases, only the 2nd harmonic of the fundamental frequency can be detected; this depends on the acoustic damping of the headphone, and probably also on microphone location. In Figure 6, measurements of the TF from 10 human subjects are presented. The TF was measured on a closed-back Beyerdynamic DT 770 PRO

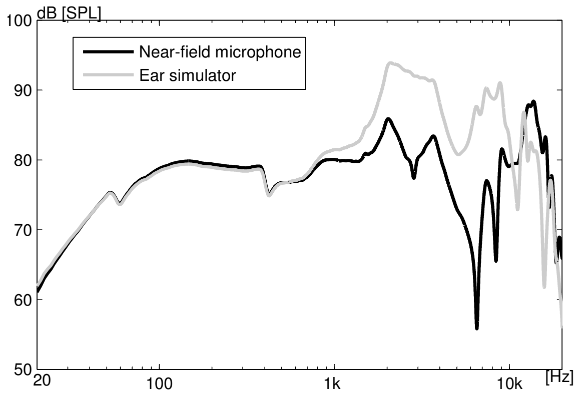

Fig. 5: Measurements of a circumaural open-back headphone on KEMAR.

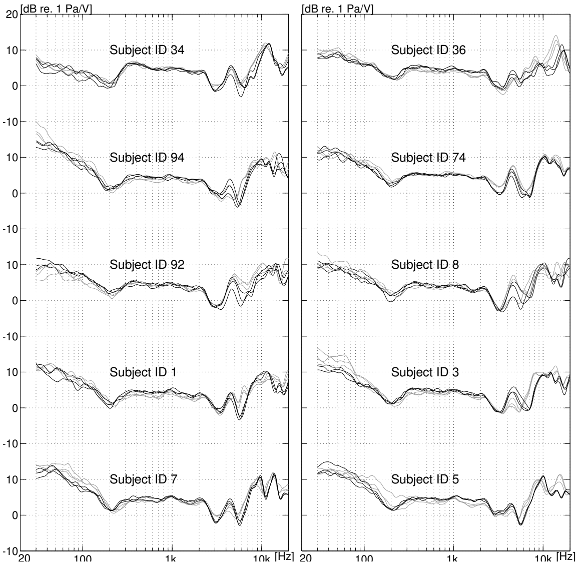

Fig. 6: Transfer functions from 10 human subjects on DT 770 PRO headphone, with near-field microphone. Light gray, left ear. Black, right ear.

over-ear headphone. Each subject was asked to reseat the headphones three times. The same type of microphone described in 2.1.1 was attached to the protection grid at approx. $h = 4$ mm distance from the driver's diaphragm. In this case, differing from the modified headphone presented in Figure 3, it was verified that the surface area S around the headphone driver was in fact a porous material that reduced the main resonance of the cavity. However, the 2nd harmonic of the fundamental frequency can be detected around 4 kHz on all subjects. The measurements also show that there were variations in the frequency response, especially at low frequencies.

2.2 Equalization Strategy

As shown in Section 2.1.1 the advantage of being able to measure the same sound pressure level at the end of the ear canal and at the driver is that it may be possible to monitor how well the headphone is coupled to the listener. Also, an individualized headphone equalization can be performed from 20 Hz to 2000 Hz to improve the sound quality of the reproduced audio by the headphones.

2.2.1 Target Response

The equalization target for a headphone can be an arbitrary frequency response. However the electroacoustical characteristics of the headphone's driver can limit the performance of the equalization. Olive et al. [7] proposed a target frequency response based on measurements of a loudspeaker calibrated in a reference listening room using a dummy head fitted with an ear simulator. In subsequent studies performed by the same authors, a preferred frequency response for around-ear (AE) and on-ear (OE) headphones was updated based on subjective evaluations on human subjects [8]. The preferred curve is shown in Figure 7. In this paper, it is called the preferred target curve. In this section, a modified version with a 2nd-order high-pass filter added to the preferred target curve has been used. The headphone preferred curve has been described using 12 cascaded PEQs listed in Table 1, plus a 2nd-order low-pass filter with f_c=12266.67 Hz.

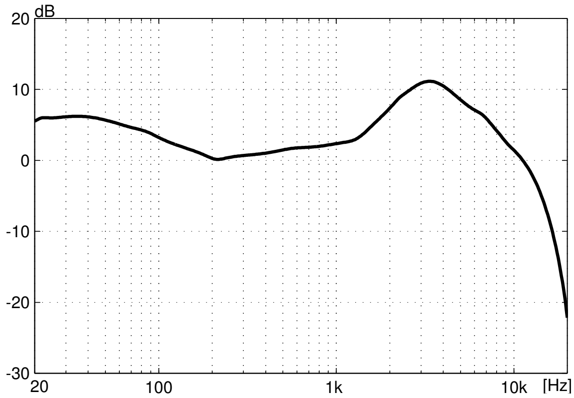

Fig. 7: Headphone preferred response curve.

Table 1: 12 PEQ (biquads) to describe the headphone preferred target curve.

f_c (Hz)	g (dB)	Q
21.46	0.86	3.61
35.50	6.13	0.61
84.94	0.51	2.41
209.94	-0.97	2.71
560.00	0.68	1.51
1258.76	-0.23	1.51
1306.24	-0.89	2.41
2258.32	0.34	4.5
3345.20	11.12	1.21
6727.04	1.01	3.31
8074.47	0.18	4.5
9175.43	-0.25	4.5

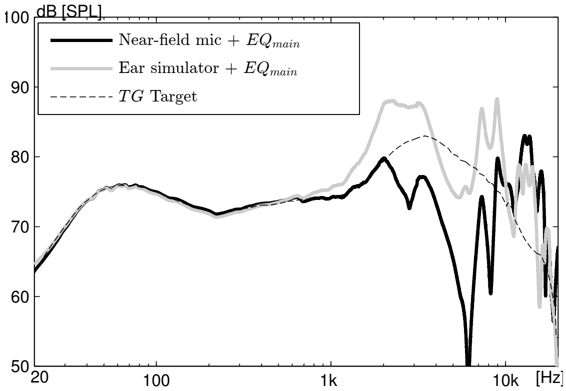

Fig. 8: Result of a personalized equalization of a circumaural closed-back headphone on KEMAR's head and torso.

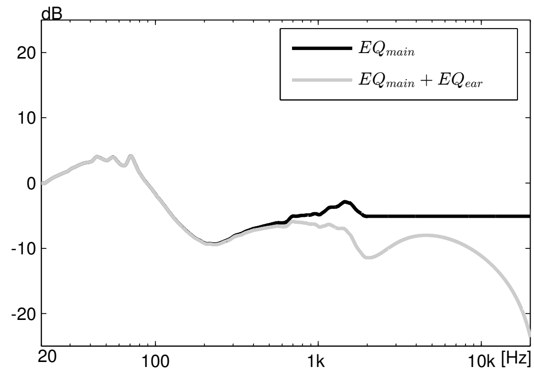

Fig. 9: Solid black line: main EQ. Solid gray line: ear EQ and the low-pass filter added to the main EQ.

2.2.2 Main EQ

Figure 8 illustrates the result of an equalization of a circumaural closed-back headphone using the near-field microphone of modified headphone shown in Figure 3. The solid light gray line represents the response equalized and measured at KEMAR's ear simulator. The dotted line represents a target frequency response, and the solid black line represents the equalized response at the near-field microphone. The method utilized to equalize the headphone response was based on the technique proposed by Ramos and López [9] using a minimum-phase filter constructed to correct the response from 20 Hz to 2 kHz. The equalization filter was constructed using infinite impulse response (IIR) filters with a 20 biquads or parametric equalization (PEQ) filters. As shown in Figure 8, the response has been corrected toward the target response up to approximately 2 kHz, and it starts to separate from the target response as the frequency increases. It also can be observed that there is a boost of approximately 6 dB to 8 dB around 2 kHz and 3 kHz after the equalization. Some other peaks are left higher than the nominal response around 7 kHz to 10 kHz on KEMAR's ear simulator after the equalization. The example in Figure 9 illustrates that the response of only EQ_{main} is equal to a constant value from 2 kHz to 20 kHz.

2.2.3 Ear Resonance EQ

The resonance of the human ear canal can fluctuate around 3–4 kHz due to its size, volume, and length for a typical human subject. The fundamental frequency of the cavity volume formed by the earphone and the human ear, including the ear canal, is located below 4 kHz [2, 3]. This suggests that one can reduce the influence of the fundamental frequency of this cavity by means of DSP and the measurement at the near-field microphone.

An alternative to compensating the response from 2 kHz to 20 kHz and matching the response to the target at the eardrum is to detect the fundamental frequency of the acoustic system formed by the outer ear, the ear canal resonance, and the earphone. From the near-field response, this can be obtained by detecting the maximum sound pressure level from 1 kHz to 3 kHz. Once the fundamental frequency f_{r1} is detected, a PEQ centered at this frequency can be applied with a gain related to the attenuation of the peak and a low Q. This PEQ can be added to the main equalization block. To reduce the high peaks at frequencies above 1 kHz, a first-order low-pass filter with a cutoff frequency at 6.5 kHz can be added as well.

There are cases where the 2nd harmonic of the fundamental is easier to detect; in such cases, the second peak in the range from 2.7 kHz to 5 kHz can be detected at $2 \times f_{r1}$. The example in Figure 10 illustrates a result of additional equalization of a circumaural closed-back headphone using the ear canal resonance and a low-pass filter. In the example, the solid gray line represents an equalized response measured at the KEMAR's ear simulator. The dash-dot line represents a target frequency response. The dashed line represents an equalized response at the near-field microphone. In the example,

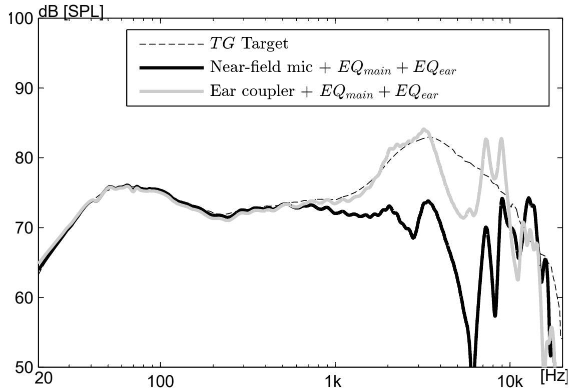

Fig. 10: Result of additional ear resonance compensation using the fundamental frequency and a low-pass filter.

the response of the headphone has been corrected up to 3.5 kHz. From 7 kHz up to 15 kHz, the equalized response is close to target. The only peaks left are the ones at 7.2 kHz and 8.9 kHz, which have very high Q and may not be recommended to be equalized. The main equalization filter is described in the following equation;

$$\begin{aligned} &EQ_{main} = (TG + TG_{norm}) - H_{near-field}, \\ &EQ_{main}(f < 20Hz) = 0, \\ &EQ_{main}(2kHz < f < 20kHz) = EQ_{main}(f2k). \end{aligned} \tag{1}$$

where TG is the equalization target, and TG_{norm} is the normalization target gain factor in dB. The ear equalization is described by

$$EQ_{ear} = PEQ_{fr1} + LP, \tag{2}$$

where PEQ_{fr1} is the parametric biquad filter at the detected fundamental frequency, and LP is the low-pass filter.

$$EQ_{final} = EQ_{main} + EQ_{ear} \tag{3}$$

In Equation 3, the final equalization filter is described. All expressions are frequency responses in dB. The frequency response of the main equalization filter and the addition of EQ_{ear} is shown in Figure 9. The equalization process is delineated in a block diagram in Figure 11.

2.3 Evaluation

In order to evaluate the proposed equalization, listening tests were performed on human subjects.

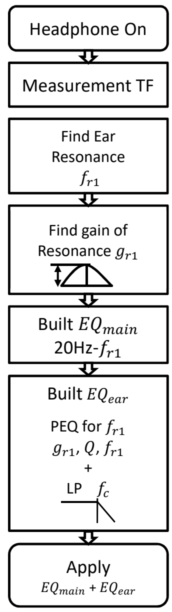

Fig. 11: Block diagram of equalization process.

2.3.1 Listeners

Eighteen listeners who were employees of Samsung Research America participated in the listening tests. The listeners ranged in age from 25 to 61 ($mean = 40$, $SD = 11$). Of these, three were female and fifteen were male. Thirteen were considered trained listeners based on their consistency in past experiments. All had been previously tested for normal audiometric hearing.

2.3.2 Program Material

Four tracks were selected for use in these listening tests. All of the tracks were imported from the original albums on CD or from high-quality online downloads (44.1 kHz 16-bit and above). Each track was edited into a 30 s or shorter clip for use in testing. A summary of the tracks can be seen in Table 2. All of the tracks used were measured for dynamic and spectral content.

2.3.3 Test Setup

To evaluate listener preference for the headphone "Personalized EQ" a double-blind preference test was run in which it was compared to the Beyerdynamic DT 770 PRO headphone ("Bypass"), and a generic "Fixed EQ" based on the headphone preferred target curve applied to this headphone. In Figure 12, the

Table 2: Table of tracks.

Artist/Genre	Song/Album/Label
Mark Ronson and Bruno Mars (MR) / Pop Funk with Male Vocal	Uptown Funk / Uptown Special/ RCA 2014 CD
Steely Dan (SD) / Rock with Male Vocal	Cousin Dupree / Two Against Nature / Giant 2000 CD
Norah Jones (NJ) / Jazz with Female Vocal	Come Away with Me / Blue Note 2002 CD
Deadmau5 (DM) / Electronic Dance with Male Vocal	For Lack of a Better Name / Mau5strap Ultra 2009

measurements of the DT 770 PRO headphone measured on KEMAR, at the ear simulator, and at the near-field microphone are shown. To ensure accuracy of the headphone preferred target curve measurements, the customized DT 770 PRO headphones were measured (five seats) on a GRAS 45CA test fixture with the customized "Welti" pinnae [10]. An average of these measurements was then equalized using the preferred target response, as shown in Figures 13 and 14 using 14 second-order IIR sections. The quality of the PEQs was limited to $Q = 4$. It was found that the headphone had a slightly lower sensitivity on the right driver. It was decided to only compensate that difference on the "Fixed EQ" since the "Personalized EQ" would correct for that difference automatically. Testing was automated using custom Max 8 and LabVIEW software which controlled the measurement of each listener, applied the "Personalized EQ" based on the measurements, randomized playback order of the audio tracks and

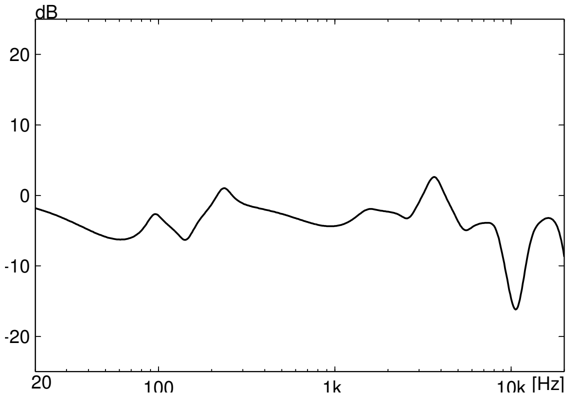

Fig. 13: "Fixed EQ" filter to fit the headphone preferred target curve.

headphones equalizations, and stored the test results. At the beginning of each test, listeners were instructed to place the headphone in a comfortable position and to click the screen to begin the measurement. A measurement of each listener using the custom headphones with the built-in MEMS microphones was made using 30 s of pink noise. After the measurement, a custom equalization curve was applied using the process described in Section 2. For the personalized filter, the resonance peak f_{r2}, was set to be searched for within the range of 3 kHz to 6 kHz, and f_{r1} was obtained by dividing f_{r2} by 2. The attenuation for the PEQ_{fr1} was set to -3 dB, and the Q was set to 1.5. The $f_{c,LP}$ of the low-pass filter LP was set to 7 kHz for all listeners. The listener measurements and equalization curve

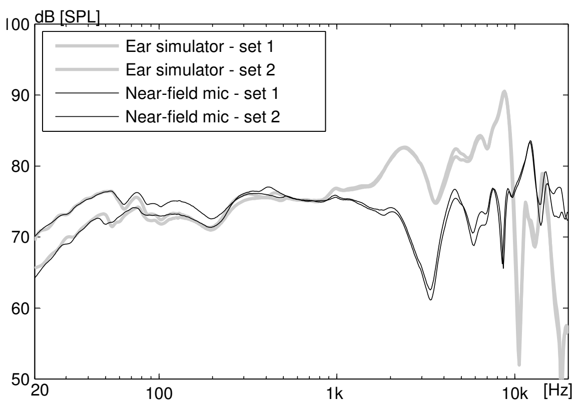

Fig. 12: Frequency response of DT 770 PRO headphone measured on KEMAR's right ear.

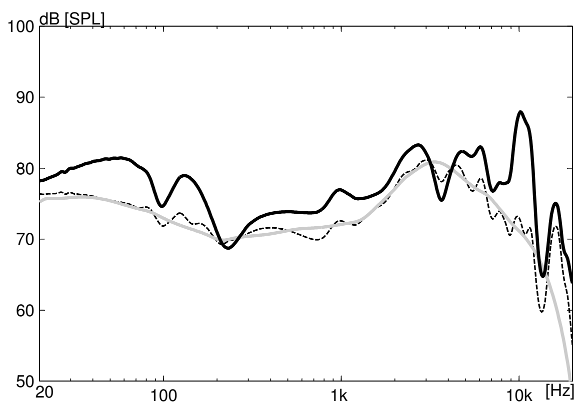

Fig. 14: DT 770 PRO Headphone measured on GRAS 45CA ear simulator. Black tick curve, no eq. Dotted curve, after eq. Light gray, target.

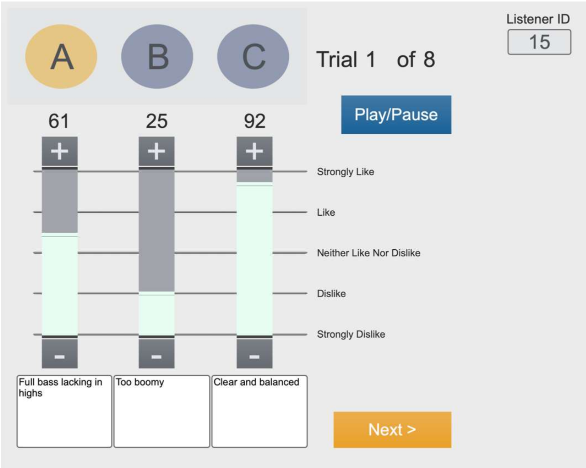

Fig. 15: Testing software GUI.

applied were both stored by the software for future analysis. Listeners were asked not to move the headphones during testing to ensure accurate results, since the positioning and leakage could change if the headphone was reseated. Listeners then proceeded to the preference test which was configured as shown in Figure 15. They were asked to rate each version of the EQ on a scale from 0 to 100 (Dislike Strongly–Like Strongly) using continuous sliders. They were also instructed that they could leave comments as desired. The test had eight trials in total, which consisted of four songs with one repeat each.

2.3.4 Data Analysis

The data was analyzed using a factorial ANOVA with repeated measured considering headphone EQ (three levels), track (four levels) and observation (two levels) to be fixed factors. Prior to analysis, the data was checked with Q-Q plots for normality of residuals and for sphericity using a Mauchly's test.

2.3.5 Signal Processing Tool

A utility tool was created in LabVIEW to measure the response of the headphone on human subjects. National Instruments hardware was used to play the test tracks and acquire the microphone signals.

Pink noise was used as stimulus to measure the headphone response on each subject as explained in Section 2.1, [6]. The input signal from each microphone was recorded for 20 seconds and a flat-top window was applied on the signal to calculate the FFT. The FFT window size was set to 200 ms which gave a resolution of 5 Hz which was appropriate considering that the resonances of interest lie between 3 kHz and 6 kHz.

3 Results

The overall outcome of the evaluation is presented in Figure 16. The results show that headphone EQ was a significant effect ($F(2, 34) = 8.6$, $p < 0.001$) with a moderate effect size of $\eta_p^2 = 0.33$. A post hoc pairwise comparison with Bonferroni correction was run and showed that the "Fixed EQ" rated significantly lower than the "Personalized EQ" and slightly lower than the "Bypass" (DT 770 PRO headphone default voicing). Track was not a significant factor and there were no significant interactions. Comments were left by nine of the listeners. 45 comments were left about the "Fixed EQ." Four listeners described it as lacking bass, three listeners found it lacked high frequencies and three believed that the midrange was too forward. 33 comments were left about the "Personalized EQ." Four listeners found that it had too much bass, while four listeners perceived it as lacking high frequencies. 32 comments were recorded about the DT 770 PRO headphone ("Bypass") and listeners tended to disagree about its attributes. In Figure 17 of Appendix A, the results of the frequency response measurements on the 18 subjects, applied "Personalized EQ" and individual targets are presented.

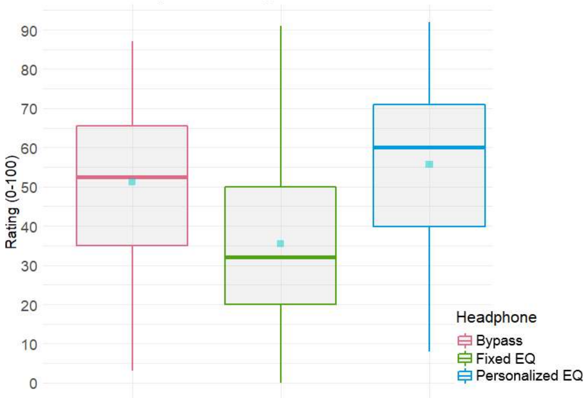

Fig. 16: Overall headphone rating results of the listening test evaluation. From left to right, "Bypass", "Fixed EQ" and "Personalized EQ."

4 Discussion

Overall results showed that the listeners tended to rate the "Fixed EQ" significantly lower than the "Personalized EQ" and moderately lower than the "Bypass". While there was a reasonable amount of variance in the listener data, 11 of the listeners consistently rated this EQ lower. One of the issues of headphone equalization is the variability of coupling which depends on different factors. Even if the headphone is well designed, anatomical differences and positioning can challenge the frequency response, at low and high frequencies. The "Fixed EQ" was based on the headphone preferred target curve which is meant to be a design goal for headphones. Although recently modified by listening tests in [8], this target addresses the response of an accurate loudspeaker in a typical listening room especially at low frequencies. Above 1 kHz it has similarities with the diffuse-field design goal. This curve is a good starting point but it may not be well reproduced on certain type of headphones. In the case of the DT 770 PRO, the equalization filter resulted in three pronounced peaks (see Figure 13). In particular, the peak with gain of 5.5 dB at 3.6 kHz, with a high Q may not fit certain subjects' geometry and might have been perceived as annoying. At lower frequencies, it may have been excessive in bass for some subjects, but for others may be perceived as weak. This is why individualized equalization is highly preferred.

5 Summary

Measurements on closed-back over-ear headphones have been performed on human subjects and a KEMAR head and torso. The transfer function has been measured using a miniature microphone attached very close to the headphone driver and used to create an individualized equalization filter. A listening test has been conducted in order to evaluate the proposed equalization strategy.

The results suggest that by using the near-field microphone in headphones, an individualized equalization can be applied, resulting in better sound quality. Future work should include a deeper investigation into a more precise compensation of the ear volume cavity and the effects on the reproduced sound quality.

Acknowledgments

The authors would like to thank Samsung Electronics and Samsung Research America for supporting this work.

References

[1] Møller, H., Hammershøi, D., Jensen, C. B., and Hundebøll, J. V., "Transfer Characteristics of Headphones: Measurements on 40 Human Subjects," in *Audio Engineering Society Convention 92*, 1992.

[2] Masiero, B. and Fels, J., "Perceptually Robust Headphone Equalization for Binaural Reproduction," in *Audio Engineering Society Convention 130*, 2011.

[3] Schmidt, S., *Finite Element simulation of external ear sound fields for the optimization of eardrum-related measurements*, Ph.D. thesis, Ruhr-Universität at Bochum, 2009.

[4] Paquier, M. and Koehl, V., "Audibility of Headphone Positioning Variability," in *Audio Engineering Society Convention 128*, 2010.

[5] Farina, A., "Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique," in *Audio Engineering Society Convention 108*, 2000.

[6] Pintelon, R. and Schoukens, J., *System Identification, A Frequency Domain Approach*, 2012.

[7] Olive, S., Welti, T., and McMullin, E., "Listener Preferences for In-Room Loudspeaker and Headphone Target Responses," in *Audio Engineering Society Convention 135*, 2013.

[8] Olive, S., Welti, T., and Khonsaripour, O., "A Statistical Model that Predicts Listeners' Preference Ratings of Around-Ear and On-Ear Headphones," in *Audio Engineering Society Convention 144*, 2018.

[9] Ramos, G. and López, J. J., "Filter Design Method for Loudspeaker Equalization Based on IIR Parametric Filters," *J. Audio Eng. Soc*, 54(12), pp. 1162–1178, 2006.

[10] Welti, T., "Improved Measurement of Leakage Effects for Circum-Aural and Supra-Aural Headphones," in *Audio Engineering Society Convention 138*, 2015.

A Appendix

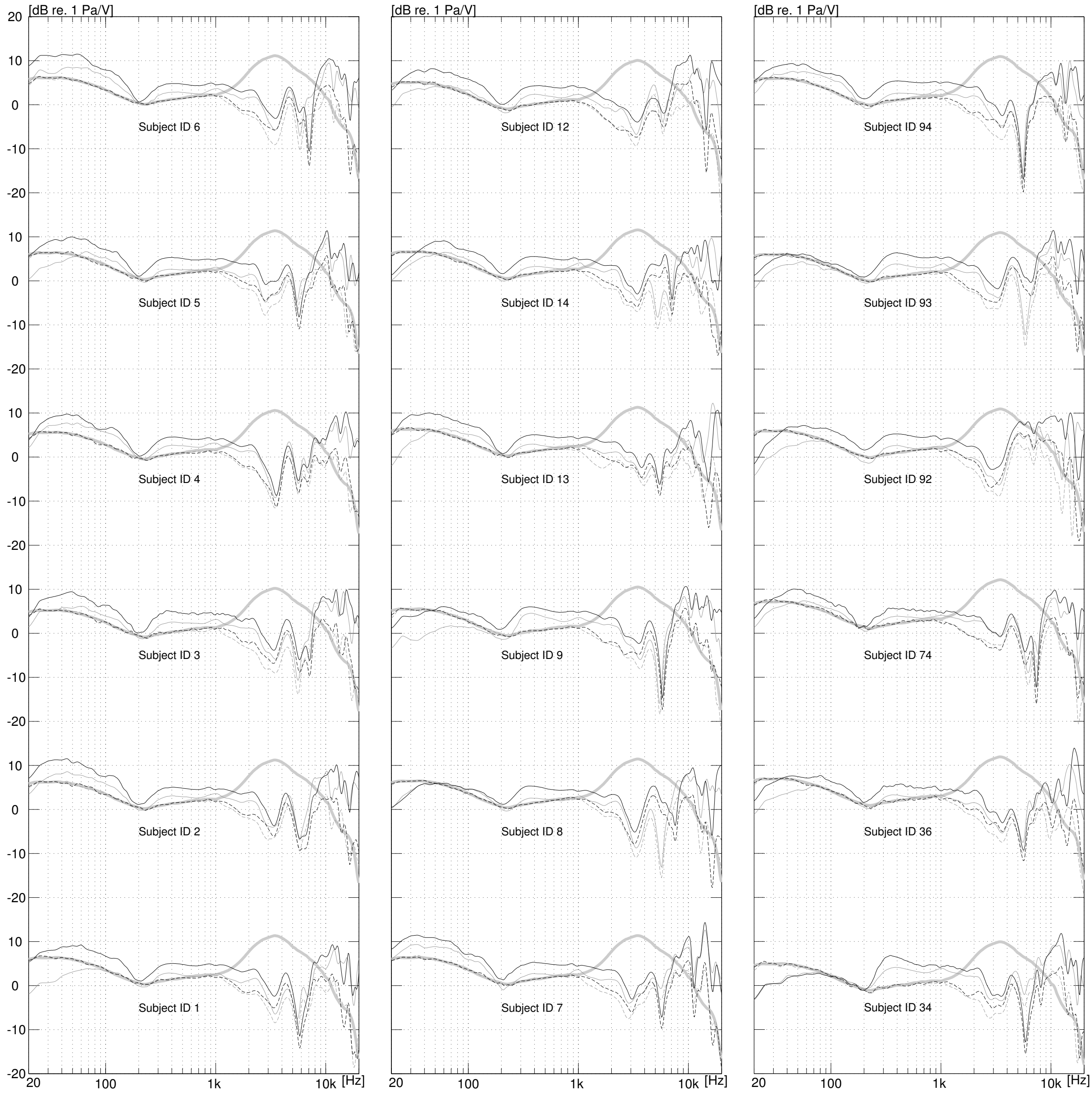

Fig. 17: Headphone transfer functions from 18 subjects. Light gray tick curve: Target. Black curve: left ear. Light gray curve: right ear. Black dotted curve: left ear + "Personalized EQ". Light gray dotted curve: right ear + "Personalized EQ".