forum acusticum 2026
8-12 September · Graz, Austria

AN OPEN-EAR CLIP-ON EARBUD RESEARCH PLATFORM FOR PERCEPTUAL STUDIES

Baptiste Fourrier[1], Benjamin Pries[1], Stephan Preihs[1], Jürgen Peissig[1]

[1] *Institut für Kommunikationstechnik, Leibniz Universität Hannover, Germany*

Abstract

Open-ear clip-on earbuds, a growing consumer product category, place a transducer against the pinna without occluding the ear canal, preserving natural ambient awareness and improving wearing comfort over traditional headphones.

Their low-occlusion design would make them attractive for research applications requiring simultaneous acoustic interaction with the real environment, such as spatial hearing or augmented listening. Yet consumer products are ill-suited for controlled experiments: their inaccessible DSP signal chains prevent full control over the audio chain, their frequency responses deviate substantially from established headphone target curves, and the absence of an acoustic seal fundamentally limits low-frequency output.

This paper presents a research platform addressing these limitations. A wired 3D-printed dual-transducer open-ear clip-on earbud prototype was developed through finite element simulations and empirical design. The dual-transducer configuration provides substantially greater headroom for bass extension via equalization. Target frequency responses are achieved using an extension of the established virtual headphone equalization method, adapted to account for elevated harmonic distortion inherent to small drivers at low frequencies. The equalized response to a flat target was validated as robust to earbud repositioning.

Keywords: *open-ear earbuds, virtual headphones, microspeaker, spatial audio*

1. Introduction

Open-ear clip-on earbuds are a recent and fast-growing class of consumer audio device. Rather than sealing the ear canal like in-ear monitors or enclosing the pinna like supra- and circumaural headphones, they clip a miniature dynamic transducer onto the pinna and radiate toward the canal entrance across a small air gap. The ear canal stays open, so ambient sound reaches the eardrum almost unattenuated. Therefore, the listener keeps situational awareness, the occlusion effect of sealed devices is avoided, and comfort over long wearing periods improves [1].

This open ear canal feature is also relevant to hearing research, where reproduced sound coexists with the surrounding acoustic scene. Augmented and assisted listening introduce virtual sources that the listener must continue to localize alongside real ones. In a recent multi-source localization study [2], for example, open headphones were used to present stimuli either as real sources over physical loudspeakers or as virtual sources rendered binaurally over the headphones; such paradigms would directly benefit from the acoustic transparency of ultra-open designs.

A controllable device is likewise needed for studying preferred headphone tonal balance, for which extensive listening tests have produced widely used target response curves [3]. Whether these findings carry over to clip-on earbuds is an open question this work does not resolve. The existing targets were derived for circumaural, supra-aural, and in-ear headphones, whose preferred curves already differ markedly at low frequencies due to their distinct ear coupling [4]. A clip-on earbud sits inside the pinna without enclosing it, a coupling unlike either and possibly more sensitive to placement. Defining an appropriate target curve would therefore require controlled listening tests, which commercial hardware cannot support: its embedded processing and wireless link block access to the raw electroacoustic path.

However, due to the lack of a sealed cavity, low-frequency output is intrinsically weak. Compensating it by equalization only would require enormous low-frequency gains that would drive the small transducer largely beyond its linear range. The virtual-headphone method that enables controlled tonal comparison, isolating coloration from comfort and expectation bias,

Corresponding author: *baptiste.fourrier@ikt.uni-hannover.de.*

was validated on large over-ear drivers for which magnitude response dominates [5], [6]; to what extent distortion stays negligible for a bass-limited driver operating in its nonlinear regime is unknown, and conventional equalization does not bound the distortion.

This paper presents a research platform built to address these obstacles. Section 2 develops a wired, 3D-printed clip-on research prototype whose enclosure is designed with the aid of a finite-element model. To widen the limited bass headroom, two transducers are combined in each earbud, raising the acoustic output available for low-frequency equalization. Section 3 then extends the automatic regularized virtual-headphone inversion of Bolaños et al. [7] with a signal-to-distortion criterion that limits equalization where it would otherwise push the transducers into excessive distortion. The resulting equalization is validated against a flat target and shown to be robust to repositioning on the pinna, providing a controllable, acoustically transparent audio path for perceptual experiments with this form factor.

2. Prototype Design

The prototype was built around the Ole Wolff OWR-1026T-32C dynamic micro-speaker, a 10 mm driver representative of those used in commercial clip-on earbuds. This transducer was originally developed for in-ear headphones with active noise cancellation (ANC), an application that requires high diaphragm excursion and output level at low frequencies. The same requirement governs the clip-on configuration, where the absence of an acoustic seal weakens the low-frequency response; transducers of this class are consequently well suited to the form factor.

Starting from the design objectives and physical constraints, the enclosure was explored with a finite-element parameter study that identified the dominant acoustic controls, then finalized through distortion-limited measurements that set the parameters governed by nonlinear effects the linear model cannot capture.

2.1. Design Objectives and Constraints

The prototype was designed to give the equalizer as much headroom as possible, particularly in the low frequencies. A sealed rear cavity cannot provide it: the rear volume required to avoid stiffening the driver, of the order of its equivalent compliance volume ($V_{as} \approx 10\,\mathrm{cm}^3$), is impractically large for an ultra-open clip-on, which must stay compact and leave the ear canal acoustically open. The much smaller cavity that would be practical here would instead act as a stiff air spring, pushing the low-frequency corner high into the audible range. The rear cavity is therefore vented, which removes this stiff loading and preserves the driver's fundamental, so usable output extends down to around 150 Hz. The rear vent and rear volume form a Helmholtz resonator that sets the first acoustic resonance. A front outlet is also required: besides protecting the driver, it directs the forward radiation

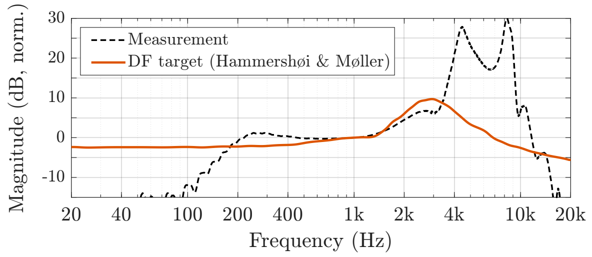


Figure 1: Diffuse-field Hammershøi and Møller [8] target versus the raw response of a single-driver prototype with a 1 mm rear vent. The first resonance at 4.5 kHz is the Helmholtz resonance of the rear vent and volume, while the second at 8 kHz is set by the front volume and outlet. Both curves are normalized to 0 dB at 1 kHz.

to the ear canal while the rear vents exhaust away from the skin, limiting destructive interference at the ear. Together with the front volume, it forms a second Helmholtz resonator that governs the high-frequency roll-off.

As a side benefit, below the first Helmholtz resonance, the vented enclosure would radiate in the far field like a dipole rather than like the monopole of a sealed box, reducing outward low-to-mid-frequency sound leakage [1].

These two resonators dominate the raw response, and shaping them is how headroom before EQ is created. As the platform is intended to support arbitrary target responses, the IEC diffuse-field curve is taken as a representative objective, its pronounced mid-to-high-frequency boost mirroring the spectral shaping of the outer ear in a reverberant field. Figure 1 compares this target with the raw response of an early single-driver prototype, which lacks low-frequency output and places its first resonance too high. Headroom is greatest when the first resonance is kept as low and the second as high as possible, flattening the response toward the target; the remaining irregularity is removed by the equalization of Section 3.

Within these aims, the form factor bounds the geometry, limiting the front volume from below by driver excursion, the rear volume from above by pinna space, and the front outlet area by the dual-transducer layout; the model below fixes each within its bound.

2.2. Finite-Element Model and Parameter Study

The enclosure was developed with a hybrid lumped-element/finite-element model implemented in COMSOL Multiphysics. The driver was represented by a small-signal electro-mechanical lumped-element model, with parameters measured using a Klippel Analyzer. This electrical equivalent circuit was coupled to the radiating FEM membrane and to an acoustic domain representing the front and rear cavities, vents, and outlet of the enclosure, with thermoviscous losses accounted for in the narrow vents. To predict the in-situ response, the enclosure was embedded in a virtual

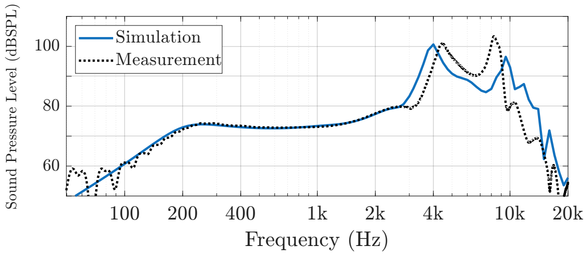


Figure 2: Measured and simulated transfer functions for a single-driver configuration with a 1 mm rear vent, measured at a 100 mVrms drive voltage.

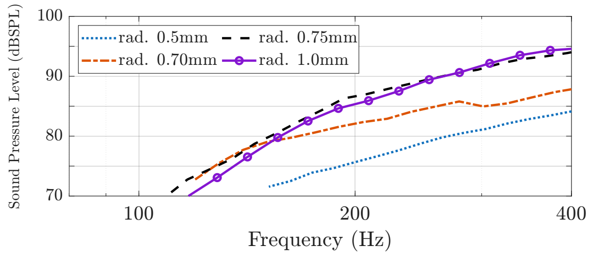
replica of the GRAS 45CA measurement fixture: a laser-scanned anthropometric pinna terminated by the fixture's high-frequency ear simulator, which is based on the IEC 60318-4 occluded-ear simulator [9] and represented in the model by its equivalent circuit [10]. The model was validated against measurements on a 3D-printed single-driver prototype, as shown in Fig. 2. The residual deviations, mainly in the resonance frequencies, likely reflect geometry mismatch between model and print, the rigid-wall assumption, and nonlinear vent flow beyond the linear model. The FEM therefore serves as a qualitative guide to resonance trends rather than a quantitative predictor of absolute frequencies, which are established by measurement.

A parameter sweep over the enclosure geometry confirmed these mechanisms and showed the first and second resonances to be largely independent of each other, so each could be placed for maximum headroom. Guided by it, the front volume was minimized to a 1 mm dust-cap clearance and the front outlet maximized (a rectangular vent equivalent to a 2.2 mm-radius circular one) to raise the second resonance, while the rear volume was set to its pinna-limited maximum of 125 mm^3 to lower the first.

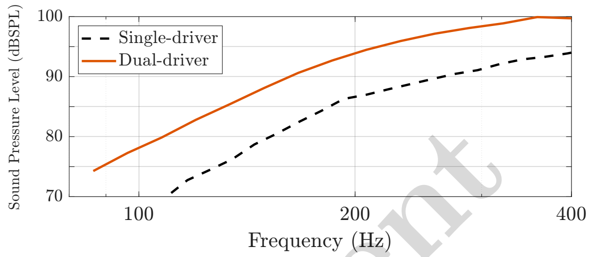
A dual-transducer arrangement, with two coherent drivers radiating in phase, adds an approximately 6 dB broadband sensitivity increase, the largest gain among the simulated parameters and especially valuable at low frequencies. Rear vents facing the skin were found to reduce low-frequency output and are therefore avoided. As the model is linear, it does not capture the nonlinear vent losses and harmonic distortion that ultimately constrain vent sizing and the single- versus dual-driver choice; those decisions were instead made from the measurements reported in Section 2.3.

2.3. Distortion-Limited Vent and Driver Selection

While a smaller rear vent lowers the first resonance, it also raises air particle velocity, producing nonlinear flow and elevated total harmonic distortion (THD), an effect not captured by linear FEM simulation. To determine the minimum acceptable vent radius, stepped-sine measurements with automatic voltage adjustment were conducted, recording the maximum SPL achievable at a 10 % THD threshold for each vent radius.



(a) Rear vent radii.



(b) Single- and dual-driver configurations.

Figure 3: Distortion-limited maximum SPL (10 % THD threshold) from stepped-sine measurements with automatic voltage adjustment. Sub-figure (a) informs rear vent sizing; sub-figure (b) motivates the dual-transducer architecture adopted for the final prototype. Axes cover the output-limited 100 Hz to 400 Hz bass region.

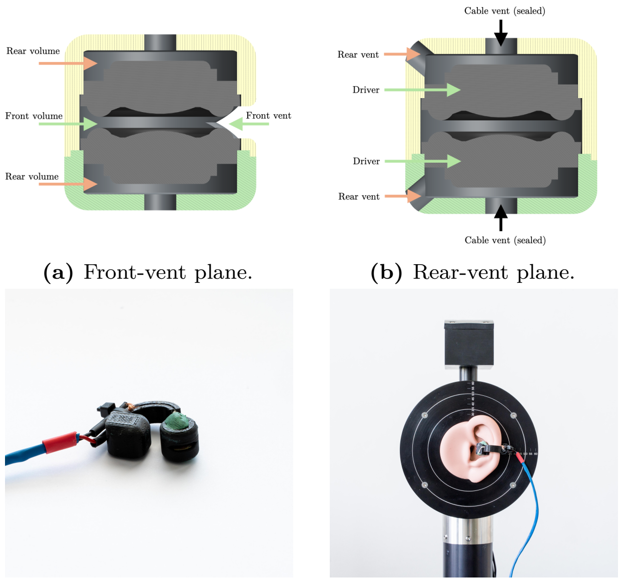
As shown in Fig. 3a, vent radii of 0.5 mm and 0.7 mm yield substantially lower maximum SPL than 0.75 mm and 1 mm, which perform comparably. A minimum rear vent radius of 0.75 mm was therefore adopted. Using the same method, single- and dual-driver enclosures were compared (Fig. 3b). The dual-driver configuration provides approximately 6 dB additional low-frequency output throughout the bass range. At 120 Hz, the single-driver reaches 70 dB SPL at the 10 % THD threshold; the dual-driver reaches 80 dB SPL under identical conditions. The dual-driver configuration was consequently selected for the final prototype.

2.4. Final Prototype

The final prototype is a 3D-printed earbud combining a rigid PLA speaker enclosure with a flexible TPU pinna bridge (Fig. 4); the material combination delivers dimensional precision at the acoustic boundaries while the compliant bridge provides secure pinna retention without risk of breakage. Both 0.75 mm rear vents are oriented at 90° to face away from the skin, minimizing bass loss. The speaker unit was assembled in two halves; precision adhesive sealing between the halves and at the enclosure seams prevents leakage between front and rear volumes.

3. Virtual Headphone Equalization

Virtual headphone equalization inverts a reference transfer function to a flat baseline onto which an arbitrary target, or a transparent binaural signal, can be imposed [3], [11]. As noted above, it relies on magnitude response being the dominant cue for perceived headphone quality, with nonlinear distortion and ex-



(a) Front-vent plane. **(b)** Rear-vent plane.

(c) Assembled prototype. **(d)** Prototype on fixture.

Figure 4: Open-ear clip-on earbud prototype. Cross-sectional views of the speaker unit along two perpendicular planes (a, b), the assembled earbud (c), and the prototype mounted on the GRAS 45CA measurement fixture (d). The speaker unit is a two-piece PLA assembly; the halves (yellow and green) are adhesively sealed at their joint. Cable vents are sealed.

cess phase contributing little [5], [6]. That premise was established on large, low-distortion over-ear drivers, not on the small open drivers used here, which the large gains needed for bass inversion may drive beyond their linear range.

This section therefore extends the automatic regularized inversion of Bolaños et al. [7], [12] with a signal-to-distortion ratio criterion that limits inversion where harmonic distortion is high, and examines how the equalized response changes when the earbud is repositioned on the pinna.

3.1. Distortion-Aware Regularized Inversion

Direct inversion of the measured headphone transfer function $H(\omega)$ is inadequate for limited-bandwidth transducers: achieving a flat response at low frequencies would require gain far exceeding the driver's physical limits. This is addressed through automatic regularized inversion [7], introducing a frequency-dependent regularization parameter:

$$H_{\mathrm{RI}}^{-1}(\omega) = \frac{H^*(\omega)}{|H(\omega)|^2 + \hat{\beta}(\omega)} D(\omega), \tag{1}$$

where $D(\omega)$ encodes a pure delay for causality and, optionally, a desired target response. In the present work, $D(\omega) = e^{-j\omega\tau}$, targeting a flat frequency response; as this platform is intended to support future perceptual studies, other targets are deferred to future work. Where $\hat{\beta}(\omega) \gg |H(\omega)|^2$, inversion is suppressed, preventing excessive gain outside the usable bandwidth or at narrow spectral notches.

The regularization parameter takes the form

$$\hat{\beta}(\omega) = \hat{\alpha}(\omega) + \sigma^2(\omega), \tag{2}$$

where $\hat{\alpha}(\omega)$ limits inversion bandwidth and $\sigma^2(\omega)$ suppresses inversion at spectral notches. In the original formulation [7], $\hat{\alpha}$ may be taken as the inverse SNR. For open-ear systems with small drivers, however, distortion levels at low frequencies can be significant even where the SNR remains favorable: SNR-based regularization alone would therefore permit inversion in ranges where harmonic distortion could be amplified to objectionable levels. This work extends $\hat{\alpha}(\omega)$ to take the more restrictive of the SNR and the signal-to-distortion ratio (SDR):

$$\hat{\alpha}(\omega) = \max\left(\frac{\lambda_1}{\mathrm{SNR}(\omega)}, \frac{\lambda_2}{\mathrm{SDR}(\omega)}\right). \tag{3}$$

The conservative default $\lambda_1 = \lambda_2 = 1$ is adopted [7], but both weights are exposed as tunable parameters: lowering them extends the usable bass range of this output-limited earbud at the cost of higher distortion, a trade whose perceptual optimum is left to future study.

The notch regularization $\sigma(\omega)$ is estimated by comparing the measured response to its half-octave smoothed envelope $\hat{H}(\omega)$, activating only where the response dips below the envelope [7]:

$$\sigma(\omega) = \begin{cases} |\hat{H}(\omega)| - |H(\omega)| & \text{if } |\hat{H}(\omega)| \geq |H(\omega)|, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

This limits inversion at notch frequencies, whose positions shift with earbud repositioning, while leaving smooth spectral regions unaffected.

3.2. Measurement Setup

The prototype was characterized with an Audio Precision system at 400 logarithmically spaced frequency points over the 20 Hz to 20 kHz bandwidth. Stepped-sine excitation at 150 mVrms was used, producing 86 dB SPL at 1 kHz at the ear simulator, representative of demanding listening levels. It was chosen because the analyzer settles to a steady-state value at each frequency, returning the fundamental, its harmonics, and the inter-harmonic noise floor from the same measurement. SNR and SDR are thus obtained directly, and reliably even where the response approaches the noise floor, without a separate noise-floor acquisition.
Unlike prior implementations limited to 10 kHz by the IEC 60318-4 coupler standard, the GRAS 45CA ear simulator, accurate up to 20 kHz, permits equalization to extend across the full audible range. Hirai et al. [12] showed that including frequencies above 10 kHz significantly improves perceptual similarity for virtual headphones.

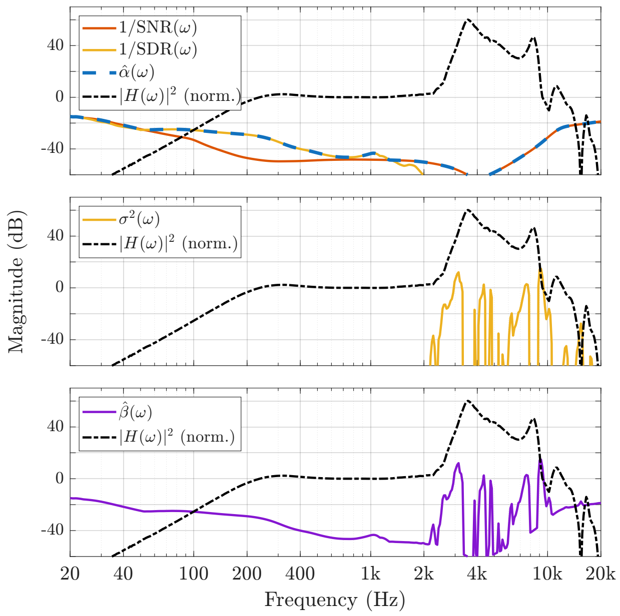


Figure 5: Regularization parameters: bandwidth limitation $\hat{\alpha}(\omega) = \max(1/\mathrm{SNR}(\omega), 1/\mathrm{SDR}(\omega))$ with its two components (top), notch regularization $\sigma^2(\omega)$ (middle), and combined $\hat{\beta}(\omega)$ (bottom). The normalized headphone response $|H(\omega)|^2$ is overlaid as a dashed curve in each panel.

3.3. Inverse Filter Design

The resulting regularization parameters are shown in Fig. 5. Between 50 Hz and 1 kHz, the SDR-based term dominates $\hat{\alpha}(\omega)$, confirming driver distortion as the limiting factor in this range and at this excitation level. Below 50 Hz, both terms contribute comparably and both substantially exceed $|H(\omega)|^2$, therefore preventing any inversion. Above 10 kHz, the SNR-based term rises as headphone output falls, limiting the inversion effort. Notch regularization $\sigma^2(\omega)$ contributes notably to $\hat{\beta}(\omega)$ at two spectral dips between 10 kHz to 16 kHz, attributed to ear canal resonances of the measurement fixture.

Since SNR and SDR differ between channels, independent equalization would produce asymmetric bass bandwidths, which would be readily perceptible. The regularization is therefore computed from the more restrictive parameters across both channels, ensuring identical low-frequency extension in each ear.

The resulting inverse filters are smoothed to 1/24-octave resolution; the minimum-phase impulse response is then derived from the smoothed magnitude spectrum and used directly as a FIR filter for convolution with audio signals.

The measured headphone transfer function $H(\omega)$ and the designed inverse filter $H_{\mathrm{RI}}^{-1}(\omega)$ are shown in Fig. 6, both normalized to 0 dB at 1 kHz. The two curves are approximately mirror images in the passband; departures from this symmetry at low frequencies and above 10 kHz reflect the regularization limiting inversion where driver output is insufficient.

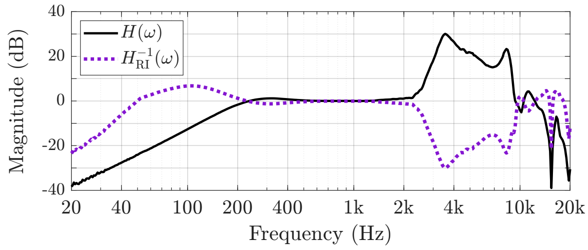


Figure 6: Measured headphone transfer function $H(\omega)$ (dashed) and designed inverse filter $H_{\mathrm{RI}}^{-1}(\omega)$ (solid), both normalized to 0 dB at 1 kHz. Departures from a mirror-image relationship at low frequencies and above 10 kHz reflect regularization-limited inversion.

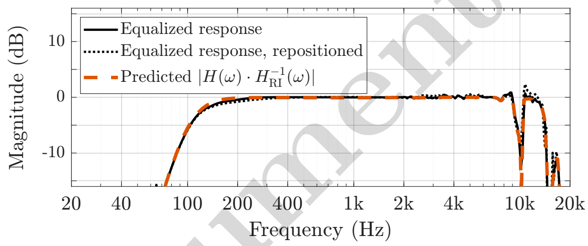


Figure 7: Equalized frequency response (solid) and predicted $|H(\omega) \cdot H_{\mathrm{RI}}^{-1}(\omega)|$ (dashed), normalized to 0 dB at 1 kHz. One representative repositioning (dotted) is overlaid.

3.4. Validation of the Equalized Response

The equalization was validated by measuring the acoustic output of the prototype on the measurement fixture in response to a pre-equalized stepped-sine signal. A custom signal with 1/24-octave frequency steps and adaptive tone durations was required, as Audio Precision does not support real-time FIR convolution. The measurement was repeated across several repositionings to assess sensitivity to fit variation on the pinna.

Figure 7 presents the measured equalized response, normalized to 0 dB at 1 kHz, against the predicted $|H(\omega) \cdot H_{\mathrm{RI}}^{-1}(\omega)|$. Above 200 Hz, the equalized response is flat within ±1 dB of the target, with variation of no more than ±2 dB between the two positions. Narrow spectral dips inherited from the original headphone response remain, as notch regularization deliberately limits inversion at those frequencies to prevent position-dependent artifacts. The dip near 10 kHz, for instance, is absent at other positions, so inverting it would introduce ringing wherever the notch does not occur.

Bass extension is limited to approximately 100 Hz by the weak low-frequency output of the prototype drivers; as discussed in Section 3.1, relaxing λ_1 or λ_2 below unity would push it lower, a penalty further softened by the markedly lower THD encountered at typical listening levels than at the characterization level.

4. Conclusion and Outlook

This paper presented a research platform for controlled perceptual studies with open-ear clip-on earbuds. A wired, 3D-printed, dual-transducer prototype was designed with a finite-element model and finalized through distortion-limited measurements, recovering about 6 dB of low-frequency headroom over a single driver. The virtual-headphone inversion was extended with a signal-to-distortion criterion that bounds equalization where aggressive bass gain would force the transducers into excessive distortion. Equalized to a flat target, the prototype matches it within ± 2 dB from 200 Hz to 13 kHz and stays robust to repositioning, apart from a narrow 10 kHz notch left uncorrected to avoid position-dependent ringing.

The result is a transparent, fully controllable audio path that suits experiments in which reproduced sound must coexist with the real acoustic scene and opens perceptual studies specific to the form factor, such as its preferred tonal balance. Indeed, target curves for circumaural and in-ear headphones already differ with ear coupling, and the clip-on's placement-sensitive coupling may call for one of its own. A second work would weigh the perceptual value of extended bass against the higher low-frequency distortion it requires.

References

[1] H. Chiba, T. Kako, and N. Harada, "Designing an Open-Back Enclosure to Mitigate Sound Leakage for Open-Ear Headphones," *IEEE Access*, vol. 14, pp. 5684–5700, 2026.

[2] S. Lutz, S. Riedel, and M. Frank, "Multi-Source Localization in the Horizontal Plane: Real vs. Virtual Loudspeaker Environments," in *Fortschritte der Akustik – DAGA 2026*, Dresden, Germany, 2026.

[3] S. E. Olive, T. Welti, and E. McMullin, "Listener Preference For Different Headphone Target Response Curves," in *Audio Engineering Society Convention 134*, Rome, Italy, 2013.

[4] S. E. Olive, T. Welti, and O. Khonsaripour, "The Preferred Low Frequency Response of In-Ear Headphones," in *Audio Engineering Society Conference on Headphone Technology*, Aalborg, Denmark, 2016.

[5] T. Welti, S. E. Olive, and O. Khonsaripour, "Validation of a Virtual In-ear Headphone Listening Test Method," in *Audio Engineering Society Convention 141*, Los Angeles, USA, 2016.

[6] S. Temme, S. Olive, S. Tatarunis, T. Welti, and E. McMullin, "The correlation between distortion audibility and listener preference in headphones," in *Audio Engineering Society Convention 137*, Los Angeles, USA, 2014.

[7] J. G. Bolaños, A. Mäkivirta, and V. Pulkki, "Automatic Regularization Parameter for Headphone Transfer Function Inversion," *Journal of the Audio Engineering Society*, vol. 64, no. 10, pp. 752–761, 2016.

[8] D. Hammershøi and H. Møller, "Binaural technique—Basic methods for recording, synthesis, and reproduction," in *Communication Acoustics*, J. Blauert, Ed., Berlin, Heidelberg: Springer, 2005, pp. 223–254.

[9] International Electrotechnical Commission, *IEC 60318-4:2010, electroacoustics — simulators of human head and ear — part 4: Occluded-ear simulator for the measurement of earphones coupled to the ear by means of ear inserts*, Standard, Geneva, Switzerland, 2010.

[10] GRAS Sound & Vibration. "Instruction manual: GRAS 60318-4 ear simulators," Accessed: Jun. 15, 2026. [Online]. Available: `https://www.grasacoustics.com/fileadmin-gras/Products/Files/manual_60318-4_ear_simulators_LI0071.pdf`

[11] S. E. Olive, T. Welti, and E. McMullin, "A virtual headphone listening test methodology," in *Audio Engineering Society 51st International Conference*, Helsinki, Finland, 2013.

[12] H. Hirai, K. Kawahara, N. Tojo, A. Hara, and K. Hamasaki, "Precise Construction Method for Virtual Headphones," in *Audio Engineering Society 6th International Conference on Audio for Games*, Tokyo, Japan, 2024.