ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

MASTER IN TECNOLOGIE AMBIENTALI PER GLI AGENTI FISICI

DPA4560 VS META RAYBAN
A BINAURAL COMPARISON

Relatore

Prof. Angelo Farina

Correlatore

Prof. Lamberto Tronchin

Presentata da

Adriano Farina

Matricola

0001150025

Sessione Aprile 2025

Anno Accademico 2024/2025

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Thanks

To my father, gone too soon

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Abstract in Italiano

Gli “smartglasses” Meta Rayban sono uno dei primi dispositivi consumer “all in one” che permettono di ottenere registrazioni binaurali. Tradizionalmente i dispositivi per la registrazione binaurale utilizzano due microfoni, posizionati nel condotto uditivo dell’ascoltatore, o nel condotto uditivo di una testa artificiale. In entrambi i casi la codifica dell’informazione binaurale dipende dalle riflessioni sul corpo e sulle orecchie. Nel caso dei Meta Rayban, invece, viene registrato il segnale di 5 diversi microfoni, tutti esterni al padiglione auricolare, e il segnale binaurale viene sintetizzato tramite beamforming, con tecniche non descritte in letteratura.

Si è dunque valutata la qualità dei segnali binaurali prodotti tramite misure di risposta all’impulso. Indossando un paio di Meta Rayban e un set di microfoni binaurali DPA4560, è stato utilizzato il metodo dell’exponential sine sweep, campionando angolarmente ogni 10°. Utilizzando i plugin Aurora sono stati elaborati i valori di IACC (Inter Aural Cross Correlation), ITD (Interaural Time Difference) e ILD (Interaural Level Difference).

Dato che i normali test in frequenza, soprattutto per quanto riguarda la riproduzione sonora, sono ampiamente disponibili, ci siamo concentrati unicamente sui parametri binaurali.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Abstract

Rayban | Meta are the second generation of smart glasses developed by Meta and Luxottica. They are one of the first mass-market all-in-one consumer devices allowing users to record and reproduce sounds binaurally. Traditionally, binaural recording systems use two microphones, one in each hearing canal, belonging either to a person or to a dummy head. In both cases, the incoming sound reflects on the body, shoulders, and ear pinnae, thus physically encoding several binaural cues.

Rayban | Meta, instead, rely on a 5-microphone array, none of which enter the ear canal, and therefore devoid of the information encoded by the pinnae. The binaural signal is obtained through a beamforming algorithm, about which nothing has been published in the literature.

For this reason, we evaluated the quality of the binaural signals through impulse response measurements. Wearing a pair of Rayban | Meta and a set of DPA4560 binaural microphones, we used the exponential sine sweep method, sampling every 10°. Using the Aurora plugins, we obtained values for IACC (Inter Aural Cross Correlation), ITD (Interaural Time Difference) and ILD (Interaural Level Difference).

As frequency response tests, especially regarding sound reproduction, are widely available, we focused on the binaural parameters only.

Introduction to binaural audio

Human hearing is intrinsically binaural, as humans perceive sound pressure with two ears. Sounds can be perceived as coming from a more or less well defined angular direction, both on the horizontal (azimuth) and vertical (elevation) planes. The process in which this angular position is perceived by the brain is called localization.

Furthermore, sounds can be perceived at different distances from the listener. Reproducing the same signal on both ears causes the listener to perceive the sound as coming from inside their head. By introducing binaural cues, an appropriate source distance can be perceived, this is called externalisation.

The position of sound sources is reconstructed by the brain using both monoaural and interaural cues. Front-back and vertical positioning are largely dependent on monoaural cues, while angular positioning mostly depends on interaural cues. [1]

On a neurological level, there are some specialised parts of the *pons varolii* that process sounds binaurally. The interaural time difference (ITD) is computed in the medial superior olive (MSO), and the interaural level difference (ILD) is computed in the lateral superior olive (LSO).

Interaural Time Difference is the difference in the time of arrival of a signal on the two ears. This is crucial in determining the lateralization of incoming sounds, and can be used to determine the azimuth. Neurologically, it is more accurately described as a difference in delays, and it is closely related to the interaural phase difference (IPD).

Audiological modelling of binaural audio

According to Durlach [2], binaural perception works through an Equalization and Cancellation (EC) model. In the equalization phase, the Inter-aural Phase Difference is "computed" by the auditory system, by temporally realigning the signals reaching the two ears. The levels are also matched between the two ears (IID, Interaural Intensity Difference, or ILD, Interaural Level Difference).

In the cancellation phase the signals reaching the two ears are internally subtracted, causing the brain to perceive the level difference between the two signals, and reducing noise, which is generally less directional, and therefore present on both ears [3].

In order to validate this model, several experiments have been conducted by audiologists over the decades, starting in the 1950s. The general experimental approach involves the creation of threshold graphs for the masking of pure tone signals with wide-band noise.

The tests involved playing wide-band noise and a pure tone over a headphone set, matching the sound pressure levels between the two ears. The test subject is asked to identify whether they can detect the pure tone under different conditions.

There were 4 scenarios:

1. Same phase on both ears, which results in an elevated masked threshold
2. 180° phase difference between the two ears, resulting in a masked threshold difference of around 10dB
3. Monoaural playback (noise and tone on just one ear, silence on the other ear), with results that are comparable to case 1
4. Noise on one ear, noise and pure tone on the other ear, which results in a significant lowering of the masked threshold.

In scenarios 1 and 3 there is no information for the auditory system to extract during the cancellation phase, as there is no coherent difference between the left and right signal.

This is consistent with studies on single-ear listening, which is in effect an impaired form of listening. For an overview of sound localization in normal-hearing and hearing-impaired listeners, see [4].

The maximum physiologically achievable ITD is around 800μs, which corresponds to a IPD of π at 625Hz, which is generally regarded as the transition point between the low frequency and the high frequency processing systems.

From [1], we find the following equations for the ITD produced by a sound source at angle θ on an ideally spherical head of radius r in a medium in which sound travels at a speed c.

$$\tau = \frac{r}{c} 2sin\theta \text{ for frequencies under 500Hz}$$

$$\tau = \frac{r}{c}(\theta + sin\theta) \text{ for frequencies above 2000Hz.}$$

Phase mechanisms are prevalent at low frequencies, where the brain follows the wavefront, amplitude mechanisms are prevalent at high frequency, where the brain follows the sound envelope.

An Interaural Coherence parameter [5] has been developed, which puts ITD and ILD in relation to each other with a single measure.

This conceptual framework was first formalized by Raleigh [6], but recent studies [7] show slight differences in the neurological working of the mechanism, compared to the historically prevalent model, while still confirming the general idea.

From all these analyses the importance of the correlation between the signals reaching the two years emerges again and again.

Binaural rendering

Binaural recordings have existed for decades, but their reliance on headphone playback have limited their commercial success until recent years. Popular recordings in the ASMR genre largely work because of binaural effects [8].

At the same time, binaural renderings of other forms of spatial audio have become popular. The general approach is to convolve the spatial signal, whether it's channel based, spherical harmonics based, or sound-object based, with an appropriate Head Related Transfer Function (HRTF). A lot of work has been done on the rapid estimation of HRTFs for the listener, for example using image-based reconstruction [9] [10] [11]. Headphone Related Transfer Functions are also used to nullify the effects of the user's headphones.

The effectiveness of methods involving HRTF-convolution at preserving the spatial perception of the scene has been evaluated in [12].

In [13] a method has been presented for the computation of room-acoustics-related binaural parameters using HRTFs and the Ramsete simulation software.

The fundamental difference between a binaural recording and a binaural rendering of a spatial audio recording is that a binaural recording is head-locked. The binaural microphones are affixed to the dummy head, or to the ears of the person performing the recording, and therefore the spatial information corresponds to that specific orientation. Conversely, a binaural rendering of a spatial recording applies binaural information in real-time to an existing acoustic scene, which allows the listener to change their orientation. This orientation is generally tracked using gyroscopes and accelerometers, and represented as a quaternion. Software capable of binaural rendering is widely available for a variety of formats, notably Ambisonics, Mach-1, and Dolby Atmos. In fact, what Apple calls Spatial Audio is really a binaural rendering of a virtual loudspeaker array, fed from a simplified Atmos decoder.

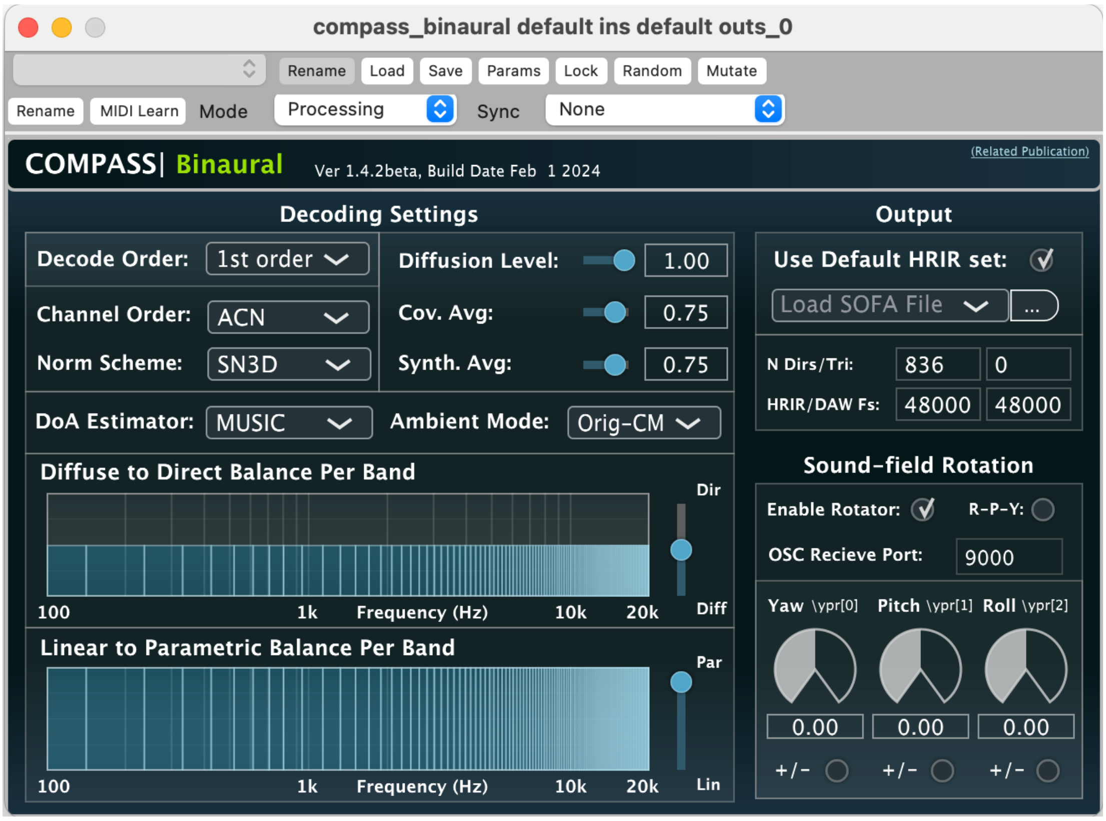


Figure 1 - Compass Binaural, a VST plugin capable of binaurally rendering Ambisonics sound-fields

This kind of software also allows for the synthesis of experimental binaural signals, which allow scientists to test the limits of the brain's decoding system. It is possible to create signals with extremely small or extremely large ITDs, despite the physical impossibility of obtaining such a stimulus in air. In fact, there is a physical limit to ITD length, trivially obtained through the diffraction formulas for soundwaves in air [1].

The comparability between multiple orientation binaural impulse responses (MOBIR) and the binaural rendering of Ambisonics impulse responses has been evaluated and measured in [14].

Furthermore, the binaural rendering of Ambisonics impulse responses has been compared to loudspeaker-based rendering in [15].

Binaural sound and reverberation

Binaural impulse responses have been employed for decades for characterizing the acoustical properties of rooms taking the spatiality into account. Historically, Ando [16] has been the most influential pioneer of binaural audio in room acoustics. Inter-Aural Cross Correlation, first defined by Ando, is the maximum value of the correlation function $\rho(\tau)$, which is described by the following equation.

$$\rho(\tau) = \frac{\int_{-\infty}^{\infty} h_r(\tau) h_s(\tau + t) d\tau}{\sqrt{(\int_{-\infty}^{\infty} h_r^2(\tau) d\tau \int_{-\infty}^{\infty} h_l^2(\tau) d\tau)}}$$

Where h_r and h_l are respectively the signals in the right and left ear. For perfectly identical signals, the result is 1, and for completely different signals it tends to zero.

Ando worked mainly in theatres, where the listener's orientation is largely static, and the focus was on completely capturing the spatiality of a room in a consistent way, as part of his efforts in developing a method for room comparison.

The use of IACC in room acoustics has been standardized in ISO3382 [17], but as pointed out in [18] there is a lack of consensus on how to interpret this data, and there is still little study on the variability of the results with misaligned dummy heads.

In fact, using the official IACC methodology requires the dummy head to face the sound source. What we are doing here is somewhat different, as we are comparing different binaural recording systems, rather than different positions or orientations in a given room.

The methodology we are using was pioneered in [19], in which IACC polar plots were first shown. There does not seem to be significant literature on the significance of the difference between the IACC polar patterns of different binaural recording systems, but intuitively, interpreting them as if they were microphone sensitivity polar patterns allows us to understand the differences between the systems under test.

In [20] a comparison between the IACC values obtained with 5 different dummy heads is performed. That comparison was performed in an anechoic chamber, and with identical source and receiver positions for all recording apparatuses. It showed significant differences in the IACC values produced by different dummy heads in the diffuse field. IACC also proved important in the vertical localisation of sounds [21].

From an audiological point of view, reverberation affects binaural parameters, and the binaural understanding of an auditory scene. Binaural loudness is higher than the corresponding monoaural loudness, but this effect is particularly pronounced for correlated binaural signals.

Reverberation reduces ILD, and it makes ITD harder to extract from the auditory stream, as successive reflections have contrasting ITDs. Temporally, the effects of reverberation on binaural technology are pretty much in line with variations in the Clarity acoustical parameter. Reverberation before the 80ms threshold acts as a reinforcement of the direct sound, while reverberation beyond the threshold acts as a masking noise, reducing the difference between the signals from the two ears, and therefore reducing the amount of information that the brain can decode. The precise time amount of this threshold is content-dependent, different kinds of stimuli behave differently. In particular, different musical genres benefit from different reverberation times, and speech intelligibility is particularly affected.

For a thorough prediction model of Speech Transmission Index (STI) under reverberation, incorporating binaural scene understanding, please see [22]

Hardware

Rayban | Meta

Introduced in 2024, the second generation of smart glasses developed by Meta and Luxottica features a 5 microphone array, that is specifically advertised as being able to provide immersive audio. To quote from the official announcement blog: *“And when you’re recording a video, the newly designed five-microphone array supports immersive audio recording, so you’ll be able to capture sounds exactly how you originally experienced them—whether in front of you, from the sides, the rear, and even above.”* [23]

There is very little publicly accessible information on the beamforming that is being performed, but we do know that it uses a Qualcomm AR1 Gen1 chipset, and white papers about sound processing on that chipset are available. There are two microphones in each arm (one facing downwards, and one to the side, hidden in the logo), and one near the nose pad.

Figure 2 - Side view, showing position of microphones – from Laurent C, Iqbal, M.Z., Campbell, A.G. Adopting smart glasses responsibly: potential benefits, ethical, and privacy concerns with Ray-Ban Stories – AI Ethics – CC-BY 4.0

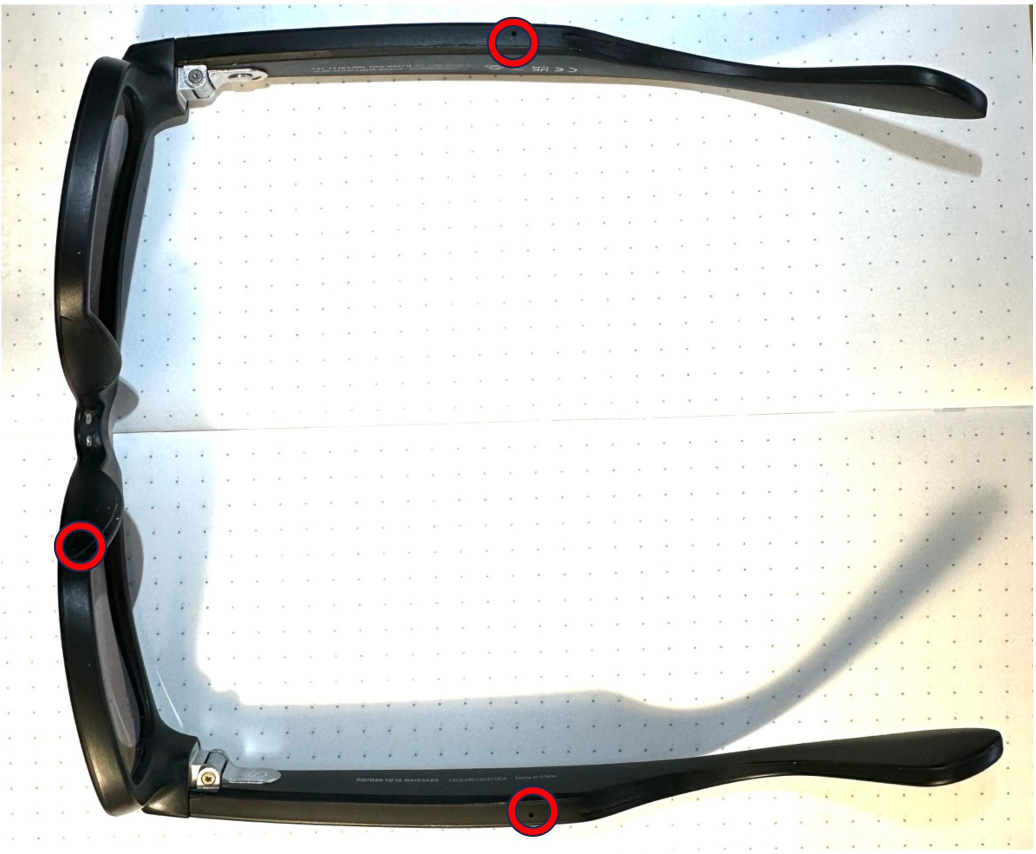


Figure 3 - Lower microphones, in the arms and near the nose pad

No specific information on the kind of microphone used is officially available, but it is virtually impossible that they are anything other than omnidirectional MEMS, given the space constraints and their commonality in the mobile electronics industry.

MEMS means **micro-electromechanical systems,** MEMS microphones are available in a wide range of varieties and form factors, but they're almost always omnidirectional and PCB-mounted. This means that they are sensitive to pressure, and not to particle velocity. Computationally, it can be easier to perform beamforming with microphones that provide some sensibility to velocity, like cardioid microphones.

DPA4560

Figure 4 - Official product photo of the DPA4560

The DPA4560 [24] is a pair of high-quality microphones, specifically designed for insertion in a user's ear canal. They are a pair of 4060 CORE Miniature pre-polarized condenser omnidirectional microphones, chosen so that they have sensitivities within +-1.5dB of each other, and mounted on a hook for easy placement.

They were connected to a "DPA Microphones MMA-A d:vice" microphone preamplifier [25], which was connected to a Samsung Galaxy S10e through a USB-C cable. The MMA-A d:vice thus operated as a USB audio card, and its input was recorded using the USB Audio Recorder Pro [26] app, at a 48kHz sampling rate in a 24-bit WAV file.

Bedrock BTB115 Advanced Talkbox

Figure 5 - Bedrock Talkbox, official product photo from the company website

The Bedrock BTB115 [27] is essentially a calibrated loudspeaker rig, specifically designed for acoustic testing. Its internal DSP system automatically reproduces sounds at the sound pressure level (SPL) specified in its settings, when measured at a distance of 1m. Its main purpose is performing STI (speech transmission index) measurements, and its shape, reminiscent of a human head, is an indication of its function.

The loudspeaker's directivity pattern closely matches that of a human speaker. It features small touchscreen displays in lieu of eyes, the left one controlling the volume, and the right one controlling the laser distance meter, which is itself placed in the nose. On the back, a large touchscreen display allows full control of the device. Test signals can be added by simply copying files onto the internal memory using the USB-C port.

Measurement methodology

A test subject was fitted with a DPA4560 binaural kit, and with a pair of Meta Rayban glasses, making sure that there was no mechanical interference between the two devices. The Meta Rayban recorded a series of videos on its own internal memory, making sure not to split measurements between different takes. Multiple takes were necessary, as videos have a maximum length of 3 minutes. Surprisingly, despite being recorded as videos, the audio files had a 44.1kHz sample rate. Even more surprisingly, according to their metadata they were stored in 32-bit. For ease of processing, they were converted to 48kHz using the "convert sample type" function in Adobe Audition.

The audio from the DPA4560 kit was recorded using an Android smartphone running USB Audio Tool Pro, as a 48kHz 24-bit wav file.

Figure 6 - The test subject wearing both recording systems

The subject sat on an office chair fitted with an angular measurement device, obtained by fixing a caliper to a circular sheet of paper, on which 10° increments had been marked. While the sheet of paper did flex vertically, it remained unperturbed horizontally.

Figure 7 - Office chair with angle marking system

Figure 8 - Detail of the angle marking system

The vertical alignment of the subject’s head was confirmed visually, using an observer (the thesis supervisor) in a fixed position and a vertical segment of the wall as reference.

The test signal was an Exponential Sine Sweep signal (20-20000Hz, 25s, with 5s of silence between repetitions) played at 64dB(A) from a Bedrock BTB115 Advanced Talkbox, positioned at 1.5m meters of height, and 1m of horizontal distance from the centre of the segment connecting the ears of the test subject.

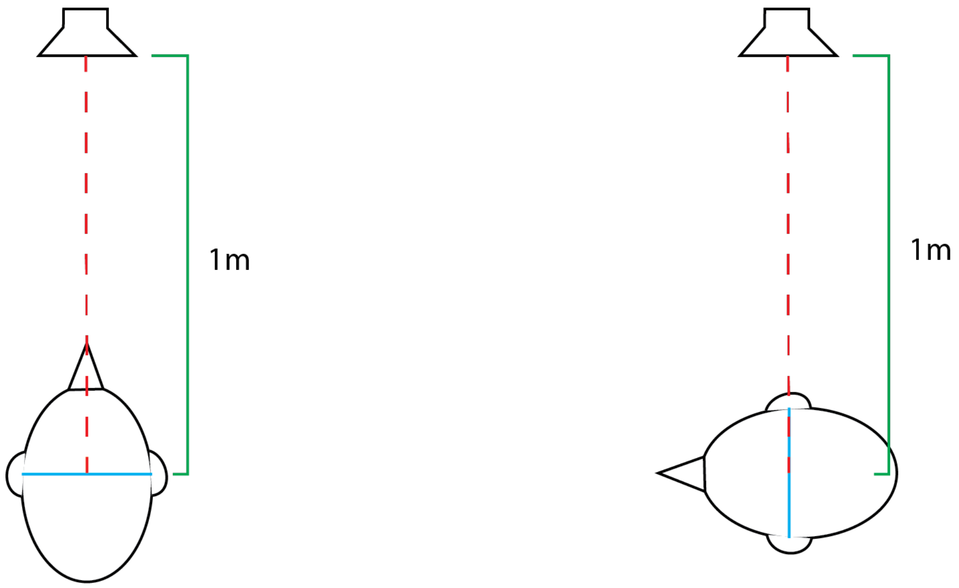


Figure 9- Geometrical alignment of the measurements

We obtained 36 binaural impulse responses for each recording system. These were convolved with the appropriate inverse sweep processed with the Acoustical Parameters plugin from the Aurora suite. Of particular interest were the binaural parameters: IACC, ITD and ILD.

Analysis of the results

The main difficulty of this study was the attempt at evaluating the differences between the two systems without resorting to user testing, which is the most common method in the literature [28]. The chosen parameters are closely related to externalisation and localisation, they rely on the same mechanisms employed by the brain, but they are still not direct measurements of human sensations. We compared proxies for the qualia, not the qualia themselves.

IACC

In the following figure, we can observe the polar plot for IACC values at different octaves with the two recording systems. We can observe that at lower frequencies both systems are approximately omnidirectional, while at higher frequencies they start beamforming more and more. The DPA4560 seems to be somewhat more regular in angular terms, with the Rayban presenting slightly backwards-facing lobes.

In the next page, we see plots with IACC values on the y-axis, and octaves on the x-axis. We can observe that in both cases the values tend to 1 under 250Hz, and they rapidly diverge at increasing frequencies. We also plotted the variance of the two systems.

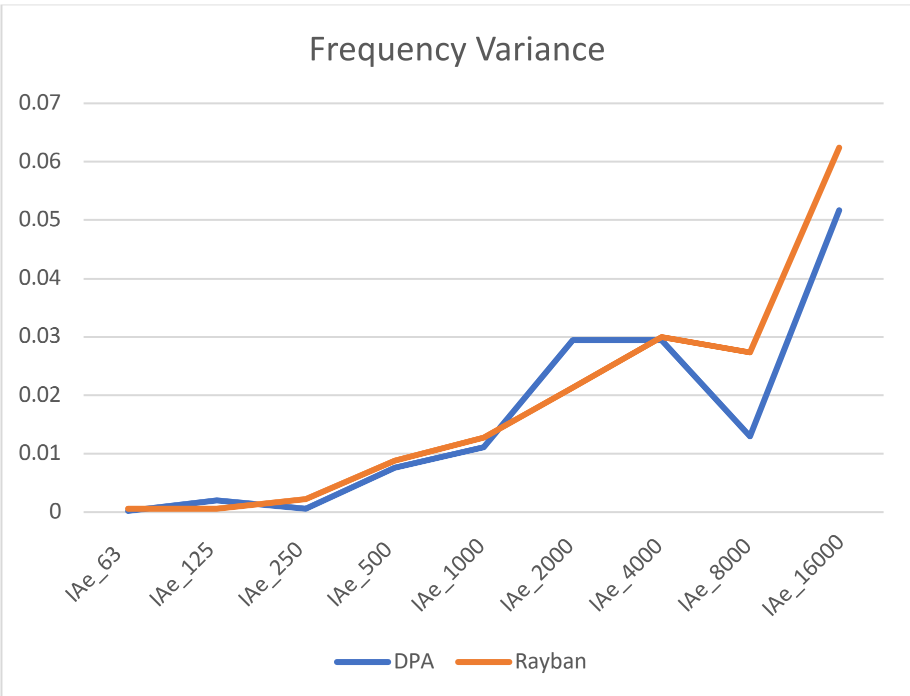


Figure 10 - IACC Frequency Variance

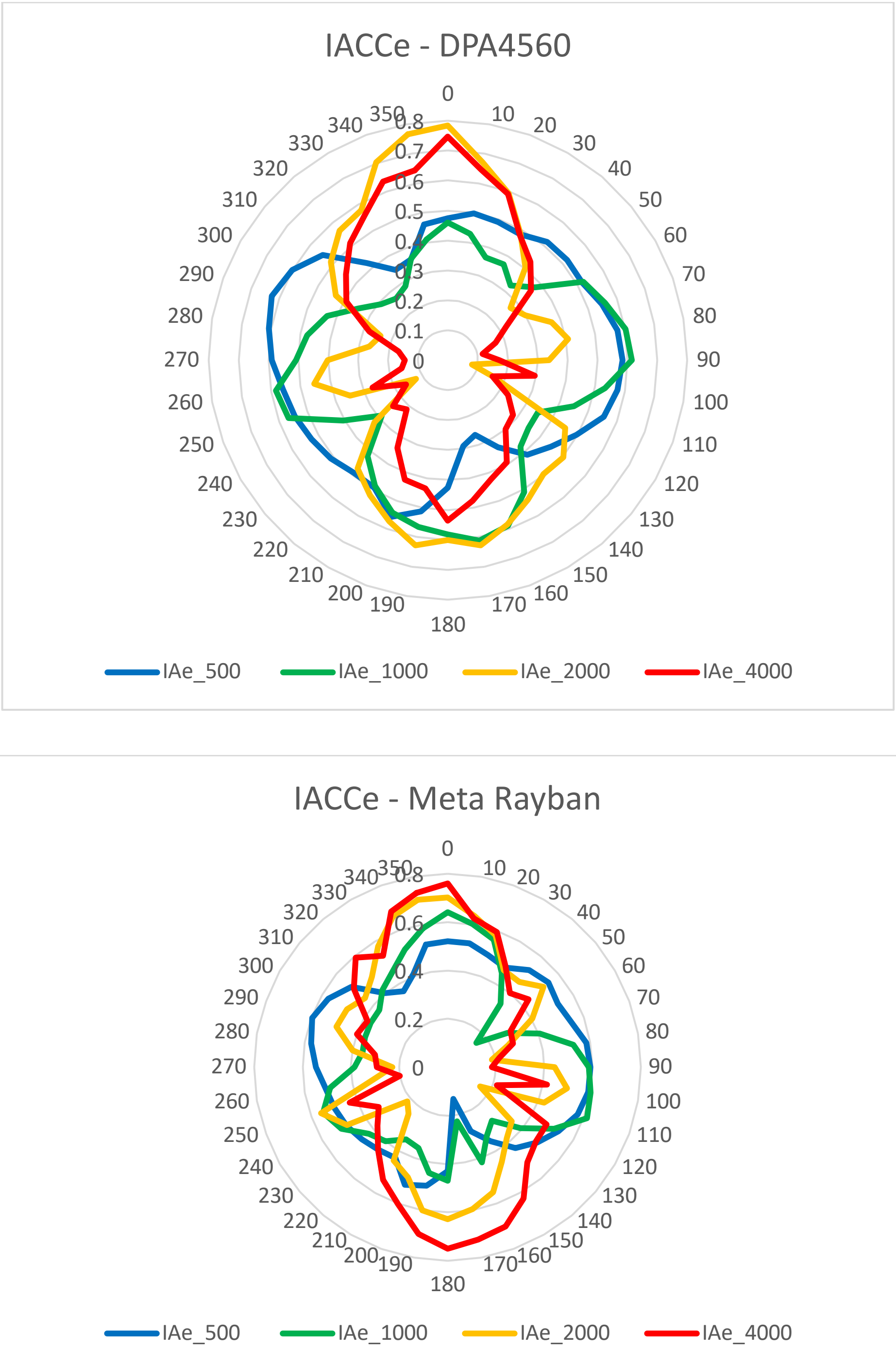
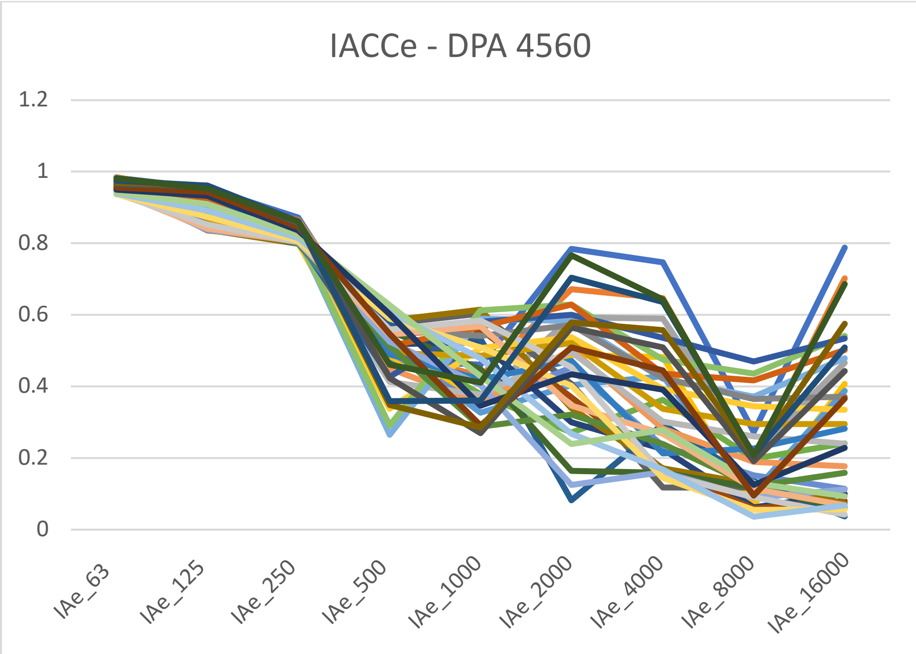


Figure 11 - Comparison of IACC polar patterns



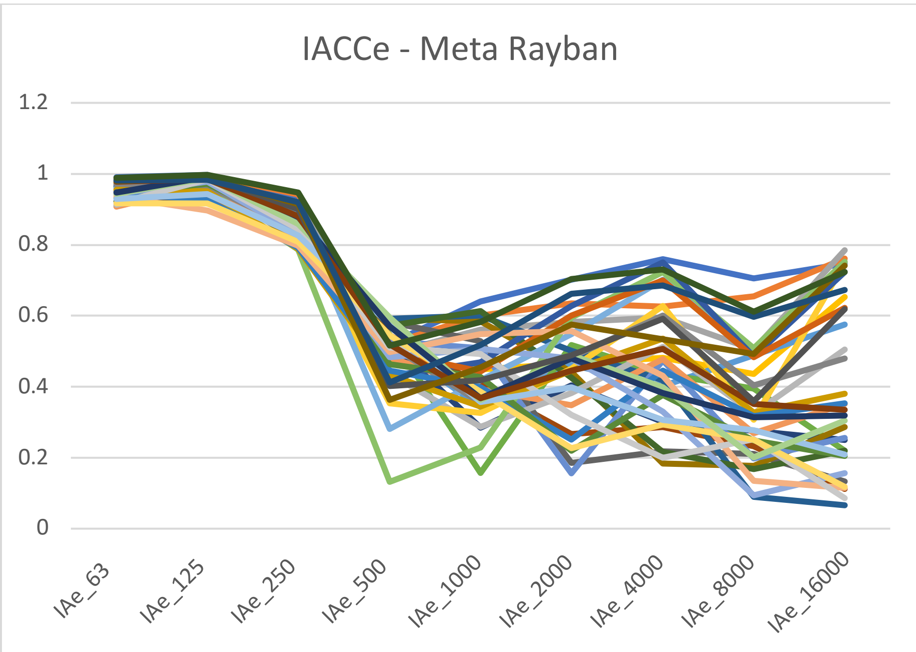


Figure 12 - Comparison of IACC per octave

ITD

In the figures below, we see a comparison between the τ_{IACC} values of the two systems. This is fundamentally the same thing as ITD, but for a clear illustration of the definition of τ_{IACC}, see Figure 5 from [29]. Both systems follow the expected sinusoidal profile.

Just like for IACC, ITD presents significant variance on a frequency basis, but very little coherent information can be gleaned by analysing variance on an angular basis.

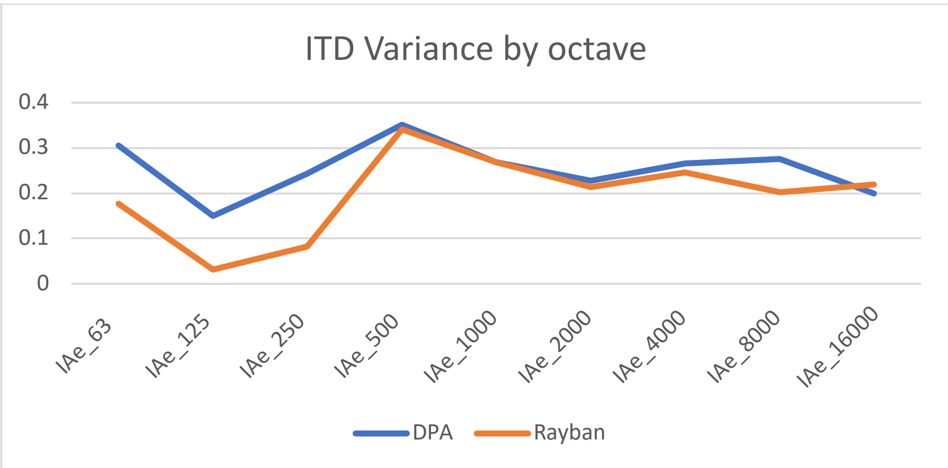


Figure 13 - ITD Variance per octave

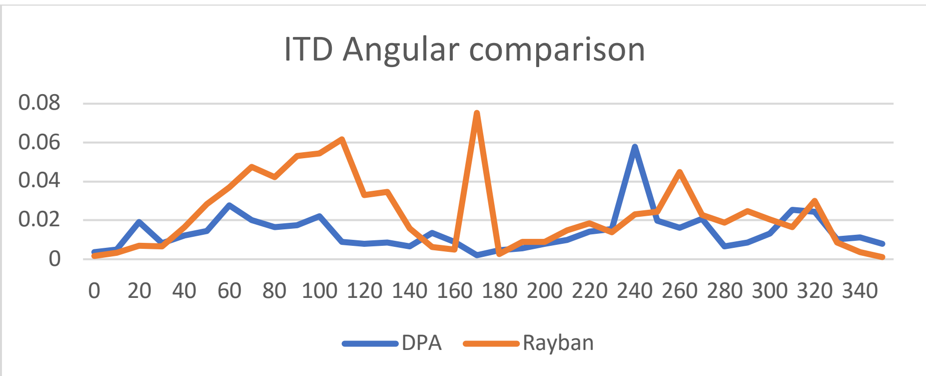
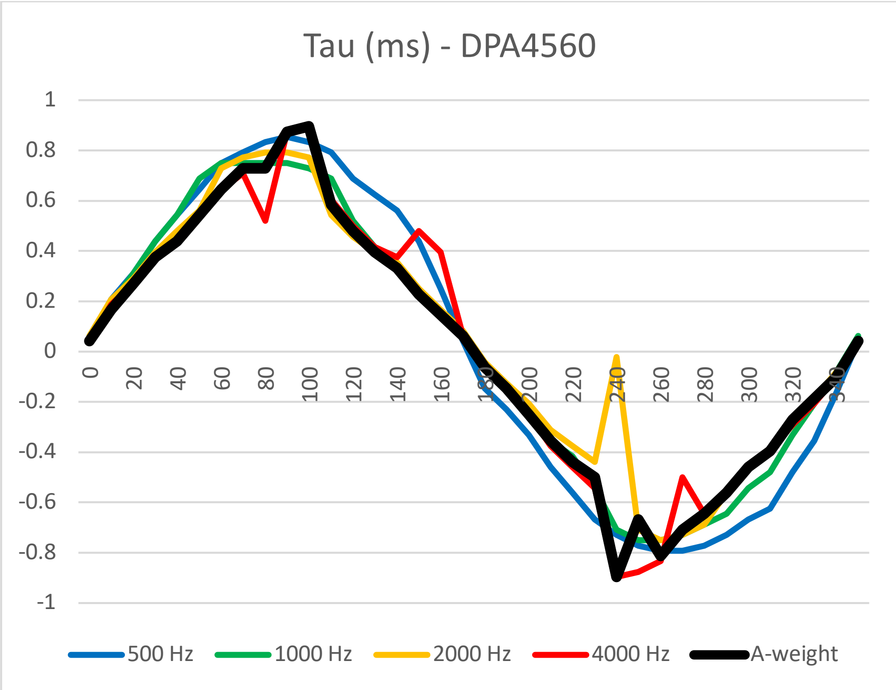


Figure 14 - ITD comparison by angle

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA



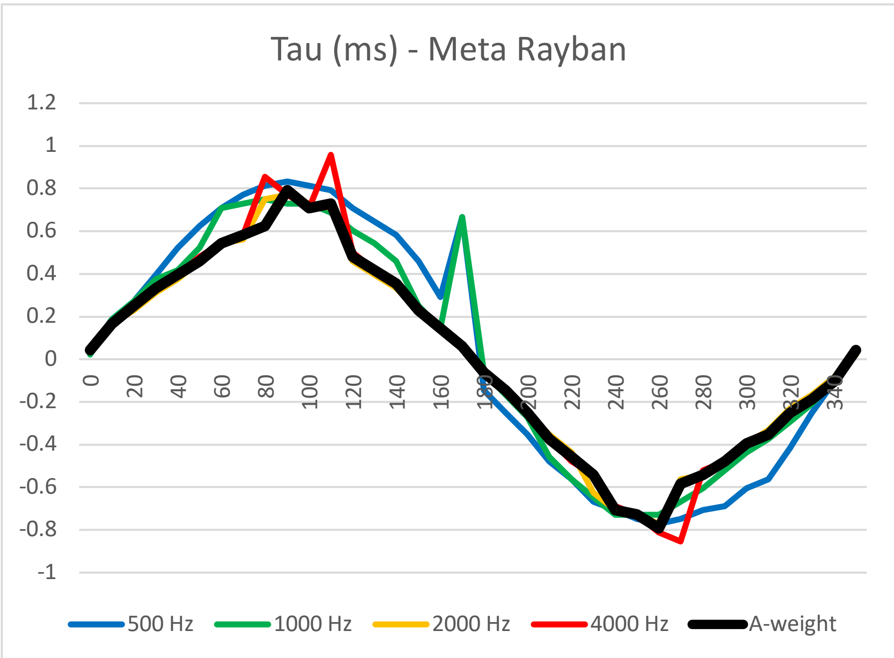


Figure 15 - Comparison of ITD values

ILD

This is the parameter where the difference between the two systems is most evident. In particular the Rayban system presents a noticeable flattening of the curve between 230° and 310°, compared to the DPA, and to the geometrically expected sinusoid-like shape. This asymmetry in the results is a serious flaw, and it should be verified more thoroughly, as it could be the result of a measurement error, however unlikely.

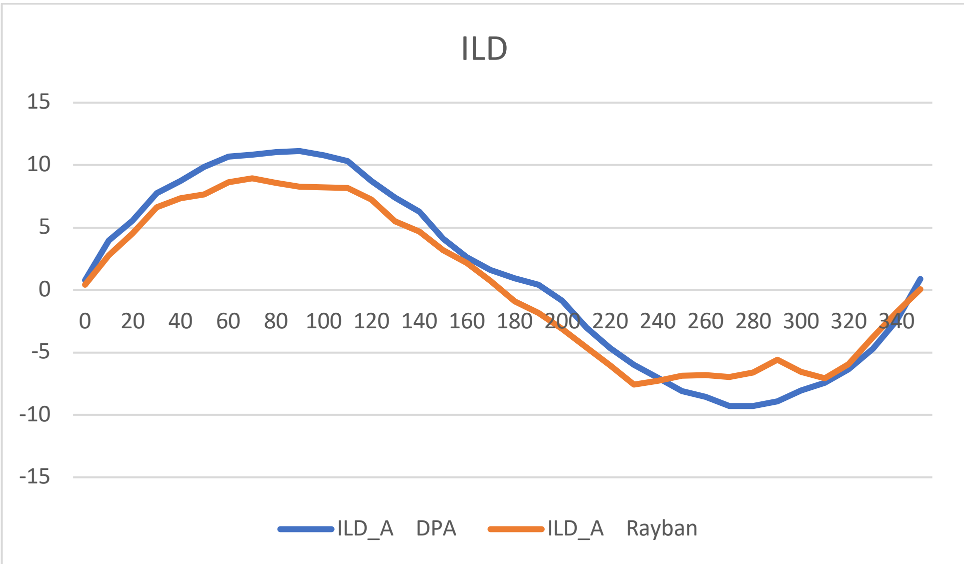


Figure 16 - Comparison of A-weigthed ILD values, by angle of incoming sound

The Meta Rayban has consistently lower values of ILD, at every angle, but especially when the sound source is perpendicular to the subject.

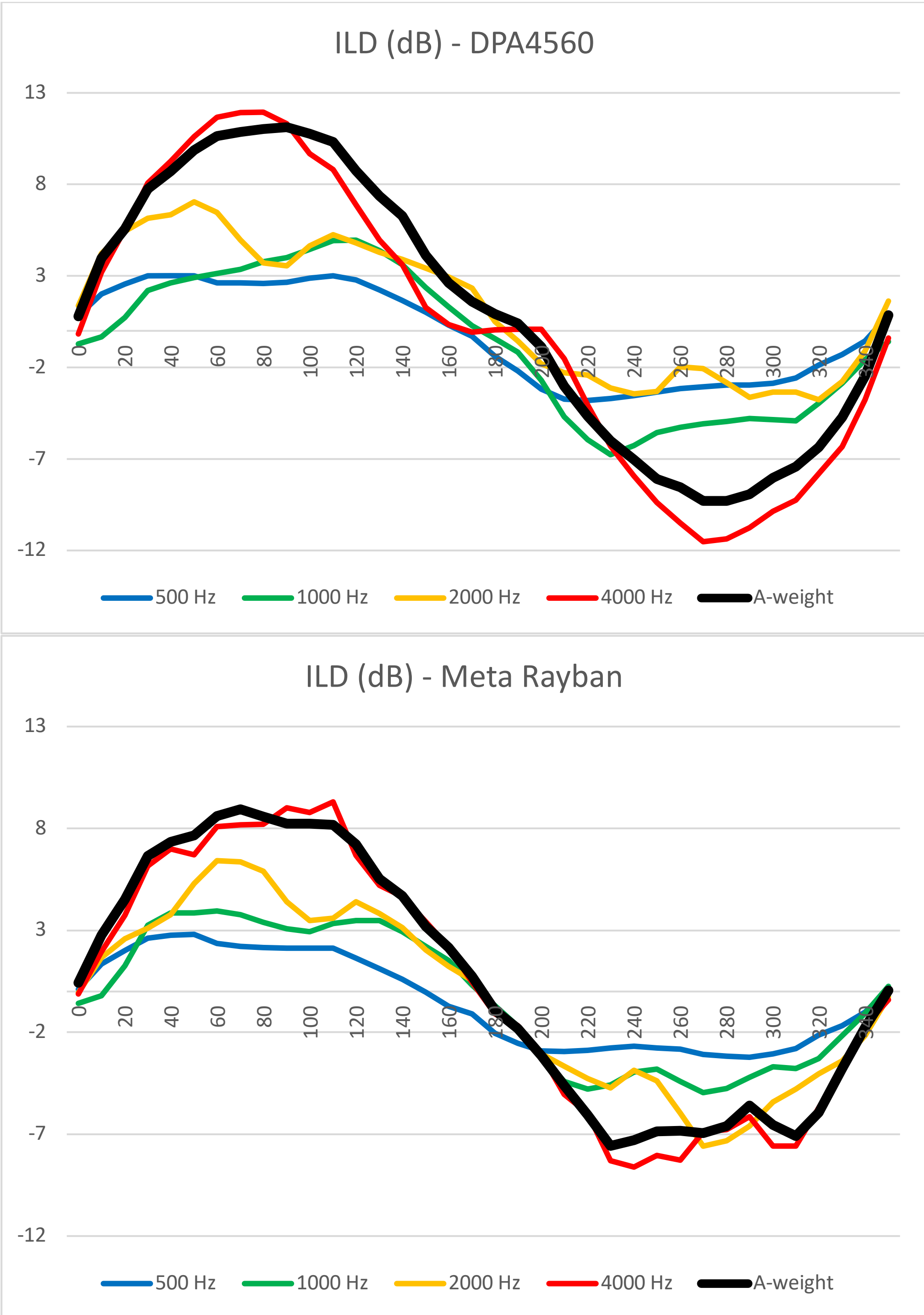


Figure 17 - Angular comparison of ILD values

Conclusions

Both systems can capture convincing binaural sound scenes, as evaluated through purely quantitative analysis.

It is plain to see that the processed parameters have some amount of variance, that is quite consistent with the existing literature, but which could be statistically mitigated if each measurement was repeated several times.

Several studies on the relationship between these quantitative measurements and psychoacoustic effects of source localization are available in the scientific literature. The comparatively lower ILD in the Rayban recording points to a slightly worse externalisation and localisation performance compared to the DPA.

However, the Meta Rayban are a playback as well as a recording system, and it would be interesting to perform further experiments, comparing their performance as binaural reproduction devices with a pair of reference headphones. Plausibly, some of the localisation cues that the Rayban does not record could be provided by the listener's very head, as they are physically present both during the recording and during playback. To test this, a dummy head should be used. The same test signal should be recorded on the dummy head, on the Meta Rayban positioned on the dummy head, on the Rayban positioned on the subject's head. It should also be recorded on the DPA microphones, positioned both on the human subject and on the dummy head. In this way, a full separation will be achieved between the effects of the recording system and the effects of the HRTF.

A paper on this subject will be presented at I3DA 2025.

Most importantly, further work is needed to validate the results by testing on actual human subjects. In particular, it would be important to evaluate the localization and externalization effects achieved by the two systems, using the techniques described in [30].

Bibliography

[1] D. Wang and G. J. Brown, “Binaural Sound Localization,” in *Computational Auditory Scene Analysis*, John Wiley & Sons, 2005, p. chapter 5.

[2] N. I. Durlach, “"Binaural signal detection: Equalization and cancellation Theory",” in *Foundations of Modern Auditory Theory*, New York, Academic Press, 1972, pp. 369-462.

[3] R. Wan, N. I. Durlach and H. S. Colburn, “pplication of an extended equalization-cancellation model to speech intelligibility with spatially distributed maskers,” *The Journal of the Acoustical Society of America*, vol. 128, no. 6, pp. 3678–3690, doi: 10.1121/1.3502458, 2010.

[4] M. A. Akeroyd, “An overview of the major phenomena of the localization of sound sources by normal-hearing, hearing-impaired, and aided listeners,” *Trends in Hearing*, no. 18, p. doi: 10.1177/2331216514560442, 2014.

[5] N. Gößling, D. Marquardt and S. Doclo, “Perceptual Evaluation of Binaural MVDR-Based Algorithsm to Preserve the Interaural Coherence of Diffuse Noise Fields,” *Trends Hearing*, 2020.

[6] L. Rayleigh, "On our perception of sound direction," *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science,* pp. 214-232, doi: 10.1080/14786440709463595, 1907.

[7] T. P. Franken, P. X. Joris and P. H. Smith, "Principal cells of the brainstem's interaural sound level detector are temporal differentiators rather than integrators," *eLife,* p. doi: 10.7554/eLife.33854, 2018.

[8] T. Koumura, M. Nakatani, H. I. Liao and H. M. Kondo, "Dark, loud, and compact sounds induce frisson," *Quarterly Journal of Experimental Psychology,* vol. 74, no. 6, pp. 1140-1152, doi: 10.1177/1747021820977174 , 2020.

[9] S. I. E. E. Limited, "Methods and systems for synthesising an HRTF". United States of America Patent 2025/0099854 A1, 27 03 2025.

[10] V. Jayaram, I. Kemelmacher-Shlizerman and S. M. Seitz, "HRTF Estimation in the Wild," in *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, New York, 2023.

[11] T. Nishino, N. Inoue, K. Taked and F. Itakura, "Estimation of HRTFs on the horizontal plane using physical features," *Applied Acoustics,* vol. 68, no. 8, pp. 897-908, doi: 10.1016/j.apacoust.2006.12.010, 2007.

[12] P. Antoniuk, S. K. Zielinski and L. H, "Ensemble width estimation in HRTF-convolved binaural music recordings using an auditory model

and a gradient-boosted decision trees regressor," *Journal of Audio Speech Music Processing,* pp. 53, doi:10.1186/s13636-024-00374-2, 2024.

[13] L. Tronchin, A. Farina, T. Hotehama, R. Shimokura and V. Tarabusi, "The calculation of Binaural Parameters by means of HRTFs and Ando's method," *Proceedings of ICA2004 - Acoustical Science and Technology for Quality of Life,* p. doi: 11585/18596, 2004.

[14] M. Zaunschirm, M. Frank and F. Zotter, "Binaural Rendering with Measured Room Responses: First-Order Ambisonic Microphone vs. Dummy Head," *Applied Sciences,* p. doi: 10.3390/app10051631, 2020.

[15] D. Thery and B. F. G. Katz, "Auditory perception stability evaluation comparing binaural and loudspeaker Ambisonic presentations of dynamic virtual concert auralizations," *Journal of the Acoustical Society of America,* pp. 246-258, doi:10.1121/10.0002942, 2021.

[16] Y. Ando, Concert Hall Acoustics, Berlin: Springer, 1985, pp. doi: 10.1007/978-3-642-69810-1.

[17] *ISO 3382-1:2009, Measurement of room acoustic parameters,* 2009.

[18] I. B. Witew, A. Lindau, J. V. Schuitmann, M. Vorländer, S. Weinzierl and D. D. Vries, "Uncertainties of IACC related to dummy head orientation," *DAGA,* 2010.

[19] A. Farina and R. Ayalon, "Recording concert hall acoustics for posterity," *24th AES Conference on Multichannel Audio,* pp.

http://pcfarina.eng.unipr.it/Public/Papers/185-AES24.PDF, 26-28 June 2003.

[20] A. Farina, P. Martignon, A. Capra and S. Fontana, "Measuring impulse responses containing complete spatial information," in *4th Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan*, Honolulu, 2006.

[21] W. L. Martens and D. Stepanavicius, "Interaural Cross-Correlation Affects Perceived Elevation of Auditory Images Presented via Height Channels in Multichannel Audio Reproduction Systems," in *The 2016 Joint Conference of The Australian Acoustical Society and The Acoustical Society of New Zealand*, Brisbane, Australia, 2016.

[22] T. Leclère, M. Lavandier and J. F. Culling, "Speech intelligibility prediction in reverberation: Towards an integrated model of speech transmission, spatial unmasking, and binaural de-reverberation," *Journal of the Acoustical Society of America,* pp. 3335-3345, doi: 10.1121/1.4921028, 2015.

[23] Meta corporation, "Ray-Ban | Meta Smart Glasses Collection: Now Available," 17 October 2023. [Online]. Available: https://www.meta.com/en-gb/blog/ray-ban-meta-smart-glasses-collection-pre-orders/.

[24] Dennis Pro Audio Microphones, "4560 Binaural Headset Microphone," 2016. [Online]. Available:

https://www.dpamicrophones.com/microphones/headset-and-earset/4560?variant=24.

[25] "MMA-A Digital Audio Interface - Flawless mobile audio," 07 09 2019. [Online]. Available: https://web.archive.org/web/20190907095155/https:/www.dpamicrophones.com/mma-a-digital-audio-interfaces.

[26] eXtream Software Development, "USB Audio Recorder PRO," 2025. [Online]. Available: https://play.google.com/store/apps/details?id=com.extreamsd.usbaudiorecorderpro&hl=en-US.

[27] Bedrock, "Bedrock BTB115 Advanced TalkBox," 2024. [Online]. Available: https://bedrock-audio.com/index.php/bedrock-btb115.

[28] K. DiPasquale, "Comparing Externalization Between the Neumann KU100 Versus Low Cost DIY Binaural Dummy Head," *Journal of the Audio Engineering Society*, 2019.

[29] D. D'Orazio, P. Guidorzi and M. Garai, "A Matlab Toolbox for the analysis of Ando's factors," in *AES 126th Convention*, Munich, 2009.

[30] J. Cubick, C. Sánchez Rodríguez, W. Song and E. N. MacDonald, "Comparison of binaural microphones for externalization of sounds," in *Proceedings of International Conference on Spatial Audio*, Graz, Austria, 2015.