

Orientation-Dependent Head-Related Transfer Functions: Effects of Source Radiation Angle on Binaural Perception

Jeremy Tsuaye

Sydney School of Architecture, Design and Planning
The University of Sydney
Sydney, NSW, Australia
jeremy.tsuaye@sydney.edu.au

Agnieszka Roginska

Steinhardt School of Culture, Education, and Human Development
New York University
New York, NY, United States
roginska@nyu.edu

Abstract—Head-related transfer functions are a fundamental concept in the field of spatial audio. However, most datasets have been measured with the sound source radiating toward the binaural recording device with less consideration given to source orientation. In contrast, this paper discusses the influence of source radiation on spatial hearing. Measurements were conducted at the 3D3A Lab to collect an orientation-dependent head-related transfer function dataset. Both head rotation and source rotation were included to capture on- and off-axis radiation behaviour. The collected impulse responses were then processed using modified diffuse-field equalisation to finalise the dataset. Interaural differences and frequency responses were analysed to examine how source radiation angle affects binaural perception. Additionally, a human subject perceptual study was designed to further evaluate the impact of sound source orientation, showing that participants identified off-axis stimuli with 95.6% accuracy, compared to 71.1% for level attenuation alone. An ILD-based pseudo-azimuth pattern was observed and discussed as a preliminary indication that source radiation angle may bias binaural level cues.

Index Terms—Spatial Audio, HRTF, Source Directivity, Signal Processing

I. INTRODUCTION

The Internet of Sounds (IoS) refers to connected systems in which networked audio devices interact to support sound-based applications and services [1]. Technologies such as binaural recording techniques have been introduced to capture spatial information for various applications [2]. In spatial hearing, interaural time difference (ITD) and interaural level difference (ILD) are described as the main binaural cues for localisation, which contribute to spatial auditory perception and localisation [3]. Due to the relationship between wavelength and head size, the primary binaural cues for auditory perception shift from one to the other, with the critical frequency of approximately 1.5 kHz, or sometimes 2 kHz [4], being a pivotal point at which the wavelength becomes comparable to the dimensions of the human head [5]. However, this is less a strict boundary and more a transitional frequency zone [6]. Information available to one ear can also provide sufficient spatial data, which is called monaural cues [7]. Monaural cues contain spectral information shaped by

anatomical features and vary between individuals, which is particularly useful for elevation detection [5], [7], [8]. These binaural and monaural cues are contained in head-related transfer functions (HRTFs), which are the frequency-domain representations of head-related impulse responses (HRIRs). A HRTF describes the filtering effect that occurs as sound waves interact with the listener's physiological features, contributing to sound localisation [6]. HRIRs are usually measured with a binaural recording device and a speaker playing a sine sweep in an acoustically controlled environment, usually an anechoic room [9]. The result contains essential spatial localisation data that can then be used for binaural reproduction [7], [9]. In [9], [10], different HRIR measurement methods are discussed, and it is noted that most conventional HRIR datasets focus on on-axis behaviour, with less attention given to source directivity. Sound localisation becomes more complicated in reality instead of in an anechoic chamber. In addition to the previous discussions, source directivity is another factor that affects the distribution and perception of sound [11]. While studies on the binaural perception of sound source directivity remain limited, existing research suggests that source directivity can affect spatial hearing [12], [13], [14], [15].

This paper focuses on the impact of sound source orientation, particularly how varying source radiation angles influence human auditory perception. The concept of orientation-dependent head-related transfer functions (OD-HRTFs) is introduced, which includes both on-axis and off-axis HRTFs achieved through sound source rotation. This contrasts with conventional HRTF datasets, which typically contain only on-axis data, where the sound source consistently points toward the listener or the binaural recording device. The measurements were conducted at the 3D3A Lab at Princeton University using the HEAD acoustics HMS II.3 LN HEC Artificial Head (SN 17032064), a binaural recording device capable of rotating 270 degrees. Although source rotation was used in this study, it is not considered necessary in all cases, as directivity can be measured separately and applied to the collected HRTFs. Hence, this paper provides a preliminary examination of the effect of source radiation angle on spatial perception.

II. METHODOLOGY

A. OD-HRTF Measurement

The OD-HRTF measurement was conducted in the anechoic room as shown in Figure 1, an environment chosen for its ability to provide a controlled setting for audio measurements. A full-band speaker, emitting sine sweeps, was the primary sound source in these measurements. The HEAD Acoustics HMS (binaural recording device) was 0.8 m in front of the sound source and maintained at the same height as the speaker (sound source) to ensure they were in the same horizontal plane. The source distance of 0.8 m was selected to represent a near-field listening condition in which source directivity was more likely to influence binaural cues. The speaker was positioned at azimuth angles ranging from -135° to 135° in 15° increments, relative to the HMS's perception. Instead of physically moving the speaker, the binaural recording device's reception angle was adjusted to simulate the change in the sound source's position. At each location, the speaker's radiation angle was adjusted from 0° to 345° in 15° increments. The radiation angle was 0° when the source was on-axis and facing the centre of the binaural recording device, with positive angles defined clockwise.

To clarify, both the binaural recording device and the speaker remained fixed in position during the experiment. The speaker changed its radiation angles, and the binaural recording device adjusted its reception angles to simulate different speaker positions. At each rotation of the binaural recording device, the speaker emitted a full sine sweep without changing angles. This process was repeated for each position (at each binaural perception angle), with the speaker completing a full-circle rotation. An exponential sine sweep (ESS) was

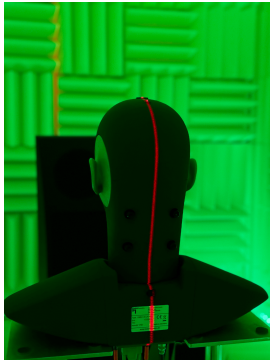


Fig. 1. Measurement setup with the HEAD acoustics HMS II.3 LN HEC Artificial Head.

generated with a frequency range from 20 Hz to 48,000 Hz, a duration of 500 ms, a pre-sweep delay of 10 ms, a post-sweep delay of 1,000 ms, and an inter-sweep delay of 200 ms. An RME M-32 D/A interface and a Bryston amplifier were used. A Plogue Bidule program generated a 1 kHz tone to test the signal flow as a trial run of the system. The BEQ interface was used to provide 200 V power to the binaural recording device and line-level outputs for further processing.

The turntable was automated to rotate a full 360 degrees at each azimuth location (effectively, at each rotation of the binaural recording device). The same Mathematica script was used to automate this process. Before each measurement, the turntable's 0° position was assigned to ensure the accurate starting angle of the sound source.

A Windows laptop was used to access the Head Acoustics GUIs: Head Control, Input Setting, and Audio Recorder. The Head Control software managed the angle and speed of the HMS rotation, while the Input Setting software configured the input gain, sampling rate synchronisation, bit-depth, high-pass filter, and equalisation methods. The Audio Recorder was not used to record the impulse responses but rather to monitor the measurement levels to avoid potential distortion and system overload.

A laser level was used to ensure that the centre of the HMS, the centre of the loudspeaker, and the centre of the room were aligned in a straight line. A wire was suspended in the anechoic chamber, connecting the centre points of the front and rear walls. The laser level projected a beam of light using the wire as a reference, and all the equipment was positioned so that the laser beam illuminated the centre points of each device. Since the speaker was fairly tall, with different components placed at various heights (elevations), the centre of the tweeter was aligned with the centre of the ears on the HMS dummy head.

B. Data Processing

One commonly used equalisation approach involves taking the microphones from the dummy head and placing them at the head's position to measure a reference impulse response [16]. However, due to the complexity of removing the microphones from the used dummy head model, no reference impulse response was measured in this study. Based on previous studies on 3D Diffuse Field Equalisation [17], a modified method was adopted in this study referred to as the 2D Diffuse Field Equalisation Compensation Method. The decision was based on the fact that the collected HRTF dataset did not include elevation variation. This method used the average on-axis raw HRTF (0-elevation) data as a reference to minimise unwanted effects such as spectral colouration. This method was applied separately to the left and right raw HRTF data. All raw HRTF data (both on-axis and off-axis) were divided by the reference average on-axis HRTF magnitude ($|H_{\text{avg}}(f)|$) to achieve EQ compensation. The method suggested in [17] involves averaging power spectra, with weighting by solid angle coverage in the case of anisotropic sampling. In contrast, $|H_{\text{avg}}(f)|$ here refers to the value computed from magnitude spectra. The original diffuse field equalisation compensation method can be found in [17], [18].

The on-axis raw HRTF magnitude denoted as $|H_{\text{on},i}(f)|$ (a total of 19 for each left and right) was used to calculate the average left and average right HRTFs as follows:

$$|H_{\text{avg}}(f)| = \frac{1}{N} \sum_{i=1}^N |H_{\text{on},i}(f)|$$

$|H(f)|$ denotes the HRTF magnitude at frequency f , and N denotes the number of raw on-axis transfer functions used to compute the reference average. Each equalised HRTF was calculated using the following equation:

$$|H_{eq,n}(f)| = \frac{|H_{raw,n}(f)|}{|H_{avg}(f)|}$$

Combining the above equations, 2D Diffuse Field Equalisation Compensation Method can be written as:

$$|H_{eq,n}(f)| = \frac{N |H_{raw,n}(f)|}{\sum_{i=1}^N |H_{on,i}(f)|}$$

where $|H_{avg}(f)|$ is the average on-axis HRTF magnitude, $|H_{on,i}(f)|$ is the magnitude of the i -th on-axis HRTF, $|H_{raw,n}(f)|$ is the magnitude of the n -th raw HRTF, $|H_{eq,n}(f)|$ is the magnitude of the n -th equalised HRTF, and N is the number of on-axis measurements.

Each transfer function in magnitude (for both on-axis and off-axis) was divided by the average magnitude to complete the equalisation process. The equalised HRIRs were then low-passed to eliminate the unnecessary information out of the microphone range [19]. Global normalisation was applied as the final step of the data processing. This step was not performed separately for the left and right channels. This step simply normalised the entire dataset based on one global maximum.

C. Perceptual Study

The main material for the perceptual experiment consisted of a series of comparisons using monophonic sounds convolved with both on-axis and off-axis HRTFs. A total of 6 stimuli (Stimuli 1, 2, 3, 4, 5, and 6) were presented, each with two audio examples (A and B). Stimulus 1A was convolved with the HRIR measured at 0° azimuth, 0° elevation, and 0° radiation, while Stimulus 1B was convolved with the HRIR measured at 0° azimuth, 0° elevation, and 90° radiation. This difference meant that Stimulus 1B used an HRIR with a different radiation angle. Stimulus 2A was identical to Stimulus 1A, whereas Stimulus 2B was an attenuated version of Stimulus 1A, with the level adjusted based on the attenuation caused by Stimulus 1B's different radiation angle. This approach was repeated for additional radiation angles. Participants were asked to select one of the following options: "Example A is off-axis," "Example B is off-axis," "Both are on-axis but with different levels," or "I can't tell the difference." This part of the survey aimed to assess participants' ability to distinguish between variations in radiation (which naturally introduces level attenuation) and level attenuation alone.

D. Rationale

In this study, fifteen participants with normal hearing ability (age 18 - 35) were recruited for the perceptual study. The number of participants was based on previous listening experiments conducted by other researchers where eleven listeners participated in the free-field signals internalisation experiment and six listeners participated in the headphone-presented signal

Stimulus	HRTF condition
1A	azi 0, ele 0, rad 0
1B	azi 0, ele 0, rad 90
2A	azi 0, ele 0, rad 0
2B	level attenuated from 2A
3A	azi 0, ele 0, rad 0
3B	azi 0, ele 0, rad 45
4A	azi 0, ele 0, rad 0
4B	level attenuated from 4A
5A	azi 0, ele 0, rad 0
5B	azi 0, ele 0, rad 165
6A	azi 0, ele 0, rad 0
6B	level attenuated from 6A

TABLE I
AUDIO STIMULI.

externalisation experiment [20]. Another research study, that investigated the perceptual impact of sound source directivity and its interaction with room acoustics, recruited twelve participants [21]. Although fifteen participants can provide meaningful output for this preliminary study, the statistical power remains limited for concluding a definitive result.

Female speech was used as the sound source for this study. In the aforementioned literature [20], different sound sources, such as full-band signals and lowpass filtered signals, have been used as the sound source for the perceptual study. High-frequency energy (energy in the 8 kHz and 16 kHz octave bands) is significant for human auditory perception [22]. Therefore, female speech signals were used to provide enough frequency information to participants. Another study shows that when the frequency bandwidth is 24 kHz, measured HRTF results in the highest correct localisation rate and the lowest front-back confusion rate among all the tests in an experiment that examined HRTF modelling using physical features [23]. This supports the use of female speech as the sound source in the perceptual study.

III. RESULTS

A. Measurement Results

1) *Magnitude Analysis:* As shown in Figure 2, the magnitude at 0° azimuth tends to be the highest at most selected frequencies, although an exception can be observed at 2 kHz. The magnitude variation caused by source rotation generally increases with frequency, which suggests that source radiation angle has a stronger influence at higher frequencies. In contrast, the magnitude differences at lower frequencies, such as 250 Hz and 500 Hz, remain relatively small, indicating that the effect of source rotation is less prominent in the low-frequency region.

Additionally, the left-channel magnitude plots are shown in Figure 3. As expected, the magnitude is generally greater on the ipsilateral side. However, consistent with the previous finding that 2 kHz behaves as an exception, the contralateral magnitude at 2 kHz is greater than the ipsilateral magnitude. This may be related to the head-shadow effect and the relationship between wavelength and head size, and is therefore of interest for further discussion, although further investigation is required to confirm this relationship.

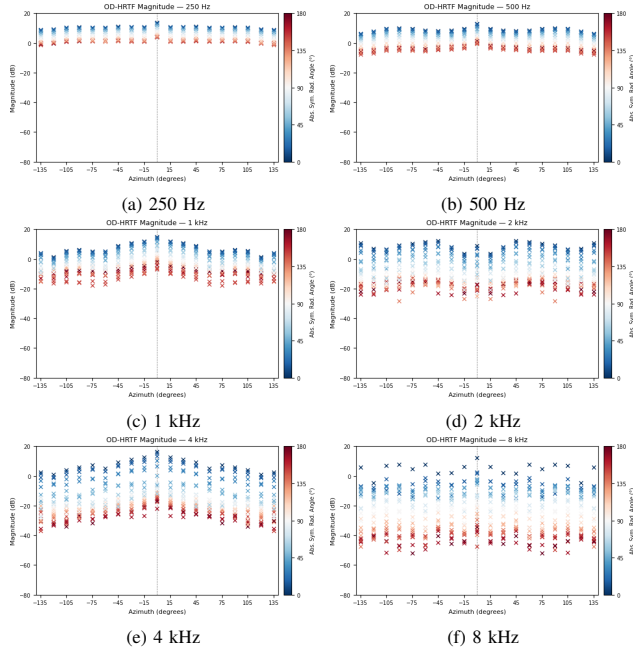


Fig. 2. Magnitude plots at selected frequencies.

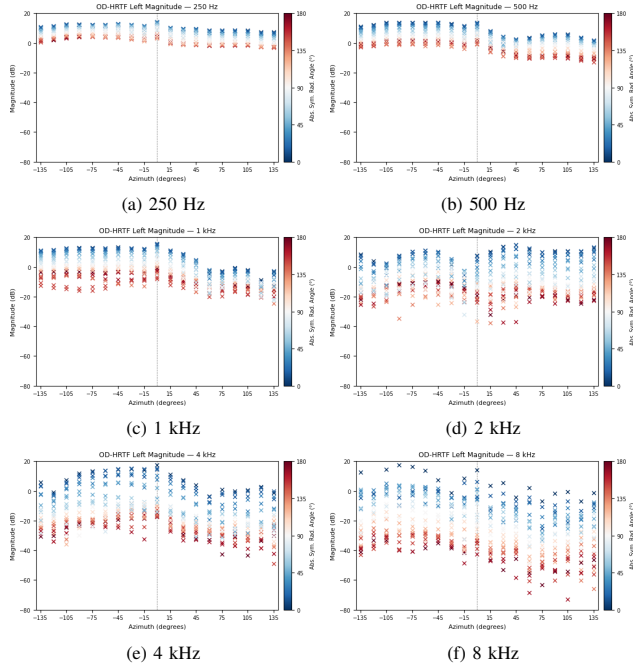


Fig. 3. Left-channel magnitude plots at selected frequencies.

2) *Pseudo Azimuth*: The interaural level difference results show a clear pattern, with smaller ILD variations across azimuth angles when the radiation angle is closer to 180° . This is physically consistent with the directivity behaviour of a non-point source. As the radiation angle deviates from on-axis, unequal levels may reach the two ears and introduce ILD even at a fixed azimuth position. A phenomenon referred to as pseudo azimuth is observed primarily in the ILD results, where the ILD at a given azimuth angle and radiation angle is similar

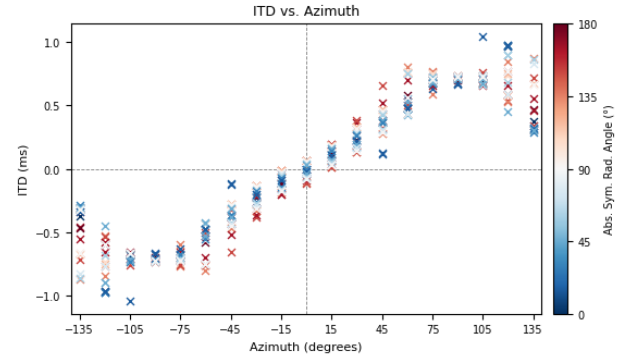


Fig. 4. Source Rotation-Induced ITD Artefacts

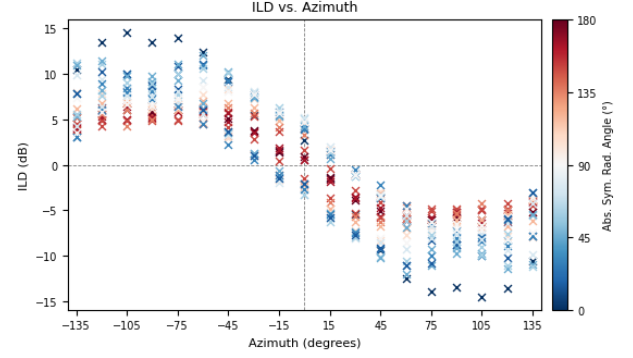


Fig. 5. ILD results

to the ILD at a different azimuth angle in the on-axis radiation case. For example, the ILD at 0° azimuth deviates from 0 dB, which indicates a lateral bias despite the source being located at the frontal position. This occurs because the directivity pattern of a non-point source produces unequal levels at the two ears as the source rotates, even when the source and listener remain at fixed positions. The pseudo azimuth effect could potentially cause left-right confusion within the range of -15° to 15° , as the ILD changes from negative to positive values (right-leading to left-leading). Although the concept of pseudo azimuth is easy to understand with the example of a sound source positioned at 0° azimuth with a radiation angle of 30° , which will theoretically produce a higher level at the left ear than the right ear, opposing the expected 0 dB ILD at the frontal position, further studies are needed to determine whether this phenomenon reflects a genuine acoustic effect or contains artefacts introduced by a slight centre shift of the source during rotation in the measurement. As the HRIRs were measured with a source distance of 0.8 m, placing the measurements in the near field, measurements with a variety of loudspeakers with different directivity patterns and distances should be conducted in the future prior to defining such a phenomenon.

B. Measurement Limitation

1) *Source Rotation-Induced ITD Artefacts*: Although the interaural time difference results show variations across different source radiation angles, these variations are likely to

have been influenced by measurement issues. Since ITD is primarily determined by the geometric path-length difference between the source and the two ears, rotating the source at a fixed position should not theoretically alter the ITD. Two possible explanations are considered. First, the speaker may not have been perfectly centred on the turntable. In this case, rotating the source would cause its acoustic centre to trace a circular path around the turntable axis. This would produce a small displacement and therefore change the geometric path length to the two ears. Second, even if the speaker itself was centred, the loudspeaker driver may not have been located at the centre of the speaker cabinet. As the source rotates, the driver would therefore also trace a circular path. This would produce a similar displacement effect. As shown in Figure 4, the ITD variations observed across radiation angle suggest that loudspeaker displacement during rotation may introduce artefacts that could be misinterpreted as genuine spatial cues. Therefore, the ITD variation observed in this study is treated as a measurement artefact and is not used as evidence for the pseudo-azimuth pattern. Future researchers performing source rotation measurements should be aware of such artefacts.

C. Perceptual Study Results

The perceptual study aims to determine how accurately participants can identify on-axis and off-axis audio stimuli, as well as level attenuation. The two types of audio stimuli used in the study are: (1) on- and off-axis stimuli, and (2) on-axis stimuli with the same amount of level attenuation as the off-axis stimuli. This comparison examines whether participants identify radiation angle variation based solely on level attenuation or through the addition of frequency characteristics.

As shown in Figure 6, the overall accuracy of the experiment is approximately 83.3%. The accuracy of level attenuation identification, approximately 71.1% ($p < 0.001$, binomial test against the 25% chance level), is lower than that of off-axis identification, approximately 95.6% ($p < 0.001$, binomial test against the 25% chance level), although 71.1% still indicates a meaningful outcome. This suggests the importance of frequency characteristics in off-axis identification, as accuracy increases when the audio stimuli include frequency information compared with level attenuation alone.

IV. DISCUSSION

This study focuses on the effect of source radiation angle on binaural hearing. Although the measurement was conducted using one turntable for source rotation and a binaural recording device with head rotation, this does not imply that it is a standardised measurement method for OD-HRTFs. Instead, OD-HRTFs can be obtained from separate directivity pattern measurements and conventional HRTF measurements. The results show several patterns, and the greater magnitude at 2 kHz on the contralateral side is of interest for discussion, as this may correspond to the influence of head size. However, future work is needed to further investigate this matter. The introduced phenomenon, pseudo azimuth, is also discussed as

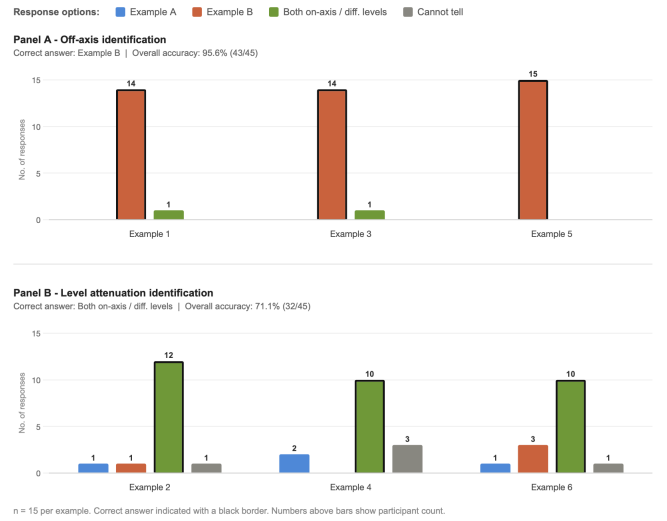


Fig. 6. Response distributions for off-axis and level attenuation identification. Correct answers are indicated by a black border. $n = 15$ per example.

a topic of interest for further studies to explore whether it may affect spatial hearing. Pseudo azimuth primarily describes the ILD introduced by source radiation angle that is similar to the ILD of a different azimuth angle with on-axis radiation. Frequency-band ILD analysis should also be explored to examine frequency-dependent patterns.

As for the limitation, the source rotation-induced ITD artefacts indicate that there may have been displacement of the loudspeaker during rotation. Source rotation should theoretically not introduce variations in ITD, as the position of the source remains fixed. However, the source centre may not have been perfectly aligned with the turntable centre, or the loudspeaker components may not have been perfectly aligned with the geometric centre of the loudspeaker. Hence, the observed ITD variation is treated as an artefact rather than evidence of the pseudo-azimuth pattern. These artefacts should therefore be considered in future measurement setups to avoid confounding effects.

The perceptual study shows higher accuracy for off-axis identification than for level attenuation. This suggests that source radiation angle can cause noticeable differences in auditory perception without reverberation in an anechoic setting. Future studies could therefore use OD-HRTFs instead of conventional HRTFs to further examine their effects on spatial auditory perception in broader applications that incorporate source directivity replication, such as 3D rendering for virtual reality environments.

V. ACKNOWLEDGEMENTS

The authors acknowledge and thank the 3D3A Laboratory at Princeton University for providing the facilities used to conduct the measurements for this study. The authors acknowledge and thank HEAD acoustics for providing the binaural recording equipment.

VI. DECLARATION OF INTERESTS

The authors declare no conflicts of interest.

REFERENCES

- [1] L. Turchet, M. Lagrange, C. Rottondi, G. Fazekas, N. Peters, J. Østergaard, F. Font, T. Bäckström, and C. Fischione, “The internet of sounds: Convergent trends, insights, and future directions,” *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11 264–11 292, 2023.
- [2] M. Schneider, “The development of dummy head microphones since 1970,” in *Audio Engineering Society Convention*, vol. 150, 2021.
- [3] E. Paulus, “Sound localization cues of binaural hearing,” *Laryngo-rhinotologie*, vol. 82, no. 4, pp. 240–248, 2003.
- [4] D. Rowan, T. Papadopoulos, D. Edwards, and R. Allen, “Use of binaural and monaural cues to identify the lateral position of a virtual object using echoes,” *Hearing Research*, vol. 323, pp. 32–39, 2015.
- [5] J. Blauert, *Spatial hearing: the psychophysics of human sound localization*. MIT press, 1997.
- [6] J. Jakka, “Binaural to multichannel audio upmix,” *Helsinki University of Technology*, vol. 29, 2005.
- [7] A. Roginska and P. Geluso, Eds., *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio*. New York: Taylor & Francis, 2017.
- [8] A. Dufour, O. Després, and V. Candas, “Enhanced sensitivity to echo cues in blind subjects,” *Experimental brain research*, vol. 165, no. 4, pp. 515–519, 2005.
- [9] J. Tsuaye, “Review: Head-related impulse response measurement methods,” in *Audio Engineering Society Convention 157*. Audio Engineering Society, 2024.
- [10] S. Li and J. Peissig, “Measurement of head-related transfer functions: A review,” *Applied Sciences*, vol. 10, no. 14, p. 5014, 2020.
- [11] J. Peng, T. Wang, and S. Wu, “Investigation on the effects of source directivity of chinese speech intelligibility in real and virtual rooms,” *Applied acoustics*, vol. 74, no. 8, pp. 1037–1043, 2013.
- [12] L. Liang and G. Yu, “The combined effect of source directivity and binaural listening on near-field binaural speech transmission index evaluation,” *Applied Acoustics*, vol. 203, p. 109198, 2023.
- [13] V. P. Sivonen and W. Ellermeier, “Directional loudness in an anechoic sound field, head-related transfer functions, and binaural summation,” *The Journal of the Acoustical Society of America*, vol. 119, no. 5, pp. 2965–2980, 2006.
- [14] S. Hoare, A. Southern, and D. Murphy, “Study of the effect of source directivity on the perception of sound in a virtual free-field,” in *Audio Engineering Society Convention 128*. Audio Engineering Society, 2010.
- [15] U. Sloma, F. Klein, S. Werner, and T. Pappachan Kannookadan, “Synthesis of binaural room impulse responses for different listening positions considering the source directivity,” in *Audio Engineering Society Convention 147*. Audio Engineering Society, 2019.
- [16] M. Zhang, W. Zhang, R. A. Kennedy, T. D. Abhayapala *et al.*, “Hrtf measurement on kemar manikin,” in *Proceedings of ACOUSTICS*, vol. 9, 2009.
- [17] V. Larcher, J.-M. Jot, and G. Vandernoot, “Equalization methods in binaural technology,” in *Audio Engineering Society Convention 105*. Audio Engineering Society, 1998.
- [18] H. Bahu, J.-M. Jot, T. Carpentier, M. Noisternig, M. Mihocic, P. Majdak, and O. Warusfel, “Towards improved consistency between databases of head-related transfer functions,” *Journal of the Audio Engineering Society*, vol. 74, no. 1/2, pp. 35–49, 2026.
- [19] HEAD acoustics GmbH, *HMS II.3: HEAD Measurement System, Basic Version with Right Ear Simulator, 3.3 Pinna & Artificial Mouth (Code 1703)*, HEAD acoustics GmbH, 2022, retrieved from https://cdn.head-acoustics.com/fileadmin/data/global/Datasheets/Artificial_Heads_Binaural_Recording/HMS-II.3-Artificial-Head-1703-Data-Sheet.pdf.
- [20] W. O. Brimijoin, A. W. Boyd, and M. A. Akeroyd, “The contribution of head movement to the externalization and internalization of sounds,” *PloS one*, vol. 8, no. 12, p. e83068, 2013.
- [21] H. Steffens, S. van de Par, and S. D. Ewert, “The role of early and late reflections on perception of source orientation,” *The Journal of the Acoustical Society of America*, vol. 149, no. 4, pp. 2255–2269, 2021.
- [22] B. B. Monson, A. J. Lotto, and S. Ternström, “Detection of high-frequency energy changes in sustained vowels produced by singers,” *The Journal of the Acoustical Society of America*, vol. 129, no. 4, pp. 2263–2268, 2011.
- [23] N. Inoue, T. Nishino, K. Itou, and K. Takeda, “Hrtf modeling using physical features,” in *Forum Acusticum 2005*, 2005, pp. 199–202.