forum acusticum 2026
8-12 September · Graz, Austria

SONIC SELVA: CALIBRATED AMBISONIC AND BINAURAL SOUNDSCAPE BASELINES OF VENEZUELAN BIOTOPES

Marc-Laurin Schimke

Department of Technical Acoustics, TU Berlin, Germany

Abstract

The preservation of highly biodiverse acoustic environments is of critical importance given the rapid global decline of natural ecosystems. This paper presents *Sonic Selva*, an open-access, spatial audio dataset [1] that captured the acoustic baselines of Venezuelan rainforests and savannas in 2024. A calibrated industrial binaural measurement system *(HEAD acoustics SQobold* with *BHS II)* and a 3rd-order Ambisonics microphone *(Zylia ZM-1)* were deployed. This paper details the methodological workflow across 62 measurement locations, incorporating 23 hours of 3rd-order Ambisonic and 14 hours of calibrated binaural recordings. The data is augmented by detailed measurement metadata and 3 hours of synchronized 360° video, providing a comprehensive contextual and visual ground truth for each measurement, as well as over 71 hours of continuous extended-duration Ambisonic recordings, capturing a full diel cycles as well as critical diurnal and nocturnal transitions.

This paper highlights the procedures used to conduct the measurements, technical challenges encountered when operating acoustic equipment in tropical extremes as well as the overall structure of the dataset. Furthermore, the application of hardware-specific noise profiles characterized in an anechoic chamber to mitigate self-noise and restore the psychoacoustic fidelity of low-level soundscapes is examined. The resulting dataset provides a robust, open-access resource for psychoacoustic evaluation, bio-acoustic monitoring, machine learning, virtual spatial soundscape reconstruction and investigations into the effects of Neotropical soundscapes on human emotion and health.

Keywords: *multimodal dataset, ecoacoustics, field-recording, spatial audio, digital heritage*

Corresponding author: *hey@marclaurin.space.*

1. Introduction

The expansion of human infrastructure and the subsequent fragmentation of natural habitats are fundamentally altering vulnerable biomes, resulting in unprecedented rates of species decline [2], [3]. This is particularly acute in Neotropical biomes, such as the Venezuelan Amazon, where intact forest loss has tripled over a decade due to logging, mining, agricultural expansion and weak environmental governance [4], [5]. Furthermore, fire-induced degradation has recently surpassed deforestation as the primary driver of Amazonian carbon emissions, impacting 3.3 million hectares in 2024 alone [6]. The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) warns that around 1 million animal and plant species are now threatened with extinction, many within decades, more than ever before in human history [7]. Consequently, endemic species are vanishing before their foundational ecology can be documented, a trajectory confirmed by the 2025 IUCN Red List assessments, which also warn that over 25% of all species assessed so far are now threatened with extinction, as all major taxonomic groups show a continuous downward trend [8].

Ecoacoustics has recognized natural soundscapes as critical indicators of ecosystem health and biophonic complexity. However, a deficiency persists within existing acoustic repositories. The vast majority of bioacoustic datasets rely on localized monophonic recordings designed for single-species identification. There remains a shortage of calibrated high-resolution, multimodal datasets that preserve the true spatial geometry of pristine acoustic environments while maintaining precise sound pressure level metrics required for scientific psychoacoustic validation and immersive acoustic reconstruction.

If we fail to archive these sites with modern spatial audio technology, future generations will be left with only monophonic or standard stereo often uncalibrated fragments, unable to replicate the immersive realism of the original biomes. The dataset also enables individuals from all backgrounds and geographical locations to study and experience authentic Neotropical soundscapes. By preserving these acoustic signatures, we ensure these environments remain accessible to the global community, regardless of geographic or temporal proximity to the tropics.

The dataset builds directly upon the historical foundations of acoustic ecology such as the *World Soundscape Project* initiated by R. Murray Schafer [9], [10]. It continues the digital heritage preservation paradigms established by David Monacchi [11]. The *Sonic Selva* dataset expands on these methodologies by adopting a strict open-access data schema coupled with an industrial calibration pipeline.

The primary objective of this paper is to introduce the architectural framework of the dataset and evaluate the methodology of spatial audio environmental baseline capturing. The text outlines the operational protocols used across 62 individual measurement locations, evaluates the physical performance boundaries of laboratory-grade acoustic apparatus when subjected to Neotropical stressors, and details the post-processing chain.

2. The *Sonic Selva* Dataset

2.1. Recording Locations

Venezuela was selected as the geographic framework for this dataset due to its exceptional biodiversity, preserving highly localized endemic ecosystems. The country's tropical rainforests and vast savannas encompass an array of distinct, structurally diverse biotopes required to capture a representative ecological and spatial audio baseline of Neotropical soundscapes. Measurement locations were chosen based on their specific habitat architectures and biological significance and altitudinal gradients. To capture pristine acoustic baselines, the deployment protocol prioritized environments with minimized anthropogenic acoustic footprints, allowing for an unobstructed documentation of complex biophonic structures and natural geophonic dynamics. The expedition spanned several distinct ecosystems, yielding recordings from the following primary regions:

Orinoco Delta
This region comprises a complex network of rivers, mangroves, and tropical wetland forests. The acoustic environment is predominantly water-driven, characterized by high amphibian density and avian communities adapted to aquatic ecosystems. Furthermore, the region provides an acoustic baseline for the anthropogenic integration of the indigenous Warao communities, documenting non-industrialized human settlements within a riparian environment.

Mount Roraima
Dominated by vast savannas and ancient flat-topped mountains known as tepuis, this region features highly isolated biotopes. These elevated plateaus act as ecological islands with distinct, endemic vegetation. Recordings conducted along the altitudinal gradient of Mount Roraima capture clear acoustic stratification, demonstrating how soundscape composition and biological activity shift with changes in elevation.

Yguraman Region
Located near the Brazilian border, the Yguraman region consists of small human settlements embedded in deep Amazonian rainforest. The soundscape here is defined by exceptionally high acoustic biodiversity. Crucially, the measurements from this region contrast the biophony of pristine forest patches with the anthropogenic noise of active deforestation and wildfires.

Kunabayna (El Abismo)
Located within Canaima National Park, Kunabayna represents a sharp ecotone where the open savanna abruptly meets the dense rainforest at a geographical fault line known locally as "El Abismo". This transitional zone supports overlapping species from both biomes, resulting in highly concentrated acoustic activity.

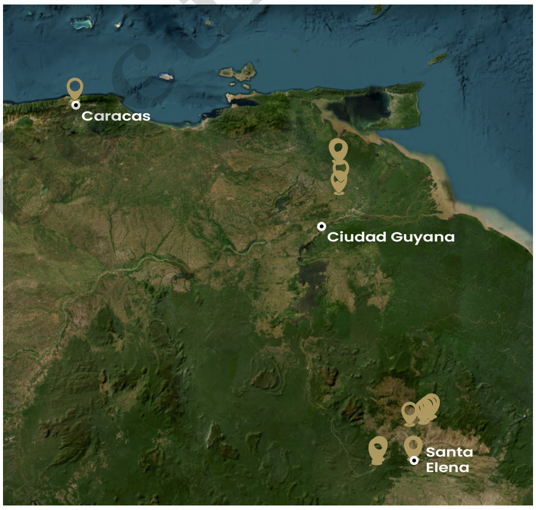

Figure 1: Each golden pin represents a measurement location.

2.2. Dataset Structure

The *Sonic Selva* dataset is systematically comprising 62 measurement points. The data set structure utilizes a hierarchical directory format, categorized first by broad geographical region, followed by specific measurement locations, and finally partitioned by data type of the measurement devices specified in Section 2.3.
To facilitate rigorous scientific cataloging and content-based retrieval, the dataset includes a comprehensive metadata file (`METADATA.csv`). This catalog documents vital contextual parameters for each measurement, including descriptive location data, environmental conditions, and specific acoustic characteristics.
An overview of the spatial audio and visual captures across the expedition regions, uniting the total number

Table 1: Overview of measurements and total recording durations across the expedition regions.

Region	Ambisonic Measurements	Binaural Measurements	360° Video Measurements	Total Measurement Location	Measurements with all devices simultaneously
Overall Project	52 (23:17:32)	47 (14:04:38)	28 (02:44:27)	62	24
Delta Region	15 (05:06:55)	14 (04:09:50)	10 (00:42:24)	15	6
Roraima Region	13 (04:14:43)	10 (01:24:43)	13 (00:32:21)	12	7
Yguraman Region	13 (03:58:39)	10 (03:36:36)	10 (00:40:21)	17	6
Kunabayna Region	15 (09:26:58)	12 (04:23:35)	10 (00:21:21)	17	4
Caracas Region	1 (00:30:16)	1 (00:29:59)	1 (00:04:30)	1	1

of measurements with their respective runtimes, is detailed in Table 1.

The dataset also includes comprehensive acoustic and psychoacoustic metadata for all binaural measurements, calculated in *ArtemiS SUITE*. Assessed parameters include percentile sound levels, extreme values and average levels. The levels are provided Z-, A- and C-weighted. Standardized psychoacoustic metrics such as loudness, sharpness, roughness, and tonality were also derived. Spectrograms of all measurements are included as well.

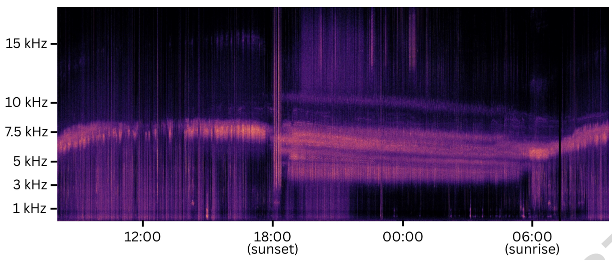

Figure 2: Spectrogram of a continuous 25-hour measurement in the Orinoco Delta. Distinct biophonic patterns and diel cycles are illustrated. These are chiefly characterized by a temperature-dependent pitch drift in insect stridulation, a pronounced attenuation of geophonic and avian spectral energy below 3.3 kHz during nocturnal periods (22:00–06:00) as well as a brief and dense broadband acoustic saturation during the dusk chorus driven by cicada activity.

2.3. Hardware Configuration

To capture the acoustic and visual ground truth of the Venezuelan soundscapes three primary devices were deployed, each serving a distinct role in the multimodal baseline documentation:

- **Binaural System:** The *HEAD acoustics SQobold* integrated mobile measurement system was utilized in combination with the *BHS II* binaural headset. This system is especially designed for aural accuracy and psychoacoustic measurements. The system allows for calibrated binaural recordings, essential for the psychoacoustic validation of the dataset [13].
- **Ambisonic System:** For 360° spatial audio capture, the *Zylia ZM-1E* was used. This spherical microphone array utilizes 19 capsules to capture full-sphere 3D soundscapes. The *ZM-1E*'s singular USB-cable interface in combination with a laptop running *REAPER* (DAW) for multichannel recording, enabled rapid battery-driven deployment in remote field conditions, while its 3rd-order Ambisonics output provides the necessary spatial resolution for immersive soundscape reconstruction [14], [15].
- **Visual Context:** An *Insta360 X3* was used to capture 5.7K monoscopic equirectangular 360° video. These visual recordings provided essential context for each measurement point, acting as a visual ground truth [16].

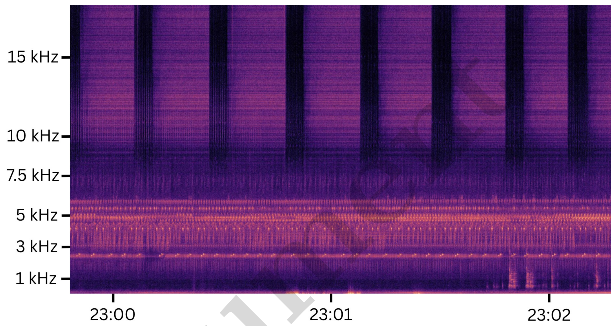

Figure 3: Spectrogram of a 2-minute excerpt from a nocturnal recording within a savanna biotope. This measurement provides a clear empirical illustration of the spectral stratification formulated by the Acoustic Niche Hypothesis [12], a structural feature observable throughout the broader dataset. Particularly within the 2–6 kHz frequency range, the dense, concurrent partitioning of distinct spectral bands by coexisting biophonic sources is strikingly apparent.

Calibration

The *SQobold* system served as the primary reference for sound pressure level calibration. Microphone sensitivities and equalization for each measurement are calibrated to a 0 dBFS reference equivalent to 103.03 dB(Z), enabling quantitative analysis of the soundscapes' absolute levels.

For the *Zylia ZM-1E*, digital pregain levels were documented for each measurement location. As detailed in Section 2.5, these metadata were utilized during the

32-bit floating-point conversion process to normalize the multichannel streams, thereby eliminating gain-related amplitude discrepancies across the dataset while preserving the integrity of the original acoustic dynamics and avoiding non-linearities.

2.4. Measurement Protocols

To ensure high-fidelity acoustic capture and consistency across the dataset, fixed measurement protocols were established during the expedition.

Figure 4: Field deployment at the Yguraman region (Measurement ID: YGURAMAN05). This image is an excerpt from the synchronized *Insta360* video captured during the measurement. The setup demonstrates the spatial relationship between the *HEAD acoustics SQobold* binaural system (worn by the operator) and the *Zylia ZM-1E* Ambisonic array, positioned to capture the acoustic baseline of the riparian biotope.

For all measurements, the *Zylia ZM-1E* was positioned at a height of 1.5 m. This elevation conforms to international acoustic measurement standards for environmental noise and soundscape assessment to represent typical human aural height [17], [18]. The binaural system was worn at natural human ear height. Whenever the topographical and biological conditions of the dense biotopes permitted, the equipment was placed at a minimum distance of 5m from major acoustically reflective surfaces to mitigate unwanted early reflections and boundary effects.

The binaural head-mounted microphones were worn by a member of the expedition crew. They would generally stand at least 5m away from the Ambisonic microphone. The 360° camera recorded only the first 2-3 minutes of a measurement because this duration usually captured the visual variety of each location. If possible the cameras were held by the person that is not recording with the head-mounted binaural microphones in order to avoid movement artifacts.

The baseline recording duration was standardized to a minimum of 10 continuous minutes per location to ensure a representative sample of the local soundscape. A total of 24 locations were recorded with all 3 devices meeting this condition.

During periods of peak biophonic activity (e.g. dawn and dusk choruses) recording durations were extended to 20 minutes or more to capture complex acoustic fluctuations and species interactions. In addition to the scheduled baseline measurements, shorter *ad hoc* recordings were conducted to document transient, acoustically significant events. Where logistical constraints allowed, repeated measurements were taken at identical locations across different days to capture meteorological and diurnal variations within the biotope.

2.5. Signal Processing

The post-processing pipeline was carefully designed to improve the psychoacoustic authenticity of the recordings while maintaining scientific integrity. The chain adhered to the following chronologically structured methodology:

Standardization to 32-bit Float

As a foundational step, all raw audio files were converted to a 32-bit float WAV format and their pregain levels were standardized to a uniform level. The 32-bit float architecture provides virtually unlimited dynamic headroom, preventing digital clipping during subsequent processing stages. Furthermore, this standardization unified the intrinsic noise floor across the dataset, significantly simplifying later noise reduction procedures.

Synchronization and BWF Metadata

Due to the lack of a centralized timecode generator among the independent recording devices, internal clocks drifted by several minutes during the expedition despite regular recalibrations. The temporal alignment was established using standardized `YYMMDD_HHMMSS` timestamps in the file names, as well as the embedding of Broadcast Wave Format (BWF) timestamps in the audio files.

Structural Editing

Initial macro-level edits involved trimming or splitting the continuous recordings to excise major, unrecoverable acoustic anomalies. This included the removal of localized disruptions such as logistical setup noise, equipment failures or human conversations.

Spectral Editing

Subsequent micro-level processing utilized spectral editing tools (iZotope RX) for the targeted removal of disturbances. In this context, "unwanted noises" are defined as extraneous acoustic events that detract from the natural biophonic and geophonic baseline. These primarily included recordist-induced foley sounds (e.g., footsteps, fabric rustling), and direct physical impacts on the microphone capsules, such as insect or leaf collisions. A conservative editing philosophy was applied to ensure that only genuinely distracting artifacts were removed, thereby avoiding potential over-processing of the natural soundscape.

Denoising Low-Level Soundscapes

Denoising was applied after structural and spectral edits. Capturing highly biodiverse but quiet biotopes presents a significant challenge, as the ambient environmental noise floor frequently drops below 30dB(A), exposing the intrinsic self-noise of the recording hardware. To transparently reduce this self-noise without compromising biophonic details or perceptive authenticity, custom noise profiles were utilized. Before the expedition, these profiles were specifically generated by recording the expedition hardware in an anechoic chamber and enabled transparent denoising algorithms to be applied without degrading delicate biophonic details.

Final Decoding (Ambisonics)
In the final processing stage, the cleaned multichannel audio streams were decoded into their respective final spatial formats using the *SPARTA*-Plugin Suite [19], yielding binaural and 3rd-order Ambisonics datasets ready for publication and analysis. To ensure compatibility with standardized spatial audio workflows, the Ambisonics files were encoded using the SN3D normalization scheme with ACN channel ordering, conforming to the AmbiX format specifications [20].

3. Challenges in Field Operation

Conducting high-fidelity spatial audio measurements in Neotropical environments presented operational and technical challenges that significantly influenced both the recording and the subsequent post-processing methodology, as well as the operational recommendations that inform the design of future expeditions.

Environmental Constraints
Extreme climatic conditions, characterized by ambient temperatures up to 38 °C and persistently high humidity, constrained the operational endurance of the expedition crew and necessitated rigorous equipment protection using dry bags and desiccant packs. The high density of insect populations posed a further physiological challenge during binaural captures: the operator, required to remain perfectly still while wearing the *BHS II* headset, inevitably introduced foley artifacts due to insect interference. Despite the geographical remoteness of the recording sites, the exceptionally low natural ambient noise floor allowed distant anthropogenic sounds — typically combustion engines — to propagate over several kilometers, occasionally contaminating the biophonic baseline.

Operational Limitations
The deployment of industrial acoustic measurement hardware in off-grid scenarios revealed several operational constraints. The boot and setup time of the recording systems proved inadequate for spontaneous ecological field recording, frequently precluding the capture of transient wildlife vocalizations; extended-duration recordings mitigated this limitation by increasing the probability of capturing such events. Robust synchronization across independent recording devices is essential for accurate multimodal alignment. Despite regular recalibration throughout the expedition, internal clock drift introduced deviations up to several minutes, necessitating manual acoustic resynchronization in post-production. Power supply via portable power banks proved well-suited to the expedition's logistical rhythm, given that recharging infrastructure was accessible every few days; for expeditions prioritizing extended-duration continuous recordings, however, higher-capacity power sources such as car batteries would be required. To minimize data loss in the event of system failure or power interruption, recording systems were configured to chunk audio measurements into 6-minute files.

Mechanical and Acoustic Vulnerabilities
Frequent deployment and teardown under field conditions, often in low-light or nocturnal environments, exposed the mechanical fragility of the hardware interfaces. The multipin LEMO connectors of the *BHS II* system proved particularly susceptible to misalignment and bending; field repairs were necessary to restore functionality. Additionally, despite the use of windshields, the microphone systems remained highly vulnerable to wind interference during savanna recordings. Even minor air movement introduced pronounced low-frequency contamination that substantially skewed the low-end spectrum. Aggressive spectral repair in post-production proved an effective but highly time-consuming remediation strategy, successfully restoring the affected measurements to a state suitable for psychoacoustic analysis.

4. Outlook

The field deployment of laboratory-grade acoustic apparatus in Neotropical extremes has yielded methodological lessons that directly inform the design of future expeditions. The foley artifacts introduced by the head-mounted binaural operator point toward a fully tripod-mounted, standalone array configuration as the preferred architecture going forward. Despite satisfying results by the *Zylia ZM-1*, selecting an Ambisonic microphone with a lower self-noise floor would further reduce the denoising burden on quiet biotopes. Ultimately, the uniqueness and ecological utility of the *Sonic Selva* dataset lies in its calibration and perceptual clarity. It provides a robust foundation for a broad range of applications: from bioacoustic monitoring and psychoacoustic research into the restorative effects of natural soundscapes including applications in urban soundscape design, to machine learning training corpora, professional media and VR production. Its calibrated, multimodal structure makes it interoperable across these domains in a way that monophonic repositories fundamentally cannot.
Most critically, the dataset constitutes a form of acoustic heritage preservation. The ecosystems documented here are subject to the accelerating pressures described in the Introduction. The dataset ensures that their biophonic complexity remains accessible, audible, and scientifically analyzable.

5. Accessibility

The *Sonic Selva* dataset is published as an open-access resource via TU Berlin [1]. The core repository includes the processed 3rd-order Ambisonics and binaural audio files, synchronized 360° video documentation, high-resolution spectrograms, psychoacoustic analysis data, the comprehensive metadata catalog, and in-depth technical documentation. Due to file size constraints, extended 24-hour continuous recordings have been archived separately and are available upon request through the Department of Audio Communication at TU Berlin. Communication details can be found in the technical documentation [1].

Acknowledgments

The realization of the project was made possible through the generous support of the *HEAD-Genuit Foundation* and *Gesellschaft von Freunden der TU Berlin e.V.* Special thanks are owed to the indigenous communities of the Warao and Yguraman, as well as local hosts in Kunabayna, for their hospitality and permission to conduct recordings on their lands. The author extends gratitude to Giulia Dàvila for her crucial contributions to co-organizing the expedition logistics, assisting in field measurements and in the spectral editing of the binaural recordings. Furthermore, the author would like to thank Dr. André Fiebig and Prof. Dr. Stefan Weinzierl from the Technische Universität Berlin for their academic and institutional support.

References

[1] M.-L. Schimke and G. Dávila, *Sonic selva: A calibrated 3rd-order ambisonic and binaural dataset with 360º video of the venezuelan rainforest and savanna*, DepositOnce, TU Berlin, 2026. DOI: 10.14279/depositonce-25755.

[2] G. Ceballos, P. R. Ehrlich, A. D. Barnosky, A. García, R. M. Pringle, and T. M. Palmer, "Accelerated modern human–induced species losses: Entering the sixth mass extinction," *Science Advances*, vol. 1, no. 5, e1400253, 2015. DOI: 10.1126/sciadv.1400253.

[3] A. Purvis et al., "Global assessment of biodiversity and ecosystem services," *IPBES Global Assessment*, 2019.

[4] P. Balvanera et al., "Status and trends – drivers of change," in *Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*, Bonn, Germany: IPBES secretariat, 2019.

[5] W. F. Laurance et al., "Ecosystem decay of amazonian forest fragments: A 22-year investigation," *Conservation Biology*, vol. 16, no. 3, pp. 605–618, 2002. DOI: 10.1046/j.1523-1739.2002.01025.x.

[6] C. Bourgoin et al., "Record-breaking 2024 amazon fires drive unprecedented carbon emissions and ecosystem degradation," *Biogeosciences*, 2025.

[7] IPBES, *Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*, E. S. Brondizio, J. Settele, S. Díaz, and H. T. Ngo, Eds. Bonn, Germany: IPBES secretariat, 2019. DOI: 10.5281/zenodo.3831673.

[8] IUCN, "The IUCN red list of threatened species. version 2025-1," International Union for Conservation of Nature (IUCN), 2025. [Online]. Available: https://www.iucnredlist.org.

[9] R. M. Schafer, *Five Village Soundscapes* (The World Soundscape Project Studies 4). Indian River, Ontario: Arcana Editions, 1977.

[10] R. M. Schafer, B. Davis, P. Harrison, B. Truax, and H. Westerkamp, *The Vancouver Soundscape* (The World Soundscape Project Studies 2). Burnaby, B.C.: Sonic Research Studio, Simon Fraser University, 1973.

[11] D. Monacchi, "The sonic heritage of ecosystems: Towards a formulation," in *SOIMA: Unlocking Sound and Image Heritage*, Part of the Fragments of Extinction project., 2015, pp. 83–89. DOI: 10.18146/soima2015.

[12] B. Krause, "The niche hypothesis: A virtual symphony of animal sounds, the origins of musical expression and the health of habitats," *The Soundscape Newsletter*, vol. 6, pp. 6–10, 1993.

[13] HEAD acoustics GmbH, *SQobold – Mobile Sound and Vibration Measurement*, https://www.head-acoustics.com/products/data-acquisition/sqobold, Accessed: 2026-06-04, 2026.

[14] Zylia Sp. z o.o., *Zylia ZM-1: Ambisonics Microphone Specifications*, Accessed: 2026-06-04, Zylia Sp. z o.o., Poznań, Poland, 2026. [Online]. Available: https://www.zylia.co/zylia-zm-1-microphone.html.

[15] Cockos Incorporated, *REAPER*, version 7.x, 2026. [Online]. Available: https://www.reaper.fm/.

[16] Insta360, *Insta360 X3*, https://www.insta360.com/product/insta360-x3, Accessed: 2026-06-04, 2026.

[17] International Organization for Standardization, *ISO 1996-2:2017: Acoustics – Description, measurement and assessment of environmental noise – Part 2: Determination of environmental noise levels*, Geneva, Switzerland, Jul. 2017.

[18] International Organization for Standardization, *ISO/TS 12913-2:2018: Acoustics – Soundscape – Part 2: Data collection and reporting requirements*, Geneva, Switzerland, Sep. 2018.

[19] L. McCormack and A. Politis, *Sparta plugin suite*, 2024. [Online]. Available: https://leomccormack.github.io/sparta-site/.

[20] C. Nachbar, F. Zotter, E. Deleflie, and A. Sontacchi, "Ambix - a suggested ambisonics format," in *Proceedings of the 3rd International Symposium on Ambisonics and Spherical Acoustics*, Lexington, KY, USA, 2011.

sonicselva.com