forum acusticum 2026
8-12 September · Graz, Austria

SOUND EXPOSURE IN CITIES: FROM ANNOYANCE TO ACOUSTIC AMENITY

Andrea Graziano[1], Camille Loyau[1], Rémy Masson[1], Luc Arnal[1,*], Boris Gourévitch[1,*]

[1] *Université Paris Cité, Institut Pasteur, AP-HP, INSERM, CNRS, Fondation Pour I'Audition, Institut de I'Audition, IHU reConnect, Paris, F-75012 France*

[*] *Authors contributed equally*

Abstract

The soundscape characterizes the acoustic environment as perceived by individuals within their context. To classify soundscapes, the categories of geophony, biophony, and anthropophony have been proposed for natural, animal, and human-produced sources, respectively. Urban environments present a unique soundscape to which inhabitants are continuously exposed, significantly impacting mental and physiological well-being. Although growing research addresses subjective judgments and general physiological responses to soundscape exposure, a detailed understanding of which specific acoustic features drive distinct physiological changes remains lacking. This gap is particularly pronounced in real-world settings. Our research addresses this methodological gap through a multimodal approach conducted in lab. By monitoring cardiac, neural, respiratory, and electrodermal activity with portable devices, we employed statistical methods to correlate sound events with physiological markers of sympathetic and parasympathetic autonomic activity. Specifically, our laboratory experiment supports the hypothesis that urban soundscapes elicit higher sympathetic (arousal/stress) and lower parasympathetic (relaxation) activity compared to natural soundscapes, while offering insights into the role of distinct acoustic elements. These findings pave the way for a deeper understanding of the interaction between acoustic and contextual elements in soundscape perception.

Keywords: *soundscape, physiology, urban acoustic environment*

Corresponding author: *boris@pi314.net.*

1. Introduction

Evolution has endowed animal species with the ability to analyze sound waves in order to interact with their environment, detect threats, and communicate. In humans, this capacity has been enriched by cultural and cognitive dimensions: all sounds of an environment—whether natural, human, or technological—play a fundamental role in emotion, memory, and well-being [1], [2], [3]. While natural sounds such as birdsong (biophony) or flowing water (geophony) are universally perceived as calming and associated with measurable health benefits [4], a growing proportion of the global population, now predominantly urban, is chronically exposed to high levels of environmental, occupational, and recreational noise. This cumulative exposure is recognized as part of the exposome, the totality of environmental influences on health across a lifetime. Urban sound exposure is unevenly distributed, reflecting social inequalities: residents near transport routes or dense urban centers are more frequently subjected to electromechanical noise (technophony) and human-generated noise (anthropophony) than people in peripheral areas. Historically, research has focused on the traumatic effects of noise on the auditory system, particularly hearing loss, which is strongly associated with isolation, depression, and elevated risk of cognitive decline and dementia [5]. Yet chronic exposure to sub-traumatic noise levels (70–85 dB SPL) also produces marked physiological and cognitive effects, including impaired communication, reduced attention and productivity [6], sleep disturbances [7], and cardiovascular dysfunction [8]. These effects are mediated not only through canonical auditory pathways but also via extra-auditory systems such as a potential reticulo-limbic network, including particularly the amygdala and hippocampus. This network regulates stress, salience, and emotional responses [9], [10] by influencing the endocrine system and the autonomic nervous system that accelerate or slow down body functions. Nature-related sounds facilitate relaxation and lower blood pressure [11], whereas acute noise increases blood pressure and heart rate [12], suggesting that short-term variations may accumulate into long-term cardiovascular risks [13].

Soundscape research has increasingly addressed such questions. Natural sound environments have been found to elicit more relaxed states as compared to urban ones, by influencing autonomic state markers of sympathetic and parasympathetic activity, including heart, respiration, and electrodermal activity [14]. Specifically, Li&Kang [15] used repeated-measures and canonical correlation analyses during laboratory exposure to four soundscapes, showing that natural sounds decreased heart rate, respiratory frequency, and depth while increasing heart rate variability (HRV) and EEG alpha/beta reactivity; in another study [16], they applied categorical principal components analysis to physiological and subjective ratings of 20 urban sound-source scenarios, identifying a primary Natural–Artificial physiological dimension aligned with perceived calmness/pleasure and a secondary Regular–Irregular dimension. However, most studies have compared broad soundscape categories, while the specific acoustic features that drive stress or relaxation—such as event temporal structure, amplitude dynamics, spectral content, abrupt acoustic changes, or exposure duration—remain poorly understood.

The present study addresses this gap by combining acoustic modelling, signal analysis, and ecological experimental conditions to examine how specific acoustic features and sound categories within urban soundscapes jointly influence human physiology and psychology. Such work has the potential to inform sustainable strategies that enhance well-being through the acoustic design of "resourcing" urban spaces, and to elucidate the brain–body mechanisms underlying emotional responses to sound.

2. Methods

2.1. Experimental design

Twenty-three participants were recruited for the study. Participants were between 18 and 50 years old, reported normal hearing, and were in good general health. Upon arrival, participants were welcomed and installed in the anechoic booth for sensor preparation and setup. All participants provided written informed consent prior to inclusion in the study. All communication and questionnaires were administered in French. Participants either reported French as their mother tongue or a high level of proficiency in French.

Physiological data were collected using a 32-channel BrainVision electroencephalogram (EEG) system with electrodes placed according to the standard

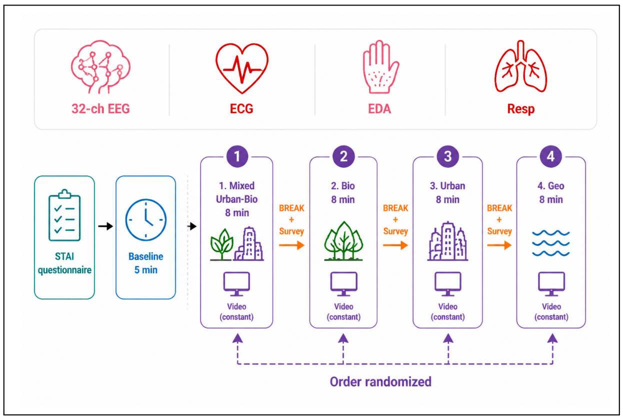


Figure 1: Schema of experimental procedure.

10–20 layout, and a Plux biosignal acquisition system recording electrocardiogram (ECG), respiration (RSP), electrodermal activity (EDA). Sensors were applied following standard procedures for each modality.

The auditory stimuli consisted of four 8-minute soundscape recordings drawn from the International Soundscape Database [17], [18], the Lion City Soundscape collection [19], and recordings from the Laboratoire des systèmes perceptifs, ENS [20]. Four main recording conditions were selected: predominantly urban sounds (Urban), predominantly biophonic sounds (Bio), mixed urban–biophonic sounds (Mix), and a condition dominated by river water flow (Geo). All recordings were originally in First Order Ambisonic format, converted to 2-channel stereo, and normalized to an integrated loudness of -23 LUFS using Audacity. Segments of 40–50 seconds were manually extracted from each recording, randomized in order, and concatenated with overlapping sine ramps to generate quasi-unique stimulus sequences for each participant. Stimuli were presented via Logitech Z207 loudspeakers at a fixed volume in the anechoic room, with loudspeakers positioned at $\pm 45°$ relative to the participant.
In addition to the sound, a video simulating a real-life walking experience was displayed on a screen. The video was identical across all sound conditions and showed a mixed landscape composed of natural elements (e.g., trees) and urban elements (e.g., buildings). Moving elements such as people and cars were kept to a minimum to reduce potential confounds.

Before the experimental blocks, participants completed the Khalfa Hyperacusis sensitivity questionnaire and the State Trait Anxiety Inventory (STAI) trait anxiety scale. Participants were then instructed to adopt a comfortable posture for listening, informed about the constant and background nature of the video, and about the presence of short questionnaires regarding the sounds they would listen to. The session began with a 5-minute baseline recording; subse-

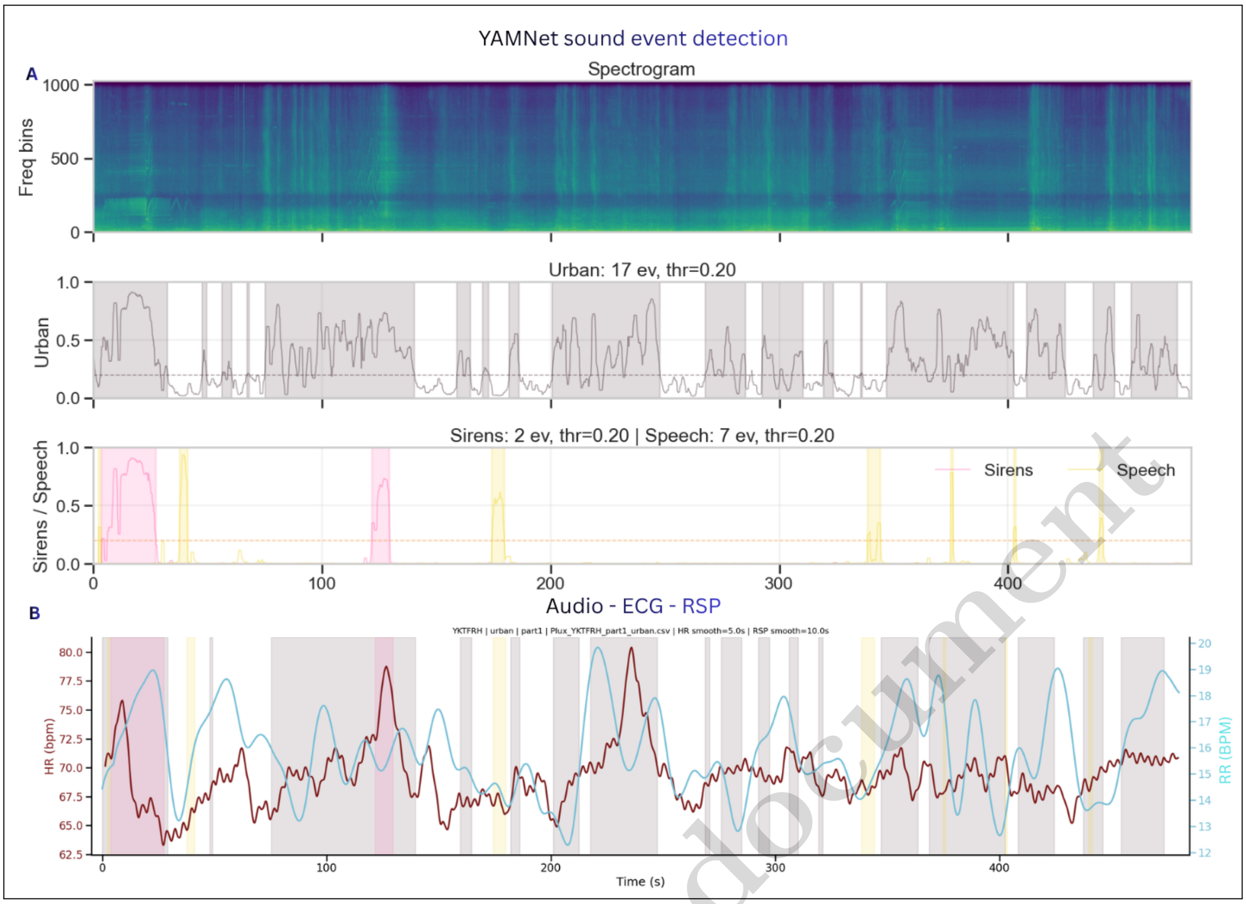


Figure 2: YamNet - ECG - RSP method example

quently, the four soundscape conditions were presented in random order. Between conditions, 5-minute breaks were inserted, during which participants completed the 6-item short-form STAI (STAI-6) and responded to ISO-validated short questionnaires [18] evaluating the acoustic characteristics of the recordings.

2.2. Data analysis

Plux physiology data for each participant was processed and analyzed using Neurokit2 library functions [21], then manually inspected. Specifically, wrongly detected ECG peaks were manually corrected. For RSP and EDA, labels were automatically created and manually checked around segments that showed anomalies either in amplitudes, peak detection, or signal-to-noise ratios. Baseline epochs were excluded and values were averaged per participant and group before being expressed as within-participant z-scores and compared across conditions and parts using paired-samples t-tests with Holm–Bonferroni correction. Linear mixed-effects models where then fitted separately for each metric; each model included fixed effects of condition and experimental segment order (or part) and a random intercept for participant, and omnibus effects were evaluated using likelihood-ratio tests.

Event-locked continous analysis (Fig2). Auditory stimuli for each participant were automatically analyzed using YAMNet [22](Fig2(A)), a pretrained sound event detection model. Five sound categories were defined (urban sounds, sirens, speech, biophony, water) by grouping the sound classes on which YAMNet was trained. The model provided time-resolved probability estimates for the presence of each sound class, therefore each category, which were then thresholded to extract sound-event time points. Thresholds were adjusted depending on the category and manually validated to optimize detection accuracy. Close events (up to 5 sec) were grouped together.

As presented in Fig2(B), for each participant, the mean of the z-scored istantaneous heart rate and respiration rate were computed in a pre-event time window of 3 and 5 s before each acoustic event, respectively, and compared with the mean heart rate in event and post-event window lasting until 10 and 20 s after each event, respectively. A similar analysis was conducted using the mean heart rate across the specific recording as reference. Mean changes relative to the pre-event window and to the recording mean were computed, and the resulting participant-level mean changes were tested against zero using one-sample t-tests and Wilcoxon signed-rank tests. Root Mean Square of Successive Differences (RMSSD) is a primary heart rate variability (HRV) metric tracking short-term

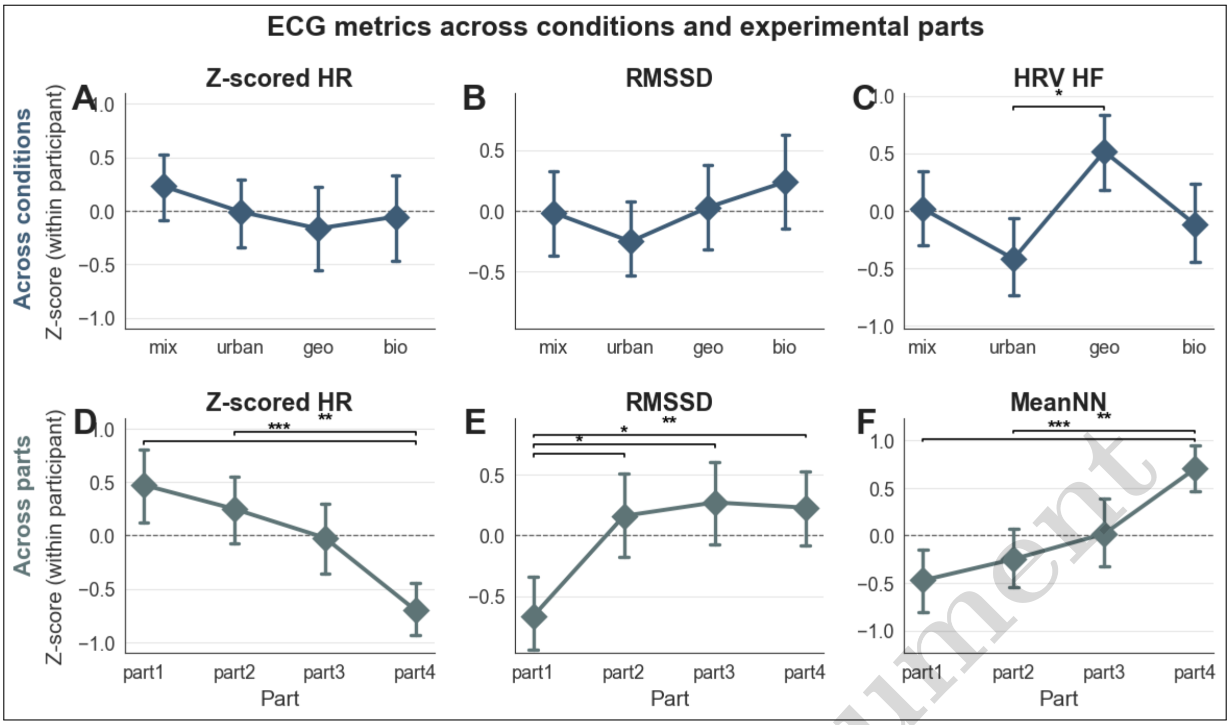


Figure 3: ECG - discrete statistics

beat-to-beat changes to measure parasympathetic activity; RMSSD event responses were aggregated and modelled with ordinary least-squares regressions including condition, sound category with the no-event baseline as the reference level, experimental part, and participant identity as fixed effects.

EEG data were automatically pre-processed using EEG Pasteurization Pipeline [23] for filtering and artifact removal and further inspected using MNE library. Preprocessed EEG epochs of 20 s were extracted for each experimental condition, and power spectral densities were estimated with Welch's method over 1–40 Hz. Each spectrum was decomposed into aperiodic and periodic components using the FOOOF algorithm [24], yielding an aperiodic offset and exponent plus band-limited periodic power (delta, theta, alpha, beta, low-gamma). Analyses were run on two sensor configurations: the average across all EEG channels and the Cz channel re-referenced to the average. Baseline epochs were excluded from condition comparisons. Epoch-level fits were aggregated per subject and retained only if the mean $R^2 \geq 0.9$ and mean model error ≤ 0.15. Omnibus effects of condition and part were assessed with Friedman tests on the subject-level means, followed by Wilcoxon signed-rank post-hoc comparisons with Holm–Bonferroni correction.

3. Results

Discrete statistics for the ECG-derived measures are shown in Figure 3. Across soundscape conditions, heart rate followed the expected pattern, showing slightly lower average levels for natural compared

Table 1: Likelihood-ratio tests (χ^2, df = 3) for condition and part effects on ECG metrics.

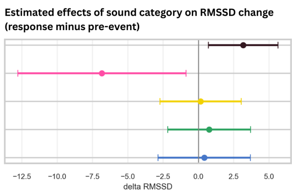
	Condition		Part	
Metric	χ^2	p	χ^2	p
HF power	16.10	**0.003**	1.27	0.867
Heart rate (HR)	0.34	0.987	17.58	**0.001**
LF power	9.52	**0.049**	1.91	0.753
LF/HF ratio	1.06	0.900	0.35	0.986
Mean NN interval	0.34	0.987	16.40	**0.003**
RMSSD	0.86	0.930	9.48	0.050

to urban sound conditions, but did not reach statistical significance (A). Parasympathetic markers such as RMSSD and high-frequency heart rate variability (HF-HRV) power, were generally higher under geophony (Geo) and biophony (Bio) than under mixed (Mix) and urban-prevalent (Urban) conditions, although in the discrete comparisons the only significant pairwise difference was between Urban and Geo (B, C) for HF-HRV. With respect to experimental part, average heart rate decreased across parts (D), while RMSSD and the mean NN interval were significantly lower in the first part, suggesting that participants could be less relaxed or still habituating during the initial video exposure (E, F). These observations were confirmed with linear mixed-effects models, as reported in Table 1.
Comparable patterns were observed for respiratory and electrodermal activity. Respiration rate showed lower trends during Geo and Bio, but condition differences didn't reach significance. For EDA, significant condition effects emerged for phasic mean

Table 2: Event-locked analysis. Participant-level changes in HR and RR across sound categories

	HR		RSP	
Sound cat.	**Δ**	***pw***	**Δ**	***pw***
Urban	0.13	0.18	0.20	**0.03**
Sirens	-0.58	0.39	-0.28	0.16
Speech	-0.51	**0.01**	0.09	0.54
Biophony	-0.07	0.57	0.08	0.09
Water	0.16	0.75	-0.03	0.80

(p=0.01) and mean SCR amplitude (p=0.03), whereas part effects for tonic mean (p=0.01).



Sound cat.	**Δ**	***pw***
Urban	**3.15**	**0.00**
Sirens	**-8.47**	**0.01**
Speech	0.06	0.64
Biophony	0.78	0.10
Water	1.07	0.16

Figure 4: Estimated effects of sound category on RMSSD changes

Continuous time locked analysis to HR and respiration rate (RR) reveald a decrease in istantaneous HR following Speech events relative to the recording mean, suggesting an orienting or social-attention response to human vocalisations (Table 2). RR, conversely, showed a modest increase during and following Urban events.
Figure 4 shows the category effects on RMSSD change. The mixed model revealed a marked decrease in RMSSD following Sirens events relative to pre-event window, consistent with sympathetic activation or parasympathetic withdrawal. Unexpectedly, Urban events were associated with a significant increase in RMSSD.

EEG spectral decomposition by condition and part is shown in Figure 5. Descriptive inspection of the bandpower spectra (A) showed distinct spectral profiles across conditions. Geo conditions showed the strongest delta-band power, usually associated with relaxation, and Urban conditions exhibited relatively higher alpha, beta, and low-gamma powers, possibly reflecting sustained attention or vigilance. (B) FOOOF analysis showed the all-channel aperiodic exponent differed significantly across conditions, $\chi^2_F(3) = 8.30$, $p = .040$, suggesting a change in the broadband $1/f$ slope between soundscapes. Specifically, Urban condition was associated with a flatter aperiodic slope, consistent with greater cortical activation or arousal, whereas Mix showed the steepest slope, suggesting a more low-frequency-dominant, possibly relaxed cortical state. Because the aperiodic exponent captures the broadband slope of the power spectrum independent of narrowband oscillations, this effect indicates that soundscapes induce global shifts in the cortical background signal beyond changes in specific frequency bands. For periodic activity, all-channel low-gamma power also varied by condition, $\chi^2_F(3) = 8.34$, $p = 0.039$. At Cz, periodic low-gamma power marginally varying across parts, $\chi^2_F(3) = 7.69$, $p = 0.053$. No other aperiodic or periodic band metrics reached significance after correction.

4. Discussion

The present study examined how natural and urban soundscapes modulate autonomic and neural activity using multi-modal wearable biosensors in a controlled laboratory setting during naturalistic video viewing. Results generally support the hypothesis that urban soundscapes are associated with heightened arousal and reduced parasympathetic tone relative to natural soundscapes.

ECG-derived markers showed a strong effect of experimental part order on heart rate (HR), root mean square of successive differences (RMSSD), and mean NN interval, likely reflecting fatigue or habituation to the constant background video across successive parts. After controlling for part effect, average high-frequency heart rate variability (HF-HRV) power—a marker of parasympathetic activity—remained significantly higher during Geophony as compared to Urban sound conditions.
Event-locked continuous analyses support these trends. Relative to recording-specific means, Speech events elicited heart-rate deceleration consistent with a social re-orienting response, which may be particularly relevant under stressful noise exposure and/or due to the isolated experimental conditions, while Urban were associated with an increase in respiration rate. Siren events strongly reduced RMSSD relative to the pre-event window, indicating enhanced arousal; because these events were the longest in duration, they are particularly suited to such analyses, as autonomic effects usually take a few seconds to manifest. Surprisingly, Urban events showed a small RMSSD increase. We hypothesize that this may reflect the broadness of the Urban events category, which includes events of very different nature and duration and may therefore be ill-suited for a temporally local analysis of RMSSD.
FOOOF-based EEG spectral decomposition further showed that soundscape condition modulates the aperiodic exponent across all channels, with Urban associated with greater cortical activation/arousal compared to the other conditions, suggesting that the broadband $1/f$ slope is sensitive to acoustic context. Periodic low-gamma power also varied with condition, and Cz low-gamma power displayed condition- and part-dependent patterns that warrant further investigation.

Several limitations should be considered. The visual environment was constant across conditions and participants, whereas the order of acoustic events

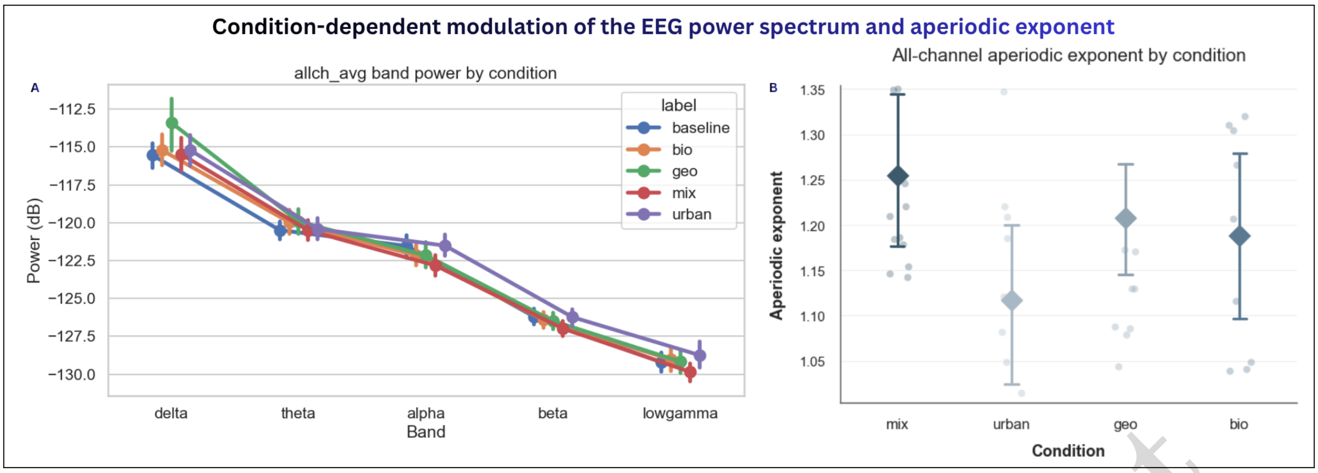


Figure 5: Condition-dependent modulation of the EEG power spectrum and aperiodic exponent

within each recording varied across participants; this may have generated audiovisual congruence or incongruence effects that are difficult to disentangle. Event categories also varied considerably in duration and specificity: Speech, Sirens, and Water grouped events of similar duration and spectral character, whereas Urban and Biophony were much broader categories and therefore probably less suited to fine-grained local analyses. Furthermore, some analyses—particularly for EDA and RSP—were limited by low signal-to-noise ratios, which could not allow to inform statistically.

In sum, soundscapes exert measurable effects on autonomic and neural markers, with both categorical soundscape typologies and specific acoustic events contributing to physiological responses. The controlled laboratory setting trades ecological validity for experimental control. Future work should combine larger samples, fine-grained acoustic feature extraction, and multimodal modelling to clarify these interactions both in the laboratory and in situ.

References

[1] R. M. Schafer, *The Soundscape*. 1993.

[2] P. Woloszyn, "L'idéo-scénarisation sonore du paysage," *Projets Paysage*, 2022.

[3] ISO, *Acoustics—soundscape—part 1: Definition and conceptual framework*, 2014.

[4] E. O. Wilson, "Biophilia and the conservation ethic," in *Evolutionary Perspectives on Environmental Problems*. Routledge, 2007.

[5] C. P. Wild, "Complementing the genome with an "exposome"," *Cancer Epidemiol. Biomarkers Prev.*, 2005.

[6] A. Mehrotra et al., "Auditory and non-auditory effects of noise on health," *Noise Health*, 2024.

[7] European Environment Agency, *Environmental Noise in Europe, 2020*. 2020.

[8] H. Kerniss et al., "Urban noise and early-onset myocardial infarction risk," *Eur. Heart J.*, 2024.

[9] L. H. Arnal et al., "The rough sound of salience enhances aversion," *Nat. Commun.*, 2019.

[10] L. H. Arnal and N. Gonçalves, "Rough is salient: A conserved vocal niche," 2025.

[11] T. Lanki et al., "Acute effects of urban green environments on cardiovascular physiology," *Environ. Res.*, 2017.

[12] S. L. Lusk et al., "Acute effects of noise on blood pressure and heart rate," 2004.

[13] C. Cai et al., "Traffic noise and heart rate variability," *Front. Psychol.*, 2022.

[14] M. Erfanian et al., "Psychophysiological implications of soundscape: A review," 2019.

[15] Z. Li and J. Kang, "Changes in physiological indicators in soundscapes," 2019.

[16] Z. Li et al., "Measuring soundscape quality with physiological indicators," *Build. Environ.*, 2024.

[17] A. Mitchell et al., *The international soundscape database*, Zenodo, 2024.

[18] A. Mitchell et al., "The ssid protocol for urban soundscape surveys," *Appl. Sci.*, 2020.

[19] K. Ooi et al., "Lion city soundscapes: A perceptually diverse dataset," *JASA*, 2024.

[20] M. Fraticelli, R. Wendling, and C. Lorenzi, UMR CNRS LSP, ENS 8248, LSP, EnRiv project, funded by MITI CNRS, 2024.

[21] D. Makowski et al., "Neurokit2: A python toolbox for neurophysiological signal processing," *Behav. Res. Methods*, 2021.

[22] TensorFlow, *Sound classification with yamnet*, 2024.

[23] R. Masson, *Eeg pasteurization pipeline*, 2026.

[24] T. Donoghue et al., "Parameterizing neural power spectra into periodic and aperiodic components," *Nat. Neurosci.*, 2020.