forum acusticum 2026
8-12 September · Graz, Austria

# A rationale for elaborating tests assessing environmental-sound perception

Christian Lorenzi[1], Karine Malysheva[2], Serhii Lytvyn[2], Volodymyr Volynets[2], Frédéric Apoux[1]

[1]*Laboratoire des systèmes perceptifs, UMR CNRS 8248, Ecole normale supérieure, PSL Univ., Paris, France*
[2]*Dept of experimental and applied psychology, Faculty of Psychology, Taras Shevchenko National Univ. of Kyiv, Kyiv, Ukraine*

This paper is a revised version of the authors' submitted manuscript that was peer-reviewed by at least two anonymous reviewers.



## Abstract

Shafiro and Gygi (2004) – who pioneered psychoacoustic research on environmental sounds – promoted a rigorous methodology for selecting appropriate stimuli to study auditory perception of this category of sounds by human listeners. Here we propose a rationale to complement this methodology, based on an ecological approach grounded in recent findings concerning the perception of natural sounds. This rationale emphasizes the importance of choosing acoustic stimuli and interpreting test outcomes based on: (i) the sounds' spectro-temporal modulation characteristics and causal properties; (ii) the role of low- vs. high-level auditory mechanisms in the perception of environmental sounds, with a fundamental distinction between biological and non-biological sounds, animal vocalisations and action sounds, and texture and non-texture sounds; (iii) context (i.e., domestic indoor, urban outdoor, and natural outdoor); (iv) the type of monitoring task to be performed by participants and their hearing-related functional character, emotional valence and importance in daily life. This rationale should help in designing test batteries that are better suited to assessing the perception of environmental sounds in healthy listeners and patients, as well as in designing rehabilitation devices and algorithms.

**Keywords**: *environmental sounds, auditory perception, test battery, ecology.*

## 1. Introduction

"Environmental sounds" – sometimes called "everyday sounds" – correspond to a poorly defined category of sounds that have only received attention from researchers and clinicians for the past twenty years or so (e.g., Marcell et al., 2000; Gygi et al., 2004). Environmental sounds are often considered to be sounds heard in everyday life that have no communicative or entertaining function, unlike speech sounds and sounds produced by musical instruments. In other words, "environmental sounds" corresponds to a catch-all category for non-linguistic and non-musical sounds. For that reason, studies attempting to evaluate the auditory capacity of [1]human listeners to perceive environmental sounds have elaborated test batteries based on disparate collections of sounds that often do not have an explicit structure. These test batteries can be divided into four broad categories, based on their objectives: i) characterizing the cues, processes and brain structures involved in the recognition of non-speech/non-musical sounds; ii) developing algorithms able to recognize automatically non-speech and non-musical sounds; iii) evaluating recognition capacities for patients with peripheral hearing impairment and effects of rehabilitation devices such as hearing aids and cochlear implants; iv) evaluating recognition capacities for patients with neurological deficits. A major methodological challenge is therefore selecting appropriate stimuli to test the ability to perceive environmental sounds. Numerous sound databases exist, and a rigorous methodology has been proposed for selecting appropriately these sounds (Shafiro & Gygi, 2004). Despite this, a large number of studies generally rely on disparate sets of sound samples selected without strong justification or theoretical basis. *As a consequence, it is unclear what these tests actually measure in human listeners and machines.* The purpose of this article is to offer some thoughts based in recent findings, aimed at supplementing the initial proposal of

***Corresponding author***: *christian.lorenzi@ens.psl.eu*

Shafiro and Gygi (2004) and providing a rationale for the development of future tests designed to explore further auditory perception of the close environment and sound-monitoring capacities.

## 2. Generation processes and features

Environmental sounds are composed of an extremely wide range of acoustic signals, ranging from tonal/harmonic sounds to inharmonic sounds and noises, narrowband to broadband sounds, and from impulsive/transient acoustic events to temporally fluctuating and steady sounds. The first characteristics of these sounds is therefore their *heterogeneity*, that results from potentially infinite acoustical disparities in terms of level, power spectrum and time, due to the process by which they are generated: these sounds are produced by biological, geophysical and technological sources and their interactions. Examples correspond to i) non-human animal vocalisations (bird, insect, amphibian, and mammal vocalisations); ii) wind, rain, thunder, streams; iii) machines. Environmental sounds are produced by the source alone (e.g., a birdsong) or result from actions and interactions between sources (e.g., impact sounds). In a pioneering study, Gaver (1993) proposed a powerful rationale to describe the generation processes of these sounds, allowing scientists to explore the perceptual processes involved in the recognition of these sounds based on the material and geometry of the sound source as well as the characteristics of the action and the agent. This framework inspired an entire research program in ecological acoustics aiming to unveil the **causal aspects** of perception, the inferential process involved in the recovery of the properties of the sound source and actions used to generate them. Importantly, this research program showed that humans are faster and more accurate at identifying the action than the material used to generate environmental sounds (Lemaitre & Heller, 2012). Some environmental sounds may be intentional and have communication meaning (e.g., birdsongs) or may be incidental (e.g., animal action sounds such as motion sounds). Communication sounds produced by a living organism equipped with a vocal production system as in the case of animal vocalisations typically convey specific and salient **spectro-temporal modulations** (e.g., Miller-Viacava et al., 2026a) to which the human auditory system is quite sensitive as they share similarities with speech and engage spectral and temporal-modulation analysis processes implemented at the cochlear and brainstem/cortical levels. However, the meaning is not intended for humans and may not be understood by naïve listeners. Other sounds with biotic or abiotic origin have **texture properties** (e.g., wind, rain, streams, insects chirping) that distinguish them from other sounds, because they result from the superposition of a large number of independent, relatively short acoustic events (e.g., bubbles). These sounds that are quasi stationary contribute to natural auditory scenes and may be perceived as well-defined, well-segregated auditory objects as such or as background (ambient) sounds. Importantly, auditory discrimination and recognition of sound textures result from the operation of a flexible, multi-second central temporal-integration process estimating summary statistics of temporal-envelope distributions within and across cochlear and modulation channels (McDermott & Simoncelli, 2011). *Taken together, these findings suggest that environmental sounds recruit different peripheral and central auditory mechanisms. Therefore, responses to these sounds by humans and algorithms should be distinguished based on their causal properties, their spectro-temporal modulation content and their texture vs. non-texture properties.*

## 3. Context and semantic information

A crucial aspect distinguishing environmental sounds is the **context** in which they have been generated. Settings include domestic, indoor (e.g., a room), urban outdoor (e.g., a street), and natural, outdoor (e.g., a forest, a park) environments. As a consequence, environmental sounds are embedded in and contribute to the organization of a "soundscape" (cf. Grinfeder et al., 2022; Lorenzi et al., 2023). As to natural settings, the context reflects ecological processes at work in the habitat such as climatic/meteorological conditions, diel and seasonal variations, water cycle, and animal behavior. Environmental sounds are therefore spatially and temporally organized acoustical patterns constrained by specific processes and shaped by the sound-propagation characteristics of the local environment where they have been generated, such as refraction, diffraction and reverberation effects. Propagation effects reflect the topography and obstacles (walls, vegetation) present in the environment such as propagation distance which often differ for domestic and outdoor settings (a few meters vs. tens/hundreds of meters). Not surprisingly, urban and natural soundscapes therefore differ in terms of their spectro-temporal modulation content, as shown by De Coensel and Botteldooren (2006).Thus, environmental-sound perception does not only involve recognition processes as discussed above but it also involves primitive auditory-scene analysis (bottom-up sequential and simultaneous grouping mechanisms) that is strongly constrained by the context and top-down, semantic processing (schema-based auditory scene analysis). This implies that context may influence environmental sound perception. For example, Gygi and Shafiro (2011) showed increased identification accuracy for sounds contextually incongruous with the background scene. Finally, complex auditory scenes can be described in terms of **global attributes** (moment of day, season, habitat, sparseness, transience, openness or navigability, etc.) that are also part of the general process of perceiving urban and natural environments and understanding the situation (Apoux et al., 2025; McMullin et al., 2024). *All in all, these findings suggest that we should expand the range of environmental sounds used for testing to domestic, urban, and natural scenes, and evaluate the ability of humans and algorithms to distinguish between these complex scenes,* *estimate their global properties and understanding of the situation. In addition, the ability to identify individual environmental sounds should be assessed for sounds in congruent or incongruent contexts.*

## 4. Early vs. late auditory processing

### 4.1 Anything "special"?

Speech and music are generally assumed to be "special" for the human auditory system. Is there any evidence that processing environmental sounds also involves specialized mechanisms and cerebral networks? Environmental sounds may be grouped or distinguished based on their **spectro-temporal acoustic features** such as overall power, harmonicity, spectral bandwidth and profile, temporal-modulation power and prominent modulation rates, correlation of temporal-envelope patterns across frequency and modulation bands, and attack and decay times (Gygi et al., 2004; McDermott & Simoncelli, 2011; Apoux et al., 2025; McMullin et al., 2024; Miller-Viacava et al., 2026b). Textures and mechanical sounds are quite different from isolated speech sounds. However, a multi-talker babble (a hubbub) is a sound texture that has "speech-like" properties. Moreover, sound textures found in musical work and animal vocalizations that may be encountered in domestic, urban and natural settings share common features with speech sounds and sounds produced by musical instruments (Thoret et al., 2021). These elements limit the possibility of finding a special character for all environmental sounds. It may be argued that environmental sounds are special because they generate specific emotional responses. In particular, natural soundscapes and especially, birds and water sounds have been shown to trigger "restorative effects" that are associated with positive emotions, reduced stress and restored attention (Buxton et al., 2021). However, musical work also triggers such emotional responses. Still, amongst environmental sounds, **biological** (living/animate) ones may have a special status for the human auditory system, as suggested by brain-imaging and behavioral studies showing segregated brain regions responding to animal vocalizations and animal action sounds (cf. Webster et al., 2017) and faster and mandatory responses to vocalizations compared to other sounds (e.g., Suied & Viaud-Delmon, 2009). A specific feature of environmental sounds is that they are continuously **monitored** by the auditory system and contribute to the perceptual representation of the immediate environment. In other words, they are not used for communication, but rather help us build a sense of "where we are" and "when we are" (Fay, 2009). Furthermore, environmental sounds serve a *variety of crucial functions*, including evaluation of resources and dangers, opportunities for action, navigation and orientation. These functions may be considered as making environmental sounds special to our auditory system. *Overall, these findings call for analyzing responses to environmental sounds by distinguishing between sounds of biological origin, involving potentially distinct brain structures for vocalizations and action sounds and other types of sounds such as textures of geophysical origin and mechanical sounds.*

### 4.2 Deficits in environmental-sound perception

Peripheral (cochlear) and central (cortical and sub-cortical) sensory mechanisms constrain the reception of some acoustical features described above and shape their low/mid-level perceptual representation. This implies that any peripheral or central deficit altering absolute thresholds (audibility) and suprathreshold auditory capacities in the amplitude, frequency and temporal domains should decrease the ability to analyze domestic, urban and natural soundscapes, assess their global attributes and identify their constituting sound sources. Consistent with this, recent audiological studies showed that auditory discrimination of natural soundscapes and recognition of environmental sounds are degraded for patients with sensorineural hearing loss and patients wearing hearing aids and cochlear implants (cf. Lelic et al., 2025). Neuropsychological studies showed that environmental-sound recognition may be degraded for patients with acquired brain-damaged (cf. Dick et al., 2016). Neuropsychological and brain-imaging (cf. Dick et al., 2016; Webster et al., 2017) evidence for domain-specific processing of environmental sounds in the auditory brain is scarce and questionable. The lack of evidence for specialized processing of environmental sounds may reflect the absence of dedicated neural networks. Alternatively, this may reflect the use of heterogeneous collections of stimuli to study cerebral responses, mixing biological and non-biological sounds arising from very different generation processes, but also natural and urban/domestic sounds with different "acoustic history" (urban and domestic sounds being fairly recent with respect to the history of homo sapiens). This may also result from the use of tasks that do not tax precisely monitoring capacities. *In summary, clinical studies on peripheral and central disorders highlight deficits in the perception of environmental sounds that may impact patients' daily lives from a functional and emotional standpoint and create risks and dangers. This requires the development of clinical tests that are easy to administer and designed in such a way that they can be used to identify areas that may be spared or show selective deficits.*

## 5. A comprehensive test battery

### 5.1 General framework

Figure 1 schematizes how environmental sounds used in a test battery could be organized according to i) context, ii) living (biotic) vs. non-living (abiotic), and ii) texture vs. non-texture properties. Fig. 1 also indicates the variety of tasks that could be performed (discrimination, recognition, magnitude scales re: global scene attributes or source attributes). Responses to each sound (accuracy, response times, rating on various magnitude scales) that would be presented to each participant either in free field or under headphones could be sorted according to the similarity of modulation features for each category of sound or causal properties. It is noteworthy that the spectro-temporal structure and human response to some of these sounds such as watercourses, rain, wind or cars and aircraft changes drastically along some dimensions such as discharge for watercourses, and strength for rain, wind, cars and aircrafts (Lorenzi et al., 2023). *Various levels* (stream discharge; wind and rain strength, etc.) should therefore be taken into account when creating this test battery.

| | Texture | Domestic (indoor) | Urban (outdoor) | Natural (outdoor) |
|---|---|---|---|---|
| Abiotic | Yes | Running water<br>Cooking/baking<br>Fire crackling | Urban water<br>Wind<br>Rain | Watercourses<br>Wind in foliage<br>Rain |
| Abiotic | No | Doorbell<br>Phone ringtone<br>Alarm clock<br>Doors, windows<br>Objects falling<br>Footsteps<br>Cutlery<br>Appliances | Emergency sirens<br>Cars<br>Aircraft<br>Metro trains<br>Pedestrian crossing<br>Construction | Rockfall |
| Biotic | Yes | Hubbub | Hubbub | Bird chorus<br>Insect chorus<br>Amphibian chorus |
| Biotic | No | Birds<br>Dogs<br>Cats | Birds<br>Dogs<br>Cats | Birds<br>Insects<br>Amphibians<br>Mammals |

**Figure 1**: Diagram illustrating the general organization of sounds used for a test battery aiming to assess auditory perception of environmental sounds.

### 5.2 Pilot study

Crucially, such test batteries should be grounded in an *ecological* approach emphasizing the *tasks* to be performed by participants and their *auditory needs* (Gatehouse, 1999). For that reason, the **functional character** (the susceptibility to prompt actions), **emotional valence** (degree of pleasantness) and **importance** to hear the sounds to be included in these test batteries are to be assessed carefully. A pilot study was performed using the collection of 38 sounds listed in Fig. 1 (15 domestic indoor sounds; 12 urban outdoor sounds; 11 natural outdoor sounds). These sounds were selected by a group of seven French and Danish hearing scientists. A first survey was then conducted on 214 healthy Ukrainian participants (172 females; 0 non-binary) aged between 18 and 40 years (mean: 21 years; SD: 5 years) to assess the functional character and emotional valence of each sound. Each participant was informed about the context (setting) before rating a given sound. The questions asked to participants are detailed in the legend of Fig. 2. Results are shown in Figs.2A and 2B. Overall, an analysis of variance (ANOVA) on the scores with Sound and Subject as factor, and a repeated-measures ANOVA with Environment (three levels) as within-subject factor showed that emotional valence and functional character vary significantly across types of sounds ($F(37, 7881)=65.01$, $p<0.001$; $F(37, 7881)=194.25$, $p<0.001$, respectively) and settings ($F(2, 426)=43.59$, $p<0.001$; $F(2, 426)=164.84$, $p<0.001$, respectively). Liquid (watercourses, rain), bird sounds and crackling fireplace were rated as the most pleasant sounds across settings, and alarm clock, objects falling, emergency vehicle siren and insect sounds as the least pleasant ones. Not surprisingly, doorbell, phone ringtone, alarm clock, pedestrian crossing warning and rockfall sounds were the most likely to prompt actions according to the tested participants. Wind and construction site sounds were the least likely to prompt actions. Interestingly, ratings of both scales were significantly but poorly correlated ($r$=-0.176, $p<0.001$), suggesting that hearing-related functional character and emotional valence are distinct attributes of environmental sounds. A complementary survey was conducted on another group of 140 healthy Ukrainian participants (113 females; 3 non-binary) aged between 18 and 40 years (mean: 21 years; SD: 4 years) to assess the effects of varying the magnitude of certain sound characteristics (wind and rain strength; car and air traffic). Two repeated-measures ANOVAs with Sound and Strength (low vs. high) as within-subject factors were conducted separately for emotional and functional ratings. The analyses revealed that all the main effects were significant ($p<0.001$) but the interaction was significant only for emotional ratings ($F(3, 423)=56.85$, $p<0.001$). Post hoc paired comparisons with Holm correction showed that pleasantness decreases with increasing strength (all $p<0.001$), with the exception of rain ($p=0.476$). In contrast, all sounds become more likely to prompt actions as sound magnitude increases (all $p<0.001$). A third survey was conducted on another group of 153 healthy Ukrainian participants (122 females; 3 non-binary) aged between 18 and 39 years (mean: 22 years; SD: 5 years) to assess the importance to hear each sound and the difficulty one would experience if that sound could not be heard. Again, each participant was informed about the context before rating a given sound. The questions asked to participants are detailed in the legend of Fig. 2. Results are shown in Figs. 2C and 2D. The analysis of variance on the scores with Sound and Subject as factors showed that importance and difficulty varied significantly across individual sounds (Importance: $F(37, 5624) = 21.00$, $p < 0.001$; Difficulty: $F(37, 5624) = 29.05$, $p < 0.001$). Repeated-measures ANOVAs with Environment as a within-subject factor showed that ratings varied significantly across settings (Importance: $F(2, 304) = 0.23$, $p = 0.798$; Difficulty: $F(2, 304) = 7.55$, $p < 0.001$). The highest-rated sounds with respect to importance were Alarm clock, Emergency vehicle siren, Rockfall, Doorbell, Objects falling. The lowest-rated sounds were Clattering cutlery, Household appliances, Doors, blinds, windows, Construction site, Insects. The highest-rated sounds with respect to difficulty were Alarm clock, Rockfall, Emergency vehicle siren, Cars and Doorbell. The lowest-rated sounds were Clattering cutlery, Wind, Water in the city (fountain, river), Amphibians (frogs, etc.), Wind in foliage. Ratings of both scales were significantly correlated ($r$=-0.459, $p$<0.001), suggesting that the two questions assess aspects related to hearing-related importance of environmental sounds. Ratings of hearing-related importance and difficulty were compared to ratings of the functional character of each sound. Only difficulty one would experience if a given sound could not be heard was significantly correlated with the functional character of this sound, that is its susceptibility to prompt actions ($r$=-0.456, $p$<0.001), indicating that auditory needs are multi-dimensional, and that importance assigned to certain environmental sounds does not necessarily reflect their "affordance".

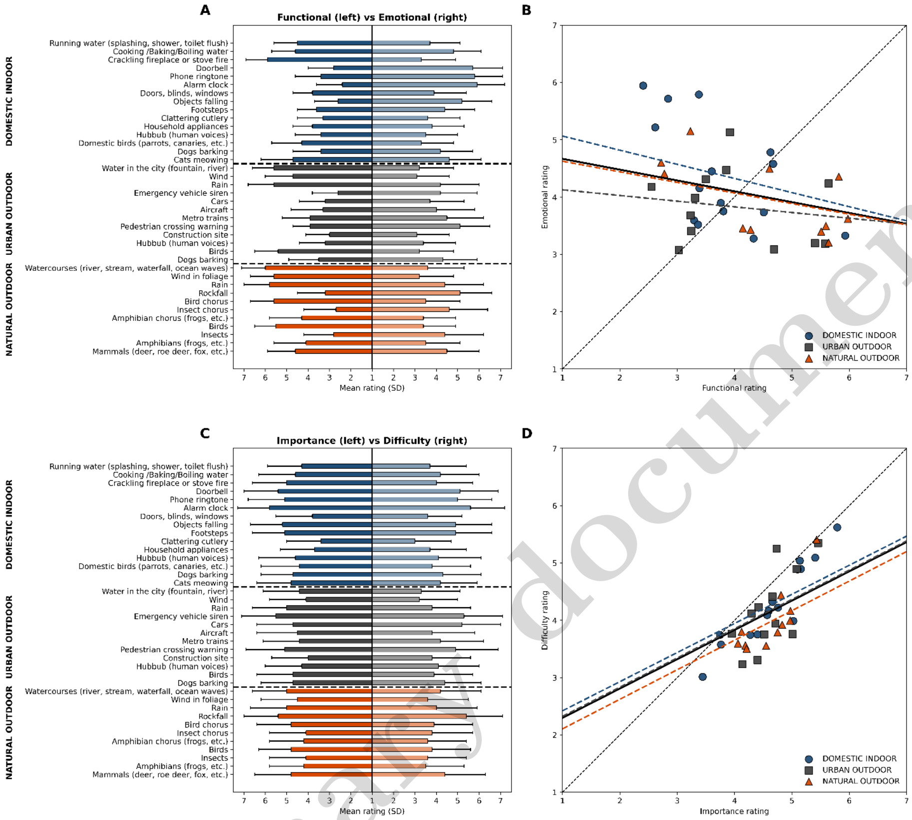


**Figure 2**: **Panels A and B:** Functional character and emotional valence for each environmental sound, assessed in 214 participants. A: Functional character [example item: "This sound prompts me to engage in specific actions". Responses collected using a 7-point Likert scale (from "1: strongly disagree" to "7: strongly agree"]. Emotional valence [example item: "This sound evokes emotions in me" Responses collected using a 7-point Likert scale (from "7: very positive" emotions to "1: very negative" emotions. B: Emotional ratings as a function of functional ratings (means across sounds), for each category of sound. **Panels C and D:** Importance to perceive each environmental sound and difficulty one would experience if that sound could not be heard, assessed in 153 participants. C: Importance [example item: "This sound …" Responses collected using a 7-point Likert scale (from "1: is not important at all" to "7: is very important"). Difficulty one would experience if that sound could not be heard [example item: "If you were unable to hear this sound …". Responses collected using a 7-point Likert scale (from "1: this would not cause any difficulty whatsoever" to "7: this would give rise to extremely serious difficulties"). D: Importance ratings as a function of perceived difficulty ratings (means across sounds), for each category of sound.

*In summary, the current sounds chosen for their relevance in everyday life differ notably in terms of their role in the well-being, functioning, safety and importance for human listeners. Moreover, their emotional valence and functional character are differentially affected by their magnitude. Assessing changes in these perceived attributes caused by cochlear or neurological damage, as well as signal processing performed by hearing aids and cochlear implants, should provide a better understanding of the difficulties encountered by clinical populations in their daily activities and their quality of life. Of course, geographical and cultural variations should be considered when designing these test batteries.*

## 6. Conclusions

Auditory tests designed to assess (i) ability to perceive environmental sounds associated with relevant situations, needs and tasks, and (ii) sensory vs. cognitive mechanisms involved in the perception of these sounds should take into account the above findings and dichotomies. This has the potential to increase the relevance of the sounds and situations studied, and provide a framework for interpreting the results obtained with a finer grain.

## References


[1] F. Apoux, N. Miller-Viacava, B. C. Moore, B. Krause, J. Sueur, and C. Lorenzi: "Auditory discrimination and identification of time of day for natural soundscapes," *Journal of the Acoustical Society of America*, vol. 158, no. 6, pp. 4734–4749, 2025.

[2] R. T. Buxton, A. L. Pearson, C. Allou, K. Fristrup, and G. Wittemyer: "A synthesis of health benefits of natural sounds and their distribution in national parks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 118, 2021.

[3] F. Dick, S. Krishnan, R. Leech, and A.P. Saygin: "Environmental sounds," in: Neurobiology of Language, pp. 1121-1138, 2016.

[4] B. De Coensel and D. Botteldooren: "The quiet rural soundscape and how to characterize it," *Acta Acustica united with Acustica*, vol. 92, no. 6, pp. 887–897, 2006.

[5] R. Fay: "Soundscapes and the sense of hearing of fishes," *Integrative Zoology*, vol. 4, pp. 26–32, 2009.

[6] S. Gatehouse, C. Elberling, and G. Naylor: "Aspects of auditory ecology and psychoacoustic function as determinants of benefits from and candidature for non-linear processing in hearing aids," in *Proc. of auditory models and nonlinear hearing instruments*, A. N. Rasmussen, P. A. Osterhammel, T. Anderson, and T. Poulsen, Eds. Holmens Trykkeri, pp. 221–233, 1999.

[7] W. W. Gaver: "What in the world do we hear?: An ecological approach to auditory event perception," *Ecological Psychology*, vol. 5, pp. 1–29, 1993.

[8] E. Grinfeder, C. Lorenzi, S. Haupert, and J. Sueur: "What do we mean by 'soundscape'? A functional description," *Frontiers in Ecology and Evolution*, vol. 10, p. 894232, 2022.

[9] B. Gygi, G. Kidd, and C. Watson: "Spectral-temporal factors in the identification of environmental sounds," *Journal of the Acoustical Society of America*, vol. 115, pp. 1252–1265, 2004.

[10] B. Gygi and V. Shafiro: "The incongruency advantage for environmental sounds presented in natural auditory scenes," *Journal of Experimental Psychology : Human Perception and Performance*, vol. 37, no. 2, pp. 551–565, 2011.

[11] D. Lelic, E. Picou, V. Shafiro, and C. Lorenzi: "Sounds of nature and hearing loss: A call to action," *Ear & Hearing*, doi: 10.1097/AUD.0000000000001601, 2024.

[12] G. Lemaitre, and L. M. Heller: "Auditory perception of material is fragile while action is strikingly robust," *Journal of the Acoustical Society of America*, vol. 131, pp. 1337–1348, 2012.

[13] C. Lorenzi, F. Apoux, E. Grinfeder, B. Krause, N. Miller-Viacava, and J. Sueur: "Human auditory ecology: Extending hearing research to the perception of natural soundscapes by humans in rapidly changing environments," *Trends in Hearing*, vol. 27, 2023.

[14] M. M. Marcell, D. Borella, M. Greene, E. Kerr, and S. Rogers: "Confrontation naming of environmental sounds," *Journal of Clinical Experimental Neuropsychology*, vol. 22, no. 6, pp. 830–864, 2000.

[15] J. H. McDermott, and E. P. Simoncelli: "Sound texture perception via statistics of the auditory periphery," *Neuron*, vol. 71, no. 5, pp. 926–940, 2011.

[16] M. A. McMullin, R. Kumar, N. C. Higgins, B. Gygi, M. Elhilali, and J. S. Snyder: "Preliminary evidence for global properties in human listeners during natural auditory scene perception," *Open Mind*, vol. 8, pp. 333–365, 2024.

[17] N. Miller-Viacava, R. Ferriere, N. R. Friedman, T. C. Mullet, J. Sueur, J. Willie, and C. Lorenzi: "Spectro-temporal modulation templates for human perception of animal vocalisations in natural soundscapes," *Journal of the Acoustical Society of America*, vol. 159, pp. 2425–2443, 2026a.

[18] N. Miller-Viacava, F. Apoux, R. Ferriere, N. R. Friedman, T. C. Mullet, J. Sueur, J. Willie, and C. Lorenzi: "Modulation statistics of natural soundscapes," *Journal of the Acoustical Society of America*, vol. 159, pp. 1263–1289, 2026b.

[19] V. Shafiro and B. Gygi: "How to select stimuli for environmental sound research and where to find them," *Behavior Research Methods Instruments & Computers*, vol. 36, no. 4, pp. 590–598, 2004.

[20] C. Suied, and I. Viaud-Delmon: "Auditory-visual object recognition time suggests specific processing for animal sounds," *PLoS One*, vol. 4, no. 4, p. e5256, 2009.

[21] E. Thoret, B. Caramiaux, P. Depalle, et al.: "Learning metrics on spectrotemporal modulations reveals the perception of musical instrument timbre," *Nature Human Behavior*, vol. 5, pp. 369–377, 2021.

[22] P. J. Webster, L. M. Skipper-Kallal, C. A. Frum, H. N. Still, B. D. Ward, and J. W. Lewis: "Divergent human cortical regions for processing distinct acoustic-semantic categories of natural sounds: Animal action sounds vs. vocalizations," *Frontiers in Neuroscience*, vol. 10, p. 579, 2017.