The Musicality of Soundscapes

Segments & Semantic Vector Synthesis

As sound ecologists and attentive listeners, we recognize that a balanced natural sonic environment offers a depth of beauty and tranquility rarely matched by human-made organized sound. The dawn chorus, the breath of a forest breeze, or the layered resonance of waves by the sea all form soundscapes of striking complexity and affective power.

Yet these natural sonic environments are not intentionally crafted for human perception. Can they, therefore, be called "music"? According to Krause's acoustic niche hypothesis (1987), such environments result from countless evolutionary decisions across the biosphere, each shaping sound to enable communication and survival. Importantly, humans are not the only species that respond to this balance. Other animals also participate and react to soundscapes in ways that suggest more than simple signaling, pointing instead to a shared sensitivity to nature's sonic events.

Musicality then, may not be exclusively a human projection, but a quality latent within ecological sound structures themselves, resonating across species through both function and form.

Is "music" merely a cognitive label applied by humans or is it also a physical phenomenon, noticeable in the patterned relations of natural soundscapes? In either case, how might compositional tools such as Segments extend this inquiry, by revealing hidden structures, reframing perception and enabling new forms of listening?

A Philosophical Tension

The conventional definition of music as an expression of human-organized sound invites a known Zen philosophical question: Does a falling tree make a sound if no one is around to hear it? If it does, can we call it music? This query reveals a tension between ontology (the nature of sound itself) and epistemology (how we perceive and interpret it), therefore it is tied to the listener's perception.

From a sound ecologist perspective, soundscapes already exhibit very pleasing musical structures combined with periodicity, rhythm, and resonance. Birdsong, whale calls, and cicada choruses display patterns and ratios that could echo human music theory (Rothenberg, 2006). A constructivist, however, might argue that "musicality" is a human projection, because our human cognition seeks order, so vibrations become scales, rhythms, and harmonies. Without a listener, sound remains just air pressure fluctuations. Storr in "The Dynamics of Creation" (1993), also suggests that seeking order within chaos is one of the vital forces promoting creativity.

From a phenomenological perspective, the question of whether soundscapes are "musical" cannot be answered by measuring their acoustic properties alone. Following Maurice Merleau-Ponty, perception is not a detached registering of external data but a lived, embodied act through which the world takes on meaning. In this view, the musicality of soundscapes does not exist as an inherent trait of natural vibrations; rather, it emerges in the experience of listening, where the listener and the sonic environment are intertwined. A waterfall's roar becomes "music" only when framed as such through perception, an idea closely related to the music perspective of John Cage as illustrated further.

This emphasis on the embodied nature of listening has been extended into sound studies by Don Ihde, particularly in "Listening and Voice: Phenomenologies of Sound" (1976/2007). Ihde argues that listening is never a passive act of reception but an active engagement with the world, which sounds are not simply "heard" but interpreted, oriented and lived through. Crucially, this active engagement means that listening constructs a bridge between perceiver and perceived - a bridge that Segments attempts to digitize but never fully control, preserving the gap as a space of creative friction. His phenomenological account shows that sound always places us in relation to space, time, and others, making listening an inherently participatory act.

From an ecological perspective, soundscapes, comprising biophony, geophony and anthrophony, reflect the balance of ecosystems (Krause, 1993). Species carve out acoustic niches, like instruments in an orchestra, for communication and survival. While this is not art in the traditional sense - what is art anyway? Is it an expression on a predetermined domain or can it also be an admirable way of living? - the analogy to musical orchestration is apparent, pointing to an aesthetic dimension that transcends human creation.

The Fragility of Soundscapes and the limit of Algorithmic Mourning

Soundscape theory often assumes continuity: that some blend of biophony, geophony, and anthrophony is always present, weaving a proto-orchestra that never fully ceases. Yet this assumption itself reveals a vulnerability. What happens when the soundscape is erased? When ecosystems collapse to the point that no voices remain, no winds stir, no rhythms survive?

In such cases, silence is not the fertile silence of Cage inviting us to listen, but a void where neither ontology nor phenomenology persists. This silence is not data; it is the absence of data and an algorithm cannot mourn what it no longer measures. This tension reminds us that sound ecology must also account for the possibility of absence - the moment when music, balance and communication are not hidden but destroyed. An example of such an environment is today the Gaza strip, where the erasure of sound becomes an erasure of life itself and the tool's visualization of "zero biophony" risks becoming a sterile dashboard rather than a call to action.

Perception as Co-Creation

Composers have long drawn inspiration from nature's sonic palette. Debussy (1862-1918) evoked water and wind in "Reflets dans l'eau" (1905), Olivier Messiaen (1908-1992) transcribed and incorporated the songs of many birds in his compositions like in "Réveil des oiseaux" (1953), while later Annea Lockwood's "A Sound Map of the Hudson River" (1989) treats rivers as complete sonic works. Finally, John Cage's "Silence" (1961) radicalized our perceived connection with the soundscape, declaring all sound as music when framed as such, thus blurring the line between music and the environment.

In soundscape composition, two distinct philosophies emerge while many exist in between. R. Murray Schafer (1977) uses environmental sounds to reflect and critique our ecological relationship, framing nature's sounds as a mirror to our roots and responsibilities. The soundscape becomes a call to listen attentively to the world's balance and fragility. Conversely, Francisco López (1997), rooted deeper in the musique concrète tradition, treats sounds as autonomous objects, stripped of context or reference. In his "absolute concrete music," a rainforest recording is not "about" the rainforest but a sonic reality to be experienced fully, emphasizing immersion over meaning. In his 1997 essay "Schizophonia vs. l'objet sonore" Lopez describes his involvement with environmental sounds as autonomous musical works, without reference, context, or manipulation.

These perspectives do not oppose but rather form a productive continuum - a field of tension where Segments operate by refusing to synthesize them into harmony. The tool's risk, however, is that it can aestheticize the very fragility Schafer mourns, turning the sound of chainsaws into a compelling texture while logging continues.

All these perspectives form a continuum: soundscapes can be ecological, symbolic, and experiential simultaneously. Natural sounds can be considered proto-musical, rich with rhythm, timbre, and texture, but become "music" - in the broader human sense - through framing, whether by listening, recording or imitation. They occupy a liminal space, neither chaos nor composition, as raw sonic matter awaiting interpretation.

Liminal Territories

The Segments application which initiated the writing of this article, is based on conceptual and spectral analysis, followed by resynthesis in both of these domains. Could a tool like Segments provide a novel perspective to soundscape composition? The answer lies equally in the conceptual part of the app - in which resynthesis becomes synonym to narrative reorganization - and to the phenomenology of the spectral characteristics of each segment. A raw soundscape is diffuse; everything unfolds simultaneously. Through segmentation and reorganization, hidden structures emerge, patterns are amplified, and unnoticed details are brought to the foreground. In this way, what was once background becomes an aesthetic focus. Transformation turns the incidental into intentional, where rhythms can be emphasized, repetitions to be shaped, tensions to be built and resolved. A storm reorganized as a crescendo and release acquires dramatic form that nature itself leaves open-ended.

Beyond narrative reorganization, the spectral resynthesis in Segments also extends perception. By slowing a bird call until it becomes a drone unveils undiscovered musical architectures embedded in the sonic fabric. In doing so, the composer creates a dialogue between nature's proto-musical patterns and experimental digital signal transformations of the spectral properties of the analyzed sounds based on the composer's aesthetics.

Yet a critical limit emerges: Segments, like most AI listening, privileges discrete, nameable events - biophony and anthrophony - while struggling with continuous, unclassifiable textures. The breath of a forest breeze, celebrated in the opening, becomes merely a categorized "wind" event rather than a continuous field of auditory experience. This disappearance reveals that Semantic Vector Synthesis is, at its core, a synthesis of classifiable relationships - leaving geophony as a resistant, untamed remainder.

In short, reorganization and resynthesis creates revelations, a way of uncovering, interpreting and amplifying the musical potential already present in the living soundscape.

AI Error and the Probabilistic Lineage in Music

Each classification or transformation in Segments carries a margin of uncertainty, a probability rather than a guarantee. Inherently, AI is not deterministic. Its training and inference rely on probability and sensitivity to tiny variations. If it were fully deterministic, it might be more predictable but also far less creative. In that respect, the bird call identified with 82% confidence may still be registered as “engine” or “cricket” by the system. In conventional deterministic engineering, this is an error; in art, it can be meaningful. To understand why, it helps to place AI within the long lineage of how music has evolved between determinism and probability.

In the Classical tradition, determinism was the ideal. A score functioned as a blueprint as every note was fixed and deviations were errors to be corrected. This pursuit of order, symmetry and control framed Western music from Bach through Beethoven (Taruskin, “Oxford History of Western Music”). By the early 20th century, cracks appeared. Schoenberg’s twelve-tone system was in theory deterministic - every note dictated by serial order - but in practice it sounded unpredictable, even chaotic, exposing the paradox between systematic logic and probabilistic perception (Whittall, “Serialism”).

John Cage’s indeterminacy reframed this tension. By tossing coins and consulting the “I Ching”, Cage deliberately relinquished control, making probability itself the compositional method. In works like “Music of Changes” (1951) and the iconic 4’33” (1952), the unforeseen was not error but essence: the environment and chance became co-composers.

Meanwhile, Pierre Schaeffer's musique concrète (1948) used recorded sounds as material. The tape medium seemed deterministic - cutting, splicing, looping - but playback introduced hiss, wow and flutter. These "errors," even unintentionally, became aesthetic resources. Later, glitch music artists like Markus Popp (Oval), embraced digital skips and corruptions as musical textures, as heard in the amazing "Re:Systemisch" (Cascone, "The Aesthetics of Failure", 2000). A piece that exemplifies the shift from music as a deterministic system to a probabilistic field is Pithoprakta from Iannis Xenakis. He incorporated stochastic processes drawing on probability distributions from the statistical mechanics of gas particles to generate "clouds" of sound. The result is a texture that feels both natural and unpredictable.

AI inherits and amplifies this domain of non-linearity, but with a new tension: while its lineage celebrates error as creative, its application in ecological monitoring demands accuracy. At the computational level, it is deterministic - same weights, same seed, same input yields the same output - but its logic is probabilistic, trained on vast datasets and designed to generalize. Each inference contains uncertainty. In an artistic context, this uncertainty is not necessarily a bug but could be a collaborator. Misclassification can create unexpected metaphors: a cicada read as "engine" transforms a jungle into an industrial landscape. Confidence scores of analysis results can be repurposed compositionally, with high-certainty inputs producing clarity while low-certainty ones blur into reverberant haze.

This is where AI-assisted composition with natural sounds acquires philosophical depth and epistemological risk. The soundscape itself is already indeterminate, emerging, shifting, unpredictable. AI, by layering its probabilistic missteps on top of nature's uncertainties, mirrors reality's messy generalization rather than distorting it into false certainty. A mislabeled cicada is not wrong but a *line of flight:* an escape from taxonomy into metaphor. Just as Cage invited us to hear chance as music and Oval turned digital collapse into texture, AI could reframe error as alternative meaning. In this sense, every session of Segments becomes an unrepeatable collaboration between three composers: nature, human and machine.

Enter Segments and Semantic Vector Synthesis

Segments is an experiment in framing, a way of leaning into the uncertainty that soundscapes already embody. Based on analysis and resynthesis, it sits at the threshold between ecological listening and aesthetic intervention, between determinacy and chance. It was built by the composer's necessities for an easily accessible, dynamic, algorithmic-driven sound synthesis, integrated with soundscape analysis. More importantly, Segments introduces an extra layer of sound synthesis, the concept of Semantic Vector Synthesis, a novel way to create musical compositions based on the meaning of sounds.

At its core, it analyzes soundscapes, distinguishing between biophony and anthrophony and rendering Krause's acoustic niches not merely as survival strategies but as audible patterns. These emergent structures become newly perceivable through segmentation. In this way, Segments does not impose music upon nature but opens a window into proto-musical relations that were already there, unnoticed.

For soundscape composition, the Semantic Vector Synthesis in Segments do not entirely arrange sonic objects in mechanical fashion but instead provide a framework for re-engaging listening itself, by shaping how a listener perceives relationships among natural, human-made and hybrid sounds. By foregrounding perception as an active process, Segments echoes Ihde's call to treat listening as relational practice, while also extending Merleau-Ponty's insight that musicality arises not in the sound alone but in the lived encounter with sound.

The spectral resynthesis functions in Segments extend this openness further. By reordering or transforming fragments of sound (segments), Segments does not aim to correct or improve the natural soundscape but to reveal latent architectures that our initial perception alone may not uncover. A cicada chorus fragmented into rhythmic pulses, a flowing river segmented into shifting tonal layers, a crackling fire magnified into percussive bursts.

The Segments application is both deterministic and chaotic. In the Render Playback section, algorithms embody various modes, based on ordering and / or chaotic behavior. The division of modes into two categories, Conceptual and Spectral, blur distinctions between analysis and invention, letting users inhabit the liminal space where error and interpretation merge. What one moment appears as pattern, the next as noise; what was "background" becomes "foreground," only to dissolve again in the following render.

Segments hold multiple music synthesis traditions at once. It aligns with Schafer's ecological critique, enabling users to foreground the fragility of the environment by exposing overwhelming anthrophony or disappearing biophony. At the same time, it resonates with Music Concrete, allowing sounds to be stripped of context, treated as autonomous presences, demanding nothing but immersion. Between these poles, it enables each listener-composer to navigate their own stance, either ecological, phenomenological, or abstract. But it refuses to let the user settle comfortably into any single stance, forcing them instead to confront the friction between them.

Segments do not settle the question, "Is nature music?" but instead it keeps the question open. It is a digital framework that makes this tension audible, an instrument promoting listening into uncertainty.

Semantic Vector Synthesis

If spectral resynthesis in *Segments* illuminates the hidden physical structures of sound, Semantic Vector Synthesis advances a different project: the construction of new, coherent soundscapes by navigating the semantic relationships the machine has learned.

At its core, Semantic Vector Synthesis begins with the semantic embeddings produced by sound classifiers. Every sound - bird, insect, engine, wind, voice - is converted into a long list of numbers called a **vector**. This vector isn't random, but places the sound inside a high-dimensional space where distance expresses the model's understanding of similarity. That is what we mean by **semantic** in this context. The spacing reflects meaning as the model has absorbed it through vast datasets. Birds cluster with birds because the model has repeatedly encountered them together.

This is how AI works at a fundamental level. It learns associations statistically and represents them as coordinates. The vector space becomes a compressed memory of the sonic world - a structured map of how natural and human-made events relate. Semantic Vector Synthesis uses this hidden intelligence to assemble soundscapes whose internal logic mirrors the ecological, cultural and perceptual relations already embedded in the data.

At this point the method clearly steps beyond classification or rearrangement. It becomes a new mode of sound synthesis, one that operates not on waveforms but on meaning. Instead of sculpting timbre directly, the composer shapes the conceptual distances and proximities learned by the model. New soundscapes emerge by reorganizing relationships inside the semantic space, producing results in a different sound domain from traditional synthesis methods.

Where spectral resynthesis reveals the micro-structure of sound, Semantic Vector Synthesis reveals its macro-structure. Which sounds belong together, which sequences feel ecologically plausible, which transitions resonate with the underlying organization of natural environments.

This semantic space is also open to other creative uses from occasional misalignments or the invention of entirely artificial sound ecologies. A semantic detour - placing a chainsaw next to soft rain - can produce unexpected however expressive results. Composers could also dive into unexplored regions of the vector space to imagine soundscapes that never existed in nature. The goal remains the meaningful ordering of sound, using the model's learned structure as a stabilizing force rather than a source of chaos. Along with coherence results, creative error and speculative invention all become reachable within the same framework.

Semantic Vector Synthesis fits naturally into the larger philosophical thread of the article. Musicality emerges not solely from perception or chance, but from a deliberate alignment between ecological structure, machine understanding and human intention. The composer collaborates with the model's internal logic, using its capacity to detect relations we overlook, while still shaping the outcome with artistic decisions. The result is something more than reconstruction. It is a method for composing with the intelligence embedded in how sounds relate across the world, both real and imagined.

How Segments Works

The Segments application is written in Javascript utilizing the Web Audio API. It is built around two neural network AI models primarily used for soundscape analysis and an Audio Render section for processing the analysis results. MediaPipe's YAMNet model listens to the world through 521 predefined sound categories - e.g., bird, car horn, applause, footsteps - then sorts these into biophonic or anthrophonic constellations. It also offers the possibility of extending its ‘perception’, through the TensorFlow model, by letting users teach the system new sounds, creating parallel user models that live in the browser and stored as JSON weights in localStorage. Therefore, each user’s Segments version is slightly different, carrying its own idiosyncratic memory if trained as such.

This dual-model structure means Segments is both fixed and open, but it also materializes a phenomenological split: YAMNet acts as the consensual, “objective” epistemology (what the world agrees a “bird” is), while the TensorFlow layer becomes your personal, embodied ontology - your listening history literally weighted into the system’s future perceptions. Together they form a two-voiced analysis tool, one as a global archive, the other as a personal interpretation of trained sounds.

Segments do not stop at identifying and naming sounds. It implements various, expandable, compositional strategies, projected through different modes in the Audio Render section. Lorenz attractors pull segments into chaotic orbits; Markov chains trace probabilistic paths; “ecosystem mutation” simulates predator-prey relations among sound categories; “narrative journey” builds arcs from calm to tension to release. Each mode is a metaphor, an interpretation of what soundscape might mean when reframed. These are not deterministic outputs but invitations to the same audio material reorganized differently each time and revealing new textures in what once felt familiar. A ‘user’ mode also provides a window for testing custom-made algorithms and extending the compositional strategies of the app.

MIDI integration broadens this openness further: labels can be mapped to MIDI notes, turning a detected birdcall into a synthesizer trigger or the thrum of traffic into modulation for a filter, pushing Segments into the instrument-building territory, treating the environment as both an archive and a score.

The visualization layer, the result tables, pie charts and frequency analyses, acts like a score. These images are not endpoints but provisional sketches of acoustic presence. They show the balance of biophony and anthrophony, but they also show the limits of listening through an algorithm: what it highlights, what it misses, what it mistakes. Each misclassification becomes not just part of the artwork but a trace of the machine’s own perceptual limits, making bias itself audible.

Through Conceptual and Spectral playback renders, Segments stitches fragments into composites, sometimes sequential, sometimes stochastic, sometimes recursive. Reverb can be applied to the whole, only to biophony or to a single chosen label. A bird call drenched in cathedral reverb becomes more than bird call; traffic stripped of its context becomes rhythm. These transformations are based on correct analysis, but they are also about speculation, imagining alternate interpretations of a soundscape.

In essence, Segments can serve as an ecological monitor, measuring habitat fidelity through the ratio of biophony to anthrophony presence and also function as an artistic collaborator, rendering stochastic, chaotic or narrative recompositions. The ratio itself becomes an existential metric - a number that can indicate either biodiversity health or the quiet violence of erasure.

Underlying all of this is a question: what does it mean to listen through an algorithm, to entrust part of perception to a system that is probabilistic, sometimes precise, sometimes wrong? Segments make that question audible.

Playback Modes

Segments compositional engine offering paradigms that organize soundscapes through narrative, chaotic, ecological and spectral principles. These paradigms extend soundscape composition into new territories, balancing conceptuality, phenomenology, ecological critique and musique concrète abstraction. The compositional playback modes reflect the dual nature of its synthesis engine: one that operates on semantic relationships between sounds and another that operates on their raw spectral, acoustic properties.

Full List of Current Modes in Segments

1. "low→high" / "high→low"
   - Function: Orders sound segments by their average frequency (spectral centroid) from low to high, or high to low.

2. "markov"
   - Function: Uses a Markov chain to sequence segments based on the probabilistic likelihood of one sound event following another, as learned from the analyzed soundscape.

3. "lorenz"
   - Function: Uses the chaotic mathematics of the Lorenz attractor to assign segments to positions in a non-repeating, fractal sequence.

4. "quantum"
   - Function: Sequences segments probabilistically based on their intensity (RMS), where louder sounds have a higher probability of being placed.

5. "tidal-rhythm"
   - Function: Organizes segments in a cyclical pattern based on their durations.

6. "euclidean"
   - Function: Distributes sound events evenly across a rhythmic grid based on Euclidean algorithms.

7. "alt" (Alternate)
   - Function: Interleaves anthrophonic and biophonic segments in an alternating sequence.

8. "label" / "exclude-label"
   - Function: Sorts and plays segments based on user-selected semantic labels (e.g., only "bird") or by excluding them (e.g., everything but "traffic").

9. "common→rare" / "rare→common"

- Function: Orders segments from the most to least frequently occurring sounds in the analysis, or vice versa.

10. "narrative-journey"
    - Function: Automatically constructs an arc by organizing segments to move from calm/low-intensity sounds, through tension/high-intensity sounds, to a point of climax, and finally to a resolution.

11. "ecosystem-simulation"
    - Function: Reshuffles sounds based on their rarity, applying "mutations" that mimic simplistic predator-prey dynamics, allowing rare sounds to occasionally dominate.

12. "user"
    - Function: A window for testing custom-made algorithms, extending the tool's compositional strategies beyond its pre-defined modes.

13. "low-emotion→high" / "emotion"
    - Function: Orders segments by their intensity or loudness (RMS), used as a proxy for emotional arousal.

14. "spectral-flow"
    - Function: Creates seamless, continuous transitions between segments based on their spectral similarity, blurring the boundaries between discrete events.

15. "rhythmic-stutter"
    - Function: Fragments a single sound segment into rapid, rhythmically repeating particles.

16. "spectral-morph"
    - Function: Blends the frequency content of two or more sounds over a user-defined duration, creating hybrid textures.

17. "formant-ladder"
    - Function: Shifts the formant resonances of a sound up or down a defined scale, independently of its pitch.

18. "spectral-stutter"
    - Function: Fragments sounds and replays them in rapid, BPM-synced pulses.

19. "formant-reorder"
    - Function: Treats the spectral frames of a segment as independent objects and reorders them.
20. "chaos-bounce"
    - Function: Uses chaotic dynamics to trigger and modulate sound segments, focusing on their textural and amplitude characteristics.

References

- Cage, J. (1961). *Silence: Lectures and Writings*. Wesleyan University Press.
- Ihde, D. (2007). *Listening and Voice: Phenomenologies of Sound* (2nd ed.). SUNY Press. (Original work published 1976)
- Krause, B. (1987). *Bioacoustics, habitat ambience in ecological balance*. Whole Earth Review.
- López, F. (1997). *Schizophonia vs. l'objet sonore: Soundscapes and artistic freedom*. Available online.
- Merleau-Ponty, M. (1962). *Phenomenology of Perception*. Routledge.
- Messiaen, O. (1953). *Catalogue d'oiseaux*. Paris: Leduc.
- Rothenberg, D. (2006). *Why Birds Sing: A Journey into the Mystery of Birdsong*. Basic Books.
- Schafer, R. M. (1977). *The Tuning of the World*. Knopf.
- Schaeffer, P. (1966). *Traité des objets musicaux*. Éditions du Seuil.
- Storr, A. (1993). *Music and the Mind*. Ballantine.
- Xenakis, I. (1956). *Pithoprakta*. Score.

Dimitris Barnias

PHD(c) HMU 2025

dbarnias@gmail.com