Rewilding Synthetic Natures — Acoustic Ecology and Machine Listening in the Technosphere

by Stefan Maier

This speculative essay explores the prospect of a new form of acoustic ecology, one where humans have spawned abstract modalites of relation to their sonic environments through computation — where technologized forms of "listening" are mapping vibrational territories that extend well beyond the limits of audibility and into realms hitherto unheard.

————————————————

Today, in the early decades of the 21st century, we can obliquely hear evidence of a strange new soundscape emerging.

This sonic environment is coextensive with the soundscape that acoustic ecology — the study of the relationship between humans and their environment through sound — has traditionally explored. It too consists of bird songs, of rushing water, and of balanced, natural acoustic niches. It too is characterized by bustling marketplaces, tinny voices from public address systems, and by bells and the various "soundmarks" that characterize human settlements everywhere. And just as acoustic ecologists have noted with alarm for decades, it too is an environment increasingly dominated by industrial noise, by the buzzing of telecommunications infrastructure, and the constant barrage of notifications from mobile computing devices.

This emerging soundscape, however, extends into sonic domains hitherto unheard: it consists as much of the familiar sounds of natural, rural and urban environments as it does sounds inaudible to human ears. It extends beyond the limits of hominid perceptual capacities, beyond the "acoustic horizon" into realms of inaudible frequencies, the electromagnetic spectrum, and ever-growing repositories of raw, noisy data as vibrational sensors proliferate everywhere, endlessly. To list but one possible subset from this ever-expanding possibility space: infrasonic elephant dirges, deep sea audio-feeds from cold-war-era hydrophone arrays, seismic data, ultrasonic biosonar, sputtering transmissions on all frequencies all at once, and ultra-compressed audio streams from billions of networked devices, all forming into an insane, incomprehensible cacophony.

But as much as the novelty of this emerging soundscape is characterized by excessive vibrations beyond the purview of the human, its most striking departure comes from how it is sensed in the first place. As smooth signals across highly disparate transmission mediums and temporalities are translated into quantized data, the Digital allows previously isolated stratas of sound to be stitched together into jarring, angular tapestries. Today, echolocating bats can be "heard" alongside the screaming of yeast cells[1] and data sonifications of black holes,[2] all thanks to the promiscuity of information and the mutability of its form. As the heterogeneous outputs of networked vibrational sensors are pieced together digitally, we can hear new sonic territories forming and expanding at an ever-accelerating rate.

But, crucially, this is only one half of the story. After all, in the original theorization of the Soundscape by R. Murray Schafer, the actual sonic features of a given environment can only be meaningfully engaged by the presence of a listener; indeed, according to Schaeffer, the soundscape itself is co-constituted by a situated listener through the act of purposeful listening. Here, the circumstances described are no different: the emerging soundscape is defined as much by what is being sensed than by "whom" is sensing it. And this is where things become truly unprecedented: the topologies of these newly-formed vibrational territories are currently being mapped and made intelligible not by humans, but by computational processes that we are proud to call "intelligent." Parsing highly-varied data streams into

---

[1] Roosth, Sophia. "Screaming Yeast: Sonocytology, Cytoplasmic Milieus, and Cellular Subjectivities."
[2] https://www.youtube.com/watch?v=ioR5np1fmEc&ab_channel=NASA%27sMarshallSpaceFlightCenter

synthetic sensory impressions — often on terms utterly foreign to the human sensorium and its attendant forms of cognition — Machine Listeners are charting both the audible and the unheard to unexpected ends. Discovering complex, sometimes incomprehensible patterns and structures across disparate scales, the overwhelming cacophony of contemporary networked environments are being carved into regimes of strange coherence. The soft, chaotic buzzing of a fly in the background that in no obvious way "sounds like" a loved one's voice might exhibit great latent congruence via some complex group-theoretical rotation according to software eavesdropping on your call. The faint spectral trace given off by a celestial body only audible to an auditor out of time, might harmonize in some serendipitous way with the Bach Aria bleeding from your headphones as you go out for an evening stroll under a starry sky. Masked by the bassy whirring of cooling fans and the incessant clicking and chirping of processors in data centers across the globe, today Machine Listeners are silently sensing a parallel soundscape into being accidently, as digitized environments are made comprehensible incomprehensibly.

However, just as this terrain emerges theoretically, the possibility of its intelligibility recedes. Make no mistake: the emerging soundscape is not for us. Even in circumstances when bleeding-edge technologies are specifically designed to imitate human listening and interface with humans directly, the incipient machinic sensate world recedes at transcendental remove — and this gap only widens as we consider the plurality of abstract computational "listenings" that are currently proliferating. Nevertheless, straining with a speculative ear, we might begin to discern some semblance of sense from the jarring reflections that are currently rendering in the cloud. From these broken impressions, an unexpected wilderness glimmers, one from which I suspect we have much to learn. Here, we might discover domains teeming with unruly potential, with multifold implications on the categories of the natural, the artificial, and on the nature of synthetic sensation itself. And just as listening to the traditional soundscape contributed to a profound awareness of anthropogenic effects on natural environments, this new soundscape may catalyze prescient reflections on artificial natures spawned in the wake of the human.

—————

Sitting at my computer, I open my personal archive of field recordings. For the last decade my public-facing artistic output has been paralleled by a steady practice in phonography. Far from the purposeful labour and conceptual reflection that goes into making official "works," this practice is far more personal and informal for me. Without any real intention in mind, while on a walk or as I travel, I simply record my environment when it strikes me as particularly compelling, beautiful, or otherwise. Surveying this small archive, I am struck by how this practice represents an almost diaristic approach to my own audition. Listening back to these artifacts of my own listening, I can hear the patterns of my ear and my subconscious habits and preferences: my early love of chaotic, industrial sonic environments, and later on for extremely quiet close-mic'd sounds from natural environments, for example. Most strikingly, I can hear evidence of the shockwave that Maryanne Amacher's work and thought made on my listening. I hear my ears gravitating towards sounds inhabiting the limits of perception — similar to what she termed the "head stretch" — as the almost ultrasonic hissing of radiators and insects, and unplaceable infrasonic rumblings feature heavily in the archive around 2015. Most of all, as I obliquely unearth this minor, idiosyncratic history of my own listening, I am struck by the intimacy and tenderness that this process evokes.

Today, these personal artifacts will form the basis of an entirely new "listening-relation" — one alien from my own, indifferent to personal-historical significance, and, most of all, to the affects these mediations conjure. I take a selection from the archive: a recording taken deep in the rainforest on Northern Vancouver Island, another from an abandoned amusement park in Berlin, and the last from the inside of a datacenter secluded deep within the Rockies. I open the command-line and begin running scripts gleaned from various institutional github repositories.

First, the recording from Vancouver Island is decomposed into what the machine deduces are the most salient streams of information contained therein. Reducing dimensionality amidst complexity, this algorithm clusters data points together that it deems to be most similar through a particularly cruel and

crude form of computational listening. I run the script and wait. An hour later, 7 streams of noisy, degraded audio emerge. I open them in my DAW and begin to listen. There is a dull high-pitched ringing that is difficult to place. A series of irregular low, sub-bass rumbles that might be wind, or distant waves, or something else entirely. A sharp, digital-sounding staccato that seems to be in conversation with the metallic ringing. A series of clicks and chirps that might be insects, or various small animals, interrupted by short, sharp bursts that might be footsteps, or the sound of branches breaking. There is a long, sustained hollow drone — an eerie spectral presence — made all the more mysterious given that listening back to the original field recording I cannot seem to identify its presence or source. Finally, the last file seems to consist of all the remaining audio, the detritus of the algorithmic analysis. It consists of bright, harsh swashes of noise from various sources — rustling branches, gusts of wind, distortion from the microphone — all sculpted and interrupted by the gaping silences left behind by the decomposition process itself. Ubiquitous in all the recordings, is a gentle, faint hissing, and fuzzy sputtering noise that bears fleeting spectral impressions from the original, as if haunted by the whole from which it was extracted.

Next, I take the recording from the abandoned amusement park in Berlin. Again, I run another decomposition script. This time, 5 streams emerge. I listen. Again, the sputtering noise is ever-present and louder this time, but behind this textural veil I can make out a series of resonant percussive sounds as if heard from a distance— perhaps far-off doors slamming. There are also what sound like faint metallic whines and screeches, and a series of deep, sub-bass rumbles. The remainder file is composed of angular, ghostly noises as before, but with an occasional recognizably human voice mixed in: a child laughing, someone calling out in German, my own voice as I comment on the environment around me.

The recording from the datacenter fares similarly. After running the script, 4 streams emerge: one composed of a noisy digital texture — perhaps an artifact from the significant ultrasound in this environment? — one of chirping processors; another of more sub-bass rumbles; and finally a long buzzing drone punctuated by occasional bursts that might be footsteps or movement. Again, there is the ever-present hissing and fuzzy bleed that cuts through all the recordings.

These sonic entities — these byproducts of a form of non-human sensation — do not just mirror my own experience back at me in an unrecognizable, uncanny form. Rather, they offer radically different ways of configuring data into synthetic sensory impressions, suggesting "modes of listening" that I could never have dreamed up myself. Listening to these recordings, I am struck by how they open up unique perceptual possibilities for understanding my environment — possibilities that would otherwise remain locked away if left to my own proclivities and habits. In particular, I'm compelled by how these machinic listenings seem to emphasize elements within the original recordings that I would never have noticed myself, and indeed, elements that remain almost inaudible to me on repeated listenings: the hollow drone on Vancouver Island, the faint screeching in Berlin, and the bassy rumblings present throughout all three recordings, for example.

But what's even more striking than the novel content that's been made audible, is the form that these recordings take — and it's here that this algorithmic processing suggests something truly compelling. Firstly, when comparing the resulting recordings with one another, we can hear evidence of what might be called the machine's idiosyncratic "taste." The recordings all share a common set of sonic elements, but the algorithm has chosen to emphasize different aspects in each case, often bafflingly so. Why are the human voices from the amusement park understood as a remainder of the decomposition process in the field recording from Berlin? Why is the faint spectral presence from Vancouver Island isolated and elevated in status despite its comparative inaudibility? And more importantly: what rubric is being used to inform this prioritization? What sensibility might be suggested here?

Furthermore, consider the ubiquitous sputtering noise that accompanies each output — the audible artifacts that render their contours jagged and fuzzy, almost fractal. These noisy edges seem to be composed of many disparate elements from the original that bleed into one another, resulting in the ghostly quality described above. Despite the focus of each stream on specific features, they seem to be

leaking into one another. I wonder: perhaps these are no longer primarily recordings of particular places anymore? Given that these files are literal renderings of the abstract, non-linear relationships between data points that have been discovered, what if it's more accurate to consider them as recordings of the process of machine listening itself? The detritus found on the edges of these idiosyncratic streams demonstrates that computational "similarity" and "likeness" are not bound to the common-sense reality that we assume when using these tools. Unless specifically designed to do so, they do follow the various gestalt principles that generally structure the auditory reality of humans. In other words, listening back to these files, what I am hearing is not just evidence of radically different ways of compartmentalizing my environment, but also insight into how these machine listenings make sense in the first place. And this seems to suggest a form of artificial listening far more relational than we might typically assume — where disparate data points form into complex networks where any given element can be connected to any other in multiple ways simultaneously.

In sum, then, these recordings do not just offer new ways of understanding my environment; they suggest novel ways of understanding what it means to listen. One one hand, they imply distinctly inhuman forms of listening — forms that we may only speculate about. But at the same time, in the process of relating to these synthetic sonic impressions, I am forced to confront how my own listening has become inextricably bound up with a form of machinic alterity. I recognize that my listening has been uniquely and irreversibly modulated by these processes. Listening back to the original recordings a newfound ambivalence sets in: what else is present here in spite of the phenomenological unity that my own listening yields? In the years to come the uniquely alienating effects of novel prosthetic listenings will increasingly be recognized as the plurality of human listenings continue to fuse with an ever-diversifying set of non-human relations that remain unknowable. And yet, despite this inherent unknowability, or perhaps because of it, I find myself strangely compelled by these recordings. There is a beauty in the strangeness of their form, in the way they make audible the limits of my own perceptual apparatus and cognitive faculties. In other words, what I find compelling about these recordings is precisely what makes them so difficult to listen to: they offer a kind of negative image of my own listening, a demonstration in sound of all the ways in which I am unable to understand my environment. They are a set of aural reflections that, in their very excess, suggest new modes and modalities of listening that recede just as they appear.

——————————

As urban, rural, and natural environments become increasingly populated by distributed, networked sensors, and the output of these varied devices are analyzed by an ever diversifying host of algorithmic listeners, we can infer that the complexity of the emerging soundscape will only increase, perhaps exponentially so. In fact, it might be more accurate to suggest that not just *one* new soundscape is emerging, but many. As Machine Listeners take on varied strategies of analysis — from the attempted modeling of human listening via computational auditory scene analysis, to "looking" at spectrograms, to parsing raw vibrational data via brute force statistical inference — different regimes of sense will emerge; and perhaps incompatibly so, resulting in a plurality of soundscapes given by different machinic sensibilities. The very same urban sonic environment in all its richness, might be rendered comprehensible on radically different terms such that we may be hard pressed to describe the resulting soundscapes as even related.

These considerations are made all the more complex, if we follow the line of reasoning that how something "thinks" — or in this case listens — is inextricable from how something senses in the first place. Different kinds of ears across the animal kingdom spawned radically different modalities of listening, and, by way of convergent evolution, a similar pattern may unfold in the interplay of code and hardware. If the capacities and limitations of specific sensors structure and pattern data irreversibly, then it's not only software that may contribute to the plurality of soundscapes: the great diversity of sensors currently proliferating may play an integral part too.

Dizzyingly, these prognostic reflections are further complicated by orders of magnitude if we consider all the possible permutations of sensors and synthetic listeners made possible by distributed, decentralized networks. Here, we can imagine an inter-assembly of heterogeneous sensor and machine listeners working in concert across time and space to indeterminate ends. And each and every permutation of software and hardware available may yield even stranger soundscapes than the last. Finally, we arrive at an unruly future that is not as distant as it might seem: where billions of synthetic ears and promiscuous algorithmic listeners — distributed from the hyper-local to the planetary, operating on timescales from individual samples to durations that are only limited by computing power — endlessly map dense, knotted webs of interlocking soundscapes from the turbulent vibrations of the Technosphere's "sonic unconscious." This unrelenting labor may give rise to incomprehensible wildernesses consisting of more-than-human relations — to synthetic natures spawned in the wake of industrial modernity.