

Endemic Machines

Acoustic adaptation and evolutionary agents

David Kadish

Supervisor: Laura Beloff
Co-supervisor: Kasper Støy
Submitted: April 30, 2021

IT UNIVERSITY OF COPENHAGEN

Abstract

Endemic Machines is an interdisciplinary investigation of the question: *What is required for a machine to adapt to a local ecosystem?* Through the medium of sound, it explores the conceptualization and design of machines that belong in an ecosystem because they evolve within and alongside it. Drawing on research in soundscape ecology, artificial life, and artificial intelligence, it builds an interdisciplinary view of machine engagement with an existing soundscape.

Building on the biosemiotic concept of a sensory world, an *umwelt*, this work formulates a basis for the machine sensing of the soundscape. In parallel, it presents work — conceptual and based in practice — that constructs a new frame for the production and understanding of machinic vocalizations in the soundscape. The term *robophony* is devised to describe the sounds of ecologically-engaged machines as they do not fit within the existing ontological framework which characterizes sound as originating from humans (*anthrophony*), (non-human) biological entities (*biophony*), or geophysical processes (*geophony*).

Additionally, the concept of *eco-technogenesis* is proposed as a way of understanding the cyclical co-evolution of technologies and ecosystems. This extends the concept of *technogenesis*, which situates technology and humanity as evolutionary partners, to the co-development of ecosystems and machines. The processes of feedback and co-creation inherent in *eco-technogenesis* provide a framework for understanding how machines can become endemic.

These theoretical contributions are scrutinized in the experimental design of an endemic machine that evolves a vocalization in an existing, real-world soundscape. Through this machine, the *Rowdy Krause*, adaptation to a local ecosystem is addressed in practice. The work delves into the nuances and complexities of adapting to a local ecosystem and the internal tensions embedded in the concept of an endemic machine.

Resumé

Endemiske Maskiner er en interdisciplinær undersøgelse af spørgsmålet: Hvad kræver det for at en maskine kan tilpasse sig et lokalt økosystem? Ved brug af lyd som medie, udforskes konceptualisering og design af maskiner der tilhører et økosystem fordi de udvikler sig som en del af og sammen med det. Ved at trække på forskning inden for lydbilledsøkologi, kunstigt liv og kunstig intelligens, bygges et interdisciplinært overblik over hvordan maskiner kan blive en del af et eksisterende lydbillede.

Ved at bygge på det biosemiotiske koncept for en følede verden, en 'umwelt', beskriver denne afhandling fundamentet for maskinopfattelse af et lydbillede. Samtidig præsenteres arbejde — konceptuelt og baseret på praktiske forsøg - som konstruerer et nyt felt for produktion og forståelse af mekaniske vokaliseringer i et lydbillede. Ordet "robophony" bruges for at kunne beskrive lydene fra maskiner der er en del af et lydmæssigt økosystem, da de ikke passer ind i eksisterende ontologiske rammer der karakteriserer lyd som kommer fra mennesker (antropomorfisk), (ikke menneskelige) biologiske entiteter (biophony), eller geofysiske processer (geophony).

Derudover, foreslåes konceptet øko-teknogenesis til at forstå cyklisk og gensidig udvikling af teknologi og økosystemer. Det er en udvidelse af konceptet teknogenesis, der opstiller teknologi og menneskeheden som evolutionære partnere, så det omfatter gensidige udvikling af økosystemer og maskiner. De processer med feedback og gensidig udvikling, som er en del af øko-teknogenesis skaber rammerne for hvordan maskiner kan opnå endemiske træk.

Teorierne er gennemprøvet ved brug af en eksperimentielt designet endemisk maskine, kaldet 'Rowdy Krause', der udvikler en vokalisering i et eksisterende, naturligt lydbillede. Denne maskine et eksempel på hvordan maskiner kan tilpasses et lokalt økosystem i praksis. Denne afhandling dykker ned i nuancer og kompleksiteterne i forbindelse med at tilpasse en endemisk maskine til et lokalt økosystem, og de interne modsætninger i selve konceptet.

Acknowledgements

The work that went into this dissertation would not have been possible without the guidance, inspiration, encouragement, and support of a network of advisors, collaborators, friends, and family. There are many people who played supporting roles in the production of this research and I want to say a heartfelt thanks to everyone who helped this project come to fruition.

To Laura Beloff, my supervisor, thank you for all of the support that you have given to me and my work over the course of this degree. You helped to integrate me into a research community through conferences, workshops, and residencies and guided the formation of the research program and kappa. Thanks as well to Kasper Støy for his advice and encouragement as a co-supervisor and to him, Sebastian Risi, and Anders Sundnes Løvlie for their collaboration on the various experiments and publications in this thesis.

I want to extend my thanks to the members of my examining committee, Louise Barkhuus, Mike Phillips, and Bo Reimer, for their roles in reviewing, discussing, and evaluating my work. For their help in preparing the kappa, thanks to Saundra Norton for her detailed proofreading and encouragement and to Mads Johansen Lassen for his translation of the abstract into Danish.

In the process of building prototypes for this research, I had help from technically-minded people from across ITU. Thanks to Andrés Faíña, Mathias Schmidt, Stig Anton Nielsen, Halfdan Hauch Jensen, Halfdan Mouritzen, Harvey Bewley, Sebastian Büttrich, and Jørn Lambertsen for answering my questions, troubleshooting electronics, helping with 3D prints, lending me tools and equipment, and teaching me new techniques.

Some courses and workshops have an outsized impact on a research project and the two weeklong summer camps at CATCH in Helsingør that

I certainly did for this work. Thanks to everyone involved in organizing those, and especially to CATCH's program manager Majken Overgaard. Special thanks to Maja Fagerberg Ranten and Megan Hines, my collaborators at those workshops and eventual co-authors, who helped shape and articulate the concept of eco-technogenesis.

Thank you to all of the members of the REAL Lab at ITU, who formed my home base within the university and with whom I spent lunches, coffee breaks, guest lectures, office parties, and Friday beers. Thanks as well to Auke Ijspeert and all of the members of BioRob who gave me an academic home-away-from-home during my stay abroad at EPFL in Lausanne, Switzerland.

The support teams at ITU were instrumental in helping me navigate my degree. Thanks to Julie Lyngsø Berg Jacobsen and Vibe Qvist Mathiasen in PhD Support as well as the IT Department for helping me get through the process with all of the tools that I needed.

The friends that I made during my time at ITU brought joy and reassurance to the Ph.D. journey. In particular, I want to thank Rosemary Lee, Jonas Jørgensen, Sarah Brin, Sarah Holmwood, Djordje Grbic, Hugo Mulder, Isabel Froes, Miguel González Duque, Teresa Almeida, and Tom Jenkins for many chats over coffee, conference adventures, and commutes. Special thanks to an old friend, Ely Lyonblum, for guidance into the world of sound and soundscapes.

Finally, I would not have arrived at this point without my family. To Kristin, thank you for giving me perspective when I am lost in the weeds, for your love and patience and support, for brainstorming ideas and answering endless questions about biology and ecology, for being excited about research, and for your drawings. Isaac and Alvin, you were not around for the beginning of this process, but I am sure glad you are here now. Thanks for adding adventures and cuddles to this time.

Thanks to my mom, who we lost in 2019, but who instilled in me a love of learning that has led here. And to my dad who introduced me to building and programming things and has always been there to support my academic adventures. And a big thank you to my siblings and extended family, cousins, aunts, uncles, and in-laws who are always curious about what I am up to and usually willing to listen to me talk about it.

Contents

Contents	vii
Acronyms	ix
1 Introduction	1
1.1 Background and State of the Art	1
1.2 Research Overview	5
2 Endemism	9
3 Sensing the Soundscape	15
3.1 Soundscape	16
3.2 Machine Sensing	17
3.3 Information Representations	19
4 Robophony in the World	25
4.1 _____phony	25
4.2 Technogeneses	27
4.3 Making some (robophonic) noise	29
5 Methodology	33
5.1 Methods	33
5.2 Experimental Platforms	36
6 Discussion	41
6.1 Implications	41
6.2 Ecology and conservation	42
6.3 Design decisions	44
6.4 Domains of endemism	47

6.5	Should machines be endemic?	49
7	Conclusions	51
	Bibliography	53
A	Additional Work	61
	Articles	65
1	Clustering sensory inputs using NeuroEvolution of Augmenting Topologies	65
2	The BioAcoustic Index Tool	69
3	An artificial life approach to studying niche differentiation in soundscape ecology	103
4	Robophony: A new voice in the soundscape	113
5	Sound as Material for Eco-technogenesis	125
6	Endemic Machines: Artificial Creativity in the Wild	133

Acronyms

ACI Acoustic Complexity Index. 17, 19, 37

ADI Acoustic Diversity Index. 17, 19, 37

AI artificial intelligence. 42, 46

ALife artificial life. 4–6, 35, 39, 48, 51, 52

ANH acoustic niche hypothesis. 4, 6, 16, 19, 30, 31, 39, 44, 46, 47, 51

BAIT the BioAcoustic Index Tool. 5, 6, 19, 36, 37, 39, 51

BI Bioacoustic Index. 19

EDA exploratory data analysis. 35–37

FFT fast Fourier transform. 18

MEMS microelectromechanical systems. 17

NDSI Normalised Difference Soundscape Index. 19, 26

NEAT neuroevolution of augmenting topologies. 29

NN neural network. 20, 30, 31

t-SNE t-distributed Stochastic Neighbour Embedding. 32

WSP World Soundscape Project. 16

Chapter 1

Introduction

What is required for a machine to adapt to a local ecosystem? This question underpins the various lines of inquiry in this dissertation. How does a machine become part of an ecology, a network of beings and energies and materials?

Machines are typically considered artifacts of the human world. In much of post-Enlightenment Western philosophy, that world is cleaved from the world of nature, which is to be either admired from afar or brought under control. This view has been critiqued and dismissed by a generation of scholars of philosophy, technology, and cybernetics, but it still dominates the design of machines that operate in ecosystems. These machines — agricultural sprayers and harvesters, sonic pest control devices, robot lawnmowers — operate with a singularity of purpose that does not account for the whole of the ecology in which they exist. They do to the ecosystem instead of becoming *part* of it.

In asking *What is required for a machine to adapt to a local ecosystem?*, this dissertation investigates the potentialities of bringing machine agency to the ecosystem. It works to integrate ecological concepts into the process and language of machine design. It attempts to reposition technology not as outside of ecosystems, a thing thrust onto them, but as a force that is an inherent part of biological life, that exists in non-human animal systems, and can be understood as a co-creative ecological force.

1.1 Background and State of the Art

The question of the role and place of machines in ecosystems is entangled in a number of fields. One approach comes from the fields of engineering

ethics and bioethics. Beginning as a brief consideration of the possible role of robots in monitoring ecosystems (Sullins, 2011), recent work has established different types of robot-ecosystem relationships and begun a discussion of the ethical implications of those robots (Wynsberghe and Donhauser, 2017; Donhauser, Wynsberghe, and Bearden, 2020). This discourse touches on some of the broader issues in AI and robotics — robot autonomy and AI safety; the environmental impact of robot manufacturing and disposal; and the predicted increase in the use of robotics in ecosystems.

However, these approaches are set in the language of engineering practice. In this framework, machines have tasks that they accomplish. They serve, augment and enhance (Wynsberghe and Donhauser, 2017). Even in the discussion of ecobots — the term Van Wynsberghe and Donhauser (2017) use for ecologically functional robots — the precise roles imagined for machines are the products of human desire and decision-making.

Though it is not explicitly acknowledged, this discourse is the product of a philosophy where humans sit apart from and above ecosystems. It renders humans as subject and machines and ecologies as object. The machines that it produces have singular goals that often make them unable to improvise, adapt, and respond to the complexities of the ecosystem they find themselves in.

A different approach can sometimes be found in the work of artists and designers. These are often explicit attempts to interrogate the conventional understanding of human-machine-ecosystem relationships as defined by their opposition to one another. These works invoke “the pattern that connects” (Bateson, 1972); they seek new ways of formulating and understanding the dynamics of human-built machines in ecology.

For example, Benitez and Vogl’s *in silico et in situ* (2016-2017) is a series of digitally fabricated installations designed to create or re-create habitats for animals. One of these artificial habitats is a 3D printed LED ring that serves as a platform for spiders to build their webs. This habitat raises interesting questions about how human intervention in ecosystems tips the scales in toward one species or another, in this case favouring the spider over its photophilic prey.

Two other artists’ work explores the creation of artificial woodpeckers. Ian Ingram’s *The Woodiest* (2010) focuses on their mating rituals while

Rihards Vitols's *Woodpecker* (2016) deals with their foraging behaviour¹. Ingram's body of work, in particular, is an exploration of the use of machines to communicate with non-human animals in novel and playful ways. Vitols's *Woodpecker* is interesting in that it has an explicit ecological and conservation impetus. The woodpeckers arise from the artist's concern for the health of the forest in the context of a decline in the woodpecker population and a simultaneous influx of insect life. The artwork is, in part, an attempt to help the trees survive in a changing climate by imitating woodpeckers to dissuade insects from approaching.

Vitols is not the only artist to approach sound in this way. David D. Dunn has produced acoustic art that engages with the dynamics of ecosystems since the work *Sonic Mirror* in 1986-1987. More recently, his work has focused on insect sonification. His recordings of bark beetles led to a collaboration with ecologists in which they discovered that it is possible to use sound to disrupt the lifecycles of the beetle — an animal that kills trees by the millions as it nests, mates, and grows in them (Hofstetter et al., 2014). To accomplish this, Dunn created an acoustic system that mixed recorded beetle sounds with synthetic audio and further manipulated the sounds to form a continuously changing soundscape to not allow the beetles to become accustomed to the sonic environment. This technique, in many ways, resembles the approach used in Dunn's earlier artwork.

Notably, many of these examples engage sound as a medium for exploring how machines relate to ecosystems. The sonic world is a rich space for this type of inquiry, and it forms the main site of the work in this dissertation. These aforementioned machines and those developed and explored throughout this Ph.D., engage with the soundscape — a term associated with R. Murray Schafer's *The tuning of the world* (1977) — which describes the breadth of an acoustic environment.

The field of soundscape ecology (Pijanowski et al., 2011a; Farina, 2014) — the source of much of the theoretical basis for the experiments in this dissertation — emerges from the study of soundscapes through the lens of landscape ecology. It is concerned with the ecological implications of sound at the landscape scale. Thus, soundscape ecology drives the consideration of the effects of work like Ingram's *The Woodiest* from the

¹Ingram's and Vitols's work is discussed in further detail in *Robophony: A new voice in the soundscape* (article 4). That article also features a partial discussion of Dunn's work discussed later in this section.

individual animals that hear the artificial mating woodpeckers to the entire network of impacted relationships in the ecosystem.

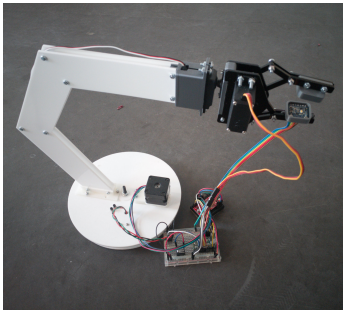
One of the concerns of soundscape ecology is understanding the dynamics of soundscape formation. Under the umbrella of soundscape ecology, several theories have been developed that offer hypotheses about how animals' vocalizations shift in relation to one another and the non-biotic portions of the soundscape. One of these theories, the acoustic niche hypothesis (ANH), has played an important role in much of the work in this dissertation.

Bernie L. Krause first proposed the ANH in his 1987 article, "Bioacoustics, Habitat Ambience in Ecological Balance." The theory holds that species tend to differentiate their calls to occur in a unique combination of space, time, and spectral frequency distribution. This process serves to minimize the chance of confusion and interference between species' signals. Under this theory, different species occupying the same physical space in a stable ecosystem might signal using the same pitch structure at different times of the day² or using different frequencies at the same times, but will not generally use the same types of tones simultaneously. This theory is a key driver of two of the main experimental platforms in this dissertation: the artificial life (ALife) experiment (article 3) and the Rowdy Krause (articles 4 to 6).

Soundscape ecology divides the soundscape into three components based on sound source. The categories — biophony, geophony, and anthrophony — were first proposed by Krause (2008) and were adopted as part of the field in its founding publications (Pijanowski et al., 2011a; Pijanowski et al., 2011b) and a subsequent book on the subject (Farina, 2014). These three categories were expanded to four with technophony in the framework used in in the broader field of ecoacoustics³ (Farina, 2018; Farina, Eldridge, and Li, 2021). This categorical breakdown of sounds and the difficulty of placing some of the works considered here within them led to the formulation of the new category called robophony in *Robophony: A new voice in the soundscape* (article 4).

²Or during the different seasons.

³Ecoacoustics, defined in 2015 by Sueur and Farina, widens the scope of inquiry of soundscape ecology to include the consideration of sound at all ecological levels — not just the landscape.



(a) Robotic arm that senses and sorts samples by colour.



(b) The BioAcoustic Index Tool prototype in the field.



(c) The Rowdy Krause installed in a community garden.

Figure 1.1: Prototypes developed and tested in the completion of this dissertation.

1.2 Research Overview

The response to the research question “*What is required for a machine to adapt to a local ecosystem?*” was built throughout the Ph.D. from the bottom-up through a series of design experiments, simulations, and prototypes followed by reflection. The experiments began with an exploration of robotic senses and meaning-making. This was followed with the design and implementation of a field experiment in machine sensing and, in parallel, a series of simulations that explored the making of sound. A final experiment that brought together the sensing and vocalization into an embodied artificially intelligent agent incorporated many of the findings from the earlier research and led to many of the conceptual contributions of this dissertation.

The research is organized around three physical prototypes (fig. 1.1) and an ALife simulation. The prototype in fig. 1.1a is the subject of *Clustering sensory inputs using NeuroEvolution of Augmenting Topologies* (article 1). That article presents a novel approach to organizing sensory information by using an evolutionary neural network to cluster unlabeled data. In the context of machine engagement with ecosystems, this is understood as a building block to meaning-making independent of a human-defined ontology.

The second prototype, the BioAcoustic Index Tool (BAIT), continues to explore machine sensing of the ecosystem, but shifts perspective from the exploration of basic meaning and grouping in article 1 to acoustic indices designed to measure biodiversity using sound. Article 2, written

for an audience of soundscape ecologists and bioacoustics researchers, presents BAIT (fig. 1.1b) as a tool on which large-scale, long-term ecological monitoring networks may be based. It details the design and testing for a system that can autonomously record acoustic indices for months and possibly years at a time for use in longitudinal studies of soundscapes and biodiversity. However, it also represents another approach to machine organization of acoustic information — one that uses expert-designed heuristics to generate meaningful information from raw data.

In parallel with the development of BAIT, a virtual experiment employing techniques from the field of ALife was underway. This experiment, detailed in *An artificial life approach to studying niche differentiation in soundscape ecology* (article 3), tested and explored some of the basic hypotheses of soundscape ecology by evolving communications strategies for virtual populations and species. In doing so, it broadens the understanding of the evolution of a species's vocalizations.

Article 3 focuses on the acoustic niche hypothesis (ANH) and demonstrates that it is possible to produce some of its predicted effects by incentivizing a population to understand members of its own species and ignore members of other species. This is significant because this reward is often intrinsic in the wild.

The final three articles (4-6) reflect on the development of the third prototype — the Rowdy Krause (fig. 1.1c) — at different stages of its conception and prototyping. The Rowdy Krause is an embodied artificially intelligent agent built to express the predictions of the ANH by evolving a vocalization and constructing an acoustic niche in an existing ecosystem. These articles use the Rowdy Krause as the basis for the creation and expansion of concepts and theories in a number of fields.

Article 4 addresses a gap that the Rowdy Krause exposes in how soundscape ecology categorizes sounds. It examines the biosemiotics of sounds in a soundscape as an alternative approach to categorizing sounds by source. In considering the Rowdy Krause alongside other artwork that produces ecologically relevant sound using machines, the article proposes a new category of sound called robophony to refer to machinic sound with ecological intent or meaning.

Similarly, *Sound as Material for Eco-technogenesis* (article 5) considers the process of co-creation that occurs between the Rowdy Krause and an existing soundscape and uses that as a basis for building up the concept of eco-technogenesis. Eco-technogenesis builds on the theory

of technogenesis, expanded from the work of Stiegler (1998) by Hayles (2012). Where technogenesis posits that technologies are not merely useful prostheses but that humans and technologies are co-evolutionary forces that shape one another in a reciprocal process, eco-technogenesis holds that a similar process can and does occur in ecosystems. In this conception of the relationships between technology and ecosystems, machines and ecosystems can adapt to one another through a process of feedback that mirrors the co-evolution of humans and technologies.

The final article, *Endemic Machines: Artificial Creativity in the Wild* (article 6), names and describes the concept that frames the dissertation. It is written after a field experiment with a finished prototype of the Rowdy Krause and reflects on the design process to formulate machines as endemic inhabitants of ecosystems. It details the production of the Rowdy Krause and frames the practice of designing endemic machines in general.

The next six chapters contextualize, clarify, and extend the insights in the articles described above. While the articles are organized expansively, beginning with the most specific moving toward the general, these introductory chapters take the opposite approach. Chapter 2 (Endemism) expands on the concept of endemic robotics defined in article 6 and describes its relationship with eco-technogenesis. Chapter 3 (Sensing the Soundscape) brings together the work on machine sensing and elaborates on how the act of sensing an ecosystem is an important step in adapting to it. The corollary to that is sound production, and chapter 4 (Robophony in the World) delves into how the production of sound is formulated throughout the work in this dissertation.

Chapter 5 (Methodology) elaborates on the methodological bases for the various lines of inquiry found within this work. Chapter 6 (Discussion) raises issues and observations that arise from the consideration of the body of work as a whole and points to directions for future work. Finally, chapter 7 (Conclusions) reviews the contributions of the thesis and summarizes the findings.

Chapter 2

Endemism

What is required for a machine to adapt to a local ecosystem? The question appears relatively clear at the outset, but upon closer examination, it becomes fuzzier. What does it mean for a machine to adapt to an ecosystem? The question is confounding even applied to biological species. Conventional notions of *native* and *invasive* have been challenged to reflect the dynamic character of ecosystems and ecosystem-ness (Rodriguez, 2006; Prévot-Julliard et al., 2011; Schlaepfer, Sax, and Olden, 2011). What has often been thought of as *the pristine state* of an ecosystem has turned out to be merely a stop along a winding path.

As the notion of an ecosystem has shifted from the pastoral to something more complex (Harries-Jones, 2008; Marris, Mascaro, and Ellis, 2013), the question of adapting to an ecosystem has become more complicated. In the first place, a species adapting to an ecosystem presupposes that the ecosystem is a static target to which the species can change itself to fit. It fails to recognize that ecosystems and species are in a state of constant co-creation.

Perhaps the question goes back even further to the very definition of an ecosystem. What is an 'ecosystem' that is being adapted to? In the seminal text *Fundamentals of ecology*, Odum (1971) describes an ecosystem as "[t]he [biotic] community and the nonliving environment function[ing] together" in "a given area" that could not exist without "the cycling of materials and the flow of energy." By that definition, a species is a part of an ecosystem simply by its existence in a place and its participation in its material and energetic flows.

If the question is considered in the negative — when is a species *not* adapted to an ecosystem — this implies that the species fails to engage

with the existing material and energetic flows. Virtually all species will cause some a shift in these flows, but a species that adapts does so without collapsing or even rapidly altering the entire network of material and energetic relationships. After the introduction of a well-adapted new species, the system is able to return to an equilibrium state that is similar to the state it existed in prior to the disturbance.

Hobbs, Higgs, and Harris (2009) introduced new terminology for thinking about these types of ecosystems in “Novel ecosystems: implications for conservation and restoration.” They devised a tripartite categorization for ecosystems as historical, hybrid, or novel. While recognizing that the term *historic* itself is somewhat vague, they define *historical ecosystems* as containing “biota and ecosystem properties” within the ecosystem’s range of variability in some previous time (Hobbs, Higgs, and Harris, 2009). A *novel ecosystem* has a species composition or functional properties that are different from any historical range, while a *hybrid ecosystem* retains many of the ecosystem’s historic properties¹ but supports a different range of species or a different geophysical environment.

What does this ecosystem classification mean for the place of machines in ecosystems? Machines are unlikely to fit into the conception of a historical ecosystem, having not been present in an ecosystem on anything approaching evolutionary timescales. They could, however, be part of a hybrid ecosystem that retains much of its historic properties and flows with some shift in its inhabitants — including the introduction of a machine. It suggests that a machine should enmesh itself in the ecosystem’s flows of material and energy to become adapted.

Endemic Machines: Artificial Creativity in the Wild (article 6) describes an approach for this type of adaptation. It borrows the concept of endemism from ecology to guide the design of machines that integrate with an existing ecosystem through behaviour.

Anderson (1994) traces the ecological concept of endemism back to the early 1800s, where it was used in a French-language science dictionary to refer to families of plants in which all of the species grow in only one country (De Candolle, 1820, p412). Since that time, it has been used in ecology to refer both to species that occur only in a small area and to

¹Such as “nutrient load, hydrology, species diversity, etc.” (Hobbs, Higgs, and Harris, 2009)

species that occur in a specific area of undefined size². (Anderson, 1994). Darwin added a requirement for a historical relationship, describing species that are endemic to a place as having been "produced there, and nowhere else in the world" (Darwin, 1869, p121)³.

Anderson (1994) notes that endemic has also been used recently in ecology to denote species that are restricted to a certain type of habitat instead of a particular geographic area, which Anderson refers to as *habitat endemism*. This definition suggests a shift from the understanding of area as a physical location to a more conceptual paradigm that links physical place by habitat type to understand a boundary for an endemic species.

In light of these varying uses of the term across research fields and time, it seems prudent to specify how *endemic* is deployed in this context. From the definitions mentioned above, one can understand endemic species as those that: are found only within a small physical area; are found only within a single country; are found only within some bounded geographic region; have evolved — and remain exclusively in — a particular place; or, are found only in certain types of habitats.

Defining endemism by political boundaries (e.g., countries, provinces, etc.) is not a useful approach here as landscapes and ecosystems cut across political borders. The size of the distribution is also not the issue of greatest concern for endemic machines; the interest is in machines that participate in an ecosystem's material and energetic flows, so it makes little difference whether the ecosystem in question is measured in square metres or square kilometres. More expansive definitions, such as habitat endemism, might be interesting for understanding whether a machine that is endemic to one place might also be considered endemic in another similar habitat but are not the main focus here.

The definition of endemism used in this dissertation and the accompanying articles has a dual focus: being bound exclusively to some geographic region, without defining a particular limit on the region's size; and the historical contingency of evolving in a specific environment, as discussed by Darwin (1869).

An entry on endemism in the 2008 *Encyclopedia of Ecology* uses language that helps to clarify the precise nature of the endemism

²Which could range from the size of a small puddle to the area of a whole continent or the entire planet.

³However, an earlier edition of the text refers only to an endemic species being "found nowhere else in the world."

discussed here. It lists six different subcategories of endemic species⁴ that help to illuminate the circumstances of the endemism. *Autochthonous endemics* evolved in the location that they now exclusively inhabit, reflecting Darwin's conception of endemism as being produced by a place. In contrast, *allochthonous endemics* evolved elsewhere but are currently found exclusively in one place. *Taxonomic relics* are the last surviving members of a once diverse group, while *biogeographic relics* were once widespread but are now confined to a small area. Finally, there are *neoendemics* and *paleoendemics*, the former which arose recently and may only be endemic to an area because it has not had the opportunity to spread further⁵ and the latter, which is an old species that may once have been widely distributed but is now confined to a particular place (Morrone, 2008).

Of these, the taxonomic and biogeographic relics and the neo- and paleoendemics describe the evolutionary journeys of species across generations. Machine lifetimes are considered on the order of years or perhaps decades, so these subcategories have little meaning and point to where the metaphor of endemism collapses for machines. Endemism in ecology is about speciation, generational adaptation, and habitat (Anderson, 1994; Morrone, 2008). Machines have models, versions, minor and major revisions. Their endemism, especially as conceived in *Endemic Machines: Artificial Creativity in the Wild* (article 6), is individual.

That assertion — that machine endemism is individual — seems contradictory at first consideration. Endemism generally deals with the distribution of a species; an individual is always limited to a particular place. However, the endemism of machines, the endemism of the Rowdy Krause, is not a general endemism. It is an autochthonous endemism. It is the endemism described by Darwin (1869) when he writes of a species “produced in a specified place and nowhere else in the world.”

For endemic machines, “produced” does not refer to the physical body. The Rowdy Krause is built from off-the-shelf electronic components that are global in origin; they are made of materials harvested from mines around the world, refined and assembled in any number of factories that are part of a global supply chain of electronic commodities.

⁴(Morrone, 2008) refers to endemic *taxa* as opposed to species, to refer to the idea that endemism can be defined at any taxonomic level (e.g., order, family, genus) and not just at the species level. For the purposes of this discussion, I'll continue to refer to species, but the ideas should be applicable on a broader level as well.

⁵As opposed to being limited by climate, geological boundary, or some other factor.

Instead, the specifics of the place produce the machine's actions. What is described here is a type of behavioural endemism in which what the machine does reflects a relationship with a particular ecosystem.

This type of co-creative process in which an ecosystem changes machine behaviour which impacts the composition of the ecosystem is described in *Sound as Material for Eco-technogenesis* (article 5). Eco-technogenesis was developed in article 5 as an extension and expansion of the concept of technogenesis (Stiegler, 1998; Hayles, 2012) from the co-evolution of humans and technologies to the co-evolution of ecosystems and technologies⁶. This process is not new, nor is it limited to human technologies⁷. However, the advent of evolution-inspired computational algorithms means that the feedback cycles inherent in eco-technogenesis can occur at a far more rapid pace than ever before.

Together, eco-technogenesis and endemism — specifically autochthonous endemism — are useful concepts for describing how machines can adapt to a local ecosystem. Endemism is an effective metaphor for the kind of the relationship that can be formed between a machine and an ecosystem. The specificity of the relationship between an endemic species that evolved in and alongside the bio- and geologic conditions of a particular place is a useful frame for thinking about how a machine must be approached to be truly of a place.

Eco-technogenesis approaches from the other direction but arrives at a similar conclusion. It comes from the human perspective, extending a notion that is concerned mainly with the co-evolution of people and technology into one where technology and ecosystems can co-evolve and co-create one another. Eco-technogenesis is a process by which a machine can become endemic. As a concept, it takes care to render the ecosystem a dynamic, creative actor in the relationship, something that is easily forgotten at the pace of digital evolution but is essential in forging long-term bonds.

⁶It is also discussed further in section 4.2.

⁷E.g., Beaver dams, termite hills, etc.

Chapter 3

Sensing the Soundscape

What is required for a machine to adapt to a local ecosystem? The previous chapter outlined a framework for this process. Endemism attempts to define what it means to adapt to an ecosystem. Eco-technogenesis describes the process by which ecosystems and technological entities can change with each other. This chapter and the next (chapter 4) examine how that happens, more specifically, delving into the sensory and motive mechanisms behind the process of eco-technogenesis and becoming endemic.

The previous chapter established that endemism — a form of adaptation to a local ecosystem achieved by evolving alongside it — can be based on behaviour. One way of focusing the work is to centre the inquiry on a particular domain of behaviour, a specific sensory field, a single pairing of inputs and outputs. In this case, sound emerged as a particularly interesting candidate.

The sonic domain is a potent space for this sort of exploration. As Cobussen (2016) points out, sound is an enveloping sense. It surrounds the producer and perceiver. The sonic bends easily around trees, through brush, and over walls. To hear a sound is to ride a wave, its pressure changing over time. In this way, it is technically straightforward to sense and produce. Working with sound requires only a microphone and speaker, electronic components which are easily actuated and recorded and that can readily create and capture complex signals.

Vision is often considered the primary human sensory modality (Colavita, 1974; Hutmacher, 2019). Sound is an essential mode of communication for many species. Auditory signals are used for everything from finding mates to warning of nearby danger and marking

territorial dominance (McKenna, 2020). The amalgamation of those biological signals and the sounds from human and geological sources combine to form what ecologists call the soundscape.

3.1 Soundscape

Sound in the environment — or the soundscape — has been studied extensively since the 1960s and is an active area of study in a number of fields today. Authors have long written about sounds in the environment, as Schafer (1977) points out extensively, but they were perhaps first brought to the fore of public consciousness as an ecological force by Carson's *Silent Spring* (1962). Carson marshalled sound — or its absence, in this case — to channel a larger conversation about environmental conservation and the hazards of overconfidence in technological interventions in ecosystems. Notably, the lack of birdsong is mainly discussed in a single chapter of the book, *And No Birds Sing*, but so powerful is the metaphor of a silent spring that it forms the title and organizing concept of the entire work.

The concept of soundscape — an auditory corollary to the concept of a landscape — was formalized by R. Murray Schafer in 1977 in the seminal work *The tuning of the world*. The concept grew in prominence through Schafer's World Soundscape Project (WSP) and the artwork and writing of WSP collaborators Hildegard Westerkamp (1974) and Barry Truax (1978).

Adopting the framework of acoustic ecology from the WSP, Bernie L. Krause proposed the acoustic niche hypothesis (ANH) in 1987, positing that the spectro-temporal plane is a finite ecosystem resource and that species tend to occupy non-overlapping niches in the timing and frequency of the sounds that they produce. Audio spectrum, in this telling, is territory just as much as a hole in a tree trunk or a particular patch of grassland, and encroachment on that territory likely leads either to adaptation or decline.

That encroachment, the projection of sounds in the same frequency and timings as a resident species in an ecosystem, can come from the arrival of a new species, but it is just as likely — or perhaps more so — to have a human origin. Krause and his collaborator Stuart Gage (2003) used the term anthrophony to describe these sounds of human

origin, along with biophony and geophony, to refer to sounds coming from biological organisms and geological processes¹.

In the first decade of the 2000s, the practice of soundscape analysis shifted with the introduction of digital recording and analysis tools. The relative ease of spectral analysis with digital tools gave rise to a set of acoustic indices such as the Acoustic Complexity Index (ACI) (Pieretti, Farina, and Morri, 2011) and Acoustic Diversity Index (ADI) (Villanueva-Rivera et al., 2011) that are used to monitor variables like biodiversity and species richness and to identify acoustic events (Abdallah, Frigui, and Gader, 2009; Towsey et al., 2014; Farina et al., 2016).

The publication of a special issue of *Landscape Ecology* (Pijanowski and Farina, 2011) and eventually a book (Farina, 2014) formalized a collection of research under a set of related sub-disciplines into the field of soundscape ecology. Human analysis of soundscapes using methods such as sound walking (Westerkamp, 1974) and cliraudience (Schafer, 1977) has remained an essential component of soundscape studies in the soundscape ecology era. However, many of the recent developments have focused on computational methods of soundscape analysis; they represent a turn toward machinic senses of sound.

3.2 Machine Sensing

Humans detect sound as pressure waves disturbing the eardrum, which transmits vibrations through the malleus, incus, and stapes² and to the hairs in the cochlea that vibrate in response to different frequencies (Wallace, 2010). The cochlear hairs translate vibration into electrical impulse to be processed in the brain.

Typically, in machines, this mechanical-to-electrical translation is conducted by a microphone (Zawawi et al., 2020). A wide range of materials and mechanisms can be used to accomplish this transformation. However, the microelectromechanical systems (MEMS) capacitive microphone, which produces an electrical signal in response to shifting capacitance between a flexible diaphragm and a fixed plate, is commonly used in small electronics applications, including those in *The BioAcoustic*

¹For example, the sounds of wind, water, rain, avalanches, and volcanoes would all be considered geophony.

²The three bones in the middle ear, also known as the hammer, anvil, and stirrup.

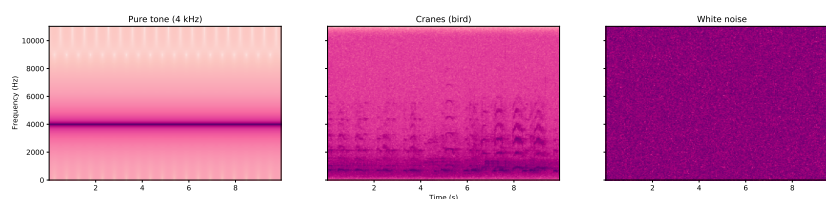


Figure 3.1: Sample spectrograms showing the frequency components of a pure 4 kHz tone (left), the call of a crane (centre), and white noise (right). The horizontal axis is time and the vertical axis is frequency, with the darkness of the image at a point indicating the intensity of that frequency component of the sound at a given point in time. (Originally in article 6)

Index Tool (article 2) and *Endemic Machines: Artificial Creativity in the Wild* (article 6).

Once sounds are converted to electrical impulses, the human auditory processing system features a hierarchy of subsystems that process sounds from low-level signals detected from the cochlea to higher-level abstractions and concepts (Kell et al., 2018). Machine auditory systems often feature a similar structure; analog electrical signals are amplified and filtered and then converted to digital representations for further processing. The process and pitfalls of digitization are well-documented (Farina, 2014; Browning et al., 2017) and will not be covered in detail here.

In machines, the journey from raw digital audio data to a higher-level, more meaningful representation of what is being heard takes many forms. Often, the first involves a shift from the time domain in which the signal is captured to the frequency domain, which more closely resembles low-level animal perception of sound (Farina, 2014). The fast Fourier transform (FFT) is a classic tool of signal analysis, and one of its variants is often used to transform an audio signal from a single value in time representing the momentary fluid pressure on a sensor to an array of numbers representing the energy in a particular frequency band at a given time.

This representation is relatively straightforward; the values produced by an FFT can be imagined as similar to the electrical response of an array of cochlear hairs at a given time (see fig. 3.1 for a visual representation of this measurement). How a machine perceives the acoustic world depends on how those values are subsequently processed.

3.3 Information Representations

The machines in this thesis utilize three distinct types of representations to engage with sensory information: *heuristic*, *semantic*, and *relational*. Heuristic representations use static, pre-defined computations to calculate a higher-level feature of a set of data; semantic representations map an *a priori* meaning to a signal which serves a reference or ground truth that an intelligent agent can learn; and relational representations describe the connections between pairs or groups of data points, which allows an intelligent agent to learn similarities and differences between sensory experiences.

The acoustic indices calculated by the BioAcoustic Index Tool (BAIT) (article 2) are found using a pre-defined algorithm and are therefore considered heuristic representations of the captured sound. The benefit of heuristic representations such as ACI and ADI is that they have a universality to them — the calculations would be the same for a particular sound regardless of the context in which it is computed. These representations can be compared across time and space.

The two acoustic indices calculated by BAIT are far from the only heuristic representations of sound. ACI is a measure of the temporal variation of a sound, and ADI is related to the spectral variation of the sound, but other acoustic indices like the Bioacoustic Index (BI) and Normalised Difference Soundscape Index (NDSI) attempt to quantify features like the intensity of sounds and the relative prevalence of anthropogenic and biological sounds. As of 2019, more than 60 acoustic indices have been developed for use in analyzing soundscapes (Bradfer-Lawrence et al., 2019).

A semantic representation of data is used in *An artificial life approach to studying niche differentiation in soundscape ecology* (article 3). The study tests the ANH by simulating the evolution of communication in two species inhabiting a virtual soundscape. Each species consists of a population of senders and receivers; the senders evolve an encoding for messages and receivers evolve the ability to identify messages from senders of their own species and decode them (see fig. 3.2 for a schematic drawing of this system). The data representation for the receivers — they are responsible for sensing the soundscape in this experiment — is semantic because each message has a ground truth meaning and this is what the receivers evolve the ability to decode.

The paper focuses on understanding the formation and dynamics of acoustic niches, but it also demonstrates one way of machines extracting

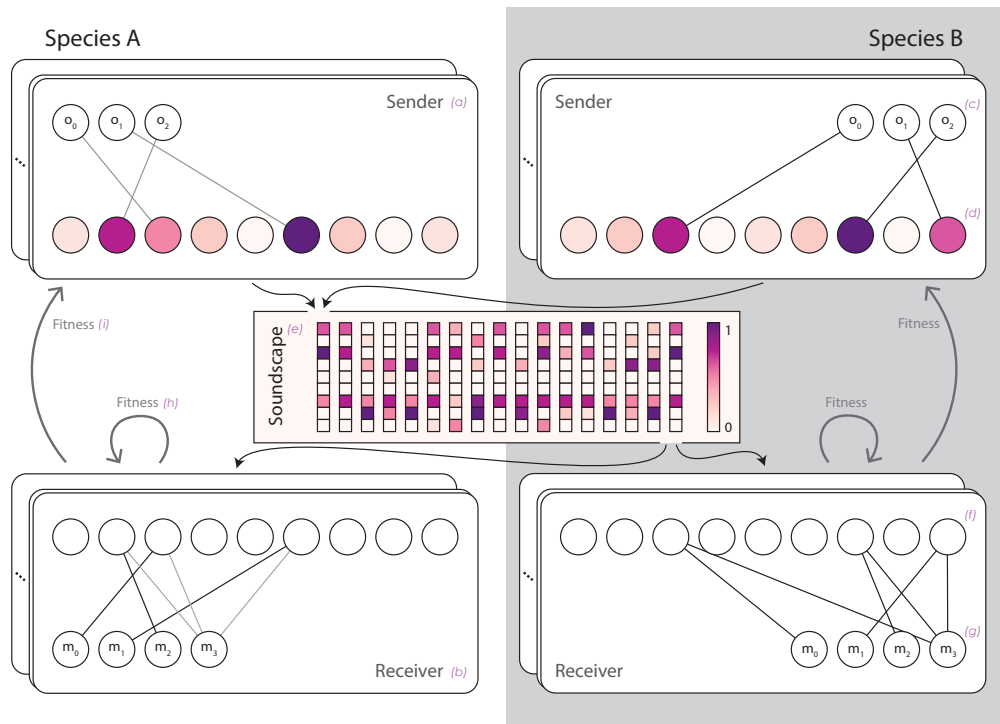


Figure 3.2: A schematic of the experimental setup in article 3. Two populations, each consisting of sender and receiver neural networks (NNs) encode and interpret messages in a simulated soundscape. Receivers of one species have to learn to ignore messages from the other species's senders while learning to decode messages from senders from their own species. This creates evolutionary pressure on the senders to differentiate their messages, and the species tend to construct their own niches. (Adapted from article 3)

meaning³ from sound. The system in the simulation is a highly simplified; the virtual soundscape is a 9-element array of floating point numbers between 0 and 1, representing the intensity of sound on 9 frequency bands. The message is far more straightforward than those that would be received in the full spectrum audio of a physical ecosystem, and the ground truth is presented as a single, absolute number.

However, the experiment demonstrates that the species form separate niches to optimize their internal communication. This simplified system in *An artificial life approach to studying niche differentiation in soundscape*

³Meaning is used here to indicate a unit of knowledge that has some useful impact in the world.

ecology (article 3) points to how semantic representations of data can help groups of machines develop a communications strategy and find an acoustic niche in a multispecies environment.

Two experiments in this thesis, *Clustering sensory inputs using NeuroEvolution of Augmenting Topologies* (article 1) and the Rowdy Krause (detailed in article 6), use relational representations to try to organize the information that they capture. Instead of processing the information and comparing it to a ground truth as in article 3, these methods are useful for establishing patterns by grouping similar sensory experiences.

These methods are interesting because they do not rely heavily on a priori human knowledge and understanding as do the first two representations. Acoustic indices are explicitly based on a human understanding of scientific analysis and methods of processing sound that reveal something about an ecosystem's underlying properties. Likewise, semantic representations require a ground truth that usually employs human arbitration of what is true. Relational representations, however, can rely almost exclusively on the sensory perception of the machine.

That is not to say that there is no human involvement. The designers will select sensors for the machine⁴, the data will be preprocessed through human-designed signal filters, and the programmer will decide which dimensions are included in the data to be clustered. However, from this point, the associations between data points are based entirely on what signals the machine can capture. There is no further need for human interpretation of the information before the machine can act on it.

This machine-centric representation of the sensory world is an important consequence of *Clustering sensory inputs using NeuroEvolution of Augmenting Topologies* (article 1). The clustering algorithm in that paper deals with visual input as opposed to acoustic data; nevertheless, the concepts therein apply to audio perception as well. Grouping similar sounds is one of the foundational processes in infants learning auditory communication (Goudbeek et al., 2005). Goudbeek et al. (2005) describe a process whereby newborns can discern between a variety of sounds wider than the phonetic categories that exist in the language(s) they are exposed to; only later do they begin to treat sets of sounds that don't have a semantic difference similarly. In other words, infants *first* learn to differentiate and group sounds and *subsequently* begin to ascertain meaning for them.

⁴Microphones and human sensing

The clustering, differentiating, grouping of experiences is then a type of pre-semiotic knowledge. It can serve as the foundation of a machine's model of the world, the basis of its *umwelt* (Hayles, 2012)⁵. As such, relational representations of sensory data serve an important function for enabling a machine *umwelt* that is distinct from the human perceptual world. Semantic representations tell a machine what *is*: This sound maps to the concept of *crow*; that one maps to the concept of *dog*. However, those distinctions are only important in a particular context. Depending on what a machine is doing, it may only need to distinguish *animal* from *vehicle*. Alternatively, the concept of *dog* is too broad for a machine that has a particular dog-related goal, and it needs to be able to differentiate *terrier* from *shepherd* or even one individual dog from another of the same species.

In declining to interpret the message concretely, algorithms using unlabelled data leave the determination of meaning and action to another process. They delay or even avoid the active assigning of meaning to signal. They allow an endemic machine to decide which distinctions are important based on what it needs to do.

The images in fig. 3.3 show the groupings determined by the robot in *Clustering sensory inputs using NeuroEvolution of Augmenting Topologies* (article 1). The robot was allowed to examine the leaves using its colour sensor⁶ and used the four light values — red, green, blue, and unfiltered — to learn three clusters of similar leaves. In fig. 3.3, some patterns are immediately discernible: many leaves of the same species are grouped together; darker green samples are mainly grouped in Cluster 0; brighter green samples seem to be grouped in Cluster 1. However, other choices are not so obvious — for example, the large leaf in the bottom of the image of Cluster 1 seems like an odd fit for that group — and a human sorting of these leaves would likely produce very different results.

This is precisely the point. The robot's sensory system is different from our own. It senses a single point of colour, whereas a human might examine the colouring, shape, texture and smell of the leaf before deciding which are most similar. What is similar in the robot's worldview is different than in a human perceptual world. A cluster-based representation of sensory information allows the robot to create

⁵Emmeche (2001) is more skeptical that a machine can have a fully formed *umwelt* akin to that of an animal.

⁶The Adafruit TCS34725 (<https://www.adafruit.com/product/1334>).

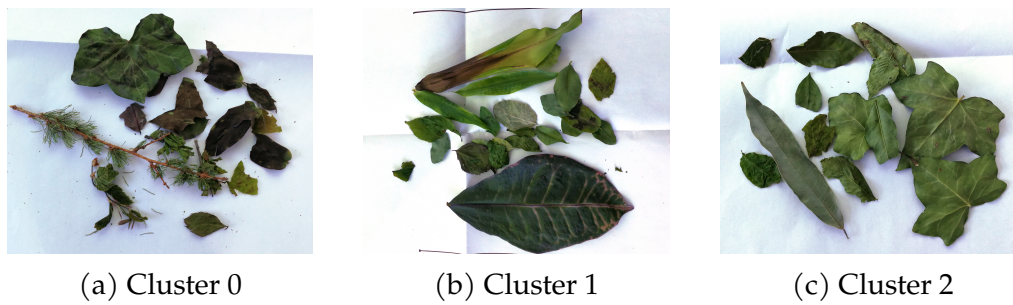


Figure 3.3: Three clusters learned by the robot in *Clustering sensory inputs using NeuroEvolution of Augmenting Topologies* (article 1).

that different experiential world without imposing the human sensory world on it.

This relational model of the sensory world can then form the basis for a machine becoming endemic to an ecosystem in the manner described in chapter 2. It can use this model to assess possible behaviours, how they relate to sensory experiences it has observed, and how they fit into a unique part of the state-space of the machine's perceptual world, a new niche.

Chapter 4

Robophony in the World

What is required for a machine to adapt to a local ecosystem? Thus far, chapter 2 has established the concept of endemism as an ecological metaphor to understand the notion of a machine fitting into an ecosystem. Chapter 3 describes the relationship between endemism and the sensory world of a machine. Nevertheless, endemism cannot be merely observational; it requires an engagement with the flows of energy and materials in an ecosystem.

As established in chapter 2, endemism is a reciprocal relationship; it requires bidirectionally. In auditory terms, this means duplex communication, hearing and speaking, sending and receiving. In other words, the ability to become part of a soundscape — to be endemic within it — depends on both the ability to perceive the soundscape and then to act upon it.

This chapter focuses on the action, the sending of signals, the speaking into the soundscape. The discussion about the role of machinic sound in ecosystems is framed by the ideas in the two papers *Robophony: A new voice in the soundscape* (article 4) and *Sound as Material for Eco-technogenesis* (article 5). The practice of designing these machines is discussed through the work in *An artificial life approach to studying niche differentiation in soundscape ecology* (article 3) and *Endemic Machines: Artificial Creativity in the Wild* (article 6).

4.1 _____phony

The roles of different sounds in an ecosystem are often understood in soundscape ecology through the framework of a tripartite categorization.

Beginning with Krause and Gage (2003), sound in the landscape was categorized as biophony, geophony, and anthrophony, though recent publications have added a fourth category — technophony (Farina, Eldridge, and Li, 2021). This division quickly became a central concept in the field of soundscape ecology. It is used to calculate acoustic indices such as the Normalised Difference Soundscape Index (NDSI), which compares the relative volumes of frequency bands that are typically anthrophony-dominated with those that are biophony-dominated (Kasten et al., 2012). It even factors in the definition of soundscape ecology in one of the papers that helped to define the field; Pijanowski et al. (2011a) describe soundscape ecology as “all sounds, those of biophony, geophony, and anthrophony, emanating from a given landscape to create unique acoustical patterns across a variety of spatial and temporal scales.”

While these categorizations are useful in many studies in soundscape ecology, they fail to capture some types of sounds. The need for a new way of considering the sounds of ecologically-active machines is the impetus behind article 4, where a new category of sound called *robophony* is defined. It argues that a machine adapted to a particular ecosystem, creating temporally layered sounds of hybrid origin is too different from the types of sounds typically considered to be anthrophony¹. It is not from a biological source, so it should not be considered biophony. Therefore it requires its own frame.

One of the ways by which robophony might be understood is by its impact on an ecosystem. In *Robophony: A new voice in the soundscape* (article 4), robophony is understood as the product of a dynamic agent, something that can shift over time and in response to the ecosystem in which it resides. The process is feedback-oriented and relies on the interplay between machine and existing soundscape. This interaction, a co-evolution of technology and ecosystem, is characterized and conceptualized in the paper *Sound as Material for Eco-technogenesis* (article 5).

¹The notion of technophony had arisen at this point but was not part of the central discourse of soundscape ecology as Farina, Eldridge, and Li (2021) was not published at the time of writing. Regardless, technophony describes the kinds of rumbling, buzzing, and humming that is generated by mechanized technologies. Robophony and technophony describe different types of sound and can be considered complementary categories.

4.2 Technogeneses

Sound as Material for Eco-technogenesis (article 5) introduces the novel concept of eco-technogenesis as the co-evolution of technologies and ecosystems. It describes how a machine can become a co-creative force within an ecosystem. In doing so, it expands Stiegler (1998)'s work — refined by Hayles (2012) — which situates technology as part of the story of human evolution, instead of as its byproduct. *Sound as Material for Eco-technogenesis* (article 5) also implies — but does not delve into — a concept of bio-technogenesis that extends technogenesis to non-human animals. This section elaborates on those concepts and discusses how the processes of bio- and eco-technogenesis inform the adaptation of machines to ecosystems.

The current understanding of technogenesis has been limited to the co-evolution of humans and technology². However, it can and should be extended to non-human animals as well. Studies of animal behaviour regularly identify species that use tools and communication techniques across generations. For example, dolphins have been shown to pass techniques for using sponges to assist in foraging by matrilineal descent (Krutzen et al., 2005), demonstrating tool use and intergenerational knowledge transfer. Evolving tool cultures have also been observed in crows in New Caledonia (Whiten and Schaik, 2007), as they cut leaves to fashion a handle for their hunting sticks. As more is gleaned about how non-human animals engage with the material world, it is less and less convincing to reserve the concept of technogenesis to human development.

Further examples can be found of animals using sonic tools in a way that demonstrates bio-technogenesis. Hardus et al. (2009) noted that populations of wild orangutans use tools to modify their kiss squeak — a sound likely used to warn off predators or other orangutans. They strip leaves from a tree and hold them in front of their mouths, lowering

²In *Technics and Time, 1: The Fault of Epimetheus* (1998), Stiegler addresses technics and technology in the context of hominization — the process of becoming human. One of the central theses of the work is that engagement and co-evolution with technology itself are what make humans human — and implicitly distinct from other animals. A great deal of recent research contradicts the notion that humans are unique in this way, as discussed in the main text. Additionally, Stiegler also engages in racist fallacies in a failed attempt to discuss conceptions of time. He quotes prodigiously from Jean Jacques Rousseau's *Discourse on the Origin of Inequality*, in which Rosseau lays a framework that considers groups of people in different states of civilization and savagery (Eze, 1997; Seamster and Ray, 2018).

the tone of their squeak and allowing them to emulate a larger animal. Palm cockatoos have been shown to use tools in courtship, with the male beating a hollow branch with a stick to attract a mate (Fitch, 2015). In these examples, tool use has fitness implications for the animals, implying that the use of the tool can be an evolutionary driver. These examples make a case for an expansion of the notion of technogenesis to non-human biological organisms. In particular, they point to the potency of biotechnogenesis in the sonic domain.

The concept of eco-technogenesis introduced in article 5 further expands technogenesis beyond a single species to entire ecosystems. Eco-technogenesis describes the process of co-becoming where ecosystems shape the development of technologies and the impacts of those technologies shape the structure and succession of the ecosystem. The article gives an example of this in the co-evolution of the tools and ecologies of modern agriculture. It also points to the emergence of beaver dams and the pond ecosystems they are associated with as an example of eco-technogenesis with no human input³.

The article emphasizes the way in which digital technologies enable a new form of eco-technogenesis — one that can be human-initiated, but ecosystem-driven. Hayles (2012) argues that programmable technologies drive technogenesis in new ways. These technologies are explicitly feedback-oriented. They are designed to be responsive, to learn and move within a changing environment.

This means that while both programmable and non-programmable technologies can be part of a process of eco-technogenesis involving human technologies, programmable technologies can engage in this process autonomously once they are deployed. Figure 4.1 illustrates these differences. Both systems in that figure involve a process of eco-technogenesis where the soundscape shifts as birds and technologies change their sound-producing patterns in response to one another. However, changing the noise produced by the non-programmable technology involves humans studying its impact on birds and redesigning the technology to be less disruptive, resulting in a new physical object — a new, quieter car. The programmable technology, conversely, is able to adapt its behaviour in situ. Once programmed, it can engage in a process of eco-technogenesis independently of the human designer.

³Beaver dams are considered to be the technological component of this relationship. See article 5 for a detailed account .

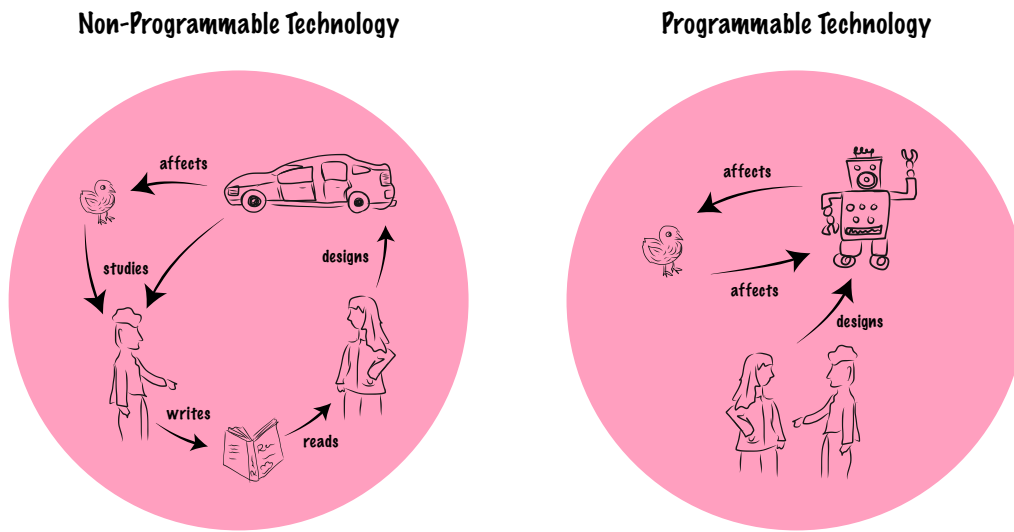


Figure 4.1: The difference between eco-technogenesis involving a programmable and non-programmable technology. Birds adjust their calls in the presence of traffic noise (Luther and Derryberry, 2012), but this has no immediate effect on the car. Eco-technogenesis might occur if an ecologist studies those relationships and then someone designs a quieter car, which has less impact on the birds. Programmable technology, such as the robot, can be designed to change in response to the environment without further human intervention.

In this way, programmable technologies can be active participants in eco-technogenesis in a form that is initiated by humans but driven forward by technology and the ecosystem. This is an important argument in *Sound as Material for Eco-technogenesis: technology, imbued with some degree of autonomy, has the potential to act ecologically while eschewing some of the traditional anthropocentric aesthetics of beauty that have driven conservation in the past. By changing itself in response to the ecosystem, it can approach “Bateson’s concept of ecological aesthetics, in which systems interact through feedback”* (article 5).

4.3 Making some (robophonic) noise

What mechanisms might machines use to evolve the production of sounds in an ecosystem? Two different approaches are attempted in this thesis, both using neuroevolution of augmenting topologies (NEAT) to

evolve neural networks. The first, detailed in *An artificial life approach to studying niche differentiation in soundscape ecology* (article 3), frames the production of signals as part of a process of evolving intra-species communication. Notably, it is undertaken in simulation and using a simplified digital code in place of sonic waves, but the principles explored there can serve as the basis for a physical instantiation of a machine. The second is mentioned in *Robophony: A new voice in the soundscape* (article 4) and *Sound as Material for Eco-technogenesis* (article 5) and described in detail in *Endemic Machines: Artificial Creativity in the Wild* (article 6) and uses soundscape ecology theory to drive the production of a vocalization.

An artificial life approach to studying niche differentiation in soundscape ecology (article 3) details a simulation that explores the mechanisms by which species adapt their communication strategies in a multispecies ecosystem. It simulates the process of two different species evolving internal communication strategies (see fig. 3.2). The experiments demonstrate that it is possible to produce the effect predicted by the acoustic niche hypothesis (ANH) — the division of a communication spectrum into distinct niches — simply by rewarding species for accurate intra-species communication. In other words, the construction of an acoustic niche arises from the benefit of communication between conspecifics — members of the same species.

Article 3 points to one possible way of formulating robotic voices to be endemic in a soundscape that mimics how animals find an acoustic niche. Though the study environment was highly simplified, it reflects one of the primary drivers of animal communication. Animals are driven to find ways of communicating with their conspecifics to attract mates, mark territory, and signal warnings (Farina, Eldridge, and Li, 2021). The better they accomplish this without interference, the more likely they are to survive, thrive, and reproduce. Acoustic niche construction is a byproduct of this process.

The second approach is discussed in detail in *Endemic Machines: Artificial Creativity in the Wild* (article 6). This approach, used in the Rowdy Krause, was designed for a lone machine operating in an existing ecosystem. The effort was framed by the understanding of the Rowdy Krause as a new species attempting to become endemic in its habitat.

The Rowdy Krause's vocalization system consists of two connected parts: a simulated vocal tract that produces audio output and an evolving neural network (NN) that controls the vocal tract to generate the desired type of sound. Each NN produced by the evolutionary process "plays"

the vocal tract and generates an audio sample, which the evolutionary process uses to evaluate the fitness of that NN.

The predictions of the acoustic niche hypothesis were used to set the fitness function — the function that evaluates the audio samples. The Rowdy Krause creates a representation of the existing soundscape by recording audio from the environment and storing a representation of which frequencies were used in each recording. The fitness function compares the frequencies used in the audio samples generated by the each NN using the vocal tract to the frequencies that are used by recordings from the soundscape. In a manner similar to the clustering algorithm in article 1, it uses the distance between frequency representations to determine which generated vocalizations are most different from the recorded samples.

As the vocalization for the Rowdy Krause evolves, vocalizations that are most different from the recorded audio samples are selected preferentially, driving the computational evolution toward unused parts of the audio spectrum. This produces the effect predicted by the ANH of a species finding a niche in a soundscape by using vacant portions of the audio spectrum.

The vocalizations that were initially produced by this process conformed to the expectations of the ANH, but they tended towards high-pitched squeals or low drones which carry little information in a biosemiotic sense. Article 6 details a small but significant set of changes that were made to the fitness function to address this. Once adjusted, the Rowdy Krause was able to evolve a set of varied vocalizations that were unique in the soundscape. A representation of these vocalizations and the recorded audio from the soundscape is shown in fig. 4.2.

Though successful in terms of emulating the predicted effects of the ANH, the Rowdy Krause and its design process raise some important issues for endemic machines: How should designers make decisions when they may be unable to sense the full range of a design's actions? How does a designer determine that a machine is doing something beneficial for an ecosystem? These are discussed further in section 6.3.



Figure 4.2: Results from the experiment with the final version of the Rowdy Krause. The image shows a representation of the Rowdy Krause's vocalizations (●) and recorded samples from the soundscape (●), plotted using a technique called t-distributed Stochastic Neighbour Embedding. The distance between points corresponds to the similarity of the spectral components of the sounds, with those closer together having a more similar spectral composition. For the Rowdy Krause's vocalizations, darker points correspond to those vocalizations evolved later in the evolutionary process. The figure shows that, with a few exceptions there is little similarity between the vocalizations produced by the Rowdy Krause and the sounds recorded from the soundscape. The Rowdy Krause appears to have constructed an acoustic niche in this soundscape. (Adapted from article 6)

Chapter 5

Methodology

The chapters of this thesis have thus far covered the formulation of a concept of machine endemism, an exploration of computational sensing of the world, and a discussion of how technological sound can be produced and understood in the context of a soundscape. These inquiries and conversations fall outside the bounds of any specific academic field. Articles 1 to 6 draw on engineering (2,6), art (4,5,6), design (2,6), soundscape ecology (2,3,4,5,6), computer science (1,3,6), artificial life (3,4,5,6), and biosemiotics (1,3,4,5,6).

As such, the work within this dissertation employs a range of research methods drawn from a similarly broad set of research traditions. This chapter focuses on the research process, including methods used and production of new knowledge, in the various projects and papers that comprise this dissertation.

5.1 Methods

Practice is an essential part of all of my research processes. Prior to this Ph.D., my previous studies were in engineering and visual arts — two fields that are highly applied and ultimately focused on producing physical instantiations of research. The importance of that creative production has not shifted in this work. All of the work in this dissertation is driven by the conception and realization of a prototype or simulation.

These prototypes and simulations (detailed in section 5.2) are used in different ways: Some are subjected to hypotheses and quantitative experimentation using classical scientific methods; others are treated in a more exploratory manner, observing the results and building knowledge

inductively. This section discusses the range of methods and their specific applications in more detail.

5.1.1 Scientific Method

Though not the primary mode of inquiry in much of this dissertation, some of the studies presented here use the classical scientific method as their main mode of inquiry. *Clustering sensory inputs using NeuroEvolution of Augmenting Topologies* (article 1) and *An artificial life approach to studying niche differentiation in soundscape ecology* (article 3) are rooted in computer science and incorporate additional insights and theories from cognitive science and soundscape ecology, respectively. Both papers make and test hypotheses¹, producing knowledge by comparing the performance of different techniques.

5.1.2 Practice-led and practice-based research

Though the two papers mentioned in the section 5.1.1 employ classical scientific methods in their studies, they are still part of the larger structure of research that stems from practice that underlies the entirety of this dissertation. Although not discussed in the paper, the study in *Clustering sensory inputs using NeuroEvolution of Augmenting Topologies* (article 1) was the result of building a robotic arm with a sensory component (fig. 1.1a) to detect colour and thinking through the process of how it could learn categories for the items that it sensed.

The role of practice in research is a topic of active discussion among interdisciplinary and arts- and design-based researchers (Candy and Edmonds, 2018; Liggett, 2020). A number of formulations of research based in practice have been articulated (Biggs and Buchler, 2008; Liggett, 2020), but this dissertation adopts the approaches to practice-based and practice-led research espoused in Candy and Edmonds's "Practice-Based Research in the Creative Arts: Foundations and Futures from the Front Line" (2018).

Candy and Edmonds define practice-based research as research where the "creative artifact is the basis of the contribution to knowledge" while practice-led research "leads primarily to new understanding about

¹*Clustering sensory inputs using NeuroEvolution of Augmenting Topologies* (article 1) is less explicit about this, but its implied hypothesis is that k-means clustering can be approximated using neuroevolution.

practice” (2018, p64). The work in this thesis engages mainly in practice-based research; new understandings about practice emerge, but they are not the primary new knowledge that results from the work. The three papers that have roots in the Rowdy Krause artifact serve as examples of this.

The Rowdy Krause was an attempt to put into practice in the physical world an understanding gained from the simulated artificial life (ALife) experiment in article 3. The concept of robophony, detailed in *Robophony: A new voice in the soundscape* (article 4), arose from realizing that the sounds produced by the artifact did not fit into the predominant conception of what composed a soundscape. This prompted research into other examples of work that also failed to conform to the existing categories and the emergence and description of a new category of environmental sound.

Similarly, a discussion of the concept of technogenesis with my eventual co-authors in the context of the Rowdy Krause revealed the gap in conceptual understanding of how ecosystems could co-evolve with technologies. This conversation resulted in the formulation of the concept of eco-technogenesis which is the basis of *Sound as Material for Eco-technogenesis* (article 5).

Endemic Machines: Artificial Creativity in the Wild (article 6) builds the new concept of endemic machines from reflections on the process of designing the Rowdy Krause. Endemic machines emerged during the process of designing the Rowdy Krause while considering the ways that the prototype evolved alongside the ecosystem. The link to ecological endemism was established by thinking through the design and programming of the machine.

5.1.3 Exploratory Data Analysis

The exploratory nature of some of the studies in this dissertation meant that for some experiments — particularly the field trial in article 2 — there was no specific hypothesis being tested at the outset. These studies used a bottom-up approach that was open to the emergence of new patterns for further investigation. The goal of data collection in these works was to generate a broad view of the observed processes in order to test the efficacy of new experimental systems. In place of testing a specific pre-formulated hypothesis, the data from the experiments was assessed using exploratory data analysis (EDA) techniques.

EDA is the process of broadly assessing collected data to identify possible patterns and trends (Jebb, Parrigon, and Woo, 2017). The main task is to generate — not test — hypotheses; the process uses a bottom-up approach with the raw data as a starting point to generate new theory (Salkind, 2010). These processes are critiqued in scientific communities when used to validate hypotheses as it is often possible to generate some statistically significant result using a dataset; however, this practice amounts to a misuse of a generally valuable strategy in data analysis and should not be seen as an indictment of the processes themselves when applied to hypothesis generation (Jebb, Parrigon, and Woo, 2017).

The clearest illustration of the types of methods associated with EDA is found in fig. 6 of *The BioAcoustic Index Tool* (article 2) and is reproduced in full here in fig. 5.1. The figure is a scatterplot matrix (Jebb, Parrigon, and Woo, 2017) of the two acoustic indices and three environmental variables measured by BAIT. Because the study was mainly intended to evaluate the efficacy of the device, it was not designed to test a specific ecological hypothesis. However, the plot in fig. 5.1 allows for the simultaneous visualization of many of the measured variables, including their distributions over time. Based on that plot and some of the other representations of the data, it was possible to offer possible explanations for the observed patterns that could be tested as hypotheses in future studies.

5.2 Experimental Platforms

The experimental platforms that were used to conduct the research in this dissertation have been discussed in previous chapters, but they are collected here to bring focus to the practice-based processes the underly the research. This includes the three physical devices that form the basis of articles 1, 2 and 4 to 6, the virtual soundscape platform used in article 3, and two prototypes that were under development but did not fit the direction of the research.

The process of developing these platforms began with a research question — which often shifted as the prototype was developed — and a rapid mockup of the system to determine its feasibility. This was followed by further prototyping, development of a testable model, and experimentation.

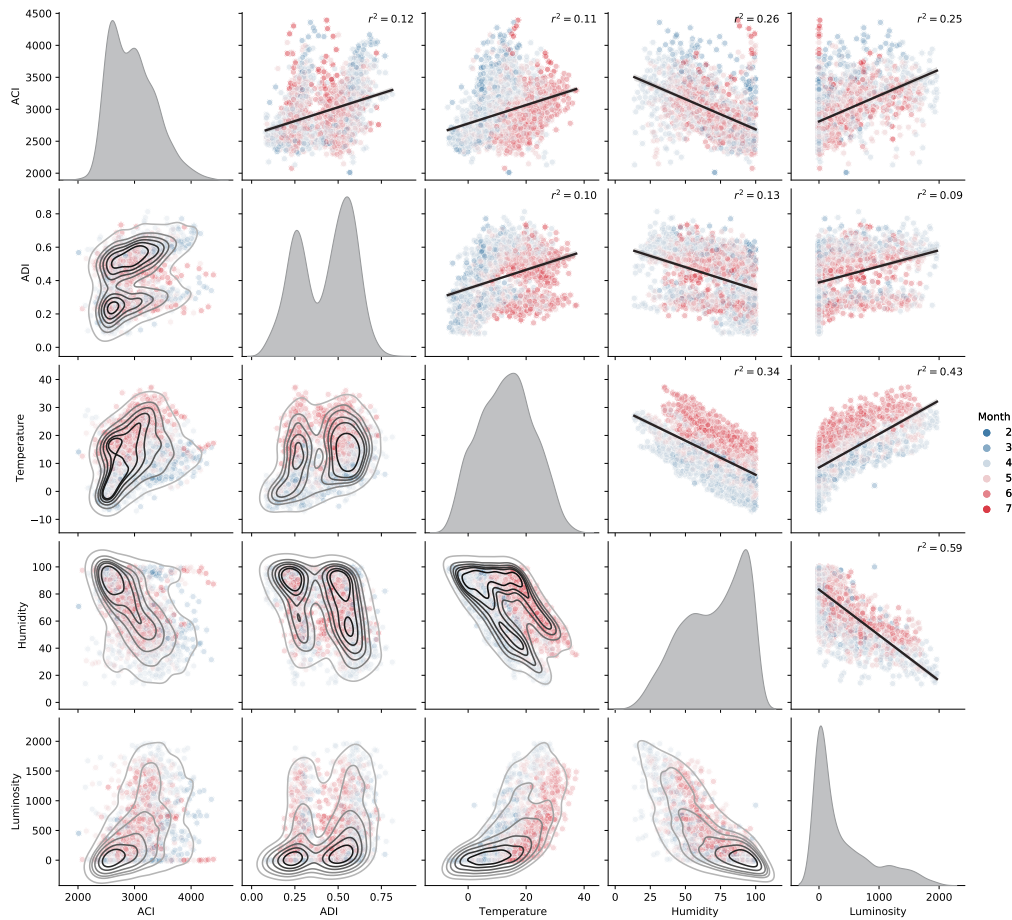
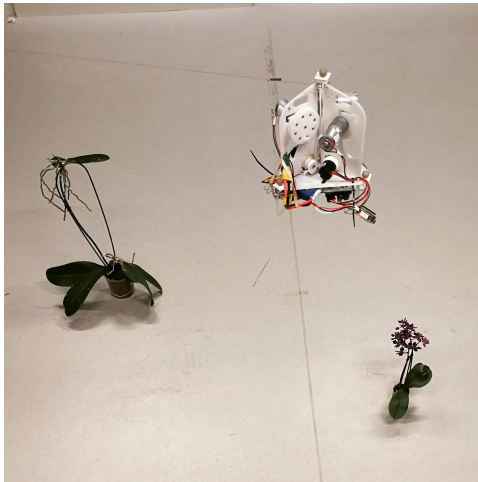
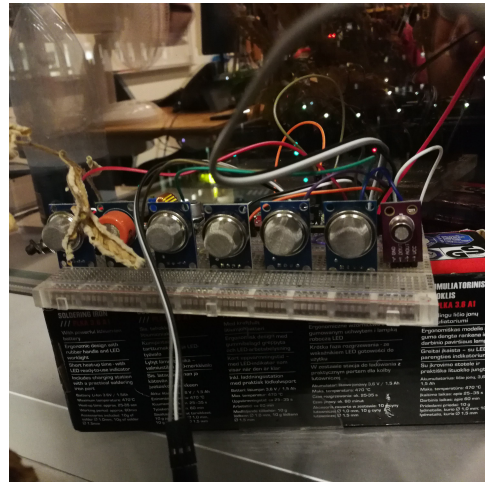


Figure 5.1: A scatterplot matrix that was part of the EDA process in article 2. It shows the acoustic indices and environmental factors measured by the BioAcoustic Index Tool (BAIT). The data is coloured by month to illustrate how the relationships change seasonally. The plots along the diagonal are density plots which — like histograms — show how each measured variable is distributed. The scatter plots above and below the diagonal show the relationships between the x - and y -axis variables; those above the diagonal are annotated with a linear regression while those below the diagonal show the density of the data using contour lines to illustrate areas of increasing density. These plots provide a helpful overview of the data gathered and allows for the discovery of new patterns and hypotheses that can be tested in later research (Jebb, Parrigon, and Woo, 2017). For example Acoustic Complexity Index (ACI) appears to be more strongly correlated with the environmental factors than Acoustic Diversity Index (ADI) does. This observation could form the basis of a new experiment to confirm and explain this apparent phenomenon. (Adapted from article 2)



(a) RUBE, a hanging robot, developed with Stig Anton Nielsen and used as part of the Beyond Digital — Towards Biological research residency.



(b) An early prototype of an electronic nose with an array of VOC sensors aligned on a solderless breadboard.

Figure 5.2: Prototypes developed over the course of the Ph.D. that were not used in the research in this dissertation.

The early prototypes and platforms began as tools for engaging some of the more basic questions related to machines in ecosystems. The latter ones were conceived as ways of asking broader questions and integrating the findings from earlier explorations.

Two prototypes — a hanging robot called RUBE and an early mockup of an electronic nose — were constructed but not included in the research discussed here. Shown in fig. 5.2, these were developed prior to the decision to focus on sound and were set aside when it was determined that they no longer fit with the direction of the research project.

5.2.1 Robotic Arm

Research question How would a machine choose to group a set of objects differently than a person?

The robotic arm (fig. 1.1a) features a single-pixel colour sensor that was used in the development of the clustering algorithm detailed in *Clustering sensory inputs using NeuroEvolution of Augmenting Topologies* (article 1). The arm detects an object in the vicinity of its gripper, closes the sensor over the object, illuminates it and measure the reflected light

with a colour sensor that detects red, green, blue, and full spectrum light. The clustering algorithm then decides to which group of objects the sample belongs, based on the clusters it has learned from the set of samples it has already seen.

A set of collected plant leaves were sampled with the robotic arm so that the algorithm could learn a cluster representation. The robotic arm was then given more leaves to see what kinds of groups it would form. As detailed in article 1, the robot and clustering algorithm emulated the results obtained using a classic clustering method — the k-nearest neighbours algorithm.

The experiments conducted on the robotic arm helped to develop the practice of using relational representations of knowledge to allow a machine to form an independent representation of its sensory experiences, as discussed in section 3.3. This formed the basis for the representation of sensory information in the Rowdy Krause.

5.2.2 The BioAcoustic Index Tool

Research question How can one measure quantitative changes in a soundscape over a span of months?

BAIT is an autonomous sensor system for measuring acoustic indices and environmental data in long-term, off-grid environments. The sensor is described in detail in *The BioAcoustic Index Tool* (article 2). BAIT was tested at a field site near Höör, Sweden for nearly six months, during which it collected over 4000 measurements of the soundscape.

The field experiments demonstrated the long-term collection of heuristic representations of the soundscape that can be used to characterize changes in the soundscape over time. Exploratory analysis of the data collected revealed some correlation between environmental factors and soundscape measurements, which warrants further exploration in future studies.

5.2.3 ALife Simulated Soundscape

Research question Does the need to communicate within a species lead to the formation of acoustic niches as predicted by the acoustic niche hypothesis (ANH)?

The simulated soundscape used for the ALife experiments is detailed in article 3. It models an ecosystem with two species trying to evolve

internal communication through a simplified soundscape. The same system was used in a later experiment (Kadish and Risi, 2020) examining adaptation to noise that is not included in this dissertation. The two-species experiment was able to demonstrate that evolutionary pressure to communicate within a species is a significant driver of acoustic niche construction.

5.2.4 The Rowdy Krause

Research question *What is required for a machine to adapt to a local ecosystem?*

The Rowdy Krause was conceived as a way of bringing together many of the lines of inquiry from earlier experiments and exploring broader questions related to the roles of machines in ecosystems. Its development is detailed in article 6, but it influenced the formation of the concepts in articles 4 and 5 as well as the organizing concept of this dissertation — endemic machines. The Rowdy Krause exists in three forms: a first virtual experiment discussed in articles 4 and 5; a second virtual experiment shown in the art exhibition at the 2020 ALife Virtual Conference; and as a physical experiment.

The process of developing, testing, and thinking through the Rowdy Krause resulted in the development of three new concepts that are key contributions of this research: robophony, eco-technogenesis, and endemic machines. The experiments with the Rowdy Krause demonstrated machine construction of acoustic niches and pointed to questions about the design and ethics of machines in ecology that are explored in article 6. Further exploration of these questions can be found in chapter 6 (Discussion).

Chapter 6

Discussion

What is required for a machine to adapt to a local ecosystem? Chapters 2 to 4 describe a process that involves becoming part of the energetic flows of a soundscape by learning its contours and evolving vocalizations focused on under-utilized portions of the acoustic spectrum.

Through virtual and embodied experiments, these chapters — and the papers on which they are based — lay the groundwork for thinking through and making a new type of machine. It is a machine that is deeply engaged in the dynamics of ecosystems, whose behaviour arises from a particular place and time in a particular ecological context.

Bridging the fields of soundscape ecology, biosemiotics, design, engineering, and computer science, the space of inquiry in the discussion of endemic machines is expansive. The previous chapters cover and add depth and context to the published articles, but some issues remain unresolved. Additionally, points were mentioned in passing that warrant further attention. This chapter focuses on those outstanding concerns to bring disparate thoughts together and open space for future work.

6.1 Implications

What are the broader implications of the concepts and prototypes developed here? In particular, how should the Rowdy Krause and the concept of endemic machines be considered outside the confines of this particular research?

The work has ramifications for an important trend in machines and ecology: the recent and ongoing proliferation of sensors and robots used in ecological contexts. Already, drones are used to monitor sensitive bird

and mammal populations, autonomous robots are used to cull invasive apex predators, and climbing robots are designed to perform tasks from inspection to pest control (Wynsberghe and Donhauser, 2017).

The way that these machines are conceived and designed matters¹. The assumptions about the world that get built into these machines affect how they impact the ecosystems that they engage. In engineering discourse, these machines tend to be understood as extensions of human agency, projecting a human-defined task beyond the reach of the human hand (Wynsberghe and Donhauser, 2017).

Endemic machines and the Rowdy Krause describe and demonstrate an alternative set of assumptions; they look toward a possible future in which machines are designed as ecological entities. In doing so, they open space to discuss what kinds of machines should be engaging with ecosystems.

6.2 Ecology and conservation

Much of the discussion space that is opened by the consideration of endemic machines concerns the fields of ecology and conservation. Endemic machines raise some of the same questions as conventional machines used in these ecology and conservation work: What is the environmental impact of creating and disposing of these machines (Wynsberghe and Donhauser, 2017)? Do narratives of their ecological benefits account for that material impact? More broadly, when should humans choose to actively intervene in ecosystems (Hobbs, Higgs, and Hall, 2013; Wynsberghe and Donhauser, 2017) and to what end?

Additional concerns arise when a machine is explicitly engaged in a co-creative evolutionary process with an ecosystem as endemic machines proposes. Some of these concerns are shared in general with practitioners of artificial intelligence (AI), who are worry about the safety of evolutionary computational systems (Lehman, 2020); others are more particular to endemic machines.

One concern is that, co-evolving with an ecosystem, the ecosystem could become reliant on endemic machines. The endemic machine could drive species towards an evolutionary dead-end instead of allowing them to adapt to changing conditions, causing problems in the future if the

¹In the words of Michelle Westerlaken (2020) channelling Haraway (2019), "It matters what designs design designs."

machine is removed. Van Wynsberghe and Donhauser (2017) raise a similar question in consideration of what they call ecobots or robots-for-ecology:

Could the failure or removal of such a robot weaken the system if said system was relying on the robot to fulfill a role/function? Is it possible that the robot will become requisite for some systems to flourish? (Wynsberghe and Donhauser, 2017, p19)

However, the concern is even more acute when considering endemic machines. Ecobots of the type described by Van Wynsberghe and Donhauser (2017) tend to have static missions, whereas endemic machines are specifically understood to co-evolve with the ecosystem. That co-evolution, while embodying an approach that is attuned to an ecosystem, also creates a greater risk for a catastrophic disturbance by the removal of the machine.

In practical terms, the removal of the machine is less a question of *if* than of *when*. Machines are bound to break down, so it is imperative to consider the impact on an ecosystem of removing a component that has — by the definition of endemic machines — been attempting to evolve into a deeply integrated part of its material and energy flows. Barring the advent of machines capable of reproduction or self-repair² — a development which would carry with it its own unique concerns — machines are destined to be temporary ecosystem inhabitants.

While this dilemma can and should be treated as a serious concern, it can also be viewed as an opportunity to pull focus to the dynamic structure of ecosystems. Species move in and out of ecosystems and ecosystems adapt to these changes. In this context, designers of endemic machines should consider planned obsolescence — not in the conventional manner of technology that breaks down after three years, but in the sense that their exit from the ecosystem is planned and built into their programming and behaviour structure. In a way, they should help the ecosystem adapt to their absence just as they learn to adapt to the ecosystem.

²Depending on the definition of a machine, these may already exist. Van Wynsberghe and Donhauser (2017) speculate that engineered plants and biofilms can be considered robots.

6.3 Design decisions

The design of endemic machines raises a question related to the discourse on human aesthetic preferences in ecology. What role should human aesthetic preferences play in the design of endemic machines? Human perception of what an ecosystem *should* be often runs counter to how ecosystems actually work (Harries-Jones, 2008). How, then, does a designer trust themselves to make decisions about endemic machines?

A further challenge arises when one considers that these machines will interact with living beings whose perceptual systems are vastly different from our own. Their senses extend beyond our own, and their *umwelten* are inaccessible to us. In work with two of the experimental platforms — the robotic arm and the Rowdy Krause — it was often difficult to tell the difference between the machines not working and them not behaving as expected. How can one design in imperceptible spaces?

These two questions are linked; how does a designer make decisions when their senses cannot be trusted to evaluate them? The approach detailed in *Endemic Machines: Artificial Creativity in the Wild* (article 6) is twofold: determine a framework for making decisions before the design process begins — in this case, designing the machine based on the principles of the acoustic niche hypothesis (ANH) — and adopting the stance that the my own aesthetic preferences should be ignored in favour of the interests of the ecosystem.

Despite the success of the Rowdy Krause prototype, reflection on the design process revealed issues with that approach and some insights into how to shift the process in future work.

6.3.1 Objectivity and the View From Nowhere

The position that it was possible to discount my aesthetic preferences in the design of the Rowdy Krause can be critiqued in a similar way as the “view-from-nowhere” positioning in journalism and the objective stance in scientific inquiry (Haraway, 1988; Young and Callison, 2021). It performs the “god trick” of pretending to be able to see all perspectives while being fixed in a single point of view (Haraway, 1988). The idea that it would be possible to remove my aesthetic viewpoint from the design process masked the important ways in which that viewpoint informs the design. It imposed a distance on my relationship to the Rowdy Krause and the ecosystem that it inhabited that was false and easily broken in the latter part of the design process.

The result is that when the time came to make aesthetic decisions, it was impossible to fully engage with the complexity and indeterminacy of my choices. When initial trials of the Rowdy Krause produced unexpected types of vocalizations, the immediate conclusion was that they were not animal-like enough and the code had to be fixed. The assumption was that my decision was serving the ecosystem's structure. However, it is possible that the sounds were indeed adapted to the ecosystem, and my expectations of the types of sounds that *should* evolve were biased. An acknowledgement of my true position in the process would have encouraged a deeper exploration of those initial vocalizations.

One approach to engaging with other species in this way can be found in the recent work of Michelle Westerlaken. In her doctoral thesis, "Imagining Multispecies Worlds," Westerlaken (2020) explores a set of design practices that she calls multispecies worlding. She engages in world-making and *thinking with* different animals (and a forest) as a way of imagining and constructing an alternative to speciesism. In doing so, she articulates the importance of care and "tracing the tell," a phrase she uses to invoke deep attention to the worlds of other beings, while acknowledging the designer's position (Westerlaken, 2020).

6.3.2 Situating knowledge

Another important realization occurred as the COVID-19 pandemic forced relocation of the development process of the Rowdy Krause from the lab to my home office and a parallel shift of the field site for testing from a community garden to my home office's balcony. Before the shift, the physical distance of the lab from the field site had enforced another type of distance in the design process — a physical separation between the intended ecosystem and the design space. Here, the soundscape that I was designing was abstract, a series of memories and short recordings.

The shift to the home office meant that work could proceed with the door open, suddenly immersed in the soundscape that the Rowdy Krause would eventually inhabit. This change concretized the work that was underway, rendered the soundscape tangible, and I found myself beginning to notice calls, noises, and vocalizations that had not previously come into focus. Though I had committed to this notion of maintaining aesthetic distance, the sudden immersion in the soundscape allowed me to begin to consider how the Rowdy Krause might resonate with what I was hearing.

This relates to Donna Haraway's (1988) concept of feminist objectivity embodied and situated in a particular context that she calls situated knowledges. The sudden closeness to the ecosystem in question focused my attention on the particularities of that ecosystem as an actor in the relationship that I was helping to form. The soundscape was no longer an abstractable entity that could be objectively engaged from a distance; I could begin to engage with the soundscape as part of the soundscape myself.

This type of situated work — driven in this case by the exigent circumstances of 2020-2021 — may be an essential component in the design of endemic machines. By engaging with the ecosystem through the design process, one can begin to attune to what Bateson (1979) calls an ecological aesthetic, or *the pattern which connects* (Harries-Jones, 2008).

6.3.3 Designing Processes

A consequence of working with artificially intelligent systems is that as a designer, the task is to craft AI's learning and evolutionary process in place of crafting the final result. One can assess the result and then adjust the process, but the focus of the design work is on the system. It matters, then, which processes are chosen as the basis of the design. This fact is perhaps obvious, but the implications become clear in consideration of two of the experiments in this dissertation.

Chapter 4 describes two different methods for evolving the communications strategies of AI agents. The first, tested in simulation in article 3, employs a real biological need — communication — as its incentive. The second, deployed in the Rowdy Krause and discussed in articles 4 to 6, motivated the evolution of sound using the expected behaviour of a species as defined by the acoustic niche hypothesis (ANH).

The simulations were able to readily demonstrate the effects of the ANH while the latter struggled to reproduce some of the anticipated behaviours and required adjustments. Part of this is due to the simplifications built into the simulation system and the added complexity of the Rowdy Krause operating in the real world. However, another plausible explanation is how the processes were designed.

The difference between the two lies in the reward structure that was designed for the evolutionary process. The species in the simulation were instructed to try to communicate; the Rowdy Krause was instructed to find an acoustic niche. In the simulation, the niche arose from the attempt

to communicate, whereas for the Rowdy Krause, what constitutes a niche was pre-defined as part of the design process.

It is worth considering how biological species construct acoustic niches. Their process resembles that of the simulation more than that of the Rowdy Krause. Species are not instructed to find a niche. Rather, their attempts to find a way to communicate lead to a niche's formation.

In the design of endemic machines, the design of the process matters. It was, in the end, possible to adjust the incentive structure to correct the way it defined an acoustic niche and ultimately encourage the Rowdy Krause to find niches within the soundscape. However, it might have been more fruitful to ask it to find a way to communicate with others like itself and allow its own conception of an acoustic niche to emerge from that process.

6.4 Domains of endemism

Much of the inquiry in this work focuses on the sonic domain, but the concept of endemic machines is broader than any one particular sense. There are many possible domains in which a machine could adapt to an ecosystem. As discussed in chapter 2, becoming part of an ecosystem entails engaging with the cycling of energy and materials within the ecosystem.

In this work, that cyclical engagement was achieved through the sensing and production of sound. In particular, the Rowdy Krause was designed to create a feedback cycle between sensed sound and the sound being produced as an explicit part of a conversation with the existing soundscape.

While the Rowdy Krause does not actually communicate with another robot, the work was designed around the impetus of communication. The process by which the Rowdy Krause produced vocalization is based on the ANH, which proposes that a species finds an acoustic niche to facilitate intraspecies communication — a finding supported by *An artificial life approach to studying niche differentiation in soundscape ecology* (article 3).

A more conventional engineering approach might question the notion of machines communicating with one another in the audible spectrum at all. Established electronic communications technologies that use radio waves — Bluetooth, WiFi, ZigBee — are secure, error-correcting, and

thoroughly tested for sending digital information between machines. In truth, these types of technologies would produce more stable and verifiable communication between two or more machines.

However, communication in that way would be inaccessible to other parts of the ecosystem. Other ecosystem actors would have no opportunity to engage with those machines — no way of intercepting, mimicking, responding, or listening to what the machines are saying. These opportunities are essential for forming the types of deep relationships that intertwine a species — or in this case, a machine — in the material and energetic flows of an existing ecosystem.

These opportunities for engagement and entanglement extend to other sensory modalities. Immediately interesting — though with significant technical challenges — is scent. While sound is mainly considered the domain of the animal kingdom³, chemical signals in the form of pheromones cross phylogenetic boundaries and are a nearly ubiquitous mode of biological communication (Jones and Bennett, 2011). The deep ecological and emotive impact of olfaction makes it an interesting target for future work in endemic machines.

Machines engaging with smell — and therefore volatile chemicals — also point to another mode of endemism. While sound is an energetic medium interacting with the ecosystem through advancing pressure waves, smell is physical and the beginning of an engagement with its material flows. Another form of adaptation to a local ecosystem could involve a machine becoming more engaged with the ecosystem's material flows.

The rapidly developing fields of soft robotics (Jørgensen, 2019) and wet artificial life (Swan, 2009; Aguilar et al., 2014) represent opportunities for the integration of biological materials into machines. Common to these fields is a shift away from the rigid metal, plastic, and silicon materiality of most digital devices. This presents an opportunity to incorporate physical material from the ecosystem into the machine. The ability to become part of metabolic and material networks would root machines to a local ecosystem in a substantial and concrete way and would be a new frontier in endemic machines.

³This is a commonly held position, though it is debated in research communities that study other branches of life. Gagliano (2013) argues that buzz pollination in plant-insect interactions constitutes a sense of sound on the part of plants. She also points to studies that indicate that plants produce and respond to acoustic signals. Further studies demonstrate that sound can induce or shift the stress responses of plants, and induce a range of other behaviours (López-Ribera and Vicent, 2017).

6.5 Should machines be endemic?

In the discussion of endemic machines in this dissertation, the question of whether machines *should* be endemic — whether it is good to make machines that adapt to and become part of ecosystems — is not directly considered; the ethics of endemic machines are only addressed briefly. The question can actually be posed as two separate questions: Are endemic machines preferable to non-endemic machines? Should machines be deployed in ecosystems?

These are daunting questions with complex ethical implications. But they need not be answered in the general terms proposed here. One of the fundamental premises of endemic machines is that the details of a specific scenario matter. These questions must be posed in the context of a specific ecosystem and a specific machine.

For example, Van Wynsberghe and Donhauser (2017) describe an autonomous underwater robot called the Lionfish Project that culls lionfish. Lionfish are voracious apex predators that can consume 80% of the fish on a reef in a single month. The Lionfish Project's robots would not be considered endemic machines — they are preprogrammed with the task of killing lionfish and do not adapt or co-evolve with the reef ecosystem — but the intent behind their mission is to preserve the biodiversity of coral reefs.

The ethics of culling a species for conservation purposes is a topic of debate in conservation biology and opinions vary about if, when, and how it should be done (Dubois et al., 2017; Hampton, Warburton, and Sandøe, 2019). Some researchers emphasize the short term improvement of the ecosystem biodiversity (Hampton, Warburton, and Sandøe, 2019), while others maintain that the first priority should be the modification of any human activities that are at the root cause of an acute ecological crisis (Dubois et al., 2017).

If one assumes that the impulse to manage the population of lionfish is ethical, the question shifts to whether it would be preferable for the project to use an endemic machine. An endemic machines-based approach could involve designing a machine to inhabit the ecosystem as an apex predator whose main prey is lionfish. As an artificial apex predator, it would have other roles in the ecosystem in addition to removing lionfish. For example, the endemic machine could be designed to account for the importance of cycling nutrients through the aquatic food chain (Schmitz, Hawlena, and Trussell, 2010; Vanni, 2002); it could

include a process for digesting the captured lionfish and distributing the nutrient material through the ecosystem.

Many conservation and restoration projects provoke questions about the root causes of an ecological problem and whether the proposed action addresses those causes or remedies a symptom. Those cases where ecological intervention is warranted do not always justify the use of a machine. However, where machines are used, there is something to be said for using ones designed to engage — through feedback and co-evolution — with the full complexity of the ecosystem they are trying to help.

Chapter 7

Conclusions

What is required for a machine to adapt to a local ecosystem? Adaptation to an ecosystem requires integration and entanglement with its material and energetic flows. A machine has to sense the system, collect experiences of the relationships therein, and form a response to the streaming energy and matter. It has to shift its behaviour in response to others and shift the behaviour of others in return. That push and pull, call and response shapes a niche, a space for the machine to call its own that is embedded within the larger network of relationships that is the ecosystem.

In the course of arriving at that response, this dissertation has yielded contributions to soundscape ecology, artificial life (ALife), engineering, and design. These include the contributions to the field of soundscape ecology of the BioAcoustic Index Tool (BAIT) — a process and prototype for the collection of soundscape data — and the demonstration of key features of the acoustic niche hypothesis (ANH) in simulation. They also include three conceptual contributions that emerged from the development of the Rowdy Krause.

The concept of **robophony** — a contribution to the field of soundscape ecology and its understanding of the sounds that comprise a soundscape — emerged from considering the biosemiotic role of sounds from an ecologically engaged machine. This addition to the ontology of ecological sounds creates conceptual space for technological entities with hybrid origins and complex networks of influences and relationships.

Reflecting on this type of sound from the ecological perspective shaped the second major contribution in this thesis: **eco-technogenesis**. Eco-technogenesis originated in consideration of sound, but expands to a more general conception of the process of technological and ecological co-

creation. As such it impacts multiple fields including ecology and ALife. The concept of eco-technogenesis suggests that, for a machine to adapt to a local ecosystem, it has to also allow a local ecosystem to adapt to it to initiate a reciprocal process.

That process of co-adaptation, of the machine developing in a way that is unique to a particular time and place, forms the third significant conceptual contribution of this thesis: **endemic machines**. The concept of endemic machines is a way of rooting the practices surrounding the design of locally adaptable machines in the analogy of ecological endemism. It implies a historical contingency and contextual specificity that grounds the machine in the particularities of an ecosystem.

The practices that form the basis of endemic machine design constitute one way of designing a machine to adapt to a local ecosystem. Future work should expand to sensory realms beyond sound, focus more intently on the ethics of endemic machines, and experiment with machines that operate on longer time scales. But the work documented and discussed here — the sensory experiments, simulations of acoustic niche construction, the Rowdy Krause as a working prototype, and the concepts of robophony and eco-technogenesis — lay a solid foundation for continued research into endemic machines.

Bibliography

- Abdallah, Ahmed Chamseddine Ben, Hichem Frigui, and Paul Gader (2009). "Context extraction for local fusion using fuzzy clustering and feature discrimination." In: *2009 IEEE International Conference on Fuzzy Systems*. Jeju Island: IEEE, pp. 490–495. ISBN: 978-1-4244-3596-8. DOI: 10.1109/FUZZY.2009.5277153.
- Aguilar, Wendy et al. (2014). "The Past, Present, and Future of Artificial Life." In: *Frontiers in Robotics and AI* 1, p. 8. ISSN: 2296-9144. DOI: 10.3389/frobt.2014.00008.
- Anderson, Sydney (1994). "Area and Endemism." In: *The Quarterly Review of Biology* 69.4, pp. 451–471. ISSN: 0033-5770. DOI: 10.1086/418743.
- Bateson, Gregory (1972). *Steps to an ecology of mind: Collected essays in anthropology, psychiatry, evolution, and epistemology*. Chicago, IL: University of Chicago Press, p. 542.
- (1979). *Mind and Nature: A Necessary Unity*. New York, USA: E. P. Dutton.
- Benitez, Margarita and Markus Vogl (2020). "Digitally fabricated aesthetic enhancements and enrichments." In: *AI & SOCIETY* 1, p. 3. ISSN: 0951-5666. DOI: 10.1007/s00146-020-00938-8.
- Biggs, Michael and Daniela Buchler (2008). "Eight criteria for practice-based research in the creative and cultural industries." In: *Art, Design & Communication in Higher Education* 7.1, pp. 5–18. ISSN: 1474273X. DOI: 10.1386/adch.7.1.5_1.
- Bradfer-Lawrence, Tom et al. (2019). "Guidelines for the use of acoustic indices in environmental research." In: *Methods in Ecology and Evolution* 10.10. Ed. by Veronica Zamora-Gutierrez, pp. 1796–1807. ISSN: 2041-210X. DOI: 10.1111/2041-210X.13254.
- Browning, E. et al. (2017). *Passive acoustic monitoring in ecology and conservation*. Tech. rep. 2, p. 75. DOI: 10.13140/RG.2.2.18158.46409.

- Candy, Linda and Ernest Edmonds (2018). "Practice-Based Research in the Creative Arts: Foundations and Futures from the Front Line." In: *Leonardo* 51.1, pp. 63–69. ISSN: 0024-094X. DOI: 10.1162/LEON_a_01471.
- Carson, Rachel (1962). *Silent Spring*. New York: Houghton Mifflin Company.
- Cobussen, Marcel (2016). *The Routledge Companion to Sounding Art*. New York, NY ; Abingdon, Oxon : Routledge, 2016. | 2016 |: Routledge. ISBN: 9781315770567. DOI: 10.4324/9781315770567.
- Colavita, Francis B. (1974). "Human sensory dominance." In: *Perception & Psychophysics* 16.2, pp. 409–412. ISSN: 0031-5117. DOI: 10.3758/BF03203962.
- Darwin, Charles (1869). *On the Origin of Species by Means of Natural Selection*. Fifth Edit. London: John Murray, p. 596.
- De Candolle, Augustin Pyramus (1820). "Geographie botanique." In: *Dictionnaire des sciences naturelles*. Ed. by F. G. Leverault. Vol. 18. Paris: Le Normant, p. 412.
- Donhauser, Justin, Aimee van Wynsberghe, and Alexander Bearden (2020). "Steps Toward an Ethics of Environmental Robotics." In: *Philosophy & Technology*, pp. 1–18. ISSN: 2210-5433. DOI: 10.1007/s13347-020-00399-3.
- Dubois, Sara et al. (2017). "International consensus principles for ethical wildlife control." In: *Conservation Biology* 31.4, pp. 753–760. ISSN: 08888892. DOI: 10.1111/cobi.12896.
- Dunn, David D. (2002). *Four Electroacoustic Compositions*.
- Emmeche, Claus (2001). "Does a robot have an Umwelt? Reflections on the qualitative biosemiotics of Jakob von Uexküll." In: *Semiotica* 2001.134, pp. 653–693. ISSN: 0037-1998. DOI: 10.1515/semi.2001.048.
- Eze, Emmanuel Chukwudi (1997). "The Color of Reason: The Idea of "Race" in Kant's Anthropology." In: *Postcolonial African Philosophy: A Critical Reader*. Ed. by Emmanuel Chukwudi Eze. Cambridge, MA, USA: Blackwell Publishers. Chap. 4, p. 374. ISBN: 9780631203407.
- Farina, Almo (2014). *Soundscape Ecology*. Dordrecht: Springer Netherlands. ISBN: 978-94-007-7373-8. DOI: 10.1007/978-94-007-7374-5.
- (2018). "Perspectives in ecoacoustics: A contribution to defining a discipline." In: *Journal of Ecoacoustics* 2.2, pp. 1–1. ISSN: 2516-1466. DOI: 10.22261/jea.trzd5i.
- Farina, Almo, Alice Eldridge, and Peng Li (2021). "Ecoacoustics and Multispecies Semiosis: Naming, Semantics, Semiotic Characteristics, and Competencies." In: *Biosemiotics 2021*, pp. 1–25. DOI: 10.1007/s12304-021-09402-6.

- Farina, Almo et al. (2016). "The Application of the Acoustic Complexity Indices (ACI) to Ecoacoustic Event Detection and Identification (EEDI) Modeling." In: *Biosemiotics* 9.2, pp. 227–246. ISSN: 1875-1342. DOI: 10.1007/s12304-016-9266-3.
- Fitch, W. Tecumseh (2015). "Four principles of bio-musicology." In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 370.1664, p. 20140091. ISSN: 0962-8436. DOI: 10.1098/rstb.2014.0091.
- Gagliano, Monica (2013). "The flowering of plant bioacoustics: how and why." In: *Behavioral Ecology* 24.4, pp. 800–801. ISSN: 1045-2249. DOI: 10.1093/beheco/art021.
- Goudbeek, Martijn et al. (2005). "Acquiring Auditory and Phonetic Categories." In: *Handbook of Categorization in Cognitive Science*. Elsevier, pp. 497–513. ISBN: 9780080446127. DOI: 10.1016/B978-008044612-7/50077-9.
- Hampton, Jordan O., Bruce Warburton, and Peter Sandøe (2019). "Compassionate versus consequentialist conservation." In: *Conservation Biology* 33.4, pp. 751–759. ISSN: 0888-8892. DOI: 10.1111/cobi.13249.
- Haraway, Donna (1988). "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." In: *Feminist Studies* 14.3, p. 575. ISSN: 00463663. DOI: 10.2307/3178066.
- (2019). "It Matters What Stories Tell Stories; It Matters Whose Stories Tell Stories." In: *a/b: Auto/Biography Studies* 34.3, pp. 565–575. ISSN: 0898-9575. DOI: 10.1080/08989575.2019.1664163.
- Hardus, Madeleine E. et al. (2009). "Tool use in wild orang-utans modifies sound production: a functionally deceptive innovation?" In: *Proceedings of the Royal Society B: Biological Sciences* 276.1673, pp. 3689–3694. ISSN: 0962-8452. DOI: 10.1098/rspb.2009.1027.
- Harries-Jones, Peter (2008). "Gregory Bateson's "Uncovery" Of Ecological Aesthetics." In: *A Legacy for Living Systems*. Ed. by Jesper Hoffmeyer. Springer Netherlands. Chap. 10, pp. 153–167. DOI: 10.1007/978-1-4020-6706-8_11.
- Hayles, N. Katherine (2012). *How we think*. Chicago: The University of Chicago Press. ISBN: 9780226321400.
- Hobbs, Richard J., Eric Higgs, and Carol M. Hall (2013). *Novel Ecosystems*. Ed. by Richard J. Hobbs, Eric S. Higgs, and Carol M. Hall. Chichester, UK: John Wiley & Sons, Ltd, p. 368. ISBN: 9781118354186. DOI: 10.1002/9781118354186.
- Hobbs, Richard J., Eric Higgs, and James A. Harris (2009). "Novel ecosystems: implications for conservation and restoration." In: *Trends*

- in Ecology & Evolution* 24.11, pp. 599–605. ISSN: 01695347. DOI: 10.1016/j.tree.2009.05.012.
- Hofstetter, Richard W. et al. (2014). “Using acoustic technology to reduce bark beetle reproduction.” In: *Pest Management Science* 70.1, pp. 24–27. ISSN: 1526498X. DOI: 10.1002/ps.3656.
- Hutmacher, Fabian (2019). “Why Is There So Much More Research on Vision Than on Any Other Sensory Modality?” In: *Frontiers in Psychology* 10, p. 2246. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2019.02246.
- Ingram, Ian (2010). *The Woodiest*. Tech. rep. www.ianingram.org.
- Jebb, Andrew T., Scott Parrigon, and Sang Eun Woo (2017). “Exploratory data analysis as a foundation of inductive research.” In: *Human Resource Management Review* 27.2, pp. 265–276. ISSN: 10534822. DOI: 10.1016/j.hrmr.2016.08.003.
- Jones, Stephen K. and Richard J. Bennett (2011). “Fungal mating pheromones: Choreographing the dating game.” In: *Fungal Genetics and Biology* 48.7, pp. 668–676. ISSN: 10871845. DOI: 10.1016/j.fgb.2011.04.001.
- Jørgensen, Jonas (2019). “Constructing Soft Robot Aesthetics : Art, Sensation, and Materiality in Practice.” PhD Thesis. IT University of Copenhagen, p. 219.
- Kadish, David (2016). “Autonomous Agents as Artistic Collaborators.” In: *The 30th Annual Conference of Society of Science, Literature and the Arts*. Atlanta, Georgia.
- (2017). “On Scale and Fields: Artistic Practice and Agricultural Machines.” In: *RE:TRACE – 7th International Conference for Histories of Media Art, Science and Technology*. Vienna, Austria.
- (2018). “Generating Diversity: Art, robots, and the future of farming.” In: *Intersections: Proceedings of the 24th International Symposium on Electronic Art*. Ed. by Rufus Adebayo et al. Durban, SA, pp. 86–92.
- Kadish, David and Laura Beloff (2017). “Hybrid hangouts : Embedded values in hybrid ecologies of plants, people and robots.” In: *First Global Radical Relevances Conference*. Espoo, Finland.
- Kadish, David and Sebastian Risi (2020). “Adapting to a changing environment: Simulating the effects of noise on animal sonification.” In: *The 2020 Conference on Artificial Life*. Cambridge, MA: MIT Press, pp. 687–695. DOI: 10.1162/isal_a_00320.
- Kadish, David, Sebastian Risi, and Anders Sundnes Løvlie (2021). “Improving Object Detection in Art Images Using Only Style Transfer [ACCEPTED].” In: *2021 International Joint Conference on Neural Networks (IJCNN)*. IEEE. arXiv: 2102.06529.

- Kasten, Eric P. et al. (2012). "The remote environmental assessment laboratory's acoustic library: An archive for studying soundscape ecology." In: *Ecological Informatics* 12, pp. 50–67. ISSN: 15749541. DOI: 10.1016/j.ecoinf.2012.08.001.
- Kell, Alexander J.E. et al. (2018). "A Task-Optimized Neural Network Replicates Human Auditory Behavior, Predicts Brain Responses, and Reveals a Cortical Processing Hierarchy." In: *Neuron* 98.3, 630–644.e16. ISSN: 08966273. DOI: 10.1016/j.neuron.2018.03.044.
- Krause, Bernie L. (1987). "Bioacoustics, Habitat Ambience in Ecological Balance." In: *Whole Earth Review* 57, pp. 14–18. ISSN: 0749-5056.
- (2008). "Anatomy of the soundscape: Evolving perspectives." In: *AES: Journal of the Audio Engineering Society* 56.1-2, pp. 73–80. ISSN: 15494950.
- Krause, Bernie L. and Stuart Gage (2003). *Testing Biophony as an Indicator of Habitat Fitness and Dynamics*. Tech. rep. Clen Ellen, CA: Wild Sanctuary Inc, p. 18.
- Krutzen, M. et al. (2005). "Cultural transmission of tool use in bottlenose dolphins." In: *Proceedings of the National Academy of Sciences* 102.25, pp. 8939–8943. ISSN: 0027-8424. DOI: 10.1073/pnas.0500232102.
- Lehman, Joel (2020). "Evolutionary Computation and AI Safety." In: Springer, Cham, pp. 181–200. DOI: 10.1007/978-3-030-39958-0_10.
- Liggett, Susan (2020). "Positioning the Arts in the Research Process: Perspectives from Higher Education." In: *Springer Series on Cultural Computing*. Springer, pp. 9–21. DOI: 10.1007/978-3-030-42097-0_2.
- López-Ribera, Ignacio and Carlos M. Vicent (2017). "Drought tolerance induced by sound in Arabidopsis plants." In: *Plant Signaling & Behavior* 12.10, e1368938. ISSN: 1559-2324. DOI: 10.1080/15592324.2017.1368938.
- Luther, David A. and Elizabeth P. Derryberry (2012). "Birdsongs keep pace with city life: changes in song over time in an urban songbird affects communication." In: *Animal Behaviour* 83.4, pp. 1059–1066. ISSN: 00033472. DOI: 10.1016/j.anbehav.2012.01.034.
- Marris, Emma, Joseph Mascaro, and Erle C. Ellis (2013). "Perspective: Is Everything a Novel Ecosystem? If so, do we need the Concept?" In: *Novel Ecosystems*. Chichester, UK: John Wiley & Sons, Ltd, pp. 345–349. DOI: 10.1002/9781118354186.ch41.
- McKenna, Megan F. (2020). "The sounds around us." In: *Physics Today* 73.1, pp. 28–34. ISSN: 0031-9228. DOI: 10.1063/PT.3.4387.

- Morrone, J.J. (2008). "Endemism." In: *Encyclopedia of Ecology*. Elsevier, pp. 1254–1259. ISBN: 9780080914565. DOI: 10.1016/B978-008045405-4.00786-2.
- Odum, Eugene P. (1971). *Fundamentals of ecology*. 3rd. Philadelphia: W. B. Saunders Company.
- Pieretti, Nadia, Almo Farina, and D. Morri (2011). "A new methodology to infer the singing activity of an avian community: The Acoustic Complexity Index (ACI)." In: *Ecological Indicators* 11.3, pp. 868–873. ISSN: 1470160X. DOI: 10.1016/j.ecolind.2010.11.005.
- Pijanowski, Bryan C. and Almo Farina (2011). "Introduction to the special issue on soundscape ecology." In: *Landscape Ecology* 26.9, pp. 1209–1211. ISSN: 0921-2973. DOI: 10.1007/s10980-011-9655-6.
- Pijanowski, Bryan C. et al. (2011a). "Soundscape Ecology: The Science of Sound in the Landscape." In: *BioScience* 61.3, pp. 203–216. ISSN: 1525-3244. DOI: 10.1525/bio.2011.61.3.6.
- Pijanowski, Bryan C. et al. (2011b). "What is soundscape ecology? An introduction and overview of an emerging new science." In: *Landscape Ecology* 26.9, pp. 1213–1232. ISSN: 0921-2973. DOI: 10.1007/s10980-011-9600-8.
- Prévot-Julliard, Anne-Caroline et al. (2011). "The Need for Flexibility in Conservation Practices: Exotic Species as an Example." In: *Environmental Management* 47.3, pp. 315–321. ISSN: 0364-152X. DOI: 10.1007/s00267-011-9615-6.
- Rodriguez, Laura F. (2006). "Can Invasive Species Facilitate Native Species? Evidence of How, When, and Why These Impacts Occur." In: *Biological Invasions* 8.4, pp. 927–939. ISSN: 1387-3547. DOI: 10.1007/s10530-005-5103-3.
- Salkind, Neil (2010). *Encyclopedia of Research Design*. 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc. ISBN: 9781412961271. DOI: 10.4135/9781412961288.
- Schafer, R. Murray (1977). *The tuning of the world*. A.A. Knopf, p. 301. ISBN: 0394409663.
- Schlaepfer, Martin A., Dov F. Sax, and Julian D. Olden (2011). "The Potential Conservation Value of Non-Native Species." In: *Conservation Biology* 25.3, pp. 428–437. ISSN: 08888892. DOI: 10.1111/j.1523-1739.2010.01646.x.
- Schmitz, Oswald J., Dror Hawlena, and Geoffrey C. Trussell (2010). "Predator control of ecosystem nutrient dynamics." In: *Ecology Letters* 13.10, pp. 1199–1209. ISSN: 1461023X. DOI: 10.1111/j.1461-0248.2010.01511.x.

- Seamster, Louise and Victor Ray (2018). "Against Teleology in the Study of Race: Toward the Abolition of the Progress Paradigm." In: *Sociological Theory* 36.4, pp. 315–342. ISSN: 0735-2751. DOI: 10.1177/0735275118813614.
- Stiegler, Bernard. (1998). *Technics and Time, 1: The Fault of Epimetheus*. Stanford University Press. ISBN: 9780804730419.
- Sueur, Jérôme and Almo Farina (2015). "Ecoacoustics: the Ecological Investigation and Interpretation of Environmental Sound." In: *Biosemiotics* 8.3, pp. 493–502. ISSN: 1875-1342. DOI: 10.1007/s12304-015-9248-x.
- Sullins, John P. (2011). "Introduction: Open Questions in Roboethics." In: *Philosophy & Technology* 24.3, pp. 233–238. ISSN: 2210-5433. DOI: 10.1007/s13347-011-0043-6.
- Swan, Liz Stillwaggon (2009). "Synthesizing insight: artificial life as thought experimentation in biology." In: *Biology & Philosophy* 24.5, pp. 687–701. ISSN: 0169-3867. DOI: 10.1007/s10539-009-9156-z.
- Towsey, Michael et al. (2014). "The use of acoustic indices to determine avian species richness in audio-recordings of the environment." In: *Ecological Informatics* 21, pp. 110–119. ISSN: 15749541. DOI: 10.1016/j.ecoinf.2013.11.007.
- Truax, Barry (1978). *The World Soundscape Project's Handbook for acoustic ecology*. 1st ed. Music of the environment series. A.R.C. Publications, p. 171. ISBN: 9780889850118.
- Vanni, Michael J (2002). "Nutrient Cycling by Animals in Freshwater Ecosystems." In: *Annual Review of Ecology and Systematics* 33.1, pp. 341–370. ISSN: 0066-4162. DOI: 10.1146/annurev.ecolsys.33.010802.150519.
- Villanueva-Rivera, Luis J. et al. (2011). "A primer of acoustic analysis for landscape ecologists." In: *Landscape Ecology* 26.9, pp. 1233–1246. ISSN: 0921-2973. DOI: 10.1007/s10980-011-9636-9.
- Vitols, Rihards (2016). *Woodpecker*. Tech. rep. vitols.xyz.
- Wallace, Mark N. (2010). *Auditory System: Structure*. Ed. by E B Goldstein. SAGE Publications, pp. 194–197. ISBN: 9781412940818.
- Westerkamp, Hildegard (1974). "Soundwalking." In: *Sound Heritage* 3.4.
- Westerlaken, Michelle (2020). "Imagining Multispecies Worlds." PhD thesis. Malmö University, p. 93. ISBN: 9789178771059. DOI: 10.24834/isbn.9789178771059.
- Whiten, Andrew and Carel P van Schaik (2007). "The evolution of animal 'cultures' and social intelligence." In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 362.1480, pp. 603–620. ISSN: 0962-8436. DOI: 10.1098/rstb.2006.1998.

- Wynsberghe, Aimee van and Justin Donhauser (2017). "The Dawning of the Ethics of Environmental Robots." In: *Science and Engineering Ethics*. ISSN: 1353-3452. DOI: 10.1007/s11948-017-9990-3.
- Young, Mary Lynn and Candis Callison (2021). "When gender, colonialism, and technology matter in a journalism startup." In: *Journalism* 22.1, pp. 3–19. ISSN: 1464-8849. DOI: 10.1177/1464884917743390.
- Zawawi, Siti Aisyah et al. (2020). "A Review of MEMS Capacitive Microphones." In: *Micromachines* 11.5, p. 484. ISSN: 2072-666X. DOI: 10.3390/mi11050484.

Appendix A

Additional Work

Some of the work that was completed during the period of the Ph.D. program was not included in the thesis. That work is listed here.

Conference Proceedings

David Kadish, Sebastian Risi, and Anders Sundnes Løvlie (2021). "Improving Object Detection in Art Images Using Only Style Transfer [ACCEPTED]." in: *2021 International Joint Conference on Neural Networks (IJCNN)*. IEEE. arXiv: 2102.06529

David Kadish and Sebastian Risi (2020). "Adapting to a changing environment: Simulating the effects of noise on animal sonification." In: *The 2020 Conference on Artificial Life*. Cambridge, MA: MIT Press, pp. 687–695. DOI: 10.1162/isal_a_00320

David Kadish (2018). "Generating Diversity: Art, robots, and the future of farming." In: *Intersections: Proceedings of the 24th International Symposium on Electronic Art*. Ed. by Rufus Adebayo et al. Durban, SA, pp. 86–92

Conference Presentations

David Kadish and Laura Beloff (2017). "Hybrid hangouts : Embedded values in hybrid ecologies of plants, people and robots." In: *First Global Radical Relevances Conference*. Espoo, Finland

David Kadish (2016). "Autonomous Agents as Artistic Collaborators." In: *The 30th Annual Conference of Society of Science, Literature and the Arts*. Atlanta, Georgia

Conference Posters

David Kadish (2017). "On Scale and Fields: Artistic Practice and Agricultural Machines." In: *RE:TRACE – 7th International Conference for Histories of Media Art, Science and Technology*. Vienna, Austria

Articles

Article 1

Clustering sensory inputs using NeuroEvolution of Augmenting Topologies

Originally published in: GECCO '18: Proceedings of the Genetic and Evolutionary Computation
Conference Companion

Clustering sensory inputs using NeuroEvolution of Augmenting Topologies

David Kadish

IT University of Copenhagen

Copenhagen S, Denmark

École polytechnique fédérale de Lausanne

Lausanne, Switzerland

davk@itu.dk

ABSTRACT

Sorting data into groups and clusters is one of the fundamental tasks of artificially intelligent systems. Classical clustering algorithms rely on heuristic (k-nearest neighbours) or statistical methods (k-means, fuzzy c-means) to derive clusters and these have performed well. Neural networks have also been used in clustering data, but researchers have only recently begun to adopt the strategy of having neural networks directly determine the cluster membership of an input datum. This paper presents a novel strategy, employing NeuroEvolution of Augmenting Topologies to produce an evolutionary neural network capable of directly clustering unlabelled inputs. It establishes the use of cluster validity metrics in a fitness function to train the neural network.

CCS CONCEPTS

• **Information systems** → **Clustering**; • **Computing methodologies** → **Knowledge representation and reasoning**; *Neural networks*; • **Applied computing** → **Engineering**; **Agriculture**;

KEYWORDS

clustering, NEAT, Calinski-Harabaz, k-means

ACM Reference Format:

David Kadish. 2018. Clustering sensory inputs using NeuroEvolution of Augmenting Topologies. In *GECCO '18 Companion: Genetic and Evolutionary Computation Conference Companion, July 15–19, 2018, Kyoto, Japan*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3205651.3205771>

1 INTRODUCTION

Grouping similar data and experiences is a fundamental building block of learning. Without the need for *a priori* information about the meaning of a particular input, clustering forms the basis for generating meaningful categorizations [3]. Given its foundational role in learning, it is perhaps surprising that few efforts have used neural networks to perform clustering and none could be found that use neuroevolution. This paper develops a technique called NEAT Clustering (NEAT-CLU) for clustering using neuroevolution.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

GECCO '18 Companion, July 15–19, 2018, Kyoto, Japan

© 2018 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-5764-7/18/07.

<https://doi.org/10.1145/3205651.3205771>

2 BACKGROUND

Clustering refers to the process of assembling unlabeled data into like groups. A number of clustering strategies are well-established: k-nearest neighbours (k-NN), k-means, fuzzy c-means (FCM), DB-SCAN, and self organizing maps (SOM) are standard tools for cluster analysis [2]. Of these, only SOMs use neural networks (NNs), generating a neural map that is overlaid on the input data. These methods share an underlying model of input data plotted on a hyperplane and the use of a distance measurement to assign a cluster.

Two existing efforts use NNs to bypass the use of a distance metric in cluster assignments and assign cluster membership as a direct output of the NN [5, 6]. They both require that data be presented in pairs with pre-determined measurements of similarity and neither employ evolutionary methods to form their networks.

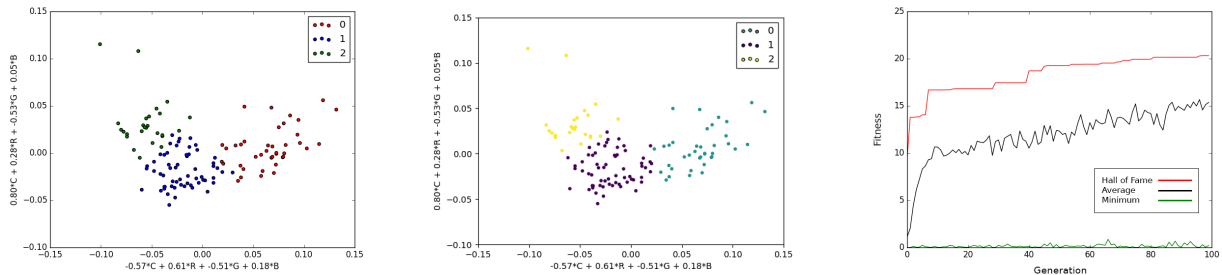
The evolutionary NN employed in this paper is known as neuroevolution of augmenting topologies (NEAT) [7]. NEAT evolves the structure and weights of its neurons in tandem, adding layers and complexity as necessary to achieve an optimal fitness. NEAT and its derivatives have been applied to a wide range of tasks but this is its first application to cluster assignment.

3 ALGORITHM

The NEAT algorithm itself is well-documented [7] and is used in its standard form in NEAT-CLU, so it is not covered in detail here. NEAT-CLU uses the raw input data, so there is no need to preprocess samples. For k clusters, the output is encoded using $\lfloor \log_2(k) \rfloor + 1$ output neurons with an unsigned step activation function. The full NN output is treated as a binary number representing the assigned cluster. When the number of binary combinations does not match the desired number of clusters, some clusters are assigned multiple binary numbers. For example, in this trial, the outputs 01 and 10 both map to the second cluster.

One of the key insights in NEAT-CLU is that clustering metrics can be a key component of the fitness function for training an evolutionary NN. The Calinski-Harabaz (CH) score is a measure of the comparison between intra-cluster variability and inter-cluster variability [1]. For a sample set of N observations, divided into k clusters with the centroid of cluster i at m_i , the CH score can be written as $\mathcal{F}_{ch} = \frac{S_B}{S_W} \times \frac{N-k}{k-1}$ where the intercluster variance S_B and intracluster variance S_W are $S_B = \sum_{i=1}^k n_i \|m_i - m\|^2$ and $S_W = \sum_{i=1}^k \sum_{x \in c_i} \|x - m_i\|^2$.

Two metrics are added to the fitness function to encourage even clustering. A demerit (\mathcal{F}_k) is assessed if fewer clusters are created than desired. Another (\mathcal{F}_n) penalizes disparity in group size. The



(a) Clusters created by the NEAT clusterer. Axes represent the two components of a PCA decomposition of the 4-dimensional input space. (b) Clusters created by a K-means clusterer. Axes represent the two components of a PCA decomposition of the 4-dimensional input space. (c) Fitness over generations in the evolutionary process that produced the clustering network.

Figure 1: Side-by-side comparison of NEAT-CLU and k-means clustering algorithms (a-b). Generational fitness (c).

Table 1: Cluster metric scores

	Calinski-Harabaz	Silhouette
NEAT-CLU	103.31	0.36
k-means	108.23	0.38

resulting fitness function can be written as $\mathcal{F} = w_{ch}\mathcal{F}_{ch} - w_k\mathcal{F}_k - w_n\mathcal{F}_n$ where w denotes a weighting constant, the demerit for deviation from the desired number of clusters (K) is $\mathcal{F}_k = K - k$ and a measure of cluster size disparity is $\mathcal{F}_n = k - \sum_{i=1}^k \frac{n_i}{\max(n)}$.

4 METHOD

This work formed part of an inquiry about how a robot could experience an ecosystem, so the experiment focused on the clustering of plant leaves. The robot was equipped with a single-pixel camera with which it R-G-B and white reflectance from different leaves. The NEAT-CLU clustering algorithm was trained on these samples and then used to sort new samples into three distinct groups. This result was compared to a k-means clustering of the same data.

5 RESULTS

Using the NEAT-CLU algorithm, the robot sorted the leaves into three clusters. The NEAT network was evolved over 100 generations with a population of 100 individuals (figure 1c). The results of NEAT-CLU are shown (figure 1a) beside the results of k-means clustering (figure 1b). NEAT-CLU and k-means produce similar results with minor differences at the cluster boundaries. NEAT-CLU has effectively learned to closely emulate a k-means clustering strategy. The cluster scores for the two methods are quite close (Table 1) though k-means fares slightly better in both evaluated metrics.

6 DISCUSSION

The slight underperformance of NEAT-CLU and its increased complexity with respect to k-means suggest that NEAT-CLU will not replace the standard clustering tools. However, NEAT-CLU can offer a degree of flexibility that is unavailable to k-means. The structure

of clusters in k-means — as well other standard clustering methods — stems from the algorithm’s clustering mechanism. K-means performs well on gaussian-distributed datasets with equal-sized clusters, but often performs poorly on data that is distributed in other ways [4]. NEAT-CLU’s clusters are structured by the fitness function, which can be easily modified to suit new datasets.

The CH portion of the fitness function could be replaced by another clustering metric, such as the S_Dbw cluster validity index [4]. This would allow NEAT-CLU to adapt readily to many differently structured datasets. This flexibility is the true advantage of performing clustering using neuroevolution. The same algorithm can be adjusted — even during the evolutionary process — to fit a wide variety of different datasets and end-goals.

ACKNOWLEDGMENTS

Thanks to Stefania Santagati, Per Nagbøl, Astrid Petitjean, Sebastian Risi, Thomas Bolander, Laura Beloff and Kasper Støy for their help with this work, and the anonymous reviewers for their feedback.

REFERENCES

- [1] T. Calinski and J. Harabasz. 1974. A dendrite method for cluster analysis. *Communications in Statistics - Theory and Methods* 3, 1 (1974), 1–27. <https://doi.org/10.1080/03610927408827101>
- [2] Adil Fahad, Najlaa Alshatri, Zahir Tari, Abdullah Alamri, Ibrahim Khalil, Albert Y. Zomaya, Sebt Fofou, and Abdelaziz Bouras. 2014. A Survey of Clustering Algorithms for Big Data: Taxonomy and Empirical Analysis. *IEEE Transactions on Emerging Topics in Computing* 2, 3 (sep 2014), 267–279. <https://doi.org/10.1109/TETC.2014.2330519>
- [3] Martijn Goudbeek, Roel Smits, Anne Cutler, and Daniel Swingley. 2005. Acquiring Auditory and Phonetic Categories. In *Handbook of Categorization in Cognitive Science*. Elsevier, 497–513. <https://doi.org/10.1016/B978-008044612-7/50077-9>
- [4] M. Halkidi and M. Vazirgiannis. 2011. Clustering validity assessment: finding the optimal partitioning of a data set. In *Proceedings 2001 IEEE International Conference on Data Mining*. IEEE Comput. Soc, 187–194. <https://doi.org/10.1109/ICDM.2001.989517>
- [5] Yen-Chang Hsu and Zsolt Kira. 2016. Neural network-based clustering using pairwise constraints. In *4th International Conference on Learning Representations*. San Juan, Puerto Rico, 1–12. arXiv:1511.06321 <http://arxiv.org/abs/1511.06321>
- [6] Federico Raue, Sebastian Palacio, Andreas Dengel, and Marcus Liwicki. 2017. Classless Association Using Neural Networks. In *Artificial Neural Networks and Machine Learning – ICANN 2017: 26th International Conference on Artificial Neural Networks, Alghero, Italy, September 11–14, 2017, Proceedings, Part II*. Springer, Cham, 165–173. https://doi.org/10.1007/978-3-319-68612-7_19
- [7] Kenneth O. Stanley and Risto Miikkulainen. 2002. Evolving Neural Networks through Augmenting Topologies. *Evolutionary Computation* 10, 2 (jun 2002), 99–127. <https://doi.org/10.1162/10635602320169811>

Article 2

The BioAcoustic Index Tool

Joint work with: Kasper Stoy

Status Accepted at *Bioacoustics: The International Journal of Animal Sound and its Recording*

BioAcoustic Index Tool: Long-term biodiversity monitoring using on-sensor acoustic index calculations

David Kadish and Kasper Stoy

IT University of Copenhagen, Copenhagen, Denmark

Abstract

Acoustic indices are valuable tools for measuring and tracking changes in biodiversity. However, the method used to collect acoustic index data can be made more effective by recent developments in electronics. The current process requires recording high-quality audio in the field and computing acoustic indices in the lab. This produces vast quantities of raw audio data, which limits the time that sensors can spend in the field and complicates data processing and analysis. Additionally, most field audio recorders are unable to log the full range of contextual environmental data that would help explain short-term variations. In this paper, we present the BioAcoustic Index Tool, a smart acoustic index and environmental sensor. The BioAcoustic Index Tool computes acoustic indices as audio is captured, storing only the index information, and logs temperature, humidity, and light levels. The sensor was able to operate completely autonomously for the entire five-month duration of the field study. In that time, it recorded over 4000 measurements of acoustic complexity and diversity all while producing the same amount of data that would be used to record 3 minutes of raw audio. These factors make the BioAcoustic Index Tool well-suited for large-scale, long-term acoustic biodiversity monitoring.

Abbreviations: **ACI** Acoustic Complexity Index; **ADC** analog-to-digital converter; **ADI** Acoustic Diversity Index; **AEI** Acoustic Evenness Index; **BAIT** the BioAcoustic Index Tool; **BI** Bioacoustic Index; **FFT** fast Fourier transform; **FPU** floating point unit; **H** Acoustic Entropy Index; **H_f** Spectral Entropy; **I²S** Integrated Inter-IC Sound Bus; **KDE** kernel density estimate; **MEMS** microelectromechanical systems; **NDSI** Normalised Difference Soundscape Index; **RTC** real-time clock; **SET** Soundscape Explorer (terrestrial).

KEYWORDS

acoustic index, soundscape ecology, biodiversity, sensor, field recorder, open source

1. Introduction

The soundscape is a rich source of information about the state and health of an ecosystem. In particular, the calculation of acoustic indices from audio recordings of a soundscape provides stable quantitative metrics to monitor ecosystems for disturbances and changes in biodiversity. However, despite rapid advances in portable electronic technologies, the methods of calculating acoustic indices have not changed since the introduction of the concept by Boelman et al. (2007) and Sueur et al. (2008b).

Taking advantage of high-performance and low power consumption of modern microcontrollers, we have developed a tool to streamline the computation of acoustic indices, enable longer-term field studies, and add context to acoustic data by also recording environmental conditions. The BioAcoustic Index Tool (BAIT) skirts the

data storage limitations of conventional audio field recorders by calculating acoustic indices in the field and storing only the index data, making it possible to leave a sensor in the field for months or even years at a time. At the same time, we have leveraged the flexibility of the onboard microcontroller to add additional sensors for light, temperature, and humidity. This means that acoustic index data can be correlated to the environmental conditions immediately surrounding the sensor. The ability to continuously monitor a soundscape over a period of months and years and to de-correlate a flexible range of environmental conditions from changes in the sonic environment has the potential to vastly improve the long-term monitoring of biodiversity using sound.

The use of sound as an indicator of biodiversity dates back at least to the publication of *Silent Spring* by Carson in 1962. As the concept of the soundscape was formalized by researchers such as Westerkamp (1974), Schafer (1977), and Truax (1978), soundscapes were recorded in analogue formats using portable stereo reel-to-reel recorders (Lyonblum 2017). By the mid-2000s digital recording and computing technologies had advanced to the point that larger-scale digital analysis of audio was possible.

These developments enabled the creation and use of the first acoustic indices for conducting quantitative analyses of biodiversity using sound (Boelman et al. 2007; Sueur et al. 2008b; Villanueva-Rivera et al. 2011) and the practice has expanded rapidly since then (Bradfer-Lawrence et al. 2019). These methods have the benefit of being able to estimate biodiversity without requiring the types of intensive surveys that have been traditionally used for assessing biodiversity (Sueur et al. 2008b; Buxton et al. 2018). The use of soundscape recordings and acoustic indices can help minimize disturbance of sensitive landscapes and reduce the cost of performing surveys while providing long-term data for assessing ecosystems.

The methods for collecting long-term soundscape data have changed little since the early days of soundscape recording. Typically, raw audio is recorded using purpose-built field audio recorders that remain in the field, untended, for anywhere from a few days to a few months (Pijanowski et al. 2011b; Pieretti et al. 2015; Gottesman et al. 2020). Commercial field recorders — such as the Song Meter from *Wildlife Acoustics* — are often used in these projects, but new, open-source tools such as AURITA (Beason et al. 2019) and the AudioMoth (Hill et al. 2019) that use electronics from the do-it-yourself (DIY) community are becoming more common as well.

Recordings are stored as high-quality audio files and collected at the end of the recording period for analysis in the lab. In the lab, recordings are preprocessed — this can include pre-filtering some audio frequencies (Towsey et al. 2014; Farina et al. 2021), removing noisy recordings (Righini and Pavan 2020), or subsampling the data in various ways (Towsey et al. 2014; Righini and Pavan 2020; Farina et al. 2021) — and then acoustic index calculations are performed.

This established method works well to capture acoustic index data, but has several drawbacks. An oft-mentioned difficulty for researchers is the sheer volume of data produced (Righini and Pavan 2020). Raw audio files require lots of storage (Bradfer-Lawrence et al. 2019) and significant data processing facilities (Towsey et al. 2014; Farina et al. 2021).

A single 15-minute uncompressed audio file, recorded at the 44.1 kHz in 16-bit stereo — a typical configuration for soundscape recordings — requires about 150MB of storage. Even with some of the largest (512GB) SD cards available, this means that a recorder is limited to about 850 hours (35 days) of continuous recording before the

Table 1.: Uses cases for different field recorders and sensors

To ...	Song Meter	AURITA	AudioMoth	SET	BAIT
record high-quality audio files for review and analysis in the lab	•	•	•	•	
log environmental data alongside acoustic indices				•	•
capture ultrasonic frequencies	•	•	•	•	
access pre-computed acoustic indices				•	•
be able to modify or hack your recorder		•	•		•
deploy sensors remotely for more than a few months					•
spend under \$100 (USD) per sensor			•		
spend under \$200 (USD) per sensor			•		•
spend under \$400 (USD) per sensor		•	•		•
spend under \$1000 (USD) per sensor	•	•	•	•	•

data must be collected¹. Researchers have also noted that archiving and processing all of that data presents its own set of challenges (Righini and Pavan 2020) and many authors describe trade-offs between the quality and depth of data they record and the storage and processing limitations they face.

The recorders that are currently used also lack a certain flexibility. They perform their assigned role of recording audio well but are limited to that particular task. Meanwhile, researchers have called for further integration of environmental data into soundscape studies to provide context for recordings and acoustic index measurements (Pijanowski et al. 2011b; Righini and Pavan 2020).

Tools like the aforementioned AURITA and the AudioMoth, as well as others like the Solo recorder (Whytock and Christie 2017), point to a way forward. These devices rely on recent developments in electronics and battery technology as well as the emergence of a vibrant DIY hardware community to create relatively cheap but powerful devices for recording the soundscape. These particular tools don't change how acoustic index data is collected, but related advances in DIY electronics have made powerful and efficient processors and sensors available to a wide community.

In particular, more microcontrollers — the small, embedded computers that power some of these sensors — are now capable of efficiently computing fast Fourier transforms (FFTs), the basis for producing spectrograms and many of the most popular acoustic indices. This enables them to perform the first stage of computation and analysis in the field, as the data is collected, instead of waiting to return to the lab with raw data. These microcontrollers have the additional benefit of being programmable — and therefore flexible in their operation — and can often connect to many peripherals including additional sensing equipment and devices for data storage and wireless communication.

There has been interest in deploying these types of technologies for ecological research. Guo et al. (2015) proposed that these types of smart sensors can improve ecological data collection by enabling continuous data acquisition and long-term operation in the field. Browning et al. (2017) and Greif and Yovel (2019) point out that onboard processing of raw data would dramatically reduce the amount of data that needs to be stored.

In the sonic domain, researchers have begun to test the use of on-sensor

¹This can be extended with recorders that feature multiple SD card slots, but one will still encounter limitations for long-term monitoring with this approach

analysis of sound. One such effort analyses acoustic data to track the grazing of cows (Deniz et al. 2017). Another uses deep learning neural networks to detect the vocalizations of different species urban bats (Balestrini et al. 2020). Other work has taken place in aquatic environments, where sound is the primary mode of communication (Baumgartner et al. 2013).

Finally, there is the Soundscape Explorer (terrestrial) (SET), developed by Luniletronik². It combines on-board computation of Acoustic Complexity Index (ACI) values with the logging of environmental data. It features two microphones — one for audible sound and one for ultrasonic detection — and can record for up to two weeks in a typical configuration³. While these features have proven useful in field studies (Farina et al. 2016; Farina and Salutari 2016; Farina 2019; Benocci et al. 2020), SET calculates only ACI and does not have solar charging capabilities, limiting its utility in long-term unsupervised studies⁴.

The BioAcoustic Index Tool is our attempt to fill that gap (see table 1). Using technologies associated with the DIY community, BAIT integrates a powerful microcontroller that calculates acoustic indices in real-time as sound is captured from the surroundings with environmental sensors that capture temperature, humidity, and light data. The resulting sensor captures acoustic index and environmental data while storing 40000 times less raw data than a conventional field audio recorder. Combined with the solar battery charging system, this allows the sensor to run autonomously and indefinitely in the field, without the need to change batteries or SD cards.

The ability to run these sensors without constant maintenance means that they are ideally suited to anchor large-scale, long-term acoustic biodiversity monitoring projects involving a grid of sensors spread out across a landscape. In the next section, we present the design of a prototype of this type of sensor system. We outline its capabilities and its drawbacks and, in later sections, discuss its applicability to existing and future studies.

2. Materials and Methods

The BioAcoustic Index Tool (BAIT) is a smart sensor that measures acoustic indices from the soundscape of an ecosystem and records those indices as well as other environmental data. Using a combination of onboard audio processing and solar power, BAIT was able to operate maintenance-free for nearly 6 months in a forest garden in southern Sweden.

The base of the sensor works in much the same way as any of the existing field recorders: it features a microphone, a processor, and an SD card. High-quality audio is captured and recorded to the SD card in much the same way as is done in the SongMeter recorders used by (Pieretti et al. 2015; Gottesman et al. 2020; Righini and Pavan 2020) and the Solo recorder used by (Bradfer-Lawrence et al. 2019)⁵.

The next step is where BAIT differs from a standard field recorder. Instead of leaving the audio files on the SD card to be collected and processed back in a lab, BAIT performs acoustic index calculations on-board, in the field. Once the acoustic indices are calculated for a particular audio file, the file is discarded and all that it

²<http://www.luniletronik.it/en/prodotto/set-soundscape-explorer-terrestrial/>

³Product specification at http://www.luniletronik.it/soundscape_explorer/.

⁴As SET is a closed-source commercial product, researchers are unable to expand its capabilities as needed.

⁵The microprocessor and microphone used in BAIT are similar to those found in the Solo recorder.

stores are the numbers representing the calculated acoustic indices⁶.

As predicted by Browning et al. (2017); Greif and Yovel (2019), performing calculations in-situ and discarding the raw audio results in a massive reduction in onboard data storage requirements. As such, storage capacity is no longer a limiting factor in the operating life of the sensor. A drawback of this approach, of course, is that it is no longer possible to reanalyze the raw audio or perform additional calculations or manual observations after the fact.

With storage capacity no longer a concern, the availability of power is now the main limiting factor for long-term operation of the sensor. A small, 2-watt solar panel powers the BioAcoustic Index Tool and charges its internal battery. While this configuration allowed the sensor to remain in the field recording data for over 6 months, it did not provide enough power or energy storage to allow the sensor to record continuously — especially in the darker, earlier months of the year. The sensor uptime is detailed in section 3.1.

2.1. Design

An important part of the BioAcoustic Index Tool is that the plans and code are open-source and therefore modifiable by researchers with specific needs. For example, knowing that there might not be enough power to carry out continuous recordings throughout the day and night, a researcher might modify the power management code to prioritize recording at dusk and dawn and only record during other times of the day if there is excess power available. They might also implement the calculation of other metrics that are useful for analyzing the soundscape.

In this section, we outline the design of BAIT to give the reader a general understanding of how the tool works. Further design and implementation details are described in section A.

2.1.1. Electronics

The electronic core of BAIT is a microcontroller, a set of sensors, and a power-management system. The microcontroller captures sound and environmental readings from the sensors and stores the environmental readings on an onboard microSD card. The audio is processed in real-time using modified versions of the algorithms to calculate ACI and Acoustic Diversity Index (ADI) on streaming data. The power management system charges the battery and informs the microcontroller to enter a lower-power mode when the battery doesn't have enough charge to run the full-scale calculations.

The microcontroller is similar to the one used in the AudioMoth (Hill et al. 2019) — an ARM Cortex M4F. In place of using a customized circuit board to run the microcontroller as is done on the AudioMoth, BAIT uses the Teensy 3.6 USB Development board which includes electronics for basic functionality like power regulation, a real-time clock (RTC), and an onboard microSD card reader (see fig. 1). The Teensy is also Arduino-compatible and BAIT's firmware is written using Arduino libraries. These factors make it relatively easy to modify the design and the associated firmware.

⁶It is entirely possible to store more than just the acoustic indices. During our evaluation of BAIT, we also recorded intermediate computations to help verify the calculations of acoustic indices after the fact. BAIT is capable of retaining data at any level — including raw audio — for verification, data, audits, or additional analysis, but doing so would negate some of the benefits discussed later on.

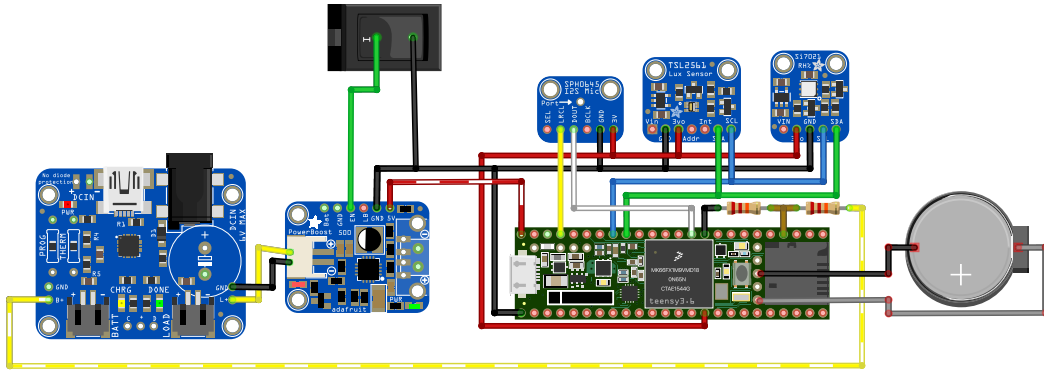


Figure 1.: Diagram of the electronic components and connections in BAIT prototype.

The sensor set includes a MEMS (microelectromechanical systems) microphone, a temperature and humidity sensor, and an ambient light sensor, all shown in fig. 1. The microphone captures soundscape data, while the other sensors enable the correlation of soundscape information with environmental data. The environmental sensors could also be used to schedule audio recordings or regulate power usage, though this is not currently implemented.

2.1.2. Firmware

The firmware⁷ for BAIT is written using the Arduino platform as well as libraries from Adafruit⁸ and PJRC⁹. It is modularly structured so that it is easy to add in the calculation of new bioacoustic and environmental measurements. It also incorporates two different power modes to enable proper charging of the batteries, while maintaining the collection of bioacoustic and environmental data as consistently as possible. The entire firmware is open-source and available on GitHub¹⁰.

The BAIT firmware controls sensor readings and data preprocessing; performs audio analysis and the calculation of acoustic indices; and manages the battery and power state of the system. The two main features of the BAIT firmware are the modified ADI and ACI algorithms and the power management system.

2.1.3. Acoustic index calculations

Two test acoustic indices were chosen for implementation in the prototype of BAIT: ADI and ACI. They were selected because both are well-documented and commonly-used in field studies. Their structure — operating on spectral representations of the audio — makes them suitable for implementation on a microcontroller that can perform FFTs. An overview of the implementation of these indices is given here, but a detailed description of the algorithm can be found in section A.2.

The ADI is an attempt to quantify the acoustic diversity of a sound. Defined by Villanueva-Rivera et al. (2011), the ADI operates between 0-10 kHz and calculates the

⁷Firmware is software that is written for embedded computing, such as the microcontroller in BAIT.

⁸adafruit.com

⁹pjrc.com, the manufacturer of the Teensy

¹⁰github.com/dkashish/BioAcousticIndexTool

Shannon entropy of the sound. This is done by dividing the frequency spectrum into 10 equal 1-kHz bands and assessing the proportion of FFT bins that contain energy above a defined threshold in each band. The Shannon index of these values is the ADI. A full calculation is shown in section A.2.1.

In the R implementation, this calculation is performed on a whole recording at once, calculating the proportion of positive bins at each frame. To efficiently calculate ADI on streaming data, BAIT collects a running sum of the number of times each FFT frequency bin exceeds the threshold along with a count of how many samples it has seen. These totals are divided at the end of the entire sample instead of at each frame, avoiding the accumulation of floating point errors over the course of the sample.

To calculate ACI in real-time on a microcontroller, it was necessary to translate the ACI algorithm into C++ but also to modify it to work with streaming data instead of a full audio file. Described in detail in Pieretti et al. (2011), the ACI represents the amount of variation of intensity of sound within frequency bands over the course of a recording fragment. It relies on the assumption that anthropogenic noises — for example, the droning of an aeroplane engine or the buzz of a factory — are often spectrally constrained and relatively constant, so it attempts to detect sounds that vary from moment-to-moment.

Normally, ACI is calculated for a complete audio file, but we modified the algorithm so that it could be computed in real-time as new audio was captured. Full details of the modified implementation can be found in section A.2.2.

2.1.4. Power modes

As storage space is no longer a limiting factor for BAIT, power is now the primary concern for the smart sensor. To conserve battery, two different power modes were designed, to maximize BAIT’s operational time.

In the main, full-power mode, BAIT has all sensors enabled and records all of the available data. Audio is captured and indices are calculated in real-time as described above. Environmental data is captured and recorded to the microSD card. However, capturing and processing audio requires the processor to be constantly active. When only environmental data is collected, BAIT is active only a small fraction of the time, meaning that it consumes less power. As such, a second, low-power mode was developed to capture only environmental and battery data. When this mode is active BAIT is unable to capture or process audio, so the data from those times is not recorded.

For the prototype, these modes are used in two different ways. In mixed-power mode, BAIT alternates between full- and low-power modes every 15-minutes to extend the basic battery life of the system. It is assumed that this still provides a reasonable temporal resolution for soundscape index measurements. Secondly, the low-power mode is engaged when the battery level falls below a set level. This allows BAIT to continue to monitor environmental parameters and be ready to return to full-power mode once the battery charge increases.

Full details of how the power modes are designed and activated are available in section A.2.3.

2.1.5. Enclosure

The enclosure for the BioAcoustic Index Tool is 3D printed and the plans are freely available online¹¹. The enclosure was designed in Autodesk Fusion 360 and features

¹¹github.com/dkadish/BioAcousticIndexTool

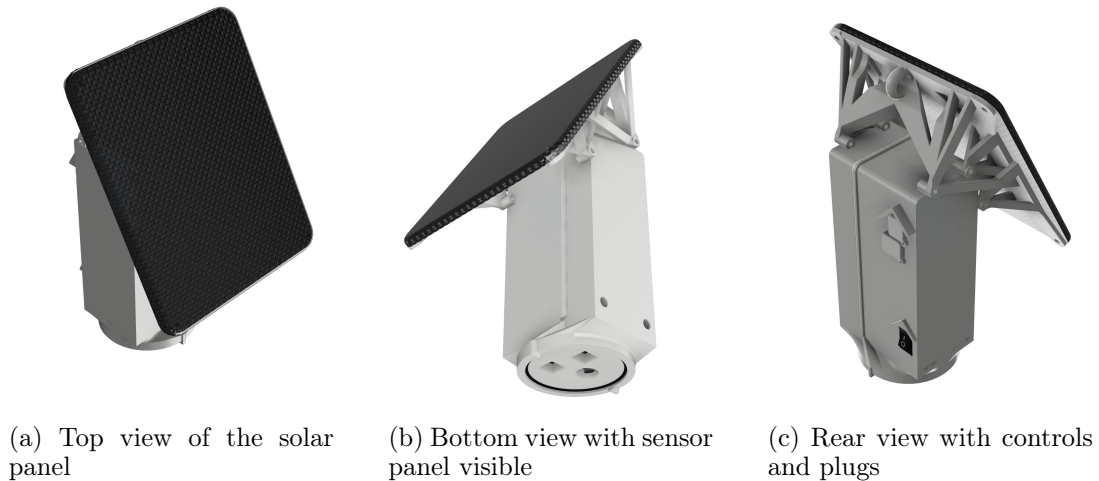


Figure 2.: Three renderings of the BioAcoustic Index Tool

a solar panel mount, an external power switch, a Micro-USB charging port, and a downward-facing sensor panel. The final design is shown in fig. 2.

The enclosure should not be considered water-proof, but with a good 3D print, it can last outside in a range of weather¹². To that end, the placement of the solar panel helps to protect the seam from heavy rain as does the placement of the sensors on the bottom of the device.

2.2. Field Experiments

The BioAcoustic Index Tool was tested in the field at a forest garden site called Holma Skogsträdgården in Höör, Sweden for approximately 5 months between February 18 and July 17, 2019. The site is an active educational forest garden set adjacent to a preschool and a forest garden teaching facility (Holma Folkhögskola), between a series of conventional farms on the outskirts of the town. A train line runs about 200m from the garden carrying local, regional, and long-distance passenger traffic as well as freight trains, and the sound of the passing trains echos loudly through the garden.

The garden itself features mixed groves of food-bearing trees, bushes, and perennial vegetables. Birds flit back and forth between the fruit trees and visit the sizeable on-site pond. Through the day, children from the preschool visit the garden to explore and classes from Holma Folkhögskola work and learn in the groves.

All of this activity provided a rich acoustic environment for testing BAIT. In addition to measuring the acoustic indices of the soundscape, the purpose of the trial was to establish the operating parameters for BAIT. This includes the following:

- Uptime: How many acoustic measurements is BAIT able to take using the available power?
- Mode power usage: What is the power consumption of each mode in the field?
- Storage requirements: How much data is actually recorded?
- Environmental data correlation: How does the environmental data collected

¹²It worked for more than 6 months at a field site in southern Sweden and was not damaged by water in that time. That said, something did appear to have built a web inside the case and a solitary bee apparently took up residence in one of the screw recesses in the exterior of the case.

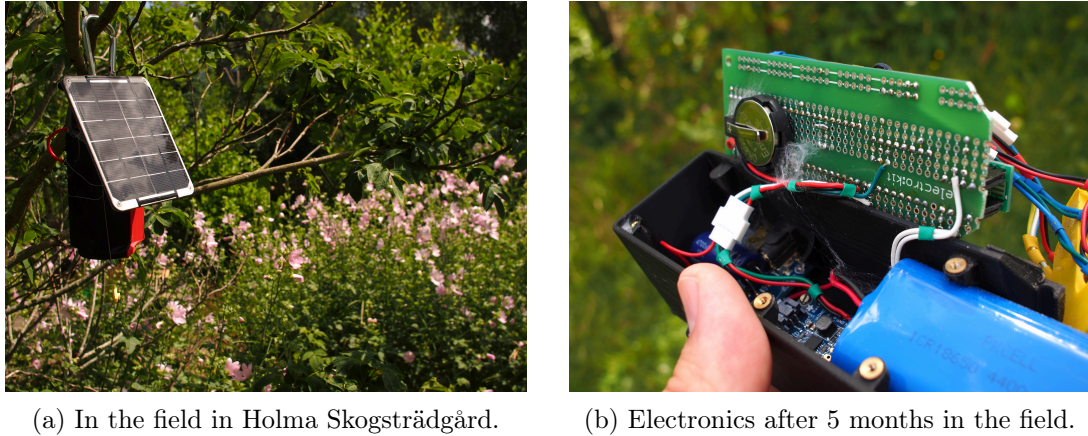


Figure 3.: Views of the BioAcoustic Index Tool during the field experiments.

correlate to the measured acoustic indices?

3. Results

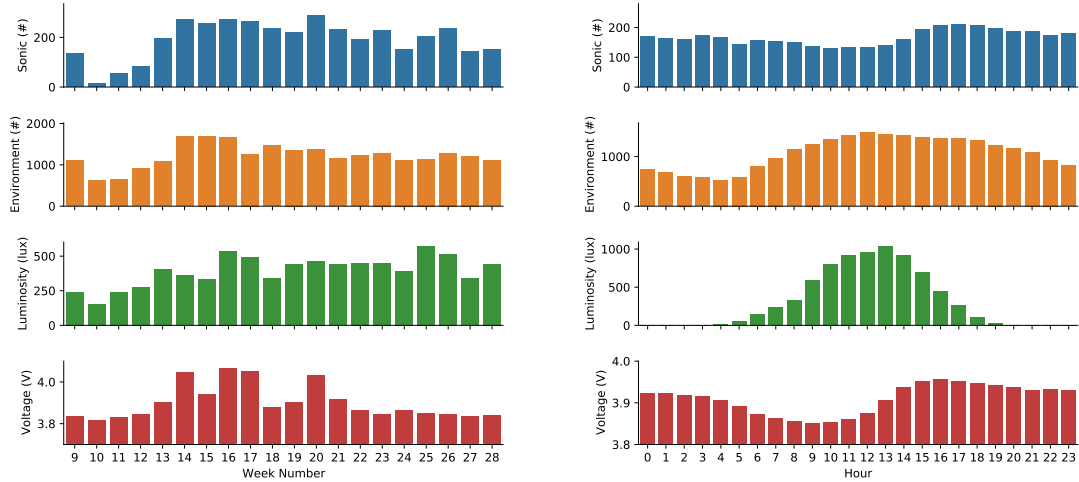
The field experiments were designed to validate the basic functionality of BAIT and to establish its operating parameters, as well as to produce a dataset that would demonstrate the types of relationships that could be explored using the sensor. In the following sections, we detail findings for operating parameters such as uptime, battery usage, and data production. We conduct a preliminary exploration of the collected acoustic index and environmental data and visualize the types of correlations that can be found in the dataset.

3.1. Uptime

Uptime refers to how many of the scheduled measurements were taken. For the prototype, acoustic index readings were generated from 15 minutes of audio every 30 minutes, so 100% uptime would correspond to 48 measurements per day for audio. Environmental readings were recorded every 5 minutes, so 100% uptime implies 288 measurements per day. The uptime depends on battery charge levels and power modes; for example, audio readings are suspended while the battery charge level is low, which causes the uptime to drop.

Uptime was calculated from the measurement counts visualised in fig. 4. Figure 4a shows the data organised by week to visualise seasonal changes in data collection. It depicts the number of acoustic index (grey) and environmental (red) readings taken as well as the average light levels (yellow, in lumens) and battery voltage (blue, in volts) for each week. The same data is shown in fig. 4b, organised instead by hour.

Overall, BAIT had an uptime of 57% for acoustic index measurements and 88% for environmental measurements. However, this varies widely by season and time of day. Seasonally, there’s a large jump in uptime for acoustic index collection between weeks 12 and 14 as average light levels increased and the battery was charged more regularly. Small variations in the average luminosity seem to correspond to large shifts in available power — and therefore the uptime — but this likely has to do with the positioning of the solar panel, meaning that the position and timing of the sunlight are



(a) Weekly grouping showing seasonal changes. Low light levels in February and March impeded solar charging, which interrupted audio readings.

(b) Hourly grouping showing diurnal changes. Midday sun drove charge levels higher which lead to increased readings in the afternoon and evening.

Figure 4.: Number of acoustic index (grey) and environmental (red) readings taken and the average light levels (yellow, in lumens) and battery voltage (blue, in volts). This visualises the uptime of the sensor and the main factors determining the sensor availability.

perhaps more important than the average level. A more regular shift can be seen in the daily cycles (fig. 4b) as midday sun charges the battery, creating a spike in readings in the late afternoon as BAIT exits low-power mode.

3.2. Battery usage

In the lab, the power consumption of the microcontroller and sensor peripherals was measured to be 284 mW in full-power mode and 16 mW in low-power mode. At the 3.7 V nominal voltage of BAIT’s battery, this translates to a current draw of 76.8 mA in full-power mode and 4.3 mA in low-power mode. Given these rates and the 4400 mAh capacity of the battery, BAIT should be able to run continuously for just over 57 hours in full-power mode and a little over 1023 hours (around 42.5 days) in low-power mode without recharging.

These values are measured under ideal, laboratory settings and should be considered an upper bound on BAIT’s battery life. For a more detailed analysis of the battery operation see Appendix B.

3.3. Storage

The onboard processing of sound data means that a 15-minute analysis of the soundscape produces mere bytes of data. Over the approximately 5 months of data collection, the sensor produced just under 15 MB of data. Though the sensor did have periods where it didn’t record, the uptime was greater than 50%, so even at full power for the entire recording period, the sensor wouldn’t have collected more than 30 MB of data. That’s roughly equivalent to the size of 3 minutes of raw audio, recorded with

standard settings. By comparison, generating the 4023 acoustic index measurements that BAIT captured using conventional recordings would have required the collection and processing of about 600 GB of raw audio.

3.4. Acoustic index

Of course, the primary task of the sensor is to capture acoustic index data. Over the course of the 5-month study period, BAIT captured 4023 measurements of each ACI and ADI. We present the collected data as it shows hourly and day-of-the-week patterns in fig. 5 with ACI shown in blue and ADI in orange.

One observes clear diurnal patterns in the measurements of both acoustic indices in fig. 5a. ACI exhibits a peak around midday with distinctive valleys around 3:00 and 19:00, while ADI has more of a plateau between 7:00 and 17:00. A study by Fairbrass et al. (2017) found that while ACI is correlated to biophony, ACI and ADI are also correlated to different types of anthrophony¹³. Given that, we expected to observe a difference between weekday and weekend patterns of the measured acoustic indices, however this is not the case in fig. 5b. The site is used during the week for teaching by Holma Folkhögskola — though there is not a constant presence there every day — and weekend events are sometimes hosted by a non-profit group that is associated with the site. It is possible that the use patterns at the site are irregular enough that there was no significant difference between human activity on weekdays and over the weekend to shift the distribution of measurements, though this issue requires further study.

Interestingly, the shape of the plot of ADI in fig. 5a resembles the average ADI values that Villanueva-Rivera et al. (2011) found on agricultural sites in the paper where they first describe the metric. There, too, they recorded small peaks in the morning and evening with relatively flat values throughout the day and night.

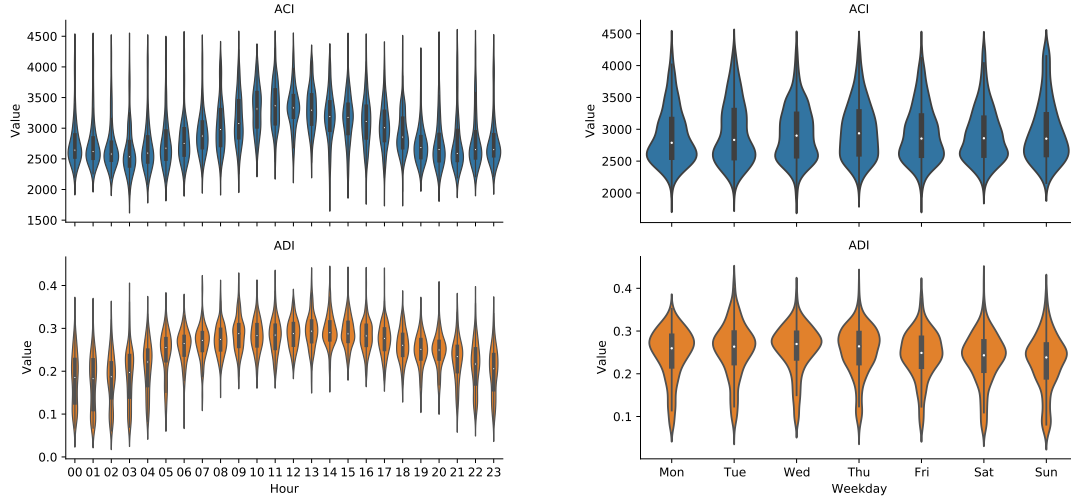
3.5. Environmental Data Correlation

In addition to calculating acoustic indices, BAIT also records environmental data that can be used to understand the acoustic information that is captured. In their introduction to the field, Pijanowski et al. (2011a) pointed out that animal behaviour as well as soundwaves themselves are often modulated by environmental variables such as weather and light conditions. While it may be possible to use forecasts and weather station data to study the effects of environmental conditions on acoustic indices, onboard sensors can give a hyperlocal view of these phenomena.

The plots in fig. 6 visualize the relationships between ACI and ADI and the measured temperature, humidity, and luminosity. The data is displayed in a scatterplot matrix, which is used to show pairwise relationships between the different dimensions of a dataset. The scatterplot matrix has 3 distinct areas: the top-left area shows scatterplots of 2 of the measured variables along with a linear regression and its r^2 value; the diagonal shows the distribution of values for a single variable using a kernel density estimate (KDE)¹⁴; and the lower-right area shows scatterplots with contour lines

¹³ACI is positively correlated to the level of anthrophonic activity — defined as the area of a spectrogram that is covered by anthrophonic sound — while ADI is negatively correlated to anthrophonic diversity which reflects the number different types of anthrophonic sound. Notably, ACI and ADI are negatively correlated to the presence of electronic sounds and vehicular noise, while human speech is positively correlated to ACI but negatively correlated to ADI.

¹⁴This is similar to a histogram, but generates a continuous plot showing the density of measurements around a particular value.



(a) Diurnal patterns of ACI and ADI. ACI shows a midday peak while ADI has more of a plateau pattern.

(b) Weekly patterns of ACI and ADI. There does not appear to be a strong pattern of difference between days of the week.

Figure 5.: Violin plots of ACI and ADI values, organized by hour and day to highlight diurnal and weekly patterns in the acoustic index measurements. Each violin shows the distribution of measurements over the category using a kernel density estimation. The violins also contain an internal box plot depicting the mean as a white dot and the quartiles as a black box.

highlighting areas of higher density. The data points and density estimates are coloured by month to help visualise seasonal changes in measurements.

In the top-left of fig. 6, the r^2 value is the square of the correlation coefficient of the linear regression and it indicates the level of interdependence of the 2 measured variables. The ACI-ADI plot (1st row, 2nd column) shows that the measured acoustic indices are relatively uncorrelated. This means that they appear to measure different aspects of the recorded soundscape and confirms the utility of having recording both metrics. Most highly correlated are the 3 environmental variables, as seen in the Temperature-Humidity, Temperature-Luminosity, and Humidity-Luminosity. Relative humidity — which is what is actually being measured — is defined in relation to temperature and days tend to be both warmer and brighter than nights, so these highly-correlated relationships are expected. Interestingly, ACI seems to be more strongly correlated to the environmental factors — especially humidity and luminosity — than temperature.

The density estimates along the diagonal show how the measurements of a single variable are distributed. As in a histogram, the x-axis shows the measured values and the y-axis depicts the relative density of measurements around that value. The ADI-ADI plot shows the distinct double-peak of the measured acoustic diversity with most of the measurement centred near ADI values of 0.2 or 0.55. ACI measurements also reveal a slight second peak, but it is much less pronounced than that of ADI.

The plots in the bottom-right side of fig. 6 show the density of points in bivariate distributions using a 2-dimensional KDE to create an overlay much like a topographical map. Darker contours outline areas of higher density, while lighter contours show more diffuse measurements. They help to reveal the manner in which variables are correlated.

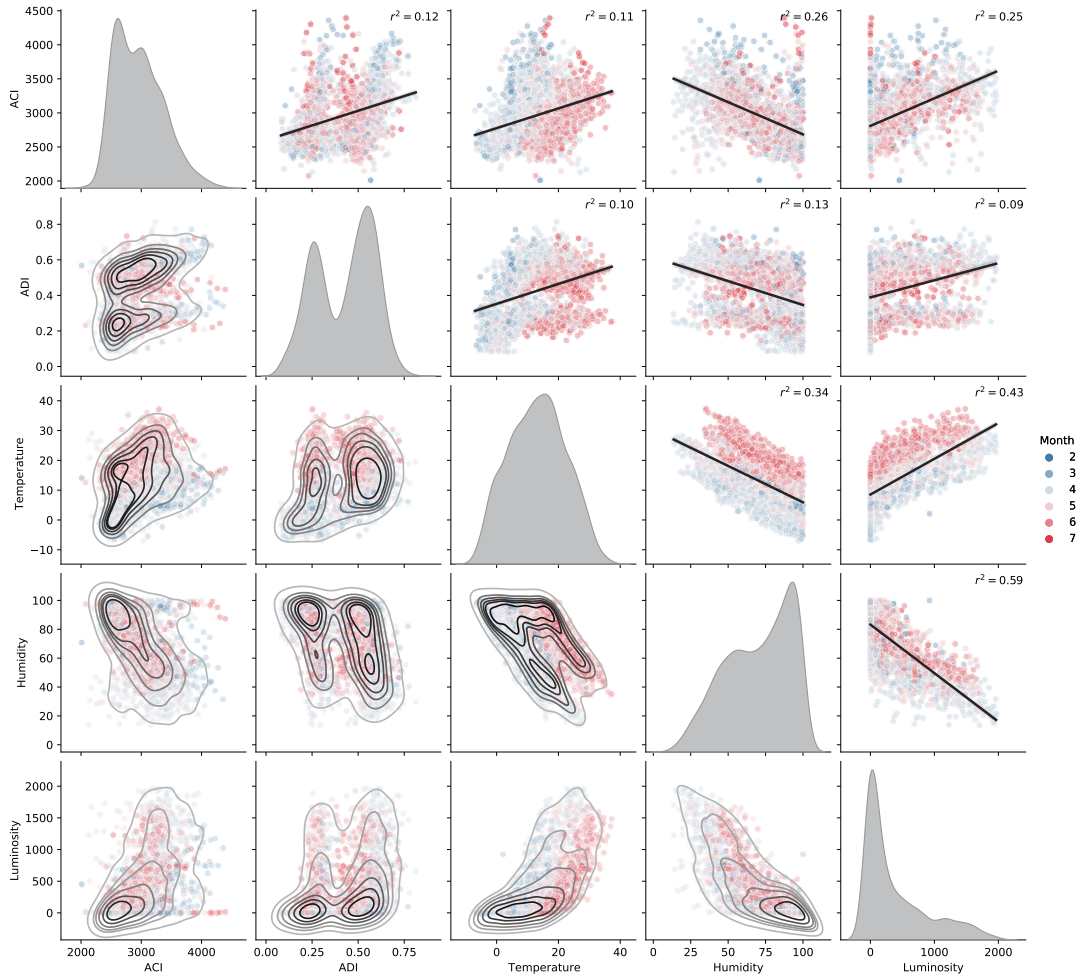


Figure 6.: A scatterplot matrix of the measured acoustic indices and environmental factors. The data is coloured by month to illustrate how the relationships change seasonally. The plots along the diagonal are density plots which — like histograms — show how each measured variable is distributed. The scatter plots above and below the diagonal show the relationships between the x- and y-axis variables; those above the diagonal are annotated with a linear regression while those below the diagonal show the density of the data using contour lines to illustrate areas of increasing density. These plots are useful as an overview of the data that has been collected. For example, the ADI-ADI plot found along the diagonal in the second column of the second row shows that ADI values cluster around two values — 0.2 and 0.55. The ACI-Luminosity plot in the first row shows that there’s a relatively strong correlation between the two measurements (compared to the other factors), which makes sense given the strong diurnal patterns seen in the ACI plot in fig. 5a. Directly below that plot, the correlation between ADI — which has a much flatter daytime curve in fig. 5a — and Luminosity is shown to be quite a bit weaker.

For example, the Temperature-ADI plot reveals an unexpected pattern. The plot shows a deep impression in between the two peaks that are centred around temperatures of about 15°C and ADI values of 0.2 and 0.55. This means that ADI values are actually most divergent at temperatures around 15°C, while outside of that range, the ADI measurements are more uniform.

These plots demonstrate the utility of collecting environmental data alongside acoustic indices measurements. Patterns in the data are evident with only the basic visualisations presented here. Further analyses could help to decorrelate the environmental measurements from the acoustic data to better understand both the impact of environmental conditions on acoustic index measurements and long-term trends in biodiversity measurements conducted under varying environmental conditions.

4. Discussion

The results in the previous section indicate how the BAIT prototype performs in the field. But what do they mean for its potential use in future field studies? What kind of performance can be expected in the field, what kinds of data can researchers expect to collect, and how might BAIT be improved to address some of its drawbacks and deficiencies and to add functionality?

4.1. Data Storage

Data storage and management are often cited as key limitations for soundscape studies (Towsey et al. 2014; Bradfer-Lawrence et al. 2019; Righini and Pavan 2020; Farina et al. 2021). In the field test, BAIT generated 4023 data points for each of the acoustic indices that it measured. Each data point was based on 15 minutes of audio which, had it been captured as raw audio using a conventional field recorder, would have required about 600 GB of storage capacity.

In contrast, BAIT produced just 15 MB worth of data over the course of 5 months in the field, capturing 2 acoustic indices twice-per-hour and 3 points of environmental data at 5-minute intervals. This represents a reduction in the data output of a sensor by a factor of 40000. At these data production rates, the size of available storage is no longer a limiting factor in the ability of the sensor to run autonomously in the field for an indefinite period.

4.2. Power

The next key limitation is the availability of power. Soundscape recordings are often conducted well away from the electrical grid and so battery power becomes essential for running recording devices. Processing data in the field does require far more power than simply recording audio files and, as such, BAIT requires a solar panel to maintain sufficient power to operate. This has both advantages and disadvantages. BAIT was able to operate over a long time (the device was still running at the end of the 5-month test period) but the uptime was intermittent, with BAIT recording environmental data in 88% of the time-periods but acoustic indices in only 57% of the scheduled times. The recording periods for the acoustic indices were biased towards the afternoon and evening as the device often lost power in the morning after a night of recording drawing

on power stored during the previous day.

4.3. Available Data

In the end, what matters most is what data is available to a researcher. The most important benefit of BAIT — the ability to capture acoustic index data without needing to store raw audio files — will be the most difficult for some researchers to accept. Without access to the raw audio, they cannot re-process the audio after having listened to the recordings, they cannot run additional analyses after the fact, and — perhaps most importantly for some — they will be unable to listen to the soundscape and interpret the recordings themselves (Righini and Pavan 2020; Farina et al. 2021). It is important to not discount what is learned qualitatively about a field site through the active listening of someone with a well-tuned ear for the details of a soundscape¹⁵.

That said, something is gained as well here. The automation of the process of generating acoustic indices saves computing time and effort, but also minimizes the opportunity for human data processing errors. In their 2018 paper on sources of errors in scientific studies, Brown et al. identify *errors of data management* as one of four major types of study error. In automating a large part of the initial data analysis, BAIT minimizes the risk of introducing errors between the capture of audio data and the calculation of acoustic indices. It means that different recordings will not, for example, accidentally be processed by different implementations of an acoustic index algorithm¹⁶ or using different parameters and settings.

In addition, the recording of synchronized environmental data has the potential to add new explanatory power to acoustic index measurements. Pijanowski et al. (2011a) describe the impact of what they call *atmospheric dynamics* on the composition of the soundscape. These conditions can have direct impacts on the measured soundscape, such as when wind and rain produce sound, but they also have indirect impacts. Animals often modulate their sound production in response to environmental factors (Pijanowski et al. 2011a) — think of cricket chirp rates responding to changes in temperature — and the actual propagation of sound also depends in part on atmospheric conditions (Ingård 1953) Pijanowski et al. (2011b) set the improvement of understanding of the relationship between environmental conditions and sound as one of six major themes in the field of soundscape ecology and the availability of a tool that records these conditions alongside soundscape data could prove to be an important step toward that goal.

4.4. Use Case

A number of recent studies have used acoustic indices — and ADI or ACI in particular — to investigate ecological questions and might have benefited from the use of a tool like BAIT (Farina et al. 2013; Farina and Pieretti 2014; Towsey et al. 2014; Bradfer-Lawrence et al. 2019; Righini and Pavan 2020; Farina et al. 2021). Some additional studies have employed the SET, which provides some of the same benefits as does BAIT (Farina et al. 2016; Farina and Salutari 2016; Farina 2019; Benocci et al. 2020).

To better understand the use cases for BAIT, it is useful to examine in detail a single study that used conventional recording methods to understand how it would be

¹⁵One possible way to mitigate this by conducting a type of mixed-mode recording is discussed in section 4.5.5

¹⁶The R packages `seewave` and `soundecology` produce different ACI values, see <https://cran.r-project.org/web/packages/soundecology/vignettes/ACIandSeewave.html>

Table 2.: Common acoustic indices and their status in BAIT

Index	Source	BAIT Status
Acoustic Complexity Index (ACI)	Pieretti et al. (2011)	working ^a
Acoustic Diversity Index (ADI)	Villanueva-Rivera et al. (2011)	working ^a
Bioacoustic Index (BI)	Boelman et al. (2007)	easily implemented ^b
Acoustic Evenness Index (AEI)	Villanueva-Rivera et al. (2011)	easily implemented ^b
BIOPHONY	Righini and Pavan (2020)	easily implemented ^b
Acoustic Entropy Index (H)	Sueur et al. (2008b)	not implemented ^c
Spectral Entropy (H_f) ^d	Sueur et al. (2008b)	easily implemented ^b
Normalised Difference Soundscape Index (NDSI)	Kasten et al. (2012)	not implemented ^c

^afully implemented and tested in the prototype

^bnot in the prototype but can be calculated using the same process as working acoustic indices

^cnot implemented, requires new processes not used in the prototype

^dSpectral Entropy (H_f) is actually one half of the calculation of the Acoustic Entropy Index, which is the sum of the spectral and temporal entropies of a signal.

changed by the introduction of BAIT. Righini and Pavan (2020) set out to characterize the soundscape of a nature reserve and compare the soundscape inside and outside the reserve using the recordings from three field sites. They used a Song Meter 3 field recorder to capture 45600 minutes (488.30 GB) of recorded audio, recording the first 10 minutes of every 30-minute period over the course of a month.

The study included qualitative and quantitative analysis, listening to recordings and viewing their spectrograms as well as calculating a set of seven acoustic indices. The study found significant differences between daytime and nighttime activity for all three sites as well as differences between the two sites within the reserve and the site outside the reserve. In addition to an analysis of the full dataset, Righini and Pavan also performed some manual filtering of the data. They listened to all of the recorded data and excluded files that featured heavy wind or rain and then compared the resulting indices to those calculated with the full dataset.

How would the study have been different if the authors had been able to use the BioAcoustic Index Tool in place of the SM3 Field Recorders? The data collection would have been similar, though they would have had the ability to leave the sensors in the field for far longer, no longer having to worry about data storage and processing capabilities¹⁷. Much of the quantitative analysis would also remain the same; of the seven acoustic indices that are calculated, two are already implemented on BAIT and three others can be easily added to the system (see table 2 for details). Two others — Acoustic Entropy Index (H) and Normalised Difference Soundscape Index (NDSI) — would require additional programming to implement on BAIT as they do not use the same underlying processes already employed by the ACI and ADI calculations. The qualitative analysis, on the other hand, would not be possible in the same manner. BAIT does not currently record audio or spectrograms — though it is possible to do so (see section 4.5.5) — so the researchers would be unable to listen to recordings and observe spectrograms in the way that they did in the study.

However, BAIT would come with one important additional benefit. Righini and Pavan write that there are unaccounted for differences between the three sites during the daytime that probably depend on environmental factors. ‘These results indicate

¹⁷In prototype testing, BAIT did not always have enough power to sample at every scheduled point, but this could be mitigated using a larger solar panel and battery and the lighter recording schedule of this study — 10 minutes every half hour as opposed to 15 in the BAIT test.

the need to have more information of environmental parameters at very local levels and thus the need to add at least light, temperature, humidity and wind sensors to acoustic recorders.’ (Righini and Pavan 2020) BAIT performs precisely this function, measuring light, temperature, and humidity on the sensor alongside the acoustic index data.

4.5. Future Developments

Even in its current form as a first prototype and proof-of-concept, BAIT can produce interesting datasets over a long period of time in the field. However, there are a range of improvements that should be undertaken to enhance current performance and add functionality.

4.5.1. Power

The power system is the most obvious candidate for some improvements. The easiest way to improve the uptime for the sensor would be to simply increase the size of the battery and the solar panel. A larger battery would allow the system to store more energy during sunny times to help eliminate the dip in recordings during the early morning hours and a larger panel could take better advantage of the available solar energy to capture more energy when it is available.

The power modes could also be adjusted for more effective operation. The system could be adjusted to prioritize certain times of the day — for example, dawn and dusk when there is often increased acoustic activity. Or it could be programmed to ensure that there is roughly even sampling of all of the times of the day so that less sampling is done during the evening when the battery is often more fully charged to save power for morning samples.

A third power mode could also be introduced that would capture but not process audio. The processing is particularly power-intensive, so when battery levels are lower or there is little sun, BAIT could capture audio, but wait to process it until reserve solar power is available and then delete the raw audio files to regain the storage space.

Additionally, there is some indication that BAIT was sampling more than necessary in the field trial. Pieretti et al. (2015) suggest that capturing audio for one of every five minutes is sufficient to accurately characterize a soundscape using ACI. Therefore it is possible to change the sensor scheduling to lower the amount and duration of audio capture and index computation, which would further extend the battery life of the sensor.

A combination of these approaches could increase the uptime for soundscape recording and improve the quality and distribution of data that is collected without intensive hardware revisions. However, on the electronic hardware side, an improved battery management system that could track current draw and power usage would be a boon to the system’s ability to self-regulate and switch between power modes. This kind of improvement should be high on the list of priorities for the next major hardware revision.

4.5.2. Wireless

There’s a benefit to the reduction of data that BAIT produces that has been alluded to, but not discussed in full. With the daily data production in kilobytes (KB), it becomes more feasible — both in terms of power requirements and cost of transmission —

to send data from remote locations back to a lab as it is being collected. Wireless transmission can be expensive both in terms of power requirements and the price of bandwidth in remote areas, so minimizing the data to be transferred is essential.

Existing acoustic monitoring systems such as Echo Box and the SAFE Acoustics monitoring network have used popular consumer wireless communications infrastructure such as WiFi and cellular networks (Balestrini et al. 2020; Sethi et al. 2020). However, remote areas where these types of devices are often deployed sometimes lack the required cellular and WiFi infrastructure. In these circumstances, it becomes necessary to transmit data by longer-range modes of communication. In practice, this would likely mean linking individual sensors by a mesh network to a common base station with a satellite or landline Internet connection. IoT wireless technologies like LoRaWAN (Vangelista et al. 2015; Margelis et al. 2015) or ZigBee (Safaric and Malaric 2006) could play this role; both are low-power mesh networking technologies designed for embedded systems.

Wireless signalling would draw some battery power, but the BioAcoustic Index Tool could collect data until it is charged to its highest power level before sending a burst of collected data back to the lab through its wireless networking system. This feature would help to make the device fully autonomous and able to operate basically without service at all, barring equipment failures. The same connection could be used to send device status updates and even potentially to adjust sampling schedules and parameters based on data observed back in the lab.

This would allow BAIT to operate as part of a large-scale, long-term, fully autonomous network of hundreds or thousands of acoustic biodiversity monitoring sensors. The ability to leave a sensor in the field indefinitely and to collect data remotely could enable entirely new types of long-term tracking studies.

4.5.3. Additional indices

For the prototype, we calculated two indices: ADI and ADI. They were chosen for their importance in the field and their relative ease of calculation. However, several other acoustic indices use similarly structured computations that would not be difficult to implement on BAIT using the structure that we have developed. These include Bioacoustic Index (BI) (Boelman et al. 2007), Acoustic Evenness Index (AEI) (Villanueva-Rivera et al. 2011), and BIOPHONY (Righini and Pavan 2020), as well as spectral entropy (Toh et al. 2005). These indices are some of the key components of the `soundecology` (Villanueva-Rivera and Pijanowski 2018) and `seewave` (Sueur et al. 2008a) R packages that are commonly used in soundscape ecology studies.

Because the code for BAIT is open-source, anyone can modify the firmware that performs these calculations and it is possible to add new indices as they are defined in the literature.

4.5.4. Environmental sensors

For the prototype, temperature, humidity, and light level sensors were chosen for inclusion in BAIT because the sensors are readily available and provide a good overview of the environmental conditions at a particular location. But there are many other sensors that are available for more detailed detection of particular parameters, depending on the needs of a particular research project.

An anemometer could be a particularly useful addition to the toolkit as wind can be a significant factor in some of the acoustic index calculations (Righini and Pavan

2020). A soil moisture or rainwater sensor could provide additional details about the hydraulic conditions at a site, and as rain is another significant source of geophonic sound that can affect calculations.

Those represent the most obvious additions to BAIT’s sensor toolkit, but one could imagine how the addition of more specialized sensors like geophones for detecting seismic events and air quality sensors for detecting vehicular emissions and forest fires might prove useful for particular studies. It would be impossible — and probably unhelpful — to exhaustively list all of the sensors that one could attach to BAIT, but the point here is to note that the system is extendible and can be modified to the specific sensory needs of a study.

4.5.5. *Mixed-mode data collection*

One of the most significant drawbacks of BAIT is that it saves no raw audio. This is a purposeful feature of the system, but it also means that there is nothing for researchers to listen to for a more experiential or qualitative impression of their field site. While this type of knowledge is seldom referred to specifically in written research, listening to the soundscape can give the researcher context and a connection to the site that the raw acoustic index data cannot provide on its own. For some, this is a crucial part of their work (Righini and Pavan 2020).

Though the prototype is set up this way, there is no reason that saving raw audio has to be an all-or-nothing proposal. It is possible to save particular samples of audio to the storage medium or even to record all possible audio and delete samples selectively to free up space as necessary. Intermediate calculations such as the raw FFT data — or FFTs averaged over time — could be stored to generate spectrograms upon collection. The system could be programmed to retain data for anomalous events that produce extreme acoustic index or environmental data to later diagnose the causes and impacts of these events.

These strategies could help to alleviate researcher concerns about the quality of data collected and can be used to perform confidence checks to confirm the accuracy of calculations. They can be used as data samples that can be examined in detail and used to illustrate the processes used for acoustic index calculations.

4.5.6. *Embedded smart sensors*

This section has so far focused on future improvements to BAIT specifically, but it is prudent to note the potential of smart sensors in general to enable new types of acoustic and ecological research. The practice of moving processing power to peripheral sensors is part of a broader trend in computing called *edge computing* (Shi et al. 2016). As microcontrollers have gotten smaller, more efficient, more powerful, and more accessible, it has become increasingly possible to perform complex computation in embedded contexts.

These shifts are occurring rapidly. A 2018 study detailed a multilevel frog detection system that performed initial data analysis on an embedded device, followed by further analysis on a cloud server (Roe et al. 2018). Only two years later, a paper proposed running a full deep learning neural network classifier designed to detect bird calls right on the sensor (Sturley and Matalonga 2020) and Balestrini et al. (2020) have produced a network of sensors with an embedded deep learning-based bat detector and classifier. These solutions use embedded computers (the Raspberry Pi and Intel Edison), but it is actually possible to run some neural networks on microcontrollers that use a fraction

of the power of even these lightweight systems (Falbo et al. 2020).

As these trends continue, it should be possible to perform more accurate species and event detection as well as advanced index calculation on-site. These changes will open new opportunities for acoustic population and biodiversity surveys as well as long-term monitoring of ecosystems.

4.6. Conclusion

The BioAcoustic Index Tool is a shift from the conventional field audio recorder and is unlikely to replace them where researchers are interested in performing in-depth analysis of a particular soundscape. However, it would be a boon to a project interested in the calculation of acoustic indices over a broad spatial and temporal field. The ability to generate acoustic index data at a large number of sample points over a long study time could enable new kinds of soundscape surveys that track patterns over months, years, and decades. The inclusion of synchronized environmental data gives researchers the tools for better understanding of how environmental conditions modulate the soundscape and impact measured acoustic indices. And the use of an open platform for BAIT enables researchers to extend the platform with new sensors and calculations as needed.

A tool like BAIT makes it possible to envision the creation of permanent acoustic biodiversity monitoring networks featuring tens, hundreds, and even thousands of sensors spread across a landscape. With wireless connections, these networks could generate a high-resolution overview of shifting biodiversity levels. It would be possible to measure seasonal and annual changes in biodiversity and better understand how environmental factors contribute to acoustic measurements of biodiversity as well as changes in biodiversity itself. The fully-automated pipeline that produces acoustic index data at the sensor would allow ecologists to focus on the interpretation of the acoustic index data instead of the process of its collection and computation.

The approach is indicative of a coming shift in the collection, processing, and analysis of acoustic data and soundscape recordings. BAIT and future sensors like it have the potential to move the first level of ecological data processing from the lab to the field and, in doing so, easing the process of collecting and analyzing data about soundscapes.

Acknowledgement(s)

Field experiments were conducted in the Holma Skogsträdgård in Höör, Sweden. Thanks to all of the staff, volunteers, and students there for their help getting the experiment running, and especially Nils Lindblad, Johanna Johansson, and Hanna Jönsson. Thanks to the REAL and IxD Labs at ITU and particularly Mathias Schmidt for the time spent discussing the sensor electronics and enclosure printing. The authors would also like to thank Laura Beloff for her supervision and Kristin Aleklett for her help with editing and for lending an ecologist's eye to the writing.

Disclosure statement

The authors declare no potential conflict of interest.

Data Availability statement

The data that support the findings of this study are openly available on figshare.com. The raw data collected from the sensor is stored at <http://doi.org/10.6084/m9.figshare.13607198> and processed data used for analysis and to generate plots in the paper is stored at <http://doi.org/10.6084/m9.figshare.13607132>.

References

- Balestrini M, Gallacher S, Rogers Y. 2020. Moving HCI Outdoors: Lessons Learned from Conducting Research in the Wild. In: McCrickard DS, Jones M, Stelter TL, editors. *Hci outdoors: Theory, design, methods and applications*. Springer International Publishing; chap. 4; p. 83–98.
- Baumgartner MF, Fratantoni DM, Hurst TP, Brown MW, Cole TVN, Van Parijs SM, Johnson M. 2013. Real-time reporting of baleen whale passive acoustic detections from ocean gliders. *The Journal of the Acoustical Society of America*. 134(3).
- Beason RD, Riesch R, Koricheva J. 2019. AURITA: an affordable, autonomous recording device for acoustic monitoring of audible and ultrasonic frequencies. *Bioacoustics*. 28(4).
- Benocci R, Brambilla G, Bisceglie A, Zambon G. 2020. Eco-acoustic indices to evaluate soundscape degradation due to human intrusion. *Sustainability (Switzerland)*. 12(24).
- Boelman NT, Asner GP, Hart PJ, Martin RE. 2007. Multi-Trophic Invasion Resistance in Hawaii: Bioacoustics, Field Surveys, and Airborne Remote Sensing. *Ecological Applications*. 17(8).
- Bradfer-Lawrence T, Gardner N, Bunnefeld L, Bunnefeld N, Willis SG, Dent DH. 2019. Guidelines for the use of acoustic indices in environmental research. *Methods in Ecology and Evolution*. 10(10).
- Brown AW, Kaiser KA, Allison DB. 2018. Issues with data and analyses: Errors, underlying themes, and potential solutions. *Proceedings of the National Academy of Sciences*. 115(11).
- Browning E, Gib R, Glover-Kapfer P, Jones K. 2017. Passive acoustic monitoring in ecology and conservation. Report No: 2.
- Buxton RT, McKenna MF, Clapp M, Meyer E, Stabenau E, Angeloni LM, Crooks K, Wittemyer G. 2018. Efficacy of extracting indices from large-scale acoustic recordings to monitor biodiversity. *Conservation Biology*. 32(5).
- Carson R. 1962. *Silent Spring*. New York: Houghton Mifflin Company.
- Deniz NN, Chelotti JO, Galli JR, Planisich AM, Larripa MJ, Leonardo Rufiner H, Giovanini LL. 2017. Embedded system for real-time monitoring of foraging behavior of grazing cattle using acoustic signals. *Computers and Electronics in Agriculture*. 138.
- Fairbrass AJ, Rennert P, Williams C, Titheridge H, Jones KE. 2017. Biases of acoustic indices measuring biodiversity in urban areas. *Ecological Indicators*. 83.
- Falbo V, Apicella T, Aurioso D, Danese L, Bellotti F, Berta R, De Gloria A. 2020. Analyzing Machine Learning on Mainstream Microcontrollers. In: Saponara S, Gloria AD, editors. *Applications in electronics pervading industry, environment and society*. Springer, Cham; chap. 12; p. 103–108.
- Farina A. 2019. Acoustic codes from a rural sanctuary: How ecoacoustic events operate across a landscape scale. *Biosystems*. 183.
- Farina A, Pieretti N. 2014. Sonic environment and vegetation structure: A methodological approach for a soundscape analysis of a Mediterranean maqui. *Ecological Informatics*. 21.
- Farina A, Pieretti N, Morganti N. 2013. Acoustic patterns of an invasive species: the Red-billed Leiothrix (*Leiothrix lutea* Scopoli 1786) in a Mediterranean shrubland. *Bioacoustics*. 22(3).
- Farina A, Pieretti N, Salutari P, Tognari E, Lombardi A. 2016. The Application of the Acoustic Complexity Indices (ACI) to Ecoacoustic Event Detection and Identification (EEDI) Modeling. *Biosemiotics*. 9(2).

- Farina A, Righini R, Fuller S, Li P, Pavan G. 2021. Acoustic complexity indices reveal the acoustic communities of the old-growth Mediterranean forest of Sasso Fratino Integral Natural Reserve (Central Italy). *Ecological Indicators*. 120.
- Farina A, Salutari P. 2016. Applying the Ecoacoustic Event Detection and Identification (EEDI) Model to the Analysis of Acoustic Complexity. Report No.:
- Gottesman BL, Francomano D, Zhao Z, Bellisario K, Ghadiri M, Broadhead T, Gasc A, Pijanowski BC. 2020. Acoustic monitoring reveals diversity and surprising dynamics in tropical freshwater soundscapes. *Freshwater Biology*. 65(1).
- Greif S, Yovel Y. 2019. Using on-board sound recordings to infer behaviour of free-moving wild animals. *Journal of Experimental Biology*. 222(February).
- Guo S, Qiang M, Luan X, Xu P, He G, Yin X, Xi L, Jin X, Shao J, Chen X, et al. 2015. The application of the Internet of Things to animal ecology. *Integrative Zoology*. 10(6).
- Hill AP, Prince P, Snaddon JL, Doncaster CP, Rogers A. 2019. AudioMoth: A low-cost acoustic device for monitoring biodiversity and the environment. *HardwareX*. 6.
- Ingård U. 1953. A Review of the Influence of Meteorological Conditions on Sound Propagation. *The Journal of the Acoustical Society of America*. 25(3).
- Kasten EP, Gage SH, Fox J, Joo W. 2012. The remote environmental assessment laboratory's acoustic library: An archive for studying soundscape ecology. *Ecological Informatics*. 12.
- Lyonblum EZS. 2017. An Omnivorous Ear: The Creative Practice of Field Recording [Doctor of Philosophy (PhD)]. University of Cambridge.
- Margelis G, Piechocki R, Kalessi D, Thomas P. 2015. Low Throughput Networks for the IoT: Lessons learned from industrial implementations. *IEEE World Forum on Internet of Things, WF-IoT 2015 - Proceedings*.
- Pieretti N, Duarte M, Sousa-Lima R, Rodrigues M, Young R, Farina A. 2015. Determining Temporal Sampling Schemes for Passive Acoustic Studies in Different Tropical Ecosystems. *Tropical Conservation Science*. 8(1).
- Pieretti N, Farina A, Morri D. 2011. A new methodology to infer the singing activity of an avian community: The Acoustic Complexity Index (ACI). *Ecological Indicators*. 11(3).
- Pijanowski BC, Farina A, Gage SH, Dumyahn SL, Krause BL. 2011a. What is soundscape ecology? An introduction and overview of an emerging new science. *Landscape Ecology*. 26(9).
- Pijanowski BC, Villanueva-Rivera LJ, Dumyahn SL, Farina A, Krause BL, Napoletano BM, Gage SH, Pieretti N. 2011b. Soundscape Ecology: The Science of Sound in the Landscape. *BioScience*. 61(3).
- Righini R, Pavan G. 2020. A soundscape assessment of the Sasso Fratino Integral Nature Reserve in the Central Apennines, Italy. *Biodiversity*. 21(1).
- Roe P, Ferroudj M, Towsey M, Schwarzkopf L. 2018. Catching toad calls in the cloud: Commodity edge computing for flexible analysis of big sound data. *Proceedings - IEEE 14th International Conference on eScience, e-Science 2018*.
- Safaric S, Malaric K. 2006. ZigBee wireless standard. In: *Proceedings ELMAR 2006*; jun. IEEE. p. 259–262. June.
- Schafer RM. 1977. *The Tuning of the World*. vol. 1st. New York, N.Y.: Random House.
- Sethi SS, Ewers RM, Jones NS, Signorelli A, Picinali L, Orme CDL. 2020. SAFE Acoustics: An open-source, realtime ecoacoustic monitoring network in the tropical rainforests of Borneo. *Methods in Ecology and Evolution*. 11(10).
- Shi W, Cao J, Zhang Q, Li Y, Xu L. 2016. Edge Computing: Vision and Challenges. *IEEE Internet of Things Journal*. 3(5).
- Sturley S, Matalonga S. 2020. PANDI: A Hybrid Open Source Edge-based System for Environmental and Real-Time Passive Acoustic Monitoring - Prototype Design and Development. 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology, IRASET 2020.
- Sueur J, Aubin T, Simons C. 2008a. Seewave, A Free Modular Tool for Sound Analysis and Synthesis. *Bioacoustics*. 18(2).
- Sueur J, Pavoine S, Hamerlynck O, Duvail S. 2008b. Rapid acoustic survey for biodiversity

- appraisal. PLoS ONE. 3(12).
- Toh AM, Togneri R, Nordholm S. 2005. Spectral Entropy As Speech Features. In: Proceedings of PEECS. p. 60–65. 1.
- Towsey M, Wimmer J, Williamson I, Roe P. 2014. The use of acoustic indices to determine avian species richness in audio-recordings of the environment. *Ecological Informatics*. 21.
- Truax B. 1978. *The World Soundscape Project's Handbook for acoustic ecology*. 1st ed. A.R.C. Publications. Music of the environment series.
- Vangelista L, Zanella A, Zorzi M. 2015. Long-Range IoT Technologies: The Dawn of LoRa. In: Atanasovski V, Leon-Garcia A, editors. *Future access enablers for ubiquitous and intelligent infrastructures*. (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering; vol. 159). Cham: Springer International Publishing; p. 51–58.
- Villanueva-Rivera LJ. 2015. Indices and ecoinformatics tools for the study of soundscape dynamics [Doctor of Philosophy]. Purdue University.
- Villanueva-Rivera LJ, Pijanowski BC. 2018. *soundecology: Soundscape Ecology - R package*; mar.
- Villanueva-Rivera LJ, Pijanowski BC, Doucette J, Pekin B. 2011. A primer of acoustic analysis for landscape ecologists. *Landscape Ecology*. 26(9).
- Westerkamp H. 1974. Soundwalking. *Sound Heritage*. 3(4).
- Whytock RC, Christie J. 2017. Solo: an open source, customizable and inexpensive audio recorder for bioacoustic research. *Methods in Ecology and Evolution*. 8(3).

Appendix A Design

The design of the BioAcoustic Index Tool is described in section 2.1, however specific implementation details are important for the reproduction of BAIT. Here, the specific electronic components, algorithms, and programming strategies are detailed to make it possible to build and extend BAIT for future studies.

A.1. Electronics

The BAIT design is based around the Teensy 3.6, which features an ARM Cortex M4 processor. The processor includes a floating point unit (FPU), which allows it to perform calculations with floating point numbers in a relatively efficient and accurate manner. The remainder of BAIT is divided into a power management system and a sensor system.

The primary sensor is the microphone, an Integrated Inter-IC Sound Bus (I²S) microelectromechanical systems (MEMS) chip-based microphone that mounts directly onto a PCB. The microphone chip captures sound with a flat response curve in the 100 Hz to 10 kHz range and digitizes it before sending it forward to the microcontroller over an I²S bus. The SPH0645 was selected for its cost-effectiveness and the ease of connecting it to the system given the pre-digitized signal that it produces, however, its linear response range of 100-10k Hz might be a limiting factor to others interested in using the tool. Fortunately, it is possible to replace this device with an external microphone for sample collection, if a different frequency response is desired.

In addition to the microphone, the BAIT features a set of environmental sensors that can gather data that can help to provide context to the bioacoustic indices and result in a deeper understanding of the patterns of biodiversity (Pijanowski et al. 2011a). Two sensors are engaged in this environmental data collection and they capture light levels, ambient temperature and relative humidity. The Si7021 from Silicon Labs measures

both ambient temperature and relative humidity and gives digital readings in degrees Celsius and percentage. The TSL2561 measures the intensity of the ambient light in lux.

A full list of the electronic components used in the prototype and their cost is available in table 3.

A.2. Firmware

The firmware for BAIT is written using the Arduino platform as well as libraries from Adafruit¹⁸ and PJRC¹⁹. It is modularly structured so that it is easy to add in the calculation of new bioacoustic and environmental measurements. It also incorporates two different power modes to enable proper charging of the batteries, while maintaining the collection of bioacoustic and environmental data as consistently as possible.

The major contribution of the BAIT firmware is the translation of the algorithms for the calculation of two bioacoustic indices into C++ and their transformation from offline post-processing algorithms into code that runs efficiently online and in real-time. The ACI and ADI were implemented in this way and each required different modifications.

The full firmware is available on GitHub at github.com/dkedish/BioAcousticIndexTool.

A.2.1. ADI calculation

Defined in (Villanueva-Rivera et al. 2011), the ADI operates between 0-10 kHz and calculates the Shannon entropy of the sound by dividing the frequency spectrum into 10 equal 1-kHz bands and applying Equation 1, where p_i is the proportion of sound in frequency band i .

$$H' = - \sum_{i=1}^S p_i \ln p_i \quad (1)$$

Note that this proportion p_i is understood as the proportion of FFT bins within the frequency band i that are above a defined threshold. What results is a measure of the diversity of the soundscape in terms of how the sounds are spread across the frequency spectrum over the period of measurement.

In the R implementation, this calculation is performed on a whole recording at once. It calculates a spectrogram for the entire file and then collects the overall level in each band for use in calculating diversity. The BAIT does not have the luxury of a complete sound file, so it instead collects a running sum of the power in each frequency bin along with a count of how many samples it has seen. p_i is tracked frame-by-frame, calculating a running sum of the number of bins with values above the threshold (P_i) and tracking the number of frames seen (N). When the ADI value is calculated for a length of time, Equation 2 is applied to avoid accumulating floating point errors.

$$p_i = \left(\sum_{i=0}^N P_i \right) / N \quad (2)$$

¹⁸adafruit.com

¹⁹pjrc.com, the manufacturer of the Teensy

Now, Equation 1 can be applied as in the original implementation and the ADI can be calculated for a longer span of time, emulating the way that a whole sound file is processed.

A.2.2. ACI Calculation

The transformation of the ACI calculation is somewhat more complex. Described in detail in Pieretti et al. (2011), the ACI represents the amount of variation of intensity of sound within frequency bands over the course of a recording fragment. It relies on the assumption that anthropogenic noises — for example, the droning of an aeroplane engine or the buzz of a factory — are often spectrally constrained and relatively constant, so it attempts to detect sounds that vary from moment-to-moment.

Over the course of a user-defined temporal step (j)²⁰, the difference in intensities from between samples (d_k) at a particular frequency bin (Δf_l) is calculated as

$$d_k = |I_k - I_{k+1}| \quad (3)$$

These differences are summed over the entire temporal window and divided by the total observed acoustic intensity over that period as in Equation 4, resulting in a measure of the ACI for a particular frequency bin and temporal step ($ACI_{j,l}$).

$$ACI_{j,l} = \frac{\sum_{k=1}^{n-1} d_k}{\sum_{k=1}^n I_k} \quad (4)$$

These measurements are added up for all q frequency bins and all m temporal steps in a recording to determine the total ACI for the audio clip as

$$ACI_{tot} = \sum_{l=1}^q \sum_{j=1}^m ACI_{j,l} \quad (5)$$

To perform this calculation efficiently, BAIT retains a running tabulation of the total ACI (ACI_{tot}), the sum of the difference between samples in the same frequency bin (D), the total acoustic intensity in the same frequency bin (I). It also stores the previous intensity measurement for each band (I_{k-1}) so that the difference (d_{k-1}) can be calculated. At the end of each temporal window, these values can be reset except for ACI_{tot} , which is retained and saved to a file at the end of the recording period.

The conversion of these scripts from post-processing calculations to ones that can be performed on streaming data saves a great deal of data and enables the processing of sound on the microcontroller without taxing its memory resources.

A.2.3. Power Modes

The BAIT has two power modes that are switched between automatically as the system's battery charges and discharges. The main power mode measures acoustic and atmospheric data and is active when the battery is charged over 3.7V, as measured by an onboard voltage divider. In this mode, audio is captured and realtime calculations are done to log ACI and ADI.

²⁰These are also referred to as clumps in (Farina et al. 2016). BAIT defaults to 30s.

Table 3.: Parts list and prices for the prototype of BAIT

Part	Price (USD)
Solar Charger v2	17.50
PowerBoost 500	9.95
4400 mAh Li-Ion Battery	19.95
Si7021 Temperature and Humidity Sensor	8.95
TSL2561 Light Sensor	5.95
I2S MEMS Microphone	6.95
Teensy 3.6	29.95
Solar Panel (2W)	29.00
22k Ohm resistor	0.75
SD Card (16 GB)	9.95
Coin cell mount	0.95
Coin cell battery	0.95
On/Off Switch	0.95
Total*	141.75

*Total cost for the purchased breakout boards and components. Does not include the cost of the 3D printed enclosure, breadboards, and consumable parts such as screws, wire, etc.

These calculations, however, are quite power-intensive and inhibit the charging of the battery when they are engaged. To maintain data logs, but allow the battery to charge when it is low, a low power mode is engaged below the threshold of 3.7V, in which the BAIT records only environmental data and turns the acoustic systems off to save energy. Since the atmospheric measurements are intermittent, the microcontroller can sleep in between measurement cycles, drastically reducing the power requirements²¹.

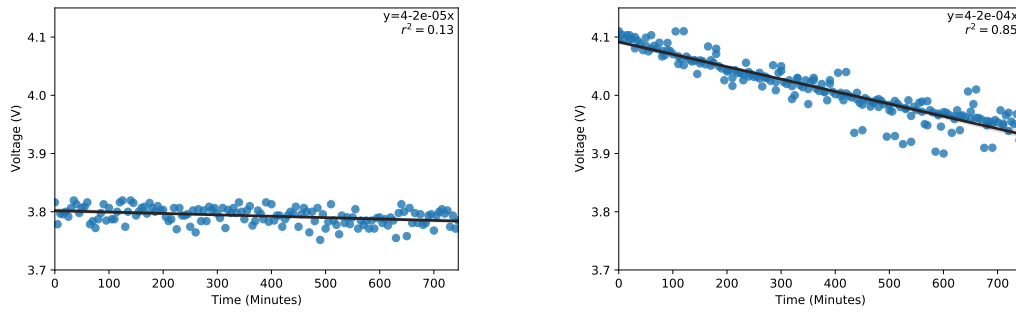
The sleep cycle is also modulated by the measurement process. ACI and ADI are calculated in 15-minute segments on the BAIT²². If low-power mode were to engage during the middle of a 15-minute measurement cycle, all data gathered before the invocation of low-power mode would be wasted. To avoid this pattern, BAIT inhibits the application of low-power mode during acoustic index calculation. After the cycle has completed and useful data are collected, then the BAIT is allowed to sleep for a cycle.

A.3. Enclosure

The enclosure is 3D printed from polylactic acid (PLA) on an Ultimaker 2 printer. It features a detachable sensor panel with cut-outs for the 3 sensors, a port for power delivery, and an external power switch. The case has mount points for a 3D-printed solar panel mount and hanging system to suspend it from a tree in the study environment. Internally, the enclosure has mount points for the power management PCBs and a cradle for the battery pack that powers the system. All of the screw points are augmented with metal heat-set inserts to strengthen the screw points. Designs were done using Autodesk’s Fusion 360 software and sliced on Ultimaker’s CURA. Design files are available in the project’s GitHub repository at github.com/dkadish/BioAcousticIndexTool.

²¹As opposed to the acoustic measurements which require continuous, intensive calculations.

²²Emulating 15-minute recordings of soundscapes.



(a) Low-power mode voltage drop. The trend line shows a drop of 0.02 mV/min (1.5 mV/hour) in low-power mode.

(b) Mixed-power mode voltage drop. The trend line shows a drop of 0.20 mV/min (12.3 mV/hour) in mixed-power mode.

Figure 7.: Voltage drop under different power modes when there is no sunlight charging the batteries via the solar panel.

Appendix B Battery Life

Section 3.2 describes a laboratory-based analysis that establishes the upper bound on the battery life of BAIT without recharging using the solar panel. However, there are serious limitations to these calculations. The battery capacity is negatively impacted by both high and low temperatures, so the capacity in field situations is likely to be less than 4400 mAh. These measurements also exclude the voltage boost electronics used to convert the 3.7 V supply from the battery to the 5 V supply expected by the Teensy 3.6. The boost electronics operate at 90% efficiency²³, meaning some power is lost in the conversion. As such, these numbers should be treated as an upper limit on the possible performance of the battery.

To establish more realistic operating parameters, it was necessary to gather data in the field. This presented another challenge as the prototype cannot directly measure power consumption. The battery voltage measurements used for switching between power modes can reveal something about the battery's state. However, it is important to note that the discharge profile of the lithium-ion battery pack is highly nonlinear so these analyses can also only provide an estimate of the power consumption of BAIT.

Voltage change in the two power modes was estimated by analysing the drop in voltage during times when the luminosity was near-zero and therefore the system was not being charged by the sun. In mixed-power mode, where audio is being captured and processed, the battery voltage was falling at a rate of about 12.3 mV/hour. In contrast, in low-power mode, the battery voltage was falling at a little over a tenth of that, 1.5 mV/hour. The data behind this calculation is shown in fig. 7.

The battery can be charged to 4.2V and mixed-power mode is engaged until it reaches 3.7V, so without any solar charging, the sensor can read and process audio for at least 40 hours²⁴. It is also possible to increase the operational time by switching to a larger solar panel or battery. A larger capacity battery would allow the sensor to store more power when the sun is shining to increase the time that it could run without charging. And a larger panel could collect more energy from the available light in order to charge the battery faster and more often.

²³Product specification and datasheet available at <https://www.adafruit.com/product/1903>

²⁴Again, this is approximate as the discharge curve for lithium-ion batteries is non-linear.

Appendix C ACI Algorithm Verification

The ACI algorithm has a number of different implementations that produce slightly different results. The implementations in the R packages `seewave` (Sueur et al. 2008a) and `soundecology` (Villanueva-Rivera and Pijanowski 2018) have differences in their outputs, which are documented in the notes for the `soundecology` package at <https://cran.r-project.org/web/packages/soundecology/vignettes/ACIandSeewave.html>. There, Villanueva-Rivera attributes the differences to slight variations in the implementation of the clumping argument, j .

Here, the implementation in BAIT is compared to the `soundecology` implementation to verify the accuracy of the approach to calculating ACI as a running sum. The test is conducted using audio recordings that were used in a comparison of the ACI results from `soundecology` and another implementation of the ACI algorithm, a plugin for the WaveSurfer software called `SoundscapeMeter.1.0.14.05.2012` (Villanueva-Rivera 2015). The recordings were resampled to 44.1 kHz (originally 48 kHz) to meet the requirements for the code to play WAV files from an SD card in the Teensy Audio Library.

The material — including code, audio, and data files — used for these tests is available on figshare at <http://dx.doi.org/10.6084/m9.figshare.14445348>.

C.1. FFT

The spectrograms produced by the microcontroller in BAIT and the `spectro` function of the `seewave` package²⁵ are slightly different. This is shown in detail in fig. 8, which plots the two spectrograms along with the differences between the two, once the values are normalized²⁶. This means that for the same audio, BAIT will inevitably produce a different result from `soundecology` (and therefore other implementations as well). For this reason, caution should be used when comparing ACI results computed using different methods.

C.2. ACI Computation

To bypass the difference in FFT implementations and verify the remainder of the algorithm, we used a modified version of the `acoustic_complexity` function from `soundecology` which calculates ACI from FFT values saved in a CSV file instead of from a raw audio file²⁷. To generate the CSV file, BAIT runs test code²⁸ that computes the ACI of a sound file while recording the FFT values to its SD card as they are computed.

Here, a second discrepancy between the two implementations is clear in the clumping procedure — the same site as the difference between the `soundecology` and `seewave` versions of the algorithm. The `soundecology` implementation of ACI calculates a

²⁵This is used to produce an FFT representation of the sound for analysis in the `soundecology` package.

²⁶Note that the BAIT spectrogram shown here has a change from the implementation used in the BAIT prototype field experiment. The default setting for the microcontroller's FFT library averages together 8 readings to produce a single temporal value for each frequency. This went unnoticed prior to the field experiment, so the data in the testing of the prototype used this setting. This has been corrected in the latest version of the code (version 0.2 on Github at <https://github.com/dkadish/BioAcousticIndexTool/releases/tag/0.2>) so that no averaging of the FFT readings is done.

²⁷The R notebook containing that test code is available on figshare at <http://dx.doi.org/10.6084/m9.figshare.14445348>.

²⁸The code that generates the CSV is available at https://github.com/dkadish/BioAcousticIndexTool/blob/0.2/firmware/bait/src/test_save_spectro.cpp.

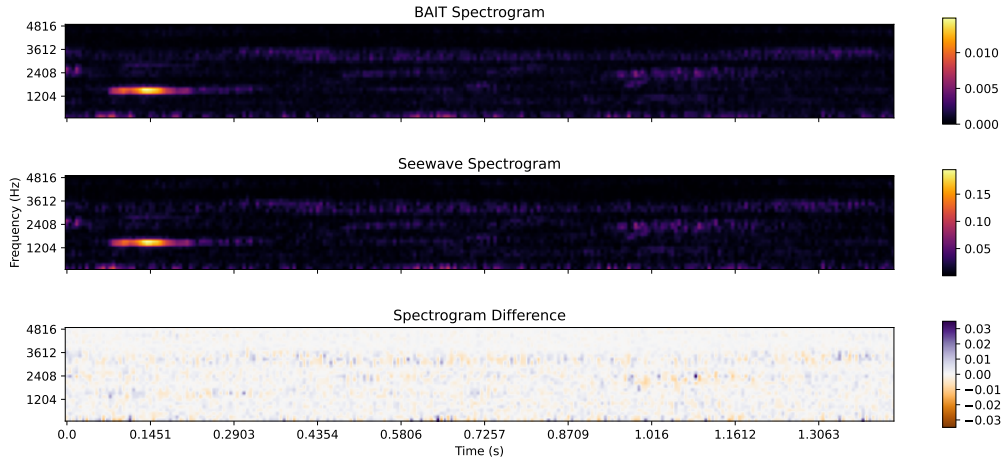


Figure 8.: Spectrograms of the lower frequencies from the first portion of 6.wav. The top image is the spectrogram generated by BAIT, the middle image is the spectrogram generated using `seewave`, and the final image is the difference between the two (normalized) spectrograms.

variable called `I_per_j` at the beginning of its computation, which is the number of temporal frames per cluster. This fixed variable is calculated as the integer (floor) of j , the number of clumps divided by Δ_{tk} , the time per frame. Each clump, then, is calculated from a fixed number of FFT readings.

BAIT, however, operates in real-time and therefore does not have the ability to look back over a fixed sound file and determine a static number of bins per clump. The number of FFT readings per clump is controlled by the timing of the microcontroller. If the clump time is set for 5 seconds, the clump rolls over once 5 seconds have passed. This leads to slight variations in the clump size if the FFT frame rate does not divide evenly into the clump time.

To account for this, the test code on BAIT also records the ACI_j value for each cluster j as well as the number of FFT readings processed at the end of the frame. That number of FFT readings is then used in the modified version of the `acoustic_complexity` function from `soundecology` so that its clumps are calculated on the same number of frames. This allows for a direct comparison between the ACI calculations of BAIT and the `soundecology` package.

Figure 9 shows the total ACI value for each clump (ACI_j) calculated by BAIT and the modified `acoustic_complexity` function from the `soundecology` package. The slight discrepancy in each value is caused by floating-point calculation errors. The FFT values from the microcontroller are saved to CSV with a 6-digit decimal precision, which leads to rounding errors when those numbers are imported into R. Additionally, the microcontroller computes floating point numbers at single (32-bit) precision, while R uses double (64-bit) precision leading to further minor differences in the results.

However, it is clear from the plot in fig. 9 that the algorithm implemented on the microcontroller in BAIT is the same as the one implemented in `soundecology`. The total ACI values that are calculated using that process — 598.18 from `soundecology` and 597.61 from BAIT — differ only by rounding errors within the calculation.

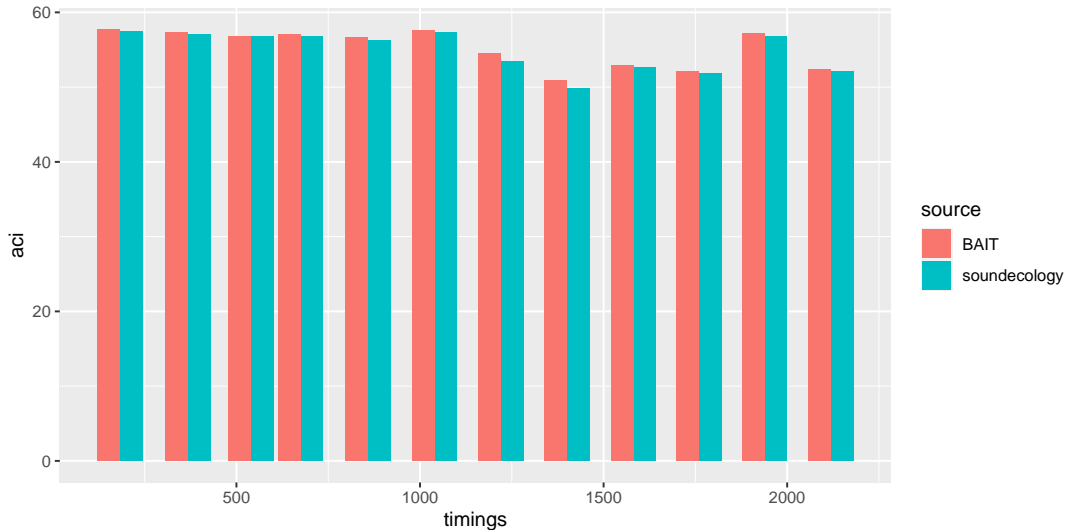


Figure 9.: ACI totals from each clump (ACI_j) in the calculation on the test file.

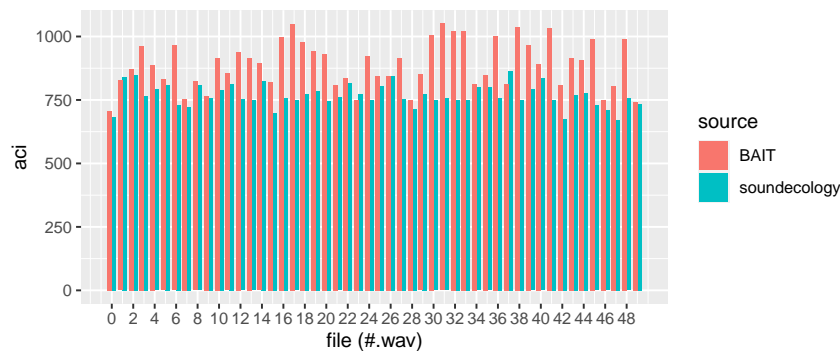


Figure 10.: ACI totals from each of the 50 test files in the first set of test data at <https://doi.org/10.6084/m9.figshare.1036395.v1>.

C.3. Comparability

Though the cluster-by-cluster computation is the similar, implementation differences between the real-time ACI computation on BAIT and the file-based computation on a computer will produce different results for the same sound. Therefore, it is inadvisable to directly compare ACI results obtained with BAIT with those obtained by recording sound and computing the ACI using implementations in R.

An examination of the waveforms of selected audio files reveals a pattern. Figure 11 shows 3 files where the ACI values computed using BAIT and the `soundecology` R package were similar and 3 where there were large differences in the values. From this sample, it appears that similar results were produced for sounds with higher amplitudes.

This can be shown formally. The median of the amplitude envelope (M) is a measure of the amplitude of a sound over an entire recording. The Shapiro-Wilk test shows the distributions of the difference between calculated ACI values and the median of the amplitude envelopes to not be normally distributed ($p = 8.2e - 3$ and $p = 2.4e - 3$ respectively), so correlation is tested using Spearman's rank-order correlation. The test

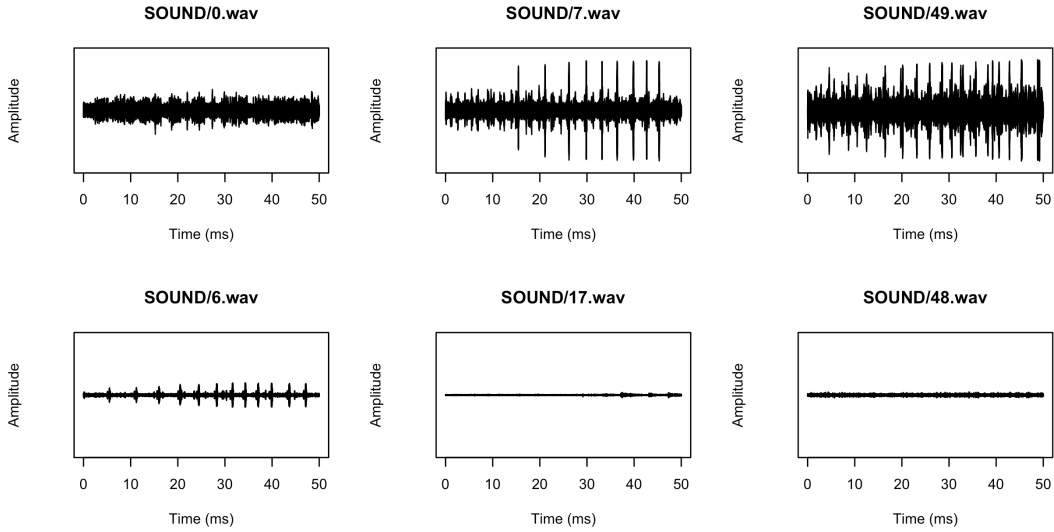


Figure 11.: Waveforms of a selection of the audio files from fig. 10. The ACI values for the audio files in the top row were similar when calculated with BAIT and the soundecology R package, while those in the lower row differed more significantly.

shows a strong negative correlation ($\rho = -0.95$, $P = 2.2e-16$), meaning that the larger difference in computed ACI values occurs in conjunction with quieter sounds.

This can likely be attributed to difference in precision discussed in section C.2. Quieter sounds — those with lower median amplitudes — produce smaller values in an FFT. These smaller signals amplify the precision errors that arise due to the 16-bit operation of the microcontroller, leading to greater differences in the final ACI calculation.

Appendix D FFT Filtering

Farina et al. (2016) introduce two modifications to the original ACI algorithm. They discuss the presence of artifacts in the FFT caused by microphone noise, analog-to-digital converter (ADC) errors, and introduced electronic noise and how these artifacts can cause inaccurate measurements of acoustic complexity.

The first modification is the addition of fixed-value threshold to eliminate spurious pulses from the FFT matrix. Any FFT values below the threshold are discarded and replaced with 0 in the FFT.

A second modification then eliminates these erroneous values from the overall ACI calculation. In the calculation of d_k (eq. (3)), the absolute difference between adjacent values in the FFT matrix, the calculation is treated as 0 if either of the values is 0. The modified version of eq. (3) is shown in eq. (6).

$$d_k = \begin{cases} 0 & \text{if } I_k = 0 \text{ or } I_{k+1} = 0 \\ |I_k - I_{k+1}| & \text{else} \end{cases} \quad (6)$$

This modified version of the ACI is used in the latest version of SoundscapeMeter (2.0). The option to perform this type of filtering has also been added to the latest

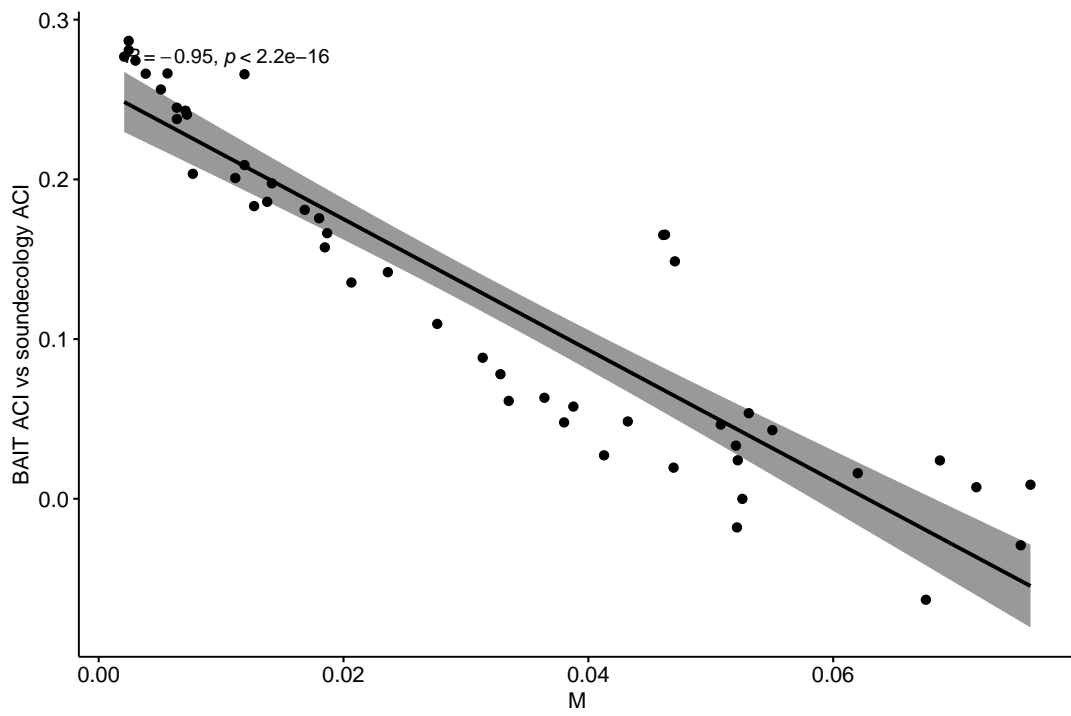


Figure 12.: Spearman's rank-order correlation for the difference in calculated ACI and the median of the amplitude envelope of a sound file. The two have a strong negative correlation ($\rho = -0.95, P = 2.2e-16$) indicating that differences in computation are associated with quieter sounds.

version of the BAIT firmware, found at github.com/dkedish/BioAcousticIndexTool, for compatibility. It is enabled by setting the `doFilter` and `doDiscardAdjacentZeros` flags to `true` when instantiating the `ACI_TemporalWindow` class in the main function.

Article 3

An artificial life approach to studying niche differentiation in soundscape ecology

Originally published in: ALIFE 2019: The 2019 Conference on Artificial Life

Joint work with: Sebastian Risi and Laura Beloff

Note The original article published in the ALIFE 2019 conference proceedings included an erroneous image in Figure 3a. This was corrected in a revised version of the paper uploaded to arXiv (<https://arxiv.org/abs/1907.12812v1>). That revised version is included in the dissertation.

An artificial life approach to studying niche differentiation in soundscape ecology

David Kadish and Sebastian Risi and Laura Beloff

IT University of Copenhagen, Copenhagen, Denmark

davk@itu.dk

Abstract

Artificial life simulations are an important tool in the study of ecological phenomena that can be difficult to examine directly in natural environments. Recent work has established the soundscape as an ecologically important resource and it has been proposed that the differentiation of animal vocalizations within a soundscape is driven by the imperative of intraspecies communication. The experiments in this paper test that hypothesis in a simulated soundscape in order to verify the feasibility of intraspecies communication as a driver of acoustic niche differentiation. The impact of intraspecies communication is found to be a significant factor in the division of a soundscape's frequency spectrum when compared to simulations where the need to identify signals from conspecifics does not drive the evolution of signalling. The method of simulating the effects of interspecies interactions on the soundscape is positioned as a tool for developing artificial life agents that can inhabit and interact with physical ecosystems and soundscapes.

Introduction

Artificial life experiments have become important tools for exploring biological phenomena. In particular, they have allowed researchers to study the relationships between evolutionary processes and ecological theories (Aguilar et al., 2014), like the emergence of interspecies relationships like mutualism and parasitism (Watson et al., 2000).

One area of ecology that has received little attention thus far from artificial life (ALife) studies is soundscape ecology. The field of soundscape ecology has been formalized by researchers over the past decade (Pijanowski et al., 2011a), building on earlier conceptions of the soundscape (Schafer, 1977). One of its foundational theories is the acoustic niche hypothesis (ANH) (Krause, 1987), which applies the concept of ecological niches — the distribution of resources that are used by a species in an ecosystem (Pocheville, 2015) — to the soundscape.

The manuscript is slightly amended from the published version to correct an error in Figure 3a. The figure in the published version plots the messages from a different run of the simulation than the one shown in Figure 3b. This in no way changes the results of the study.

This experiment tests the proposed mechanisms for the formation of these niches in a virtual soundscape in order to understand how species change vocalizations in response to one another. It models the behaviour of two species in a virtual ecosystem and tracks how their calls shift through the audio spectrum in response to different evolutionary pressures. Through the experiment, evolutionary pressure to communicate within a species is found to play a significant role in the formation of acoustic niches.

In examining the emergence of communication between artificially evolved species, this study draws from a body (Arita and Koyama, 1998; Wagner, 2000; Sasahara and Ikegami, 2007) of ALife-based studies of communication including the work of Floreano et al. (2007) in emergent communication between robotic agents. However, it is distinct from these previous studies in its focus on the effect of the emergent communication on the ecological phenomena of niche differentiation.

The main contributions of this study are the development of a simplified model of a soundscape for the purpose of rapid experimentation and in-depth analysis of population-soundscape dynamics, and the demonstration of the ANH on this model.

Background

In the physical world, the concept of soundscape — the collection of the acoustic features of a landscape — has roots and influences in a diverse array of academic fields (Lyonblum, 2017). It grew initially out of the arts and cultural studies work of Westerkamp (1974), Schafer (1977), and Truax (1978), but has since expanded into the sciences. In the field of ecology, the soundscape is considered an important ecological resource and its composition is thought to indicate the diversity and stability of the ecosystem (Pijanowski et al., 2011b). Though the field of soundscape ecology was only proposed relatively recently (Pijanowski et al., 2011b), the application of ecological principles to the study of soundscape has a longer history. Notably, the concept of ecological niches was first introduced in the context of sonic resources by Krause

(1987) as the acoustic niche hypothesis.

Acoustic niche hypothesis (ANH)

The acoustic niche hypothesis expands the concept of ecological niches to the spectro-temporal plane of the soundscape. Krause proposed that, in the same way that niche differentiation leads to species making use of the range of physical resources available in an ecosystem, species tend to differentiate their use of an ecosystem's sonic resources. This differentiation, according to Krause, occurs spectrally in the sonic frequencies that animals use for vocalization and temporally in the time-based patterns of their sounds. The theory holds that older, more mature ecosystems should show a greater degree of differentiation between the auditory niches that long-established species occupy.

The ANH describes the result of acoustic differentiation, but Endler (1992) proposed the primary mechanism for this evolutionary driver: sexual selection based on a mate's ability to hear a call and the ability to maintain territory. In this formulation, vocalizations and auditory receptors have co-evolved to maximize the reception of signals from members of one's own species (conspecifics), while minimizing interference from members of other species (heterospecifics).

This type of spectral differentiation has been observed numerous times in the wild: in the calls of certain species of frogs (Feng and Schul, 2007); in the buzzing of cicadas (Sueur, 2002); and in the overall division of a soundscape in Borneo between a series of birds, gibbons, and accompanying insects (Krause, 2008). However, it has proved difficult to experimentally probe the formation and division of spectral niches, due to the lengthy timescales that would be required to allow evolutionary processes to progress (Miller, 1995) and the complexity of the systems and soundscapes that are encountered "in the wild" (Wheeler et al., 2002).

ALife approaches

Where ecological phenomena have been difficult to experimentally investigate, researchers have proposed that ALife approaches can be a mode of inquiry that allows for the manipulation of particular conditions and the rapid collection of large quantities of data about a simulated ecological system (Miller, 1995). In 2018, Eldridge and Kiefer proposed synthetic acoustic ecology (SAC) as a toolset for exploring questions in the field of soundscape ecology using ALife methods in virtual ecosystems. Their study examined one of the assertions of ANH (Krause, 1987) — that one can identify the maturity of an ecosystem by examining its acoustic signature. Using a multi-agent system model, they demonstrated that patterns emerge in two common acoustic indices that indicate the stability of a model ecosystem.

Niche differentiation mechanisms

The study in this paper uses a virtual soundscape to test hypotheses in soundscape ecology, building on the work of Eldridge and Kiefer (2018). While Eldridge and Kiefer's study focused on the verification of acoustic biodiversity metrics, this study examines the mechanisms that breed interspecific diversity and intraspecific convergence in the vocalizations of communities in a soundscape. In particular, it is designed to test Endler's hypothesis (1992) that the ability to identify vocalizations from members of the same species drives acoustic niche differentiation.

The acoustic niche hypothesis posits that soundscapes niches are differentiated on both spectral and temporal levels, so that species ensure that their calls are isolated in both frequency and time. In order to simplify the modelling and analysis and to allow for a deeper examination of the effects of differentiation, this study focuses only on the spectral component of this differentiation.

Approach

The experimental setup for testing the drivers of acoustic niche differentiation consists of a set of evolving populations and a soundscape that they communicate within. The experiment tests two hypotheses: the alternative hypothesis (H_1), that acoustic niche spectral differentiation is driven a need to identify signals from potential mates or territorial rivals of the same species; and the null hypothesis (H_0) that spectral differentiation in acoustic niches is not driven by the need to distinguish the species of the signaller.

In order to facilitate rapid experimentation and ease the analysis of the emergent signalling systems, the experiments use a simplified, discretized model of a soundscape instead of a full-spectrum, temporally-varying acoustic space. Sounds are modelled as 9-bit vectors that represent the use of 9 available frequency bands in an instantaneous signal. These simplifications allow the repetition of the experiments many times with a large number of generations and individuals, such that results reflect general trends in the dynamics of these systems rather than the peculiarities of any single simulation. The entire system is illustrated in Figure 1 and described in detail in the sections below. Lettering in brackets refers diagram labels in Figure 1.

Populations

In soundscape ecology in the physical world, the communicative process is often assessed in two parts: sender and receiver. Every individual, of course, is both sender and receiver, but the processes experience different evolutionary pressures; "[n]atural selection favors signals that elicit a response in the receiver that increases or maintains the fitness of the sender" (Endler, 1992). The same is true in reverse, such that the sender and receiver of a particular species evolve alongside one another, but with slightly different driving forces.

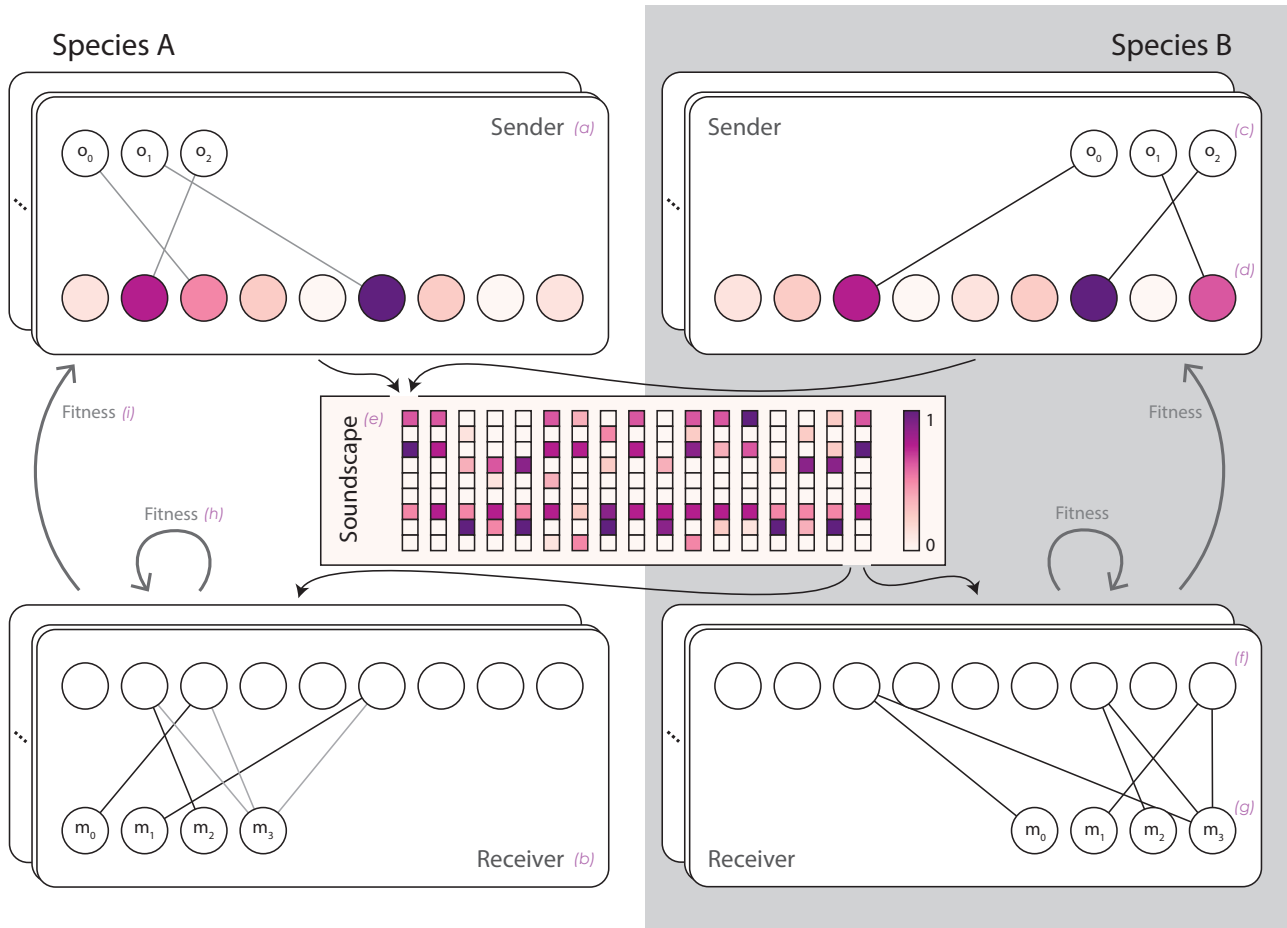


Figure 1: The experimental setup. Senders (a) encode a 3-bit message ($o_{0..2}$) into the 9-band soundscape (e) using a neural network with 3 inputs (c) and 9 outputs (d). Receivers (b) “hear” encoded messages from all species’ senders and predict the original message ($m_{0..2}$) and whether the message originates from a conspecific (m_3) using the 9 inputs (f) and 4 outputs (g) of their neural networks. Sender fitness (i) depends on how well conspecific receivers identify their species and decode their messages. Receiver fitness (h) depends on how well they identify the species of all senders and how well they decode messages from conspecifics.

The populations in this experiment are modelled as artificial neural networks, which are optimized with the neuroevolution of augmenting topologies (NEAT) algorithm (Stanley and Miikkulainen, 2002). NEAT models individual phenotypes as neural networks with a fixed number of inputs and outputs and an evolvable internal structure and connectivity. This allows the population to begin with simple neural structures and to evolve complexity as necessary to achieve the task.

For the experiment presented here, each species actually consists of two NEAT populations: a population of senders (a in Figure 1) and a population of receivers (b in Figure 1). The populations communicate over a simulated soundscape (e) that consists of 9-bit vectors, interpreted as acoustic frequency bands which can be used by senders to transmit messages. The senders encode 3-bit messages ($[o_0, o_1, o_2]$) into a representation in the 9 frequency bands using their 3-input (c), 9-output (d) neural network structures. The structure of 2 species encoding 3-bit messages into a 9-band soundscape allows for the development of relatively complex messaging while allowing the soundscape to remain undersaturated as each species could theoretically communicate in only 3 of the 9 bands. The frequency bands form the inputs to the 9-input (f), 4-output (g) receiver neural networks. The first 3 outputs ($[m_0, m_1, m_2]$) of the receiver network are its estimation of the original message and the final output represents the receiver's prediction of whether the message comes from a conspecific ($m_3 \geq 0.5$) or from a heterospecific individual ($m_3 < 0.5$).

The soundscape (e) is shared among species but messages are received serially in order to decouple timing effects; therefore, each receiver "hears" messages from the senders of all of the present species, but receives them one at a time. Additionally, any spatial arrangement of the individuals is not considered as part of this experiment, so each receiver "hears" the signals from every sender at the same "volume" with no attenuation due to a distance or set of obstacles between them.

Fitness

In a communicative process, the evolutionary pressure on senders and receivers is related but differs in some crucial aspects. The fitness functions used in this experiment reflect these differences. Since communication for mating and territorial maintenance is hypothesized to drive acoustic differentiation (Endler, 1992), the sender is indifferent to how its messages are interpreted by receivers from other species. The receiver, however, processes all messages regardless of their origin; it has to learn how to differentiate messages from conspecifics from those of heterospecifics.

Following this reasoning, the fitness of the sender (i) is formulated to reflect how well its message is understood — or correctly decoded — by the receivers of its own species; it does not depend on how the receivers of the another species

process its messages. The fitness of the receiver (h) reflects both how well it is able to distinguish the species of the sender *as well* as whether it is able to correctly decode the message.

The ability of a receiver to perform these two tasks — identifying messages from conspecifics and decoding messages — is formulated into components of the the fitness function as f_s (species identification fitness, Equation 1) and f_d (message decoding fitness, Equation 2). m is the decoded message where the first three components ($m_{0..2}$) are message as decoded by a receiver. The fourth value output by the receiver (m_3) determines whether the receiver has identified this message as coming from a conspecific ($m_3 > 0.5$) or from a member of another species. The original message is a three-bit string represented by o_i .

$$f_s(m) = \begin{cases} f_{adj}(1 - |1 - m_3|) & \text{if same species} \\ f_{adj}(1 - |0 - m_3|) & \text{if different species} \end{cases} \quad (1)$$

$$f_d(m) = 3 * \prod_{i=0}^2 f_{adj}(1 - |o_i - m_i|) \quad (2)$$

To achieve the desired fitness formulations, these equations are applied in different ways for senders and receivers by adjusting the enabling/disabling coefficients e_s and e_d in Equation 4. For each message produced, a sender's fitness is based on the interpretation of the message by all receivers *from its own species*. Equation 4 is applied for each receiver from the sender's species with $e_s = 1$. The value of e_d depends on whether the species is identified incorrectly ($e_d = 0$) or correctly ($e_d = 1$). If the species is incorrectly identified, then the interpretation of the message is of no consequence, which is why the fitness of the message decoding is ignored.

Receivers "hear" messages from the senders from both species and their ability to identify and ignore messages that are not from their species is an important component of their fitness. For each message that a receiver "hears", f_s is calculated as part of its fitness ($e_s = 1$). If the receiver correctly identifies that a message originated from a member of its own species, it receives an additional score for decoding the bits of the original message (f_d) and a bonus multiplier (f_b) for correctly identifying multiple bits ($e_d = 1$), as described in Equation 4.

$$f_b(e_s, N) = \begin{cases} 1.0 & \text{if } e_s = 0 \\ \prod_{i=0}^N \left(\frac{i}{10} + 1\right) & \text{if } e_s = 1 \end{cases} \quad (3)$$

$$f_t = (e_s f_s(m) + e_d f_d(m)) * f_b(e_s, N) \quad (4)$$

One detail that requires some explanation is the adjustment function (f_{adj}) applied to the fitness equations for species identification (f_s) and message decoding (f_d).

The results that these equations evaluate are treated as binary in the operation of the system but the receivers produce output as decimal numbers between 0 and 1. If the receiver outputs $m_3 = 0.6$ for a message from a member of its own species, the consequence is no different from $m_3 = 1.0$ — the receiver has correctly decided that the message should not be ignored. However, an application of Equation 1 without f_{adj} would result in quite different fitnesses for the two outputs. Equation 5 creates a sharp rise in the fitness, centred around a value of 0.5 without producing a discontinuity, which was found to create an effective fitness landscape for the evolutionary process.

$$f_{adj}(x) = \frac{1}{2} (\tanh(8.0 * (x - 0.5)) + 1) \quad (5)$$

Null model and hypothesis

The model used to test the null hypothesis (H_0) uses a modified formulation of the fitness functions. The null hypothesis is that the need to identify messages from members of the same species *does not* play a role in niche differentiation. Therefore receivers are assumed to be able to know *a priori* which messages come from senders of their own species and no fitness is assigned for the task of species identification in this null model.

In the null version of the model, this results in the receivers only processing messages from members of their own species and ignoring messages from the other species. Senders and receivers are evaluated with the fitness function in Equation 4 with $e_s = 0$.

Results

We ran simulations of our ecosystem with senders and receivers for two species. Each population consisted of 50 individuals and the simulation was run for 300 generations. The results discussed here are averages and standard deviations from 20 independent simulations. Additionally, the results from a representative example simulation are highlighted in figures and throughout this section in order to discuss specific features of an individual simulation.

For each simulation, we generated spectrograms that mirror the type of chart that is often presented in studies of soundscapes (Krause, 1987; Pijanowski et al., 2011a), except that the x -axis of these plots represents generations instead of real-time auditory signals. These diagrams, such as the one seen in Figure 2, show how the two species's use of the frequency bands shifts from generation to generation. The initial populations's encoded messages are randomly distributed across the 9 frequency bands, but the signals converge over the course of the first 50 to 100 generations into a subset of bands used primarily by one species. In this example, after an initial series of about 100 generations, both species show consistent use of 3 bands — 0, 2, and 4 for *Species A* and 1, 5, and 7 for *Species B* — for the remainder of the simulation. *Species A* develops and then eventually

scales down the use of band 3 and band 8, but *Species B*'s use of 1, 5, and 7 remains remarkably stable through most of the latter 200 generations.

The spectral plots provide a useful visual representation of the divergent signals, but the actual level of separation can be quantified further and visualized in another manner. Figure 3a shows a mapping of the high-dimensional messages to two-dimensions using t-distributed Stochastic Neighbour Embedding (t-SNE), plotted for particular generations of interest. The encoded messages generated by senders from the two species can be seen to rapidly separate into clusters from an initial state of near-total overlap. This can be further examined in the plot below the cluster maps (Figure 3b) which shows the silhouette score for the clusters over the course of generations. The silhouette score is used in the evaluation of clustering algorithms and is a measure of the density of clusters (Rousseeuw, 1987), where a score of 0 indicates overlapping data and a score of 1 indicates dense and well-separated clusters. The rapid rise of the silhouette score here indicates the splitting of the spectrum audio spectrum between the senders in relatively few generations.

The plot shows the average and standard deviations of the silhouette scores from the series of 20 trials of H_1 (dark grey) alongside the silhouette score from the specific run from which the clusters in the plot above were derived (pink). In addition, it shows the average and standard deviation of silhouette scores from 20 trials of the null model H_0 (light grey). A test of the hypotheses using Welch's t-test — because the variance of the samples cannot be assumed to be equal — reveals that the difference between the two models is significant after generation 4 ($P < 0.01$), with an average P-value of 15×10^{-5} for latter 295 generations.

While the null model does produce a level of clustering of the species' messages, this is to be expected as a result of the selection of frequency bands on which to communicate. However, in the null model, this selection is not competitively driven by the presence of the other species. In H_1 , the receivers of the two species drive their senders towards diverging frequency bands as their fitness increases with their ability to identify messages from their own species and reject those from the other.

We also examined the actual performance of the species with regard to their ability to recognize and decode messages from their conspecifics. Figure 4 shows the scores of the senders and receivers from a species over the course of 300 generations. On average, the proportion of messages that are correctly identified as being from members of the same or other species (red) rises sharply in the first generations before steadying near 80%. The proportions of bits that are correctly decoded and messages that are fully decoded correctly are slower to rise, but continue to do so throughout most of the evolutionary process.

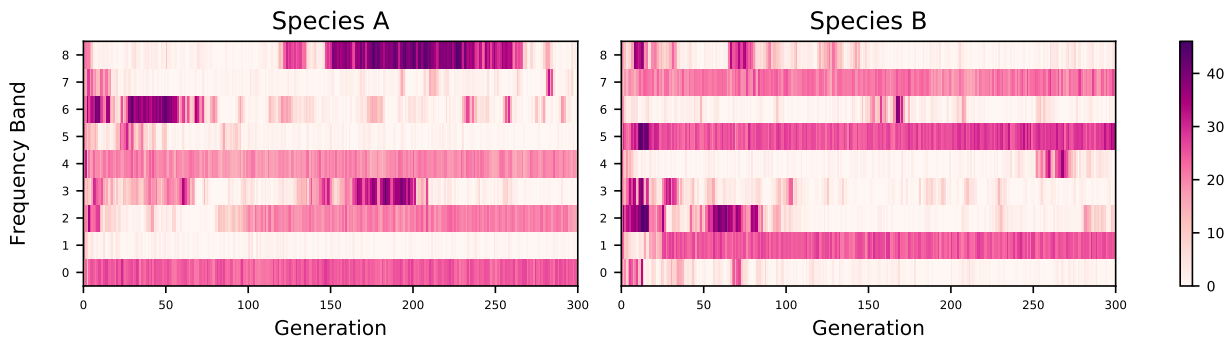


Figure 2: The use of the 9 frequency bands by the population of senders in the example simulation of H_1 . The graph shows the changing use of the frequency bands over 300 generations. In the first generations, both species’s signals are spread across the 9 bands such that the signals from the two species overlap. These signals converge rapidly to a smaller subset of the available bands. By about the 50th generation, there is little overlap between the two species — *A* uses mainly bands 0, 4, and 6 while *B*’s signals are concentrated on 1, 5, and 7 — though there is some use of bands 2 and 3 by both species. In this example, *Species B* uses band 2 heavily but intermittently until just before the 100th generation, when it ceases almost all activity on the channel and *Species A* begins to make consistent use of it for the remainder of the simulation. By the generation 300, both species have converged to the near-exclusive use of 3 channels: 0, 2, and 4 for *Species A* and 1, 5, and 7 for *Species B*.

Discussion

The results presented in the previous section demonstrate that it is possible to drive spectral differentiation in the acoustic signature of an agent through an impetus to communicate with other members of the same species. An analysis of the distance between intraspecies messages and interspecies messages shows a significant difference between the test of the alternative hypothesis (H_1) and the null hypothesis (H_0), as seen in Figure 3b. Moreover, Figure 2 provides a visual reference for the division of the spectrum in a selected simulation of the alternative hypothesis (H_1). The spectrum has been split between the two species after the first 100 generations, such that *Species A* primarily makes use of bands 0, 2, and 4 while *Species B* relies on bands 1, 5, and 7. It is interesting to note that, in the first 100 generations, band 2 is used mainly by *Species B*, however this changes around generation 90 as *Species A* begins to use the band regularly. Once *Species A* establishes regular use of the band, *Species B* never returns to it with any stability for the remainder of the simulation.

In models of the null hypothesis (H_0), the two species occasionally achieve a level of differentiation of their messages, however this occurs only by chance. In both models, species tend to converge to the primary use of roughly 3 of the 9 available channels for communication. Three channels is the fewest that can be used to encode the three-bit message and it is often the easiest solution for the evolving neural networks to find. However, in the null case, the channel selection is not driven by competition between the species, only by cooperation within a species. This lack of competition often leads to overlapping channel selections, which in turn, is responsible for the lower silhouette scores

for the null models (Figure 3b).

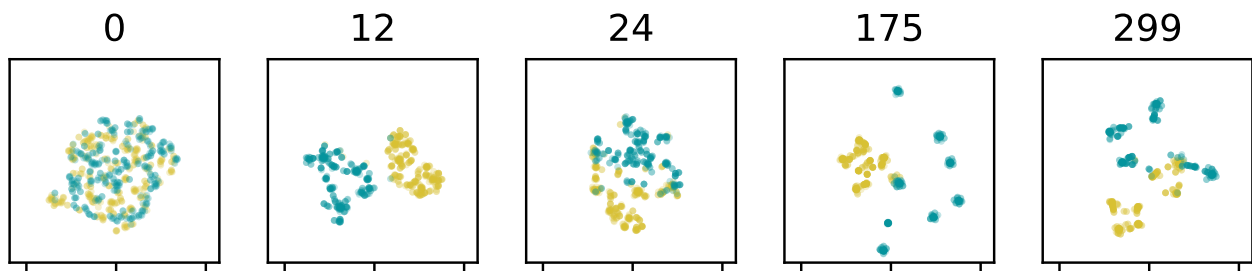
While these results cannot be taken as confirmation of the proposed mechanism of the ANH, they demonstrate that the mechanism is plausible. The drive to produce signals that are identifiable and understandable to members of one’s own species within the finite resource that is a soundscape results in the formation of acoustic niches for vocalizing species.

This study also demonstrates the efficacy of a highly simplified model in demonstrating the plausibility of a particular mechanism for the formation of patterns within a soundscape. It compliments the work of (Eldridge and Kiefer, 2018), which explores the way that common acoustic indices respond to changing populations and signals, and presents another application for a synthetic acoustic ecology. Together with other types of computational studies of soundscapes (Eldridge and Kiefer, 2018), this paper lays the foundation for a method of rapidly interrogating evolutionary acoustic processes. In addition to providing insight into ecological studies, research in this area can also be used to inform the development and analysis of evolutionary acoustic agents live “in the wild” and interact with biological ecosystems.

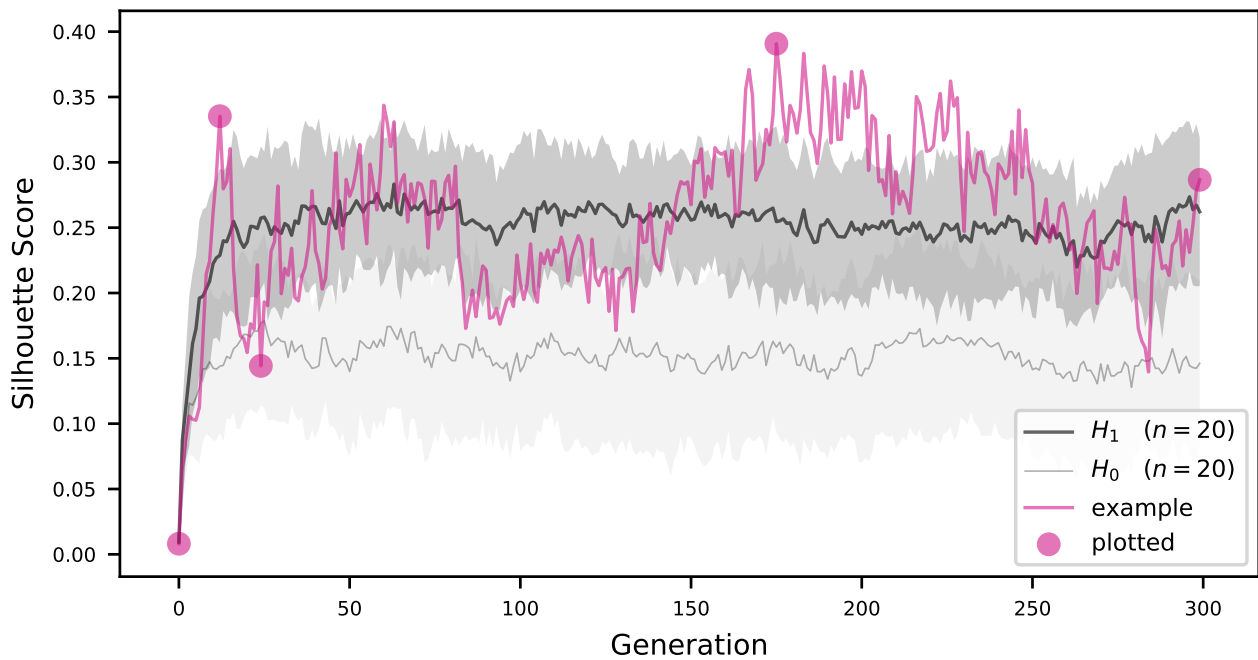
Conclusion

Though the experiment presented here is based on a highly simplified model of a physical ecosystem, it demonstrates that it is possible to rapidly and repeatedly test some of the basic principles of soundscape ecology. As predicted, the experiment was able to demonstrate the important role of intraspecies communication in the partitioning of the acoustic resources of an ecosystem.

This has important implications for the development of



(a) Cluster diagrams of messages in selected generations, mapped to 2 dimensions using t-SNE. Each point is a message generated by a sender in a single generation (labeled above the plot) of the simulation. The selected generations are marked on the silhouette score plot below with pink circles. The different colours represent messages originating from members of the two different species. The messages in generation 0 are scattered randomly from both species as the initial neural network connections for the senders are randomly generated. The messages rapidly converge to two clusters by generation 12. However, these clusters are still evenly spaced internally, as the initial selection pressure is mainly to differentiate messages between the two species. In later plots, for example in those from generations 175 and 299, smaller clusters form within the messages from a single species as the senders from each species converge on representations for particular bits and messages. This clustering drives the increasing bit and total scores in Figure 4.



(b) Silhouette score of the encoded messages, grouped by species, over the course of 300 generations. Scores reflect the validity of the message clusters when grouped by species, averaged over 20 runs of the simulation, and plotted with the standard deviation in the background. An example of an individual run is also plotted (pink) and the generations of that run that are plotted in the cluster diagram above are noted. The difference between the alternative hypothesis (H_1) and the null hypothesis (H_0) is significant ($P < 0.01$) after generation 4. The average P-value after generation 4 is 15×10^{-5} .

Figure 3: Cluster validity scores over 20 runs of the simulation. Message clusters are shown above for selected generations of an example run.

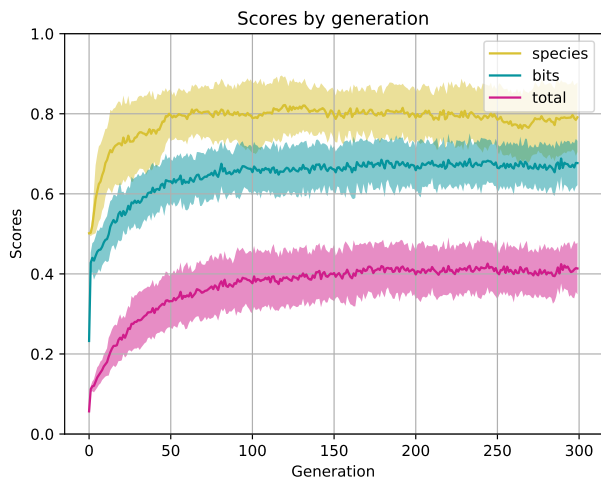


Figure 4: Performance of a species, showing the receivers' ability to identify the species (yellow), and the rate at which sender-receiver pairs were able to correctly identify individual message bits (blue) and the entire message (pink).

hardware-based ALife agents for the production of sound in a physical, hybrid ecosystem. It suggests that, if one of the goals of that agent is to identify a niche for itself in the soundscape, it is important to co-evolve the auditory production with auditory perception to drive the vocalizations into an empty portion of the spectrum.

In a broader sense, this experiment sets out the foundation for a method of testing ideas for hardware-based agents in software simulations to understand the possible dynamics once they are released in the field. It grounds the inquiry into a complex phenomenon with a concrete example that solidly demonstrates the theoretical basis for a physical experiment through repetition and statistical analysis on a scale that is difficult to achieve in the field. And it demonstrates the feasibility of a key theory in soundscape ecology.

Acknowledgements

Thanks to the REAL Lab, especially Kasper Støy, Rosemary Lee, Djordje Grbic, Miguel Gonzalez Duque, Mads Johansen Lassen, and Niels Justeen for their suggestions.

References

Aguilar, W., Santamaría-Bonfil, G., Froese, T., and Gershenson, C. (2014). The Past, Present, and Future of Artificial Life. *Frontiers in Robotics and AI*, 1:8.

Arita, T. and Koyama, Y. (1998). Evolution of Linguistic Diversity in a Simple Communication System. *Artificial Life*, 4(1):109–124.

Eldridge, A. and Kiefer, C. (2018). Toward a Synthetic Acoustic Ecology: Sonically Situated, Evolutionary Agent Based Models of the Acoustic Niche Hypothesis. In *The 2018*

Conference on Artificial Life, pages 296–303, Cambridge, MA. MIT Press.

Endler, J. A. (1992). Signals, Signal Conditions, and the Direction of Evolution. *The American Naturalist*, 139:S125–S153.

Feng, A. S. and Schul, J. (2007). Sound processing in real-world environments. *Springer Handbook of Auditory Research*, 28(Bregman 1990):323–350.

Floreano, D., Mitri, S., Magnenat, S., and Keller, L. (2007). Evolutionary Conditions for the Emergence of Communication in Robots. *Current Biology*, 17(6):514–519.

Krause, B. L. (1987). Bioacoustics, Habitat Ambience in Ecological Balance. *Whole Earth Review*, (57):14 – 18.

Krause, B. L. (2008). Anatomy of the soundscape: Evolving perspectives. *AES: Journal of the Audio Engineering Society*, 56(1-2):73–80.

Lyonblum, E. Z. S. (2017). *An Omnivorous Ear: The Creative Practice of Field Recording*. Doctor of philosophy (phd), University of Cambridge.

Miller, G. (1995). Artificial life as theoretical biology. How to do real science with computer simulation. *Cognitive Science Research Paper*, 378(5):1–33.

Pijanowski, B. C., Farina, A., Gage, S. H., Dumyahn, S. L., and Krause, B. L. (2011a). What is soundscape ecology? An introduction and overview of an emerging new science. *Landscape Ecology*, 26(9):1213–1232.

Pijanowski, B. C., Villanueva-Rivera, L. J., Dumyahn, S. L., Farina, A., Krause, B. L., Napolitano, B. M., Gage, S. H., and Pieretti, N. (2011b). Soundscape Ecology: The Science of Sound in the Landscape. *BioScience*, 61(3):203–216.

Pocheville, A. (2015). The Ecological Niche: History and Recent Controversies. In *Handbook of Evolutionary Thinking in the Sciences*, pages 547–586. Springer Netherlands, Dordrecht.

Rousseuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65.

Sasahara, K. and Ikegami, T. (2007). Evolution of Birdsong Syntax by Interjection Communication. *Artificial Life*, 13(3):259–277.

Schafer, R. M. (1977). *The tuning of the world*. A.A. Knopf.

Stanley, K. O. and Miikkulainen, R. (2002). Evolving Neural Networks through Augmenting Topologies. *Evolutionary Computation*, 10(2):99–127.

Sueur, J. (2002). Cicada acoustic communication: potential sound partitioning in a multispecies community from Mexico (Hemiptera: Cicadomorpha: Cicadidae). *Biological Journal of the Linnean Society*, 75(3):379–394.

Truax, B. (1978). *The World Soundscape Project's Handbook for acoustic ecology*. Music of the environment series. A.R.C. Publications, 1 edition.

Wagner, K. (2000). Cooperative Strategies and the Evolution of Communication. *Artificial Life*, 6(2):149–179.

Watson, R. A., Reil, T., and Pollack, J. B. (2000). Mutualism, Parasitism, and Evolutionary Adaptation. In Bedau, M., editor, *Artificial life VII : proceedings of the seventh International Conference on Artificial Life*, page 564. MIT Press.

Westerkamp, H. (1974). Soundwalking. *Sound Heritage*, 3(4).

Wheeler, M., Bullock, S., Paolo, E. D., Noble, J., Bedau, M., Husbands, P., Kirby, S., and Seth, A. (2002). The View From Elsewhere: Perspectives on ALife Modeling. *Artificial Life*, 8(1):87–100.

Article 4

Robophony: A new voice in the soundscape

Originally published in: RE:SOUND 2019 – 8th International Conference on Media Art, Science, and Technology; August 2019

Robophony: A new voice in the soundscape

David Kadish
IT University of Copenhagen
itu.dk
davk@itu.dk

Soundscape studies typically distinguish between three sources of sound: biophony, sounds produced by animals, plants, and other biological entities; geophony, sounds produced by non-living features such as water or wind; and anthrophony, sounds produced by humans and human technology. Recent developments in the study of hybrid ecosystems and ecological robotics challenge these categorizations. A series of four soundscape interventions are considered, which lead to the proposal of a new category: robophony. These interventions — robots and autonomous digital agents — operate in feedback loops with the existing bio-, geo-, and anthrophony in the ecosystem. The properties that emerge from these cases: site specificity, hybrid sourcing, and layered temporality form the basis of this new category of sound in the soundscape.

robophony, soundscape ecology, robots, NEAT

1. INTRODUCTION

The categorization of sound in a soundscape is necessarily imperfect. Sounds are fluid, overlapping and — in the age of digital reproduction — often removed from their sources and therefore difficult to group. And yet, the categorization of sound is an important part of the analysis of a soundscape. It helps researchers to understand how a sound enters a soundscape, what role it plays in structuring the sonic environment, and how it relates to other parts of the soundscape (Pijanowski, Villanueva-Rivera, et al., 2011). A set of thoughtfully constructed categorizations can help researchers and practitioners identify patterns and discuss the features of a soundscape in a meaningful way.

Schafer (1977) presented the first taxonomy of sound objects in his 1977 text “The tuning of the world”. The labels were based on the catalogue headings in a collection of literary descriptions of sounds gleaned from written documents that his team had gathered. The headings were compiled as entries were added to the catalogue, so that the taxonomy is more the result of a generative process than a deliberate organization. They capture the whole set of catalogued sounds into a three-layer hierarchy, with the top layer dividing sound objects into the supercategories of natural sounds, human sounds, sounds and society, mechanical sounds, quiet and silence, and sounds as indicators.

The field of soundscape ecology, which developed in part out of Schafer’s work, uses a different system

of categorizing sound sources. Bernard Krause, one of the pioneers of the field, used a tripartite division of biophony, geophony, and anthrophony to group sounds in a 2003 technical report on soundscape as an indicator of ecosystem health with Stuart Gage. The categories of geophony and anthrophony was split further in Krause’s “Anatomy of the soundscape: Evolving perspectives” (2008). Here, he used electromechanical, physiological, controlled, and incidental as the four subcategories of anthrophony, covering anthropogenic sounds from repetitive mechanical from whirring motors to the crunching of leaves underfoot as a person walks through the woods, while the subcategories of wind, water, weather, and geophysical forces subdivide geophonic sounds into more precise groups.

These categories are operationalized in different ways by Krause and his contemporaries. They are deployed in the production of a framework for understanding the dynamic relationships within a soundscape and as way of understanding which types of sounds dominate over different temporal and spatial configurations (Pijanowski, Farina, et al., 2011). They also help ecologists to measure overall ecosystem health (Pijanowski, Villanueva-Rivera, et al., 2011), understand the effects of different sounds of aspects of animal life (Pijanowski, Villanueva-Rivera, et al., 2011), and establish the relational dynamics between the different categories of sound (Gasc et al., 2017).

The model of these relationships in Pijanowski, Farina, et al. (2011) gives specific attention to the feedbacks between the whole soundscape and these categories of sound. This describes the way in which, for example, anthropogenic additions to the soundscape sounds can mask bird calls, prompting them to shift, thus prompting a secondary soundscape change.

One thing missing from this model, however, is a way of accounting for the pace of these feedback mechanisms. If one wants to understand the dynamics of a soundscape — which is a key goal of soundscape ecology — it is important to be able to discuss the speed at which different types of sound shift in response to other shifts in the soundscape.

It also places little emphasis on the biosemiotics of the sounds that are entering the soundscape. Under the current model, the sound of a car engine and the playback of recorded soundscapes over a loudspeaker are both treated as anthrophony, but the two sounds would likely be perceived in wildly different ways by inhabitants of an ecosystem, human or nonhuman.

This does not require a wholesale re-imagining of the system currently in use. Bio-, geo-, and anthrophonic sounds are largely internally consistent in terms of their feedback and response rates. However, a new set of actors is emerging that requires its own category in this formulation.

Learning robotic systems and autonomous agents, particularly those designed to engage directly with the sonic ecosystem, have the ability to change their behaviour in a manner similar to the ways that animals shift their sonic outputs in response to environmental changes. However, robots operate on sped-up timescales, shifting behaviours in ways that are unencumbered by established group dynamics and highly-specified morphologies. This paper proposes a new category of ecological sound, *robophony*, to capture this emerging class of sound objects.

What follows is a summary of the history of the contemporary categorizations of soundscape and their place in a dynamics-based conception of these categories. Examples are brought to motivate the creation of the category of robophony, going into detail about two cases which represent the primary motivators. Finally, the category of robophony and the issues that remain with its instantiation are discussed.

2. BACKGROUND

All categories represent an imperfect flattening of an idea space. At their best, however, categorizations and taxonomies of objects and phenomena can bring clarity and new analytical perspectives to a field of study. Since the formal definition of the soundscape in Schafer's *The tuning of the world* (1977), two main taxonomies have been used to group sounds. Schafer's own categorizations were presented in Chapter 9 of his book, while those used by Krause were developed as part of an understanding of soundscape ecology beginning in the late 1990s.

2.1. Schafer's Taxonomy

Schafer dedicates an entire chapter of his *The tuning of the world* to the subject of "classification". He notes that — depending on one's perspective — sounds might be categorized according to their acoustic, psychoacoustic, semiotic-semantic, or emotional-affective qualities. He first describes a system of classification according to physical characteristics — the duration, frequency, dynamics, internal fluctuations, mass, and grain — before noting that these describe isolated sound events, thus limiting its utility in the study of soundscape ecology.

Most pertinent to this discussion, however, is Schafer's subsequent description of a taxonomy based on the semiotic and semantic content of a sound. This taxonomy emerged from a process of cataloging written descriptions of sound. The World Soundscape Project (WSP) team had engaged in a lengthy process of collecting what Schafer calls earwitness accounts and sorting them into a catalogue. Schafer's taxonomy is drawn directly from the categories and subcategories that emerged from this study.

The taxonomy consists of six primary categories: *natural sounds*, *human sounds*, *sounds and society*, *mechanical sounds*, *quiet and silence*, and *sounds as indicators*. These categories — with the exception of *quiet and silence* — are broken into secondary and often tertiary subcategories. In total, 46 secondary subcategories and an extensive set of tertiary subcategories serve to categorize every sound that the WSP catalogued from written sources.

Because the taxonomy is derived from the cataloging process, it doesn't have a pre-conceived organizing principle. It is, however, biased by the types and origins of literary sources that are chosen as some of its subcategories reveal. There is, for example, a *town soundscapes* subcategory of *sounds and*

society which features a single tertiary subcategory: *Britian and Europe, etc.*

It is also important to note that the categories are based on a semiotic approach to sound. Specifically, the approach is anthroposemiotic (Emmeche, 2007), prioritizing the human understanding of meaning from sound, as evidenced by subcategories such as *eating* (Schafer, 1977). This approach is useful for creating categorizations of sound for humans, but is less relevant to the categorization of sounds in soundscape ecology. In ecological terms, a biosemiotic approach that accounts for the interpretation of sonic signs by nonhumans in the ecosystem is crucial. A human semiotic category does not necessarily have any meaning to other ecosystem inhabitants. The study of soundscapes in an ecological context necessitates a new taxonomy.

2.2. The Krause-Gage Taxonomy

The field of soundscape ecology was not formally defined until a special issue of the journal *Landscape Ecology* in 2011, but the taxonomic framework used by soundscape ecologists has roots in much older work. From their introduction to the field, (Pijanowski, Farina, et al., 2011) use a tripartite system for grouping sounds by origin: geological, biological, and anthropogenic. These terms have roots in a 1997 book by Krause and a 2003 report by Krause and Gage.

Unlike Schafer, Krause and Gage are deliberate in the purpose behind the construction of these categories in their 2003 report. In an analysis of the relationship between acoustic indicators and ecosystem states, they point out that the “classification will assist in identifying the introduced elements that may cause stress or change not otherwise noticed by traditional visual evaluation”.

The notion of what is “introduced” provides a clear link to work in landscape ecology and conservation such as (Wiens, 2009). It sets up the later development of soundscape conservation (Dumyahn and Pijanowski, 2011), in which the authors discuss the need to preserve “natural sounds”.

For their part, Krause and Gage are clear about which sounds they consider to be natural: “In most environments today, soundscape signatures are comprised of two natural components, biophony and geophony, and a probable human component that includes the third, anthropony” (2003).

This characterization immediately raises some significant conservation questions. Are the biophonic sounds produced by introduced species themselves not considered introduced? Are they “natural”

in this analysis? What about the reproduction of biophonic sounds through loudspeakers and playback devices? If a recorded bird song is played over a speaker within their native range, is that considered biophonic or anthroponic? Is it something introduced or is it a part of the natural soundscape?

2.3. Hybrid Ecosystems

These are important underlying questions in soundscape ecology. There are objects and sounds that blur these lines between the categories in the Krause-Gage taxonomy. They complicate the identification of “introduced” and “natural” elements in the soundscape. These works challenge the understandings of soundscape and landscape conservation that underpin the basic sonic taxonomy of soundscape ecology.

Ecologists have already begun to acknowledge these new ecological coalescences. Hobbs, Higgs, Hall, et al. (2014) propose that, instead of describing them as natural or anthropogenic, ecosystems can be described as *historical*, *hybrid*, and *novel* to varying degrees. This reflects the understanding that few ecosystems — including vast tracts of forest that are often regarded as pristine, like the Amazon (Roosevelt, 2013) — are free of human influence and that the state of an ecosystem is in constant flux (Hobbs, Higgs, and Harris, 2009). The notion of a historical ecosystem considers whether an ecosystem is currently situated within a dynamic range that correlates with the past or whether human activity has fundamentally shifted the ecosystem’s composition.

An ecosystem where, for example, a particular species has gone extinct or a new species has been introduced is perhaps no longer historical, but could retain many of its historical dynamics with some human intervention. Then, it is considered hybrid: not exactly within its former range, but retaining much of its former character or significant properties. Novel ecosystems have undergone a wholesale change in composition — perhaps they have lost an irreplaceable keystone species or it is the site of a rehabilitated mine — so it may not be possible to reconstitute its former makeup.

Conservation priorities might be set differently for each type of system such that maintenance of a historical state may be the priority in a historical ecosystem, hybrid and novel ecosystems can be managed to conserve a general ecosystem function or a dynamic biodiversity. In the context of hybrid and novel ecosystems, anthropogenic sound — like other anthropogenic interventions — might play a role in supporting or reconstructing the ecosystem and its

soundscape. These types of interventions would be complex and fraught, but the anthropogenic sounds that could be used might not have the kinds of mostly negative connotations that Krause tends to associate with the anthrophonic sounds of car engines and airplane overflights.

2.4. Ecobots

What could produce the kinds of sonic interventions that might contribute to the soundscapes of hybrid and novel ecosystems? The introduction of a new biological species — either intentionally or unintentionally — would likely make an impact on the soundscape. But this discussion focuses on another potential sonic actor: robots.

In their (2017) paper, Wynsberghe and Donhauser discuss a category of robots that act in ecological manner which they call ecobots. These are robots whose design purpose is to perform an ecological function, such as helping to manage the proliferation of an overpopulated species or consuming toxins from a contaminated ecosystem.

The concept of ecobots is important here because they stand apart from what Wynsberghe and Donhauser (2017) call robots-in-ecology. These are robots that happen to be operating in ecosystems, but that aren't necessarily responsive to the ecosystem or performing an ecological role. Any sounds that these robots-in-ecology produce, such as the sound of the rotor of a drone that is hovering over an ecosystem, would fall into the same anthrophonic category as that of a ATV engine.

Ecobots, however, have a different relationship with the ecosystem. Their actions are responsive to changing conditions in the ecosystem. They become part of the ecosystem's flows and feedbacks in a manner that is in some ways similar to that of an animal. Their contributions to the soundscape are not quite the same as those of a robot-in-ecology. Their sounds carry different meaning.

3. ROBOTS IN THE SOUNDSCAPE

This section explores these provocations through a series of four case studies, the last of which is a work in progress by the author. Each of the cases discussed here has a digital component operating with some degree of autonomy, producing sound in an outdoor ecosystem and could be considered to be ecobots.

3.1. David Dunn

David Dunn is a composer and researcher whose work has often involved computer-mediated sonic

engagements with ecosystems. Two works are of particular interest here: *Sonic Mirror* from 1986-1987 and *Autonomous Systems* from 2003-2005.

Both works are attempts to engage soundscapes, using computers as part of a sonic feedback mechanism. Their roots can be traced to the rise of cybernetics and the availability of portable computing technologies, and they are steeped in cybernetic concepts such as feedback, complexity, and emergence.

Sonic mirror can be understood as an attempt to insert computational node within the larger ecosystem. "The original concept was conceived as a stationary cybernetic sound sculpture capable of processing acoustic data within an outdoor environment. Eventually the sculpture might function as an autonomous system structurally coupled to its surrounding environment in a manner that might allow for 'learning' between components" (Dunn, 2002).

Autonomous Systems is something of a follow-up to *Sonic Mirror*. In this work, the soundscape is recorded, processed, and replayed in a repeating, cyclical process. As the work records and replays its own shifting of the soundscape, the animals in the ecosystem engage it and shift their behaviour, and thus the behaviour of the machine as well.

These works are interesting as early examples of digital sonic systems operating in a feedback loop with a whole ecosystem. They bring a cybernetic approach to interventions in the soundscape of an ecosystem and, in doing so, create systems that are simultaneously reliant on human design and invention and bio- and geophonies. The sounds that they makes are not neatly human or non-human, digital or analog. The are ontologically uncertain.

3.2. Ian Ingram, The Woodiest (2010)

Artist Ian Ingram's *The Woodiest* (2010)¹ is an "auto-erotic, hermaphroditic, all-in-one woodpecker love-bot" according to the accompanying artist statement. It is inspired by the mating ritual of the Pileated Woodpecker, which consists of some initial drumming by an individual — a form of long-distance communication — followed by drum-tapping by the pair, once the second bird has been attracted to the site (Kilham, 1979).

Ingram (2010) notes that this appears to be akin to foreplay on the part of the birds; as such, *The Woodiest* is a robotic system that can engage in the entire act on its own. The robot listens for territorial drumming by a biological

¹The Woodiest documentation: <https://vimeo.com/16213036>

woodpecker and, upon hearing this drumming, the male artificial woodpecker subsystem responds with its own territorial drum. The female subsystem then responds with a courtship drum-tap, completing the encounter from a sonic perspective. The robot simultaneously engages under-explored areas of robotics for animal pleasure as well as robotic self-pleasure.

In video documentation of the work, the robot's drumming sounds quite similar to that of the woodpecker that it responds to, with the addition of the whine of the servo motor before and after the drumming and the layering of the vibration motor sounds on top of the rapping of the artificial beak on the hollow tree. The robot is notably unable to produce the higher-frequency vocalizations that the woodpecker demonstrates.

These differences mark the robot as an imposter to the human listener. But it is unclear if the woodpeckers — or other species in the ecosystem — experience a semiotic difference between *The Woodiest* and a mating pair of biological woodpeckers. This raises a question of audience for the taxonomic system under examination. Whose perception of the sound is important? In some sense, the sound is anthroponic. Ingram designed and built *The Woodiest* from human mined, fabricated, and assembled components. He decided the parameters within which *The Woodiest* would make noise and also chose to install *The Woodiest* on a particular tree in a particular forest at a particular time. And yet, *The Woodiest* makes no sound unless a woodpecker triggers it. The sound is necessarily collaborative. It is both bio- and anthro- and also somehow neither.

3.3. Richard Vitols, *Woodpecker* (2016)



Figure 1: One woodpecker from Vitols' *The Woodpecker* (2016). Used with permission of Richard Vitols.

Ingram's are not the only artificial woodpeckers to be found outdoors. The *Woodpecker*² is a series of

²Woodpecker documentation: <https://vimeo.com/180702278>



Figure 2: A broken woodpecker from *The Woodpecker* (2016). Used with permission of Richard Vitols.

30 artificial woodpeckers, installed in a forest near Dusseldorf in 2016 by artist Richard Vitols. The titular woodpeckers are, according to the artist, an attempt to rescue the local trees from a coming infestation of insects.

The inspiration for the work emerges from the intersection of two phenomena: the first is the ability of insects to respond to the sound of woodpeckers and the growing detrimental effect of insect infestations on the global forest canopy; the second is the projected decline of bird populations. This led Vitols to wonder if a robotic woodpecker impersonator might be able to help the trees survive an insect onslaught.

The woodpeckers themselves are simple robots, constructed from a basic microcontroller, a solar panel and battery pack, and a solenoid (linear) motor that hits the tree to produce the woodpecker sound. They were installed in the forest for four weeks, knocking on trees to produce pecking sounds when their batteries were charged. They were examined once per week and some that were broken — such as the one in Figure 2 — were removed and retired.

The solenoid motors don't produce the rapid beats of the vibration motors in Ingram's work. But these artificial woodpeckers aren't attempting to perform a woodpecker mating call. They are simulating the sounds of a woodpecker foraging and the sequential hammering they produce is difficult to distinguish from an actual woodpecker.

Woodpecker is unique, in part, due to its materiality. Few roboticists discuss their work in terms of breakdown and decay, yet Vitols takes explicit note of the robots that have broken or been destroyed (see Figure 2). This points to two different temporal cycles in digital producers of sound. There is the working lifecycle and the material lifecycle. The working lifecycle can be as short as a few days or weeks in

this case, but the copper and silicon that form the electrical circuits that power the robot won't decay for many biological generations.

At the same time, these digital actors are unable to self-repair or reproduce. They have no way of continuing their existence beyond their individual bodies. Without human intervention, their direct temporal mark on the soundscape is fixed in a manner quite distinct from the biological and geological sources of sound.

3.4. The rowdy krause (2019-)

The rowdy krause is a work-in-progress to evolve a voice for a new endemic species within an ecosystem. The robot's vocalization is based on Krause's (1987) acoustic niche hypothesis (ANH), which contends that animals differentiate their class across the frequency spectrum in a manner analogous to the differentiation of resource use in the traditional understanding of ecological niches.

The rowdy krause begins by listening to the existing soundscape. It listens in particular for interesting sounds that rise above the background din of the soundscape. Within those sound events, it pays attention to which parts of the frequency spectrum are most heavily used and which are left open for a new species to occupy.

In order to encourage the emergence of a complex, but biologically-plausible voice, the robot uses a mammalian vocal tract simulator to produce sound. It is able to control the voicebox, throat, tongue, lip and nasal cavity of a virtual vocal tract called Pink Trombone³. This creates a context for the robot to produce sounds that could be biological in origin.

The vocal tract is controlled by a neural network that is evolved using neuroevolution of augmenting topologies (NEAT) (Stanley and Miikkulainen, 2002). NEAT allows for the evolution of both the structure and the weighting of the neural network, so that it can add complexity as needed. The neural network takes a representation of the last sound that it made as input — so that it can “hear” itself — and outputs the next shape of the trachea, epiglottis, tongue, and lips, as well parameters to control the nose and voicebox according to the current network structure.

Once the neural network-controlled vocal tract produces a vocalization, the NEAT evolutionary process assesses how well this sound fits into the unused frequencies in the soundscape. Over time this process is repeated for many different evolved voices. Neural networks whose voices better match the predictions of the ANH are selected more often

³Neil Thapen's Pink Trombone: <https://dood.al/pinktrombone/>

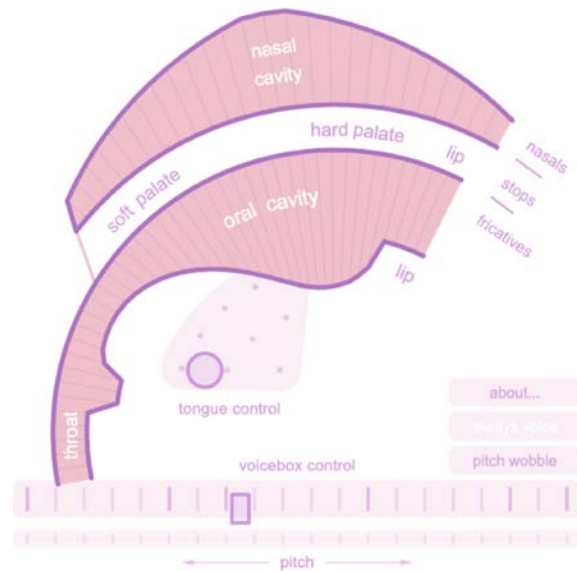


Figure 3: A screen capture of the Pink Trombone vocal tract simulator. Used with permission of Neil Thapen.

for reproduction in the evolutionary process and are more likely to pass their traits to the next generation of neural networks. This leads to a series of vocalizations that make increasingly better use of the least used frequency bands in the soundscape⁴.

The rowdy krause continues to listen to the soundscape as it shifts in response to changing seasons as well as the presence of the new robotic voice. It continually incorporates what it hears into its understanding of the sonic structure of its ecosystem, so that the vocalizations that it produces are always shifting in response to changes in the soundscape.

Currently, the rowdy krause has been tested in the lab using recorded soundscapes to drive the evolutionary process. However, field trials are planned for a peri-urban forest garden ecosystem, as well as a set of urban garden environments.

The rowdy krause represents another dimension in the discussion of robophony. Unlike Dunn's work, it is not a reprojection of past sound events back into the soundscape. It is also not an analog percussive sound like the two artificial woodpeckers. This voice is entirely synthesized, and yet it maintains some of the characteristics of biophony. Its generative frameworks — the vocal tract, neural network, and evolutionary process — reference biological objects and phenomena. It is also explicitly responsive to the soundscape in a way that mimics the theoretical response of a biological species.

⁴A demonstration of this process is available online at <https://vimeo.com/359044847>

The rowdy krause is clearly not biological. It has a human builder and programmer and can be installed or removed on a whim. Its actions are bounded by human decisions about the constraints of the vocalization system and the neural network that controls it and yet it is relatively free from human control within those bounds.

Because its calls are new to the ecosystem, the biosemiotic significance of them is unclear. It is possible that the sounds it produces, in the perception of some of the inhabitants of an ecosystem, is close enough to that of a potential predator, mate, or prey to carry some meaning. But the human semiotics of the rowdy krause are equally unclear. Is it considered noise on par with that of an engine? Or is it more akin to bringing a dog into an ecosystem?

4. ROBOPHONY

The four cases discussed in the previous section are distinct from one another, but each points to a type of hybridity that exists somewhere between the categories of anthrophony and biophony. Sonic Mirror and Autonomous Systems introduce the notion of a dynamic soundscape component that uses a system of feedback between anthropogenic and bio-geological actors to produce a dynamic sonic component that is hybrid in origin. The Woodiest's sonic emanations raise questions about the biosemiotics of robotically-produced sound. Woodpecker adds to this the temporal dynamics of a soundmaking technical object and distinguishes these robots further from their biological analogs in their inability to heal and reproduce and their sensitivity to the element and predation. Finally, the rowdy krause prompts the review of questions about the origins of biologically-informed, computationally generated sound and the semiotics and biosemiotics of an introduced digital species.

The hybrid nature of these objects points to the need for a new category of in the taxonomy of sounds in soundscape ecology: robophony. Broadly, robophony is the set of sounds produced by ecological robots in a soundscape. Informed by the cases in the previous section, it encompasses the following distinguishing features: site specificity, hybrid sourcing, layered temporalities.

4.1. Site Specificity

The notion of site specificity is borrowed from mid-21st century art and is often — though not always — invoked in the context of monumental land art works such as Christo and Jeanne-Claude's *Surrounded Islands* (1983) or the many ecological interventions of Helen and Newton Harrison. Here, site specificity

refers not to the physical object itself, but to the particular sounds that it produces.

In fact, the sound-producing objects themselves in the four cases are for the most part not specific to a particular site. Dunn's computer equipment for Sonic Mirror and Autonomous Systems could be set up almost anywhere, as could the physical body of the rowdy krause. The two woodpecker-based works are designed to sit on a tree, but they were not built for the particular trees — or even the particular species of trees — that they ended up on. They feature adjustable straps so that they can be mounted on a tree of the artist's choosing after the fact.

However, the sonic output of all of these works is a result of the particular configuration of their surroundings. Dunn and Kadish's work are most explicit in this regard. Dunn's work uses the live soundscape as the raw material for the computational modulation and reprojection. The soundscape is part of the cybernetic system that produces new sound. It does not exist without the specificity of its place. In the rowdy krause, the existing soundscape is an extra step removed from the eventual sound that is produced, but it is no less a part of the work. The rowdy krause's voice is evolved to fit the ecosystem's soundscape and is a result of the combination of existing biophony, geophony, and anthrophony of the site.

These works could inhabit another site, but they would at that point be new works. Their voices would be different, would evolve differently, and their robophonic contribution to the soundscape would be change.

This holds to a lesser, but not inconsequential degree for The Woodiest and Woodpecker. The Woodiest's sonic projections are a product of its sight insofar as it exists in dialogue with resident woodpeckers. Its pecking sequence is activated only in response to another woodpecker's call. As such, the particularities of its actions can be understood as the response to a particular community.

Woodpecker is not responsive to external sounds, but its own sound — like that of The Woodiest — is not entirely of itself. The source of sound in both of these works is the percussive action of the robot meeting the tree. In fact, the main source of sound in both of these robots is the vibration of the tree or branch. Its materiality, its age, texture, and solidity, all determine the resulting sound. These sounds are specific: to the tree, to the branch, to the meeting of biological matter and technological artefact.

As devices built by humans, it is tempting to categorize the sounds that these works produce

as anthrophony. They would likely be understood as electromechanical or controlled sounds under Krause's taxonomy (1987), both subcategories of anthrophony. But site specificity is a trait more often associated with biophony. In fact, Krause's acoustic niche hypothesis (1987) along with another foundational theory of soundscape ecology — the acoustic adaptation hypothesis which concerns the adaptation of animal calls to a particular physical environment (Morton, 1975) — explicitly frame biophony as site specific. Neither of these categories are quite sufficient to capture the acoustic impact of these new ecological actors. Robophony is therefore positioned as originating from devices constructed by humans, but with sound that it embedded in a local context and specific to a place.

4.2. Hybrid sourcing

The Krause-Gage taxonomy is source-based — it categorizes sounds according to the object that produces the sound. However, the source is not always clear. When a recording of a bird call is played back into an ecosystem, is the source biological or technological and therefore anthropogenic?

This question implicitly refers to a phenomenon that Schafer (1977) terms *schizophonia*. The sound is separated from its source, creating a disconnect from its original location and meaning. If a bird produces an alarm call, it signifies that there is danger for themselves, other members of their species, and possibly for others as well. When that call is recorded and played back, the sign has — in biosemiotic terms — lost its object (Emmeche, 2007).

This describes the situation of *The Woodiest* and *Woodpecker*. Though neither work features the playback of recorded sound, both sign the presence of a woodpecker through sound without the actual presence of that woodpecker as an object. *The Woodiest* signals a biological woodpecker that a mating pair is present and that this is their territory, but were the bird to claim the territory for themselves, there would be no consequence normally associated with infringing on a fellow woodpecker's territorial claim. *Woodpecker* similarly alerts insects to the presence of a predatory woodpecker without the possibility of capture and consumption.

The origin and biosemiotics of the other two cases are perhaps even more confounding. In *Sound Mirror*, *Autonomous Systems*, and the rowdy krause, the existing soundscape in an ecosystem acts as the base material for the sound that is produced. Dunn's work processes that soundscape and reprojects the result back into the ecosystem. The sound is partially bio-, geo-, and anthrophonic in origin, depending

on the composition of the soundscape that the works record during their operation. But there are another set of anthropogenic forces that act on it: the hardware system of computer, microphone and speaker; and the software system of algorithms that modulate and process the incoming sound. It cannot be said to be purely anthrophonic, nor is it bio- or geophonic.

The rowdy krause evolves something to actively differentiate itself from what it hear — but the material from which it differentiates itself is likely a mix of biophony, geophony, and anthrophony, depending on the ecosystem that it inhabits. It continues to evolve its voice in relation to the soundscape as the soundscape shifts in response to its presence. This feedback is most likely to occur in the domains of bio- and anthrophony as geological sounds are not likely to change in the short term in response to the rowdy krause. This means that the call of the rowdy krause is also hybrid in origin. It may be primarily generated by the anthropogenic computational system, but it is a result of all three of the Krause-Gage phonic taxa.

For all of these works, the framework for their sounds are set by their human designers. But they would not exist or would exist differently without the bio- and geophonic components of the soundscapes that they inhabit. Their sources are distributed between biological, geological, and anthropological sources and they therefore fit neatly into none of those categories. This points to the position of robophony as a set of sounds with hybrid sources that fail to fit neatly into the three other categories.

4.3. Layered temporalities

The hybrid sourcing of robophonic sound also gives way to a complex network of temporalities. It is important to understand the dynamic timescales of each of the sources of sound in a soundscape, because these timescales impact the rate of feedback between sources.

Patterns of geophony tend to shift on geological timescales — though anthropogenic activities are accelerating these as well — meaning that the sound of rain, for example, is not immediately responsive to most biological changes in the landscape. Biophony and anthrophony operate on a range of timescales. Evolutionary changes operate on relatively slow cycles that are depending on the lifespan and mating rates of the species in question. But behavioural changes can occur much more rapidly.

Computational timescales can be sped up even further, calculations taking place at the microsecond scale, and rendered without a perceivable delay

for many biological audiences. The material flows of computational technologies, however, tend to operate on more geological timescales. Biomass breaks down quickly, but silicon and copper remain in the landscape for many biological generations.

Sound Mirror and Autonomous Systems perform this kind of near-instantaneous computational temporality. But their performances also induce rapid — yet still comparatively slow — engagement from nearby biological actors. The performance mixes the instantaneous computational timeframe with the behavioural timeframe of biological entities.

The rowdy krause does not directly incorporate existing sounds into its own vocalization, but it does perform computational evolution at a pace far faster than biological evolution is able to act. In its offline version, it can simulate multiple generations — birth, life, mating, and death — of about 20 individuals in under 15 minutes. The version that will eventually inhabit an outdoor ecosystem will evolve more slowly than this, but still at a pace unavailable to biologically evolving species.

Woodpecker serves as a reminder that although robotic material flows are geologically slow, their functional flows can still be quite rapid. Figure 2 shows a robotic woodpecker that was destroyed by the elements after less than four weeks in operation. Robots like these are sensitive to moisture, cold, and animal attacks, meaning that although the materials may persist for many generations, they may not contribute to the soundscape for very long.

4.4. Future robophonies

The four cases that are presented above together make a case for the addition of robophony to the taxonomy of sound in the soundscape. But there are plenty of examples of technologies that do not quite qualify as robophony at the moment, but could in the near future.

Many of these are sold as sonic pest control devices. They play high pitched sounds to deter bats, mice, and mosquitos. One plays tones in the 400-1000Hz range in an attempt to drive moles away. Currently, these devices typically lack a sensory system to determine their effect on the ecosystem and whether they are “working” or when their target is nearby.

However, it is not unreasonable to imagine these things getting smarter. As the harmful effects of indiscriminate sonic output are better understood one could imagine that instead of constantly playing a static tone to deter rodents, they might begin to play varying but relevant sounds only when they detect a rodent. This type of feedback between the

rodent and the operation of the sonic agent could qualify it as robophony.

5. CONCLUSIONS

Taxonomies are inherently flawed, but they inform our understanding of a system and its component parts. The categories of sound in a soundscape have served soundscape ecology well thus far, enabling discussion about the composition of a soundscape and the impacts of human activity on the nonhuman inhabitants of an ecosystem.

But the cases presented here demonstrate that these categories are insufficient to capture the dynamics of a growing number of biologically interactive technological agents. They are human-made, but not human-driven. They are temporally complex, and they produce sound specific to their adopted habitat.

These agents, entities, actors in the soundscape require a new descriptor, one that captures these properties and, in doing so allows for the consideration of the ethics, biosemiotics, and new feedback loops that accompany their presence. This descriptor is *robophony*.

6. REFERENCES

- Dumyahn, Sarah L. and Bryan C. Pijanowski (Nov. 2011). “Soundscape conservation”. In: *Landscape Ecology* 26.9, pp. 1327–1344. ISSN: 0921-2973. DOI: 10.1007/s10980-011-9635-x.
- Dunn, David (2002). *Four Electroacoustic Compositions*.
- Emmeche, Claus (Dec. 2007). “A biosemiotic note on organisms, animals, machines, cyborgs, and the quasi-autonomy of robots”. In: *Pragmatics & Cognition* 15.3, pp. 455–483. ISSN: 0929-0907. DOI: 10.1075/pc.15.3.06emm.
- Gasc, Amandine et al. (Jan. 2017). “Future directions for soundscape ecology: The importance of ornithological contributions”. In: *The Auk* 134.1, pp. 215–228. ISSN: 0004-8038. DOI: 10.1642/AUK-16-124.1.
- Hobbs, Richard J., Eric Higgs, Carol M. Hall, et al. (2014). “Managing the whole landscape: Historical, hybrid, and novel ecosystems”. In: *Frontiers in Ecology and the Environment* 12.10, pp. 557–564. ISSN: 15409309. DOI: 10.1890/130300.

- Hobbs, Richard J., Eric Higgs, and James A. Harris (Nov. 2009). "Novel ecosystems: implications for conservation and restoration". In: *Trends in Ecology & Evolution* 24.11, pp. 599–605. ISSN: 01695347. DOI: 10.1016/j.tree.2009.05.012.
- Ingram, Ian (2010). *The Woodiest*. Tech. rep. www.ianingram.org.
- Kilham, Lawrence (July 1979). "Courtship and the Pair-Bond of Pileated Woodpeckers". In: *The Auk* 96.3, pp. 587–594. DOI: 10.1093/auk/96.3.587.
- Krause, Bernard L. (1997). *Into a wild sanctuary : a life in music & natural sound*. Berkeley Calif: Heyday Books, p. 200. ISBN: 9781890771119.
- Krause, Bernie L. (1987). "Bioacoustics, Habitat Ambience in Ecological Balance". In: *Whole Earth Review* 57, pp. 14–18. ISSN: 0749-5056.
- (2008). "Anatomy of the soundscape: Evolving perspectives". In: *AES: Journal of the Audio Engineering Society* 56.1-2, pp. 73–80. ISSN: 15494950.
- Krause, Bernie L. and Stuart Gage (2003). *Testing Biophony as an Indicator of Habitat Fitness and Dynamics*. Tech. rep. Clen Ellen, CA: Wild Sanctuary Inc, p. 18.
- Morton, Eugene S. (Jan. 1975). "Ecological Sources of Selection on Avian Sounds". In: *The American Naturalist* 109.965, pp. 17–34. ISSN: 0003-0147. DOI: 10.1086/282971.
- Pijanowski, Bryan C., Almo Farina, et al. (Nov. 2011). "What is soundscape ecology? An introduction and overview of an emerging new science". In: *Landscape Ecology* 26.9, pp. 1213–1232. ISSN: 0921-2973. DOI: 10.1007/s10980-011-9600-8.
- Pijanowski, Bryan C., Luis J. Villanueva-Rivera, et al. (Mar. 2011). "Soundscape Ecology: The Science of Sound in the Landscape". In: *BioScience* 61.3, pp. 203–216. ISSN: 1525-3244. DOI: 10.1525/bio.2011.61.3.6.
- Roosevelt, A.C. (Dec. 2013). "The Amazon and the Anthropocene: 13,000 years of human influence in a tropical rainforest". In: *Anthropocene* 4, pp. 69– 87. ISSN: 22133054. DOI: 10.1016/j.ancene.2014.05.001.
- Schafer, R. Murray (1977). *The tuning of the world*. A.A. Knopf, p. 301. ISBN: 0394409663.
- Stanley, Kenneth O. and Risto Miikkulainen (June 2002). "Evolving Neural Networks through Augmenting Topologies". In: *Evolutionary Computation* 10.2, pp. 99–127. ISSN: 1063-6560. DOI: 10.1162/106365602320169811.
- Vitols, Rihards (2016). *Woodpecker*. Tech. rep. vitols.xyz.
- Vitols, Rihards and Regine Debatty (2016). "The Woodpecker: Could fake birds save our forests?" In: *We Make Money Not Art* August.
- Wiens, John A. (Oct. 2009). "Landscape ecology as a foundation for sustainable conservation". In: *Landscape Ecology* 24.8, pp. 1053–1065. ISSN: 0921-2973. DOI: 10.1007/s10980-008-9284-x.
- Wynsberghe, Aimee van and Justin Donhauser (Oct. 2017). "The Dawning of the Ethics of Environmental Robots". In: *Science and Engineering Ethics*. ISSN: 1353- 3452. DOI: 10.1007/s11948-017-9990-3.

Article 5

Sound as Material for Eco-technogenesis

Originally published in: RE:SOUND 2019 – 8th International Conference on Media Art, Science, and Technology; August 2019

Joint work with: Megan Hines and Maja Fagerberg Ranten

Sound as Material for Eco-technogenesis

Megan Hines
Department of Art
Stony Brook University
2224 Staller Center for the Arts
Stony Brook, NY 11794
United States
megan.hines@stonybrook.edu

David Kadish
Robotics, Evolution, and Art Lab
IT University of Copenhagen
Rued Langgaards Vej 7
DK-2300 Copenhagen S
Denmark
davk@itu.dk

Maja Fagerberg Ranten
Department of People and
Technology
Roskilde University
Universitetsvej 1, 10.2,
DK-4000 Roskilde
Denmark
mranten@ruc.dk

This paper delineates the conceptual outcomes from a two-week intensive cross-disciplinary conversation between an art historian, an interaction designer, and an artist/engineer. With the aim of applying the concept of technogenesis to an exploration of sound as material for art and design, we consider sound as a material force within an ecosystem. Through this lens, sound produced by either life- or technological-forms allows us to consider the ecological impact and potential meanings of generated sound. Drawing on biosemiotics, we propose that the co-evolution of sound, technology, and environments, what we call eco-technogenesis, demands relational, and thus ethical, thinking. The rowdy krause, an autonomous sonic agent, designed by Kadish to identify and inhabit an acoustic niche within an ecosystem, serves as a case study for thinking through eco-technogenesis.

Biosemiotics. Ethics. Soundscape ecology. Sound objects. Technogenesis.

1. INTRODUCTION

This article is the conceptual outcome of a two-week intensive research- and practice-based collaboration between Megan Hines, an art historian, Maja Fagerberg Ranten, an interaction designer, and David Kadish, an artist/engineer. Drawing on conversations that took place at Catch: Center for Art, Design, and Technology, Helsingør, and the RE:SOUND Conference, Aalborg, we discuss sound as a material and the ecological impact and potential meanings of generated sound. Approaching our topic from a cross-disciplinary lens, we posit that the increasing presence of generated sound within ecosystems demands expanding technogenesis to the non-human realm. We propose bio- and eco-technogenesis as tools to consider the relationality of generated sound.

Technogenesis, as developed by Bernard Stiegler (1998) and expanded on by Katherine Hayles (2012) describes human evolution as occurring as a result of interactions with the environment and the available tools or technologies within that environment. Technogenesis is a co-constituted process of becoming, an intertwining of the past, present, and future trajectories of the techno-human ensemble. In this framework, technologies drive shifts in human genetic, epigenetic, and

developmental traits. Though the process is non-teleological, the entities as they currently exist could not have existed independently and are rendered meaningless without one another. Stiegler (1998) traces a line from the emergence of bipedal mobility and the freeing of the hands to the beginnings of tool use and the development of a co-constitutive relationship between technics and human genetic and epigenetic changes. He argues,

“the prosthesis is not a mere extension of the human body; it is the constitution of this body *qua* ‘human’” (1998, p. 152).

Hayles (2012) focuses on contemporary technogenesis, aiming to describe the unprecedented change and amplified feedback loops that digital technologies have activated in our environment and ourselves.

We argue here that this process of technological becoming can and should be applied to other biological forms as well as to ecosystems. The inherently relational quality of sound provides an excellent medium through which to introduce ecological thinking to the concept of technogenesis. Soundscape ecology, sometimes called ecoacoustics or soundscape studies, is the study of sounds in an ecosystem. Its history predates its naming. For example, ornithologists took advantage

of sound recording technologies as soon as it was plausible to bring them into the field in the late nineteenth century. These studies concentrated on single bird songs in order to preserve and analyse them. Soundscape ecology incorporates individual sounds into a whole soundscape, a term coined by Canadian composer R. Murray Schaffer (1969). Schaffer's concept of a soundscape was groundbreaking for including sounds made by humans, nature, and machinery, whether electric or mechanical (Ibid., 5-6). The soundscape was a landmark contribution to soundscape ecology because it moved the study of sonic relations beyond the confines of music. Furthermore, the soundscape pushed the study of sound into a wider realm that included non-human and non-organic listeners and creators. To accept that sound acts as an ecological force implies that its effects go beyond human aesthetics and touch on material relations within an environment.

1.1 Cross-disciplinary Backdrop

The topics of technogenesis and soundscape ecology are necessarily cross-disciplinary, spanning fields as wide as sound studies, ecology, evolutionary biology, anthropology, and media studies. As representatives from three different fields of study, we spent two weeks together in Denmark in August 2019: the first week as participants at the workshop *Artistic and Curatorial Practices in the Age of Technogenesis* at Catch in Helsingør; and the second week at the 2019 *Media Art Histories* conference in Aalborg titled *RE:SOUND Sound, Media and Art - Theories, Histories, Practices*.

Throughout the two weeks, we discussed what is at play when we consider sound as material within an ecological soundscape. How do we define and perceive sound? How do nonhumans perceive sound? Moreover, what happens when one perceives sound that is computationally generated? By acknowledging that sound acts as an ecological force, we discuss the effects of sound beyond human aesthetics within an ecosystem.

The work-in-progress explored at the pre-conference workshop, Kadish's the *rowdy krause*, serves as a subject for our conversation. We begin our discussions through the presentation of three conversation themes, followed by a conversation between us based on our respective practices. We posit that viewing technogenesis from a cross-disciplinary lens can broaden the perspective on sound as material within ecological soundscapes. The two conversation themes are sound as computational material and eco-technogenesis.

1.2 The Rowdy Krause

The rowdy krause is an autonomous sonic agent that is designed to inhabit an ecosystem and find an acoustic niche for itself within that ecosystem's soundscape. The work is currently in progress (see Figure 1), but the rowdy krause is already able to perform its search for a niche using a recorded soundscape, while future iterations will perform this in real-time, embedded in an ecosystem.

The rowdy krause begins its search for a niche by listening to the soundscape in which it is placed. It makes note of the different sounds that it hears and analyses their spectra so that it can start to determine which frequencies are least used. It continues listening and revising its understanding of what already exists in the soundscape throughout the process of evolving its own voice.

To make sound, the rowdy krause makes use of a computational model of a mammalian vocal tract so that the sounds that it produces have something in common with sounds that one might hear from a biological creature. The simulator was developed by Neil Thapen and is called *Pink Trombone*.¹ The rowdy krause controls the simulator using an evolutionary neural network that uses a process called the neuroevolution of augmenting topologies (NEAT), meaning that the structure and weighting of the neural network both undergo evolution (Stanley, 2002).

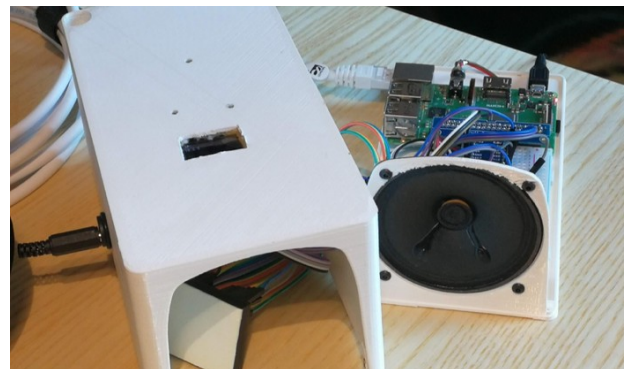


Figure 1: The rowdy krause's physical instantiation in progress. The rowdy krause's computation occurs on a Raspberry Pi single board computer, and it senses its surroundings using a microphone, temperature and humidity sensor, and light detector, while engaging with the world through an amplified speaker.

Every time the neural network produces a vocalization by controlling the vocal tract, the vocalization is compared to all of the sounds that the rowdy krause has already heard in the ecosystem's soundscape. The neural network is assigned a fitness based on how different the sound it produces is from the existing soundscape and the networks with the best fitness are more likely to pass on traits to the next generation of neural networks in the evolutionary process. Over time, the rowdy krause tends toward finding a voice that occupies a unique niche within its soundscape.

While the parameters for the rowdy krause's voice are selected by Kadish, the actual voice arises from the particular composition of the ecosystem that it inhabits. At the same time, the projection of the evolving vocalizations into the ecosystem adds to the soundscape and affects other ecosystem inhabitants. Their responses and the new voice are part of the ecosystem's continual evolution, and also drive the evolution of the rowdy krause's vocalization, an example of eco-technogenesis in action.

2. SOUND AS COMPUTATIONAL MATERIAL

Sound made by humans and machinery ranges from recordings, digitized sounds and algorithmically-generated sounds. Typically, humans author content; the composer, artist, designer, or developer uses sound as material in their technologically-aided practice. The idea of sound as immaterial or ephemeral has changed as the technological possibilities for the recording and manipulation and sound increased. Composer Pierre Schaeffer was the first to refer to an "object sonore" or sound object (Schaeffer, 1966). As audio technologies progressed, recorded sound objects no longer referred to a sound source, but existed as objects in their own right (Chattopadhyay, 2017).

Joseph Klett has extended the materiality of sound to the sonic object setting, the place of interaction between material sounds and their interpreters (Klett, 2014). Besides labelling sound as social as it occurs among bodies, Labelle (2006) addresses sound as spatial. Sound is always more than one place and performs with and through space. Sound is a relational phenomenon: from immaterial to material, from inside one's thoughts to others', and through space (Labelle, 2016, p. xi).

Søndergaard (2019) describes sound as the perfect material for experimental practices because of its "difficulty" as time-based, immaterial and fugitive. He adds that sound is nothing without being experienced. He describes sound as an artist's material and the first media art practice:

"...and because of the invention of technologies that make it possible to liberate sound from its source, sound even became the first technologically emerging artistic material; as such, sound art could be seen as the first media art practice, historically." (Søndergaard, 2019, p. 96).

Cox (2013) elaborates on sound having a sonic ontology. He states that whereas the ontology of "matter" privileges sight and touch, the invisible, intangible and ephemeral objects of smell, taste and hearing exist in the shadows compared to solid materials. He concludes that sound has a different ontology and materialism:

"...a conception of being and matter that can account for objecthood better than an ontology of objects can account for sounds." (Cox, 2013).

According to Cox (2009), sound instead affirms an ontology of flux, where objects are replaced by events; a sonic philosophy of sound as flux, event and effect.

The same can be stated about computational material. It too can be described as an intangible matter and as an ontology of flux. We are bodily affected through, with, and by computational material, and no longer differentiate between the subjective inside and technology from the outside. Rather than dividing the human-technology relation into a matter of a foreground and a background we can unpack Hayles' (2012) claim that digital media and the actions of computers are embodied, that technical objects have agency or the potentiality of computational material as autonomous.

When Hayles expanded the concept of technogenesis, the idea that humans and technics coevolve, the focus was on contemporary digital technologies. She relates technical beings versus embodied living beings and focused on the interfaces between programmable machines and humans:

"...the actions of computers are also embodied, although in a very different manner than with humans. The more one works with digital technologies, the more one comes to appreciate the capacity of networked and programmable machines to carry out sophisticated cognitive tasks, and the more the keyboard comes to seem an extension of one's thoughts rather than an external device on which one types." (Hayles, 2012, p. 3).

Thus, embodiment takes the form of an extended cognition where larger networks beyond the desktop computer are entangled with human agency and thought. Hayles clearly described technical objects as having agency. In the sense that technical objects are agents of complex temporalities, in evolutionary terms, they are repositories of change. (Hayles, 2012, p. 85)

Carvalhais (2010) addresses the autonomous quality of programmable media beyond their creators and users:

"A system's autonomy can be regarded not only as an amount of control that is conferred to or appropriated by it but also as a transfer of some amount of agency to the system." (Carvalhais 2010, p. 421).

He compares procedural systems to biological systems and claims that computational material too can be complex. With reference to Herbert A. Simon he describes an inner and outer environment of a system as the substance and organization of the

artefact and the operating surroundings (Carvalhais, 2010, p. 634). Manovich (2001) makes a related distinction between the cultural layer and the computational layer as the distinction between the interface and what the computer reads.

Computational material is potentially autonomous. Computational media and sound are not just mediators, but also programmable, and thus potentially autonomous beyond their makers and users/listeners. The notion of sound as material and computational material is interesting in relation to the rowdy krause. Is the rowdy krause autonomous, and who is the author of the voice of the new species? What is the role of the computational material in relation to the rowdy krause? In what follows, we situate potentially autonomous technological forms as material forces within ecosystems and ask how an example such as the rowdy krause impacts and is impacted by the relationality of sound within its environment.

3. ECO-TECHNOGENESIS

Eco-technogenesis describes the co-becoming of ecosystems and technologies. Hayles asserts,

“technical objects embody complex temporalities
enfolding past into present, present into future”
(2012, p 86).

While Hayles applies this idea to human development, here we apply it to an understanding of the past and future development of ecosystems. The contemporary farm is a good working example of eco-technogenesis (Mazoyer, 2006). The typical large wheat monoculture found in mid-western North America, for example, did not arise overnight, but is the result of tens of thousands of years of co-evolutionary development involving technical ensembles, human societies, and ecosystems. When the first hunter-gatherers scattered seeds from their food in a known location or the first metalworkers forged a plow, their goal was not to produce unbroken hectares of wheat monoculture. And yet, that ecosystem is only possible because of the advent of seed harvesting and metal plows, the selection and evolution of new crops and countless other technological and biological changes.

Examples of the co-evolution of humans, technology, and ecosystems abound, but what about non-human eco-technogenesis? This is a more difficult proposition to consider, given the difficulty of observing the slow changes continually modified by feedback loops between animals and environments occurring on evolutionary and geological timescales. Scholars have pointed to habitat-modifying species like the beaver as an example. Beavers are best known for cutting trees to dam rivers, creating larger pools of water within a river ecosystem. The technical object of the dam is,

on its own, a significant factor in the formation of these dammed river ecosystems. It has an impact on the plant and animal communities in the ecosystem (Rybczynski, 2007), reshaping the network of interactions and relationships in the landscape. Dams become an essential component of the development of the ecosystem. Considering the evolutionary history of beavers provides evidence that this behaviour evolved through technological and evolutionary reinforcement. Natalia Rybczynski argues the building of dams evolved from simpler behaviours like cutting and feeding on smaller vegetation (Ibid., 2007). It implies a long, slow, multigenerational process in which dams began as smaller debris and eventually took the form of larger constructions, thereby reconstituting the ecosystem and reinforcing the behaviour. Beavers' activities are implicated in the evolution of fish and amphibians in dammed rivers as well as the defences of nearby tree species (Ibid., 2007). Beaver dams can be considered a canonical example of eco-technogenesis.

3.1 Sound and eco-technogenesis

Returning to soundscape ecology, we argue sound provides a domain in which to think through the possibility of eco-technogenesis. In fact, this idea is not entirely new. The intermeshing of sound, technology, and ecology exploded into public consciousness in 1962 with the publication of Carson's landmark book *Silent Spring*, in which she revealed the devastating effects of industrial chemicals on bird populations by pointing to gaps in the sound spectrum as bird species went extinct. Carson's method of measuring ecological health using sound predated yet paved the way for the acoustic niche hypothesis (ANH) (Krause, 1987) and the acoustic adaptation hypothesis (AAH) (Morton, 1975). The ANH is based on empirical observations that suggest sounds produced by species vocalizing within an ecosystem tend not to interfere with one another, creating a partitioning of the acoustic range (Sueur and Farina, 2015, 495). Relatedly, the AAH argues animal-generated sounds have adapted to their particular habitat, taking into account the properties of the landscape and plant life to maximize sound dispersal (Sueur and Farina, 2015, 495). The ANH and AAH provide the basis for thinking in terms of the co-constitution of sound and landscape. Based on Carson's and Krause's work, an ecosystem's soundscape became an indicator of its diversity and health. Loss of diversity across the acoustical space indicated an imbalance and an infiltration of human activity in the ecosystem. For example, Krause observed that smaller parks in the American Northwest established by lumber companies and characterized by a monoculture of young pines displayed a shocking lack of biodiversity evidenced by large gaps in the acoustical space (Krause, 1993).

Soundscape ecology allowed for thinking of the intermeshing of sound, technology, and ecology and the technogenesis of soundscapes, while preserving the nature/culture divide. Having defined eco-technogenesis, we now focus on the possibility of technogenesis across human, non-human, and non-organic or technological systems. Throughout the course at Catch, we considered autonomous agents capable of creative sound generation through case studies and in practice. Our aim here is to situate generative sound practices within local ecologies. Kadish's rowdy krause, developed over the course of his doctoral studies and presented for the first time at Catch, provides a case study. Given that we have accepted sound as a material force in an ecosystem, we conclude by speculating on the ethics of an eco-technogenesis that includes non-organic sound generation.

There are a number of automated or robotic sound-generating systems that are designed specifically to engage with ecosystems in a dynamic and generative manner. David Dunn's *Sonic Mirror* (1986) is an early example of this type of work. Dunn's *Sonic Mirror*, which he considers a sound performance, arose from an interest in rules and systems for processing and reproducing soundscapes (Dunn, 2013). *Sonic Mirror* involves the recording of sound in an ecosystem by an autonomous computer system, the processing and modulation of that recording, and the subsequent reprojection of the sound into the ecosystem. Dunn, who was trained as a composer, observed that the ecosystem inhabitants began to engage with the recording and playback system, and noted the ecosystemic nature of sound:

"The song of a bird is not just grist for compositional manipulation; it is a code of signification not only between members of that particular species, but also for the extended fabric of mind that forms the biobehabitat within which that species resides" (Dunn, 2013, p. 100).

Dunn's statement parallels our own argument, that eco-technogenesis demands that non-organic self-organizing systems be considered as relational elements within an ecosystem.

In 2001, biologist Claus Emmeche speculated, "Does a robot have an *Umwelt*?" (Emmeche, 2001). The biologist Jakob von Uexküll's concept of *Umwelt* is defined as an organism's perceptual world (1909, 1992). From the concept of *Umwelt*, it follows that an organism acting within a perceptual environment relies on signs to interpret the material world and is a communicative being, even if one does not accept this as proof of high-level reasoning. Emmeche speculated autonomy and self-organization would be necessary qualities for occupying an *Umwelt*, and that robots could therefore have an *Umwelt*. In response, biologist Winfried Nöth offered the example of a robot successfully moving around its

environment and argued even if it did have an experience of *Umwelt*, it would be impossible for humans to know anything about that experience (Nöth, 2001). In 2012, Hayles referring to Uexküll's famous example of a tick (Uexküll, 1992) agreed, arguing if an animal as simple as a tick could have an *Umwelt*, surely a spatially-aware robot could. (Hayles, 2012, p. 249n4). Each of these examples stresses vision and proprioception as markers of semiosis. We argue here just as a robot moving around its environment has an *Umwelt*, an artificial intelligence system occupying an acoustical niche has an *Umwelt*.

The rowdy krause meets this threshold by listening to its environment, identifying the least utilized portions of the audio spectrum, and evolving a voice to fill those frequencies. It performs the acoustic niche hypothesis as an autonomous technological species, its *Umwelt* framed by its primary sensory organ: its microphone. As its vocalizations evolve to fill previously empty portions of the soundscape, how do existing inhabitants of the ecosystem perceive its calls? Its interference in their own communication is minimal by design, but the rowdy krause's calls are nevertheless part of the acoustic environment.

3.2 Biosemiotics as a basis for an ethics of generated sound in eco-technogenesis

Technogenesis belies the idea of infiltration and instead presents the problem of intermeshing and relationality. Relationality forces the question of ethics, a code that governs behaviour within a social setting. Hoffmeyer, Kull, Tønnensen, Beaver, and Hendlin have all approached ethics from the perspective of biosemiotics, the production and interpretation of signs in the living world beyond human language. (Hoffmeyer, 1993; Kull, 2001; Tønnensen, Beaver and Hendlin, 2015). We argue here that the creative use of sound requires an exploration of the ethics of sound generation within an ecosystem, an occurrence that only promises to increase as techniques for sonic manipulation continue to develop. By thinking through eco-technogenesis, sounds regardless of origin contribute for better or worse to a biosemiotic ecosystem, an environment within which the generation and interpretation of sounds as signs evolve in co-constitution with one another regardless of origin, whether biological, natural, or technological. Given this description, preserving the nature/culture divide proves likely to be a fruitless enterprise in the face of ongoing technological development. As sound-generating autonomous agents increasingly become part of the soundscape, it is crucial we consider how to design responsible inhabitants. This idea follows Cox's argument that sound art provides an opportunity to explore the ontogeny of sound (Cox, 2009). But whereas in Cox's argument, sound art points to a realm of

sound outside human sensory experience, in our estimation, sound not only points to a realm of sensory experience outside the human, it acts as a source of potential various meanings to an ecosystem's inhabitants. Even if these sounds are generated by non-living self-organizing systems, their potential to occupy an *Umwelt* makes them "morally considerable" (Beever and Tønnessen, 2015, p. 45).

Part of the appeal but also the challenge of biosemiotic ethics is its insistence on the existence of systems of meaning that reside outside of human experience (Hoffmeyer, 1993; Kull, 2001; Tønnessen, Beever and Hendlin, 2015). Traditionally, however, conservationist rhetoric has relied on anthropocentric notions of beauty to drive home the necessity of caring for ecosystems (Harries-Jones, 2008). Like beautiful works of art in a museum, it was reasoned, landscapes should be preserved for future generations to enjoy. Even Carson relied on this line of thinking by lamenting the loss of bird songs for human ears (Carson, 1962). Recently, it has become apparent that these human-centred values are not only questionable in their efficacy, but also potentially harmful to the ecosystems they try to protect (Harries-Jones, 2008). In emphasising stasis over dynamics, the idea of beauty as applied to ecosystems betrays their essence. Peter Harries-Jones argued Bateson's idea of an ecological aesthetics, which leaves behind human-centred ideas of beauty, has finally started to gain traction, if without full appreciation for its biosemiotic implications (Ibid.). Bateson's concept of ecological aesthetics, in which systems interact through feedback, provides a foundation for thinking of ecosystems as composed of layers that add up to something greater than themselves (1972). In other words, ecological aesthetics is a type of holistic thinking that resists reductionism.

4. CONCLUSION

Throughout this article, we considered sound generated by an autonomous agent as contributing to the co-constituted becoming of ecosystems and technologies. Based on our collaborative work and discussions, we offer insights into sound as computational material, eco-technogenesis, and the possibility for an ethics of sound generation based on biosemiotics in this new ecological paradigm.

We do not present an exhaustive account of the subjects but invite other scholars and practitioners to continue the conversation and further explore the subject of technogenesis in relation to sound and ecosystems. Finally, the thorny questions of ethical approaches to eco-technogenesis will continue to present themselves in the future, especially as the possibilities for technological sound generation

increase. Here we offer a starting point for considering ethical relationality from the perspective of biosemiotics.

Acknowledgments

Artistic and Curatorial Practices in the Age of Technogenesis was a joint project between Catch, the IT-University of Copenhagen and Aalborg University. Thank you to the academic supervisors Morten Søndergaard Laura Beloff and Catch supervisor Majken Overgaard. Megan Hines would like to thank the Stony Brook University Graduate Student Organization for their financial support.

5. REFERENCES

- Bateson, G. (1972) *Steps to an Ecology of Mind*. Ballantine, New York.
- Beever, J. and Tønnessen, M. (2015) 'Justifying Moral Standing by Biosemiotic Particularism', *Zeitschrift für Semiotik* 37 (3-4), pp. 31-53.
- Carson, R. (1962) *Silent Spring*. Houghton Mifflin, Boston.
- Carvalho, M. (2010). "Towards a Model for Artificial Aesthetics. Contributions to the Study of Creative Practices in Procedural and Computational Systems." Doctoral thesis. Universidade do Porto.
- Chattopadhyay, B. (2017). 'Beyond Matter: Object-disoriented Sound Art', *Seismograph*. <<https://seismograf.org/fokus/sound-art-matters/beyond-matter-object-disoriented-sound-art>> (retrieved 09 November 2019).
- Cox, C. (2009) 'Sound Art and the Sonic Unconscious', *Organised Sound* 14 (1), pp. 19—26.
- Cox, C. (2013). 'Sonic Philosophy', *ARTPULSE Magazine*. <<http://artpulsemagazine.com/sonic-philosophy>> (retrieved 09 November 2019).
- Dunn, D. (2013) 'Nature, Sound Art, and the Sacred', in David Rothenberg (ed.), *The Book of Music and Nature: A n A nthology of S ounds, Words, T houghts*. Wesleyan University Press, Wesleyan.
- Emmeche, C. (2001) 'Does a robot have an *Umwelt*?', *Semiotica* 134 (1/4), pp. 653—693.
- Harries-Jones, P. (2008) 'Gregory Bateson's 'Uncovery' of Ecological Aesthetics', in Hoffmeyer, J. (ed.), *A Legacy for Living Systems: Gregory Bateson a s P recursor t o B iosemiotics*. Springer, Dordrecht.
- Hayles, N. K. (2012) *How We Think: Digital Media and Contemporary Technogenesis*. University of Chicago Press, Chicago.

- Hoffmeyer, J. (1993) 'Biosemiotics and ethics.' in Witoszek, N. and Gulbrandsen, E. (eds.), *Culture and Environment: Interdisciplinary Approaches*. Centre for Development and the Environment, Oslo.
- Klett, J. (2014) 'Sound on Sound: Situating Interaction in Sonic Object Settings', *Sociological Theory* 32 (2), pp. 147--161.
- Krause, B. L. (1987) 'Bioacoustics, Habitat Ambience in Ecological Balance', *Whole Earth Review* 57, pp. 14--18.
- Krause, B. L. (1993) 'The Niche Hypothesis: A hidden symphony of animal sounds, the origins of musical expression and the health of habitats,' *The Explorers Journal*, (Winter), pp. 156-160.
- Kull, K. (2001) 'Biosemiotics and the problem of intrinsic value of nature', *Sign Systems Studies* 29, 1: pp. 353--365.
- Labelle, B. (2006). *Background Noise: Perspectives on Sound Art*. Continuum, London.
- Manovich, L. (2001) *The Language of New Media*. MIT Press, London.
- Margulis, L. (1970) *Origin of eukaryotic cells; evidence and research implications for a theory of the origin and evolution of microbial, plant, and animal cells on the Precambrian earth*. New Haven: Yale University Press.
- Mazoyer, M., & Roudart, L. (2006). *A History of World Agriculture: From the Neolithic Age to the Current Crisis*. Earthscan, London.
- Morton, E. S. (1975) 'Ecological Sources of Selection on Avian Sounds', *The American Naturalist*, 109(965), pp. 17--34.
- Nöth, W. (2001) 'Semiosis and the *Umwelt* of a robot', *Semiotica* 134 (1/4), pp. 695--699.
- Rybczynski, N. (2007) 'Castorid Phylogenetics: Implications for the Evolution of Swimming and Tree-Exploitation in Beavers', *Journal of Mammalian Evolution* 14(1), pp. 1--35.
- Schaeffer, P. (1966) *Traité des objets musicaux*. Éditions du Seuil, Paris.
- Schaffer, R. M. (1969) *The New Soundscape*. Berandol Music Limited, Ontario.
- Søndergaard, M. (2019) 'Exscribing the Archive: Curating the Unheard (in Scandinavia)'. in Weibel, P. (ed.) *Sound Art*. MIT Press, Cambridge.
- Stanley, K. O. and Miikkulainen, R. (2002) 'Evolving Neural Networks through Augmenting Topologies', *Evolutionary Computation*, 10(2), pp. 99--127.
- Stiegler, B. (1998). *Technics and Time 1*. Stanford University Press, Stanford.
- Sueur, J. and Farina, A. (2015) 'Ecoacoustics: the Ecological Investigation and Interpretation of Environmental Sound', *Biosemiotics* 8, pp. 493--502.
- Tønnessen, M., Beever, J. and Hendlin, Y. (2015) 'Introducing Biosemiotic Ethics', *Zeitschrift für Semiotik* 37 (3-4), pp. 3-12.
- Uexküll, J. (1909) *Umwelt und Innenwelt der Tiere*. Julius Springer, Berlin.
- Uexküll, J. (1992). 'A Stroll Through the Worlds of Animal and Men' in Uexküll, T. (ed.) *Semiotica* 89 (4), pp. 273--391.

¹ Thapen, N. <https://dood.al/pinktrombone/> (retrieved 09 November 2019).

Article 6

Endemic Machines: Artificial Creativity in the Wild

Status Abstract accepted for *Transformations: Journal of Media, Culture & Technology*, Issue 36: Artificial Creativity

David Kadish

Bojana Romic

Artificial Creativity, Transformations #36

30 April 2021

Endemic Machines: Artificial Creativity in the Wild

Abstract

Artificial creativity is often applied in attempts to mimic or eclipse human creativity with machines. However, as a creative force that is not bound to human experiences, it can act as a way of approaching non-human or more-than-human creative forces from a new perspective. This paper develops a concept of endemic machines to describe a process of engaging the creativity of an ecosystem through a machine that adapts with that ecosystem. A case study detailing the design and testing of an endemic machine called the Rowdy Krause helps to ground the concept of endemic machines in practice.

The discourse of artificial creativity tends to be centred around human creativity. It features explorations of whether artificial intelligences (AIs) can emulate artists, whether a computational process can be considered creative, and what artificial creativity means for human creative practice. Much as R.U.R, the original conception of a robot, is a humanoid android, we strive to see our own creativity in and through artificially creative machines.

But creative AIs also offer the opportunity for a different kind of exploration: the exploration of a creativity that is expressly *not* human. Artificial creativity can act as a lens through which to explore not only machine creativity, but other forms of creativity to which we have no experiential access. In particular, this article focuses on non-human — henceforth referred to as more-than-human — forms of biological creativity: the creativity of ecosystems and evolution.

The article outlines endemic machines, a set of practices and principles relating to machines that engage creatively with ecosystems through co-evolution. These are examined in greater depth through the presentation of a case study that details the design and testing of an endemic machine called the Rowdy Krause.

CREATIVITIES

Much of the focus of artificial creativity has been on the arts and this has resulted in a broad range of creative technologies from computational musical composition (Kroos) to robot painters (Cohen) and interactive sculptures (Chan et al.). What these artificially creative works share is their audience; their creativity is for human enjoyment. Furthermore, the creativity that they display is modelled on human creativity. They are the result of attempts to imbue machines with the types of creativity that we, as humans, perform.

This type of work has made important contributions to the understanding of artificial intelligence and human creativity and, in some cases, pushed the boundaries of the creative fields with which it engages. But to limit explorations of artificial creativity to these kinds of tasks would be to limit a fundamentally non-human creativity to trying to emulate and extend humanity. Robots, artificially intelligent agents, and autonomous machines are built by humans, modelled on humans, and often do work for humans, but they are in fundamental ways, not human. They are other-than-human, or perhaps more-than-human (Abram).

A more-than-human creativity affords the capacity to interact with *other* more-than-humans in new and perhaps interesting ways. Robots can sense the world with an expanded — or constrained — set of sensory apparatuses that immediately creates a perceptual world to which humans don't have direct access. Artificially intelligent agents are unencumbered by some of the innate ways that humans think, leaving them open to relating to other more-than-humans in ways that we are fundamentally incapable.

Due to that openness, number of efforts to use AI to understand animal communication are already underway. For example, an initiative called CETI¹ aims to use AI to better understand whale communication (Andreas et al.). Other efforts have produced promising results translating the ultrasonic vocalizations of rodents (Coffey, Marx, and Neumaier).

However, current artificially intelligent robotics situated in biological ecosystems tend to be programmed for productivity rather than creativity (e.g. Bergerman et al., Oberti and Shapiro). In terrestrial ecosystems, vision-based AI is used in precision agricultural robots to apply spot-adjusted treatments of water, fertilizer, and pesticides in order to regularize plant growth and maximize yields (C. Yang, Sui, and Lee). Underwater, deep learning systems can assess fish behaviour, estimate their size, and assess water quality in fish farms (X. Yang et al.). These practices impose an industrial regularity on landscapes, flattening their complexity and muting the creativity of the ecosystem.

The creativity of ecosystems resides in the complex web of relationships and dependencies that perpetuate the flow of materials and energies. It is the combined creativity of the evolutionary processes shaping the species that compose the living community of the ecosystem (Gould). This is the virtually boundless creativity that results in Darwin's "endless forms" (1859), the incredible diversity of the tree of life.

The creativity of evolution is a particular creativity. It is distributed, with creative contributions arising from many sources and interactions (Beatty, "The Creativity of Natural Selection? Part I: Darwin, Darwinism, and the Mutationists" "The Creativity of Natural Selection? Part II: The Synthesis and Since"). It is recursive, looping back on itself in spirals of feedback and response. And it is non-teleological; there is no end goal, though there may be

1

some directionality (Dobzhansky Beatty, “The Creativity of Natural Selection? Part I: Darwin, Darwinism, and the Mutationists”).

In this creativity of evolution and ecosystems lies an opportunity for artificial creativity. It is an opportunity to engage, as a more-than-human creativity with a creative force that is itself more-than-human. A chance to engage ecosystems with an openness that is partially liberated of human ideas of what an ecosystem should be, what form it should have. There is the possibility of forming new types of relationships with more-than-human inhabitants of an ecosystem that lie outside the realm of human sensory perception and human desire. The possibility of approaching ecosystems as they are, instead of as we want them to be.

TRENDS IN ROBOTICS IN ECOSYSTEMS

The aforementioned AI-based animal communication projects are just one way that researchers have attempted to develop machines with ecologies in mind. Van Wynsberghe and Donhauser defined three categories to group different kinds of environmental robots: *robots-in-ecology* are general-use robots that are used for environmental purposes such as data collection or surveying; *robots-for-ecology* are robots designed specifically for use in ecology or by ecologists; and *ecobots* are robots that are "ecologically functional" — that is, they perform some ecosystem function as opposed to merely observing or collecting.

This last category, ecobots, is of particular relevance to the discussion here. The notion of functional implies an ecologically significant role. An ecobot, therefore, is active in "the cycling of materials and the flow of energy" (Odum) that shape the ecosystem. Note, however, that there is no requirement in Van Wynsberghe and Donhauser's conception of ecobots that the machines are artificially intelligent or necessarily creative. The two examples of digital ecobots that they give — they also discuss bio-tech hybrids such as genetically engineered plants and biofilms as potential ecobots — are autonomous underwater robots designed to hunt and kill predators that

have become overabundant and are destabilizing coral reefs.

These ecobots are interesting examples of robots performing an ecosystem function. They use AI systems to detect their targets and help conservation biologists to bring the ecosystem into a state of equilibrium. They are designed to address a situation where an apex predator has become successful beyond the carrying capacity of their environment. In that sense, they are indeed engaging the material and energetic flows of an ecosystem. That engagement, however, is fixed, along with their mission.

CREATIVITY AND ENDEMISM

Though the ecobots described in the previous section have an ecological mission, the terms of their operation are still set by human designers. The concept of endemic machines proposed here is something more open-ended. It sets forth a paradigm for a type of digital engagement with ecosystems that seeks creative ways of contributing to the ecology of a place. It relies on the opacity of computational evolution — including the widespread ability of computationally evolved systems to produce results that confound their programmers (Lehman et al.) — as an entryway into the equally opaque world of biological evolution.

Endemic machines are grounded in the ecological concept of endemism. In ecology, endemism describes the relationship between a species and a particular place (Morrone). The singularity of the linkage between place and species signifies a special bond. An endemic robot, like an endemic species, is “produced in a specified place and nowhere else in the world” (Darwin).

Materially, this would seem antithetical to the way robots are produced. As artefacts of a globally-connected system of trade, robots are conglomerations of standardized parts, each manufactured in a different, highly-specialized factory from materials harvested from around the world. They are, in this sense, the very opposite of endemic.

While it is possible for robots to incorporate locally sourced materials, the discussion here focuses not on the production of the physical robot, but on the development of a robot's behaviour. The robot's physicality may be of a distributed origin, but for an endemic robot, its behaviour is learned, evolved, or otherwise produced in a specified place. Like evolution itself, the development of the robot's behaviour is uncharted; there is no defined destination.

As the endemic robot learns with the ecosystem, it engages in a process of eco-technogenesis (Hines, Kadish, and Ranten). An ecological extension to the concept of technogenesis (Stiegler, expanded by Hayles), eco-technogenesis refers to a process of co-evolution whereby an ecosystem and a technology form a shared, entangled history. Each exists as it does because of the other.

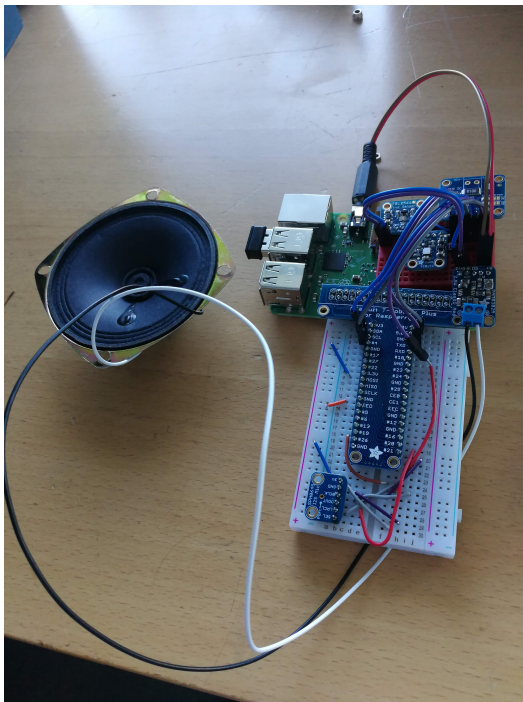
CASE STUDY: THE ROWDY KRAUSE

The concept of endemic robotics is explored in more depth through a case study of an embodied, artificially intelligent agent called *the Rowdy Krause*. The Rowdy Krause is an experiment in artificial niche construction in a biological ecosystem. The machine's goal is to create a space for itself, a niche within an existing ecosystem.

NICHES AND THE ACOUSTIC NICHE HYPOTHESIS

An ecological niche describes the collection of environments and resources that impact the lifecycle of an organism (McCormack). It encompasses their food, shelter, predators, prey, symbionts, and waste streams (Pocheville). Niche construction recognizes that, as an organism forms its own niche, it shifts the resource landscape such that the ecosystem itself changes form (Laland, Matthews, and Feldman). This opens space for new niches and shifts the adaptive pressures on other species. The act of a species inhabiting a space changes that space and has impacts that ripple through the ecosystem.

The Rowdy Krause engages with a particular part of an ecosystem and a specific part of



(a) The internal electronics.



(b) Testing at Byhaven på Sundholm.

Fig. 1. The Rowdy Krause, in development and initial testing.

its resource landscape: the soundscape. The soundscape is the collection of all of the sounds in a particular environment (Schafer) and its study in the context of ecology is called soundscape ecology (Pijanowski et al. Farina). Niche theory appears in soundscape ecology as the acoustic niche hypothesis (ANH), which treats the soundscape as a limited resource that ecosystem inhabitants can use (Krause). Much as plants compete for sunlight in a dense forest, species compete for sonic territory and construct niches in the acoustic spectrum.

The ANH hypothesizes that different species try to minimize overlap in their use of sonic resources to not confuse signals. Partitioning can occur spectrally, by using different frequencies or tones; temporally, by vocalizing at different times of the day or the year; or spatially, by moving to different locations. Of these, the Rowdy Krause focuses on spectral partitioning to find itself an acoustic niche.

DESIGNING FOR ECOSYSTEMS

One of the central questions in the practice of endemic machines is how to design a machine for an ecosystem. Even the notion of what it means to do something for an ecosystem appears to require an understanding of what an ecosystem wants which is itself problematic. It implies a teleology to the dynamics of ecosystems that runs counter to the prevailing understanding of evolution.

Two ideas ultimately served as guides in the design process for the Rowdy Krause. The first is Gregory Bateson's concept of ecological aesthetics and his notion of being "responsive to the pattern which connects". The other is Rafael Lozano-Hemmer's assertion that electronic art should have the ability to surprise the artist (Lozano-Hemmer and Ranzenbacher).

Ecological aesthetics helped to focus the process on the feedbacks and interactions that the Rowdy Krause would encounter. It drew attention to how current inhabitants of the ecosystem might perceive the sudden arrival of a new sonic agent. It focused the inquiry on to how to frame the process of listening within a soundscape and how to produce ecologically relevant sound to project back into that environment.

The idea of art surprising the artist helped to reinforce that the Rowdy Krause should not necessarily produce the sounds that met my own desires or expectations, but that its aim was to fit the fabric of the existing soundscape. Surprising vocalizations produced by the Rowdy Krause are acceptable and perhaps even valuable so long as they serve the purpose of establishing an acoustic niche. This prompt reinforced the idea that, in the process of building and programming the Rowdy Krause it was important to be able to differentiate between something broken or not working and something not working as expected, but in a manner that is still in fulfillment of its overarching goal.

PROTOTYPE

The first task in the design of the Rowdy Krause was to design the mechanism for producing sound. The intent was for the Rowdy Krause to behave as a novel animal in the soundscape, so it was important for it to have behaviours which responded to the ecosystem, but also for it to be able to evolve over time. Neuroevolution of augmenting topologies (NEAT) is a computational evolution algorithm that is well-suited to this task as it evolves a neural network structure, which can be used to drive specific behaviour. The algorithm adds complexity to the network as needed, meaning that the structure of the behaviour tends to move from relatively simple at the outset to more complex behaviours over time.

However, an artificial neural network (ANN) itself does not produce sound. It merely maps inputs to outputs through a network of artificial neurons, analogous to a brain. That brain requires some form of instrument to turn its signals into sound. Recent research has used a range of different "instruments" for this task. Some ANNs generate raw audio waveforms that can be played directly on a speaker (van den Oord et al.). Others generate audio in the frequency domain, producing spectral representations that are then converted into sound (Engel et al.). Yet other attempts use a more symbolic approach, generating musical scores or MIDI instructions that can be played on real or electronic instruments (Huang et al.).

For the purposes of the Rowdy Krause these approaches all seemed either too limited or too open. The symbolic efforts are typically used to generate music or speech, both of which are rooted in human culture. If the project of endemic machines is based on using more-than-human intelligence to interact in new ways with the more-than-human ecosystem, it would seem antithetical to limit the range of sounds to those generated by human cultures. The raw waveforms and frequency domain representations pose almost the opposite challenge: they

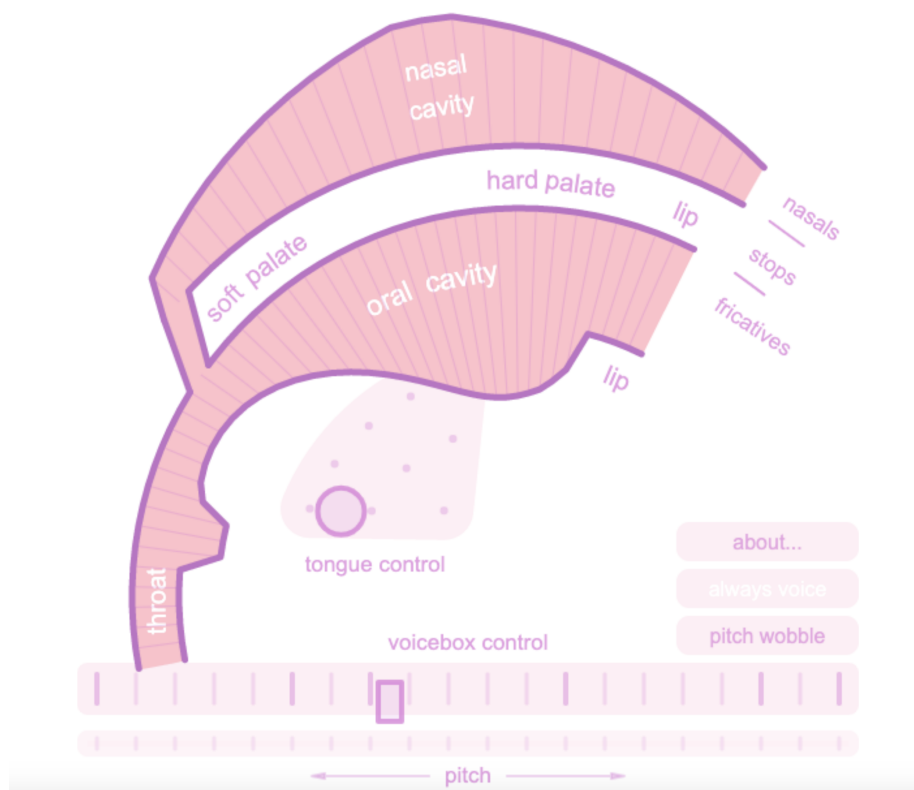
can create almost any sound imaginable, far beyond the range of what would be found in an ecosystem.

The task, then, was to identify a mechanism for generating sound that could produce a range of sounds that would not be out of place in an ecosystem, but that isn't limited to human semiosis. A search for a suitable mechanism led to the Pink Trombone (<https://dood.al/pinktrombone/>). Pink Trombone (fig. 2) is a vocal tract simulator made for touchscreen devices so that users can control a virtual tongue and palate as well as the voicebox to create human-like noises. However, it is also possible to re-code it so that the dimensions of the vocal tract can vary outside of the range of human anatomy and the tract can be controlled in ways that are not possible for a human to achieve. Together, these features formed an even balance between something that is based in biology, but not too specifically human.

To control Pink Trombone, the ANN's outputs were connected to the control points that determine the shape of the vocal tract's throat, tongue, and lips (see fig. 2). The ANN's outputs were calculated at every time step and the shape of the vocal tract was adjusted accordingly. This produced a unique vocalization for each evolved iteration of the neural network structure.

In NEAT, the process used to evolve the ANN, a fitness is calculated for each of the individuals — the different neural network structures — in a generation (Stanley and Miikkulainen). The ANNs that produce the highest fitness levels in each generation are selected to reproduce and form the next generation, driving the population towards higher fitnesses. The design of an appropriate measurement of fitness is a key component in the success of an evolutionary algorithm.

The goal of the Rowdy Krause was to find an acoustic niche in an existing ecosystem. Thus, the fitness for an ANN controlling Pink Trombone was calculated as the uniqueness of the spectral composition of the sound that was produced. In practice, this involved sampling the



© 2017 Neil Thapen

Fig. 2. The Pink Trombone interface (Used with permission). A video of the Rowdy Krause prototype evolving control of Pink Trombone in real time is available at <https://vimeo.com/359044847>.

soundscape of the ecosystem in question, performing a fast Fourier transform (FFT) to calculate its spectrum, and then creating a database of the spectral components of the soundscape samples. Sounds produced by the ANN and Pink Trombone were then compared to this database and those most different from the recorded sounds in the database were determined to be most fit.

This system was demonstrated initially at a workshop in the summer of 2019, using recorded audio from a forest garden in southern Sweden as the soundscape. This proof-of-concept demonstration was entirely virtual and offline, but the sounds that were generated were interesting enough to warrant further exploration.

EMBODIED IMPLEMENTATION

The prototype demonstrated the efficacy of evolving a neural network to control the Pink Trombone. Missing, however, was the element of feedback from the ecosystem that is an essential component of an endemic machine. The Rowdy Krause could learn from the recorded audio in the virtual versions, but the ecosystems in question had no opportunity to respond to the Rowdy Krause.

The shift from a prototype to an embodied device prompted a miniaturization of the computational components of the system. The code that had previously been executed on a laptop was now running on a Raspberry Pi — a single-board, embedded computer with vastly less processing power. The consequence of this was that code that had run previously in realtime, a key feature of a system that generates live audio, was now unable to do so. The process had to be reconsidered.

The prototype system had three subsystems: the Pink Trombone, the evolutionary system (NEAT), and a control system that linked the evolved ANNs to the Pink Trombone's interface. The combination of these three elements in this way was convenient, but highly inefficient. To enable the software to run on the embedded computer, the Pink Trombone was recoded into

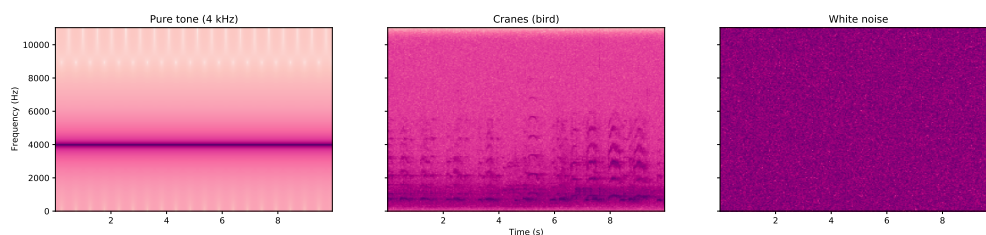


Fig. 3. Sample spectrograms showing the frequency components of a pure 4 kHz tone (left), the call of a crane (centre), and white noise (right). The horizontal axis is time and the vertical axis is frequency, with the darkness of the image at a point indicating the intensity of that frequency component of the sound at a given point in time.

Python and better integrated into the rest of the software.

The system worked — the embodied the Rowdy Krause was able to evolve a vocalization — but minor changes in the new implementation of the process meant that the sounds it produced were quite different from those of the virtual prototype. In place of the types of vocalizations produced by the prototype — a sort of rhythmic blooping that sounded vaguely like it could have come from an undiscovered primate — the sounds were more often longer and droning. This veered occasionally into an unpleasant high-pitched whine.

Consideration of the structure of the fitness function revealed the likely reason for this. The fittest sounds — those most spectrally different from the set of sounds heard in the soundscape — are likely to be pure tones on frequencies that have minimal usage (see fig. 3 for sample spectrograms). The structure of the virtual system had constrained the Rowdy Krause from finding these sounds but the implementation in the embodied version of the Rowdy Krause made it possible. The fitness function now had to be adjusted to account for the change in capability of the instrument.

From a biosemiotic perspective, the problem with droning vocalizations and pure tones is that they are informationally poor. A source that produces only tonal sounds tends to have

low Shannon entropy which, from the perspective of information theory, means that it has low information content. A tonal vocalization might be appropriate as signal of alarm, but not for general communication.

The other challenge is that, in this configuration, the Rowdy Krause had a tendency to become trapped in an evolutionary dead end. As the evolving population of neural networks found these tonal vocalizations, the populations converged towards these types of highly fit solutions. They ceased to explore the evolutionary landscape of possible vocalizations.

To address these two problems, two modifications were made to the structure of the neuroevolutionary process: The fitness function was adjusted to reward sounds with greater spectral entropy; and the vocalizations produced by the various evolved neural networks were added to the database of sounds alongside those recorded from the soundscape. The first modification encouraged the pursuit of a more temporally varied vocalization. The addition of evolved vocalizations to the database meant that the populations were encouraged to create sounds different from those they had created in the past in addition to being different from sounds recorded from the soundscape. The effect of this was to create a sort of novelty search, driving the evolutionary neural networks toward new configurations and helping them to avoid becoming stuck in a particular type of vocalization.

With these modifications, the Rowdy Krause was deployed virtually once again as an art installation at the Artificial Life virtual conference in the summer of 2020. For this version, titled Virtual Rowdy Krause - Point Pelee, vocalizations were evolved using streaming audio from a Point Pelee National Park near Windsor, Canada. A recording of that work is available online².

The embodied version of the Rowdy Krause was also tested in the field in Malmö,

2

Sweden in October, 2020. Due to travel restrictions related to the COVID-19 pandemic, the field site was moved from a community garden in Copenhagen to the balcony of my apartment in Malmö, where it overlooked a small park and playground, a busy urban road, and an active construction site for the regional hospital. A short video recording of the field experiment can be viewed online³.

VOCALIZATIONS

The two iterations of the final version of the Rowdy Krause — one evolving in a virtual soundscape from Point Pelee and the other in the real soundscape on my apartment balcony — were able to produce vocalizations. Those vocalizations were varied, occasionally sounding like a strange frog and other times more like the wind whistling over a pipe without quite producing resonance.

The embodied version of the Rowdy Krause inhabited the balcony for approximately two weeks. About a week into its residency, I heard a sound from outside and found myself unsure of whether it was coming from the Krause. Whether it was the Rowdy Krause or a bird or something else entirely, I found myself paying more attention to the soundscape outside my workspace than I had previously.

NICHE CONSTRUCTION AND ENDEMISM

It is possible to determine more precisely whether the Rowdy Krause was able to construct a niche in the two ecosystems that it inhabited. Figure 4 visualizes the results of the two experiments. The plots are two-dimensional representations of the spectral components of sounds created using a process called t-distributed Stochastic Neighbour Embedding (t-SNE). Each point represents a recorded or evolved sound and the proximity of two points reflects their similarity.

3

The darkness of the points representing evolved sounds shows when in the evolutionary process that sound was produced, with sounds from early in the evolutionary process appearing lighter.

Figure 4a shows the evolved sounds from the balcony experiment in blue and the recorded sounds from the soundscape in green. There is very little overlap between the clusters of evolved and recorded points, indicating that, for the most part, the evolved sounds were spectrally different from the recorded soundscape. The Rowdy Krause appears to have been successful in constructing an acoustic niche — represented by the cluster of blue points — that is distinct within the soundscape. A similar pattern is seen in the evolved (orange) and recorded (purple) sounds from Point Pelee in fig. 4c.

The plots can also help to address the question of whether the Rowdy Krause was able to become endemic to these two soundscapes. One of the features of endemism is particularity to a place. Figure 4b shows the evolved vocalizations from the two experiments. The points have some overlap towards the centre of the plot, but for the most part are found in two distinctive clusters. This indicates that the evolutionary processes proceeded differently in the two soundscapes. However, it is difficult to discern the degree to which this is due to the different soundscapes or different random initial conditions for the evolutionary process. It does mean that it is possible that the Rowdy Krause demonstrated a degree of endemism.

REFLECTIONS

There is a point in the description of the design process in the previous section that illustrates a central issue in the design of endemic machines. After the redesign of the Rowdy Krause's software to work on an embedded system, it was not performing in the same way it had been in the initial prototype. Moreover, the change in performance produced a result that was not in line with my expectations of what the Rowdy Krause should sound like. As a result, adjustments were made to the code.

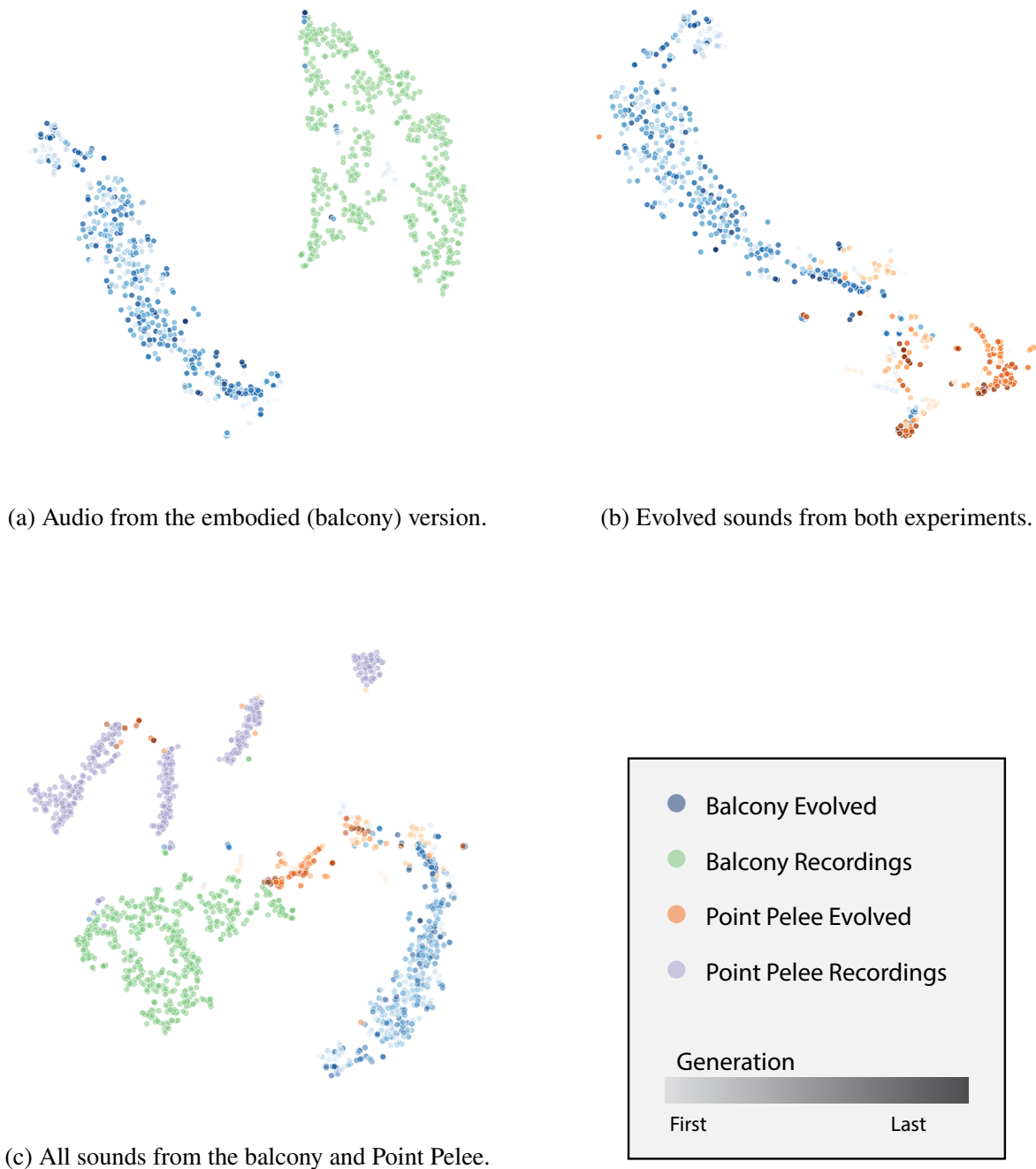


Fig. 4. Plots of the recorded and evolved sounds from the experiments at Point Pelee (virtual) and the balcony (embodied). Sounds that were evolved later are shown in darker shades. The plots use t-SNE to visualize high-dimensional data (the frequency components of each sound) on a two-dimensional plane. The space between points relates to the similarity between the sounds with sounds that are more alike being clustered together.

In retrospect, however, there is a question as to whether those adjustments should have been made. It is not clear that the changes served the ecosystem. While the new vocalizations did not meet my expectations, they may still have been ecologically relevant. Was the Rowdy Krause broken or merely being creative in a manner that was alien to me? In the design of endemic machines, this is often unclear.

Particularly in the throes of a design cycle, it can be difficult to step back and consider whether a result is correct for an ecosystem. One way of addressing this is found by returning to Bateson. An often omitted part of his quotation about "the pattern which connects" is that "[t]he pattern which connects is a meta-pattern. It is a pattern of patterns." Focusing on the result is problematic because the result is not necessarily intended for human consumption. One might instead concentrate on the processes — the patterns of patterns — to determine whether an endemic machine is working.

Works Cited

- Abram, D. *The spell of the sensuous: Perception and language in a more-than-human world*. New York: Vintage Books, 1996.
- Andreas, Jacob, et al. “Cetacean Translation Initiative: a roadmap to deciphering the communication of sperm whales.” arXiv: 2104.08614 (Apr. 2021). arXiv: 2104.08614. <<http://arxiv.org/abs/2104.08614>>.
- Bateson, G. *Steps to an ecology of mind: Collected essays in anthropology, psychiatry, evolution, and epistemology*. 1972.
- Beatty, John. “The Creativity of Natural Selection? Part I: Darwin, Darwinism, and the Mutationists.” *Journal of the History of Biology* 49.4. ISSN: 0022-5010. DOI: 10.1007/s10739-016-9456-5 (Dec. 2016): 659–684. <<https://link.springer.com/article/10.1007/s10739-016-9456-5>>20<http://link.springer.com/10.1007/s10739-016-9456-5>>.
- . “The Creativity of Natural Selection? Part II: The Synthesis and Since.” *Journal of the History of Biology* 52.4. ISSN: 0022-5010. DOI: 10.1007/s10739-019-09583-4 (Dec. 2019): 705–731. <<https://doi.org/10.1007/s10739-019-09583-4>>20<http://link.springer.com/10.1007/s10739-019-09583-4>>.
- Bergerman, Marcel, et al. “Robotics in Agriculture and Forestry.” *Springer Handbook of Robotics*. Edited by Bruno Siciliano and Oussama Khatib. Cham: Springer International Publishing, 2016. 1463–1492. DOI: 10.1007/978-3-319-32552-1_56. <http://link.springer.com/10.1007/978-3-319-32552-1_56>.
- Chan, Matthew T. K., et al. “Interacting with curious agents: User experience with interactive sculptural systems.” *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, Aug. 2016. 151–158.

Coffey, Kevin R., Russell G. Marx, and John F. Neumaier. “DeepSqueak: a deep learning-based system for detection and analysis of ultrasonic vocalizations.” *Neuropsychopharmacology* 44.5. ISSN: 0893-133X. DOI: 10.1038/s41386-018-0303-6 (Apr. 2019): 859–868. <<https://doi.org/10.1038/s41386-018-0303-6>><<http://www.nature.com/articles/s41386-018-0303-6>>.

Darwin, C. R. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. 1st. London: John Murray, 1859. 502.

Dobzhansky, Theodosius. “Chance and Creativity in Evolution.” *Studies in the Philosophy of Biology*. Macmillan Education UK, 1974. 307–338. DOI: 10.1007/978-1-349-01892-5_18. <https://link.springer.com/chapter/10.1007/978-1-349-01892-5_18>.

Engel, Jesse, et al. “GANSynth: Adversarial Neural Audio Synthesis.” *ICLR 2019*. New Orleans, USA, Sept. 2019.

Farina, Almo. *Soundscape Ecology*. Dordrecht: Springer Netherlands, 2014. DOI: 10.1007/978-94-007-7374-5.

Gould, Stephen Jay. *Ever Since Darwin: Reflections in Natural History*. New York, USA: W. W. Norton, 1977. 285.

Hayles, N. Katherine. *How we think*. Chicago: The University of Chicago Press, 2012.

Hines, Megan, David Kadish, and Maja Fagerberg Ranten. “Sound as Material for Eco-technogenesis.” *Proceedings of RE:SOUND Media Art Histories 2019*. Edited by Morten Sondergaard and Laura Beloff. Aalborg, Denmark: BISL, Aug. 2019. 1–7.

Huang, Cheng-Zhi Anna, et al. “Music Transformer.” arXiv: 1809.04281 (Sept. 2018). arXiv: 1809.04281. <<http://arxiv.org/abs/1809.04281>>.

- Krause, Bernie L. "Bioacoustics, Habitat Ambience in Ecological Balance." *Whole Earth Review* 57, ISSN: 0749-5056 (1987): 14–18.
- Kroos, Christian. "The art in the machine." *Cognitive Science and Technology*. Springer International Publishing, 2016. 19–25. ISSN: 21953996. DOI: 10.1007/978-981-10-0321-9_2. <https://link.springer.com/chapter/10.1007/978-981-10-0321-9%7B%5C_%7D2>.
- Laland, Kevin, Blake Matthews, and Marcus W. Feldman. "An introduction to niche construction theory." *Evolutionary Ecology* 30.2. ISSN: 0269-7653. DOI: 10.1007/s10682-016-9821-z (Apr. 2016): 191–202. <[/pmc/articles/PMC4922671/%20/pmc/articles/PMC4922671/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4922671/%20http://link.springer.com/10.1007/s10682-016-9821-z](https://pmc/articles/PMC4922671/%20/pmc/articles/PMC4922671/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4922671/%20http://link.springer.com/10.1007/s10682-016-9821-z)>.
- Lehman, Joel, et al. "The Surprising Creativity of Digital Evolution: A Collection of Anecdotes from the Evolutionary Computation and Artificial Life Research Communities." *Artificial Life* 26.2. ISSN: 1064-5462. DOI: 10.1162/artl_a_00319 (May 2020): 274–306. <https://doi.org/10.1162/artl%7B%5C_%7Da%7B%5C_%7D00319%20https://www.mitpressjournals.org/doi/abs/10.1162/artl%7B%5C_%7Da%7B%5C_%7D00319>.
- Lozano-Hemmer, Rafael and Heimo Ranzenbacher. "Metaphors of Participation." *Takeover: Who's Doing the Art of Tomorrow*. Edited by Gerfried Stocker and Christine Schöpf. Linz, Austria: Springer, 2001. 240–243.
- McCormack, Jon. *Creative ecosystems*. Edited by Jon McCormack and Mark D'Inverno. Berlin: Springer-Verlag, 2012. 39–60.
- Morrone, J.J. "Endemism." *Encyclopedia of Ecology*. Elsevier, Jan. 2008. 1254–1259. DOI: 10.1016/B978-008045405-4.00786-2. <<https://linkinghub.elsevier.com/retrieve/pii/B9780080454054007862>>.

- Oberti, R. and A. Shapiro. “Advances in robotic agriculture for crops.” *Biosystems Engineering* 146. ISSN: 15375110. DOI: 10.1016/j.biosystemseng.2016.05.010 (June 2016): 1–2. <<https://www.sciencedirect.com/ezproxy.library.ubc.ca/science/article/pii/S1537511016302628> % 20<https://linkinghub.elsevier.com/retrieve/pii/S1537511016302628>>.
- Odum, Eugene P. *Fundamentals of ecology*. 3rd. Philadelphia: W. B. Saunders Company, 1971.
- Pijanowski, Bryan C., et al. “Soundscape Ecology: The Science of Sound in the Landscape.” *BioScience* 61.3, ISSN: 1525-3244. DOI: 10.1525/bio.2011.61.3.6 (Mar. 2011): 203–216. <<https://academic.oup.com/bioscience/article-lookup/doi/10.1525/bio.2011.61.3.6>>.
- Pocheville, Arnaud. “The Ecological Niche: History and Recent Controversies.” *Handbook of Evolutionary Thinking in the Sciences*. Dordrecht: Springer Netherlands, 2015. 547–586. DOI: 10.1007/978-94-017-9014-7_26. <http://link.springer.com/10.1007/978-94-017-9014-7%7B%5C_%7D26>.
- Schafer, R. Murray. *The tuning of the world*. A.A. Knopf, 1977. 301.
- Stanley, Kenneth O. and Risto Miikkulainen. “Evolving Neural Networks through Augmenting Topologies.” *Evolutionary Computation* 10.2. ISSN: 1063-6560. DOI: 10.1162/106365602320169811 (June 2002): 99–127. <<http://www.mitpressjournals.org/doi/10.1162/106365602320169811>>.
- Stiegler, Bernard. *Technics and Time, 1: The Fault of Epimetheus*. Stanford University Press, 1998.
- Van den Oord, Aaron, et al. “WaveNet: A Generative Model for Raw Audio.” arXiv: 1609.03499 (Sept. 2016): 1–15. arXiv: 1609.03499. <<http://arxiv.org/abs/1609.03499>>.

- Van Wynsberghe, Aimee and Justin Donhauser. “The Dawning of the Ethics of Environmental Robots.” *Science and Engineering Ethics*. ISSN: 1353-3452. DOI: 10.1007/s11948-017-9990-3 (Oct. 2017). <<http://link.springer.com/10.1007/s11948-017-9990-3>>.
- Yang, Chenghai, Ruixiu Sui, and Won Suk Lee. “Precision Agriculture in Large-Scale Mechanized Farming.” *Precision Agriculture Technology for Crop Farming*. CRC Press, Oct. 2015. 177–211. DOI: 10.1201/b19336-6. <<https://www.taylorfrancis.com/chapters/precision-agriculture-large-scale-mechanized-farming-chenghai-yang-ruixiu-sui-won-suk-lee/e/10.1201/b19336-6%20https://www.taylorfrancis.com/books/9781482251081/chapters/10.1201/b19336-6>>.
- Yang, Xinting, et al. “Deep learning for smart fish farming: applications, opportunities and challenges.” *Reviews in Aquaculture* 13.1. ISSN: 1753-5123. DOI: 10.1111/raq.12464 (Jan. 2021): 66–90. <<https://onlinelibrary.wiley.com/doi/10.1111/raq.12464>>.