Dreaming Species

Melanie Wilson

Abstract

Dreaming Species (www.dreamingspecies.com) is an online audio visual artwork designed for headphones, conceived by artist composer Melanie Wilson (U.K.). Across three episodes, this dream-like piece opens speculative doors into a world in which human connection to animals is re-imagined using the processes of machine learning. To engage with the crisis of biodiversity loss in the U.K., Wilson collaborated with a neural net, the voices of animals at risk of extinction and human voices, to create a sonic aesthetic that stages the conflicted terrain of human/animal relations. Combining multi-part vocal music, animal and machine voice, text and binaural recording techniques, Dreaming Species is the opening creative chapter of a multi-year research process, in which Wilson is exploring the possibilities of machine learning to imagine new sonic meeting grounds for human and animal narratives.

1 Creative Team

Concept, score, text, voiceover, & sound design
Melanie Wilson

Singers
Peyee Chen, Adey Grummet, Simone Ibbett-Brown and Melanie Pappenheim

Neural Net training voices
Peyee Chen and Héloïse Werner

Neural net programmer
Dr Christopher Melen

Video artist
Akhila Krishnan

Website design, identity & additional video design
Graeme Swinton

Music facilitator
Yshani Perinpanayagam

Project producer
Fuel

Recorded at Urchin Studios, London

2 Initiating Idea

The research process that informs Dreaming Species began as part of Sound and Music's New Voices 2020/21 composer programme. Through this programme and with the support of the producing company Fuel, I began a partnership with the Practise Research in Science and Music Lab (PRiSM), based at Royal Northern College of Music, Manchester, U.K.

The impetus for this project was to make a creative response to the crisis of biodiversity. Since the industrial revolution, the U.K has lost almost half of its biodiversity, a winnowing birthed in Manchester, my place of home and creative practise. My aim was to explore the human connection with non-human species through sound and music. The research question that initiated this exploration was: can a language for sound be created that allows the listener to perceive interdependence with other species, whether poetically or actually?

This question formed through creative curiosity in and solidarity with the more than human voices we share our world with, and an interest in connecting audiences with the fragility and necessity of our bond with those voices.

I chose to co-create the beginnings of this research process with a neural net, to explore a vocal aesthetic that stages a meeting place in sound between humans and animals: a speculative voicing, within which to play. I was interested in how machine learning might 'listen' for varying relational aspects in those voices, and whilst not entirely circumventing my anthropocentric biases, might perhaps re-frame them in a way that allowed new insights or alternative paths towards encounter and composition.

2 Animals

To prepare for training the neural net, I built a cohort of recordings of vocalising animals experiencing threat, both native to the U.K. and globally. Some of these voices I recorded in the field, some I was given by the composer Sarah Kierle, and some I sourced from libraries.

The creation of this cohort raised an initial question of anthropocentric bias. Animals such as wolves, indri or song thrush have expressive vocalisations that occupy ranges more easily accessible to human voice. I assumed this might be more readily mapped with human patterns by the neural net, but not inclusive of other animals who create sound in different ways.

I therefore also included wart biter crickets, who stridulate to create sound, and horseshoe bats who echolocate. Also included were common loon, great bittern, corncrake, humpback whale, adder, natterjack toad, red squirrel, gibbons, otter, turtledove, arctic fox, elk, puffin, orangutan, orca, nightjar, honeybee, hedgehog, black crested gibbon and lynx.

Once this cohort was created, I spent time listening to these animals in detail. I then extrapolated motifs and phrases that I could notate into scored music, or translate into vocal instruction scores to share with singers.

During this process, I thought about the erasure of animal microtones when translated into an equal tempered scale for scored music, and that scale being an example of the force of western human organisation. I also thought about the scale as a working location to situate the human 'story' in the project, and a place to travel out from towards the pitch sphere of animals and their story.

3 PRiSM SampleRNN

The neural net I worked with is PRiSM SampleRNN. PRiSM Sample RNN is a project including development of the code prism-samplernn, an open-access neural audio synthesis software tool released on GitHub in June 2020 as part of PRiSM Future Music #2. Development of the software is funded by Research England, Expanding Excellence in England (E3).

The software is built upon the SampleRNN model which addresses unconditional audio generation in the raw acoustic domain, it is able to generate new audio outputs by 'learning' the characteristics of any existing corpus of sound or music. Artists are invited to collate their own datasets (or rather, sound libraries), curated for their unique creative contexts. Changing how these datasets are organised, as well as the parameters of the algorithm during the generation processes significantly change the resultant output – making these choices part of the creative process.
The software was developed in response to work by Dr Sam Salem (PRiSM Senior Lecturer in Composition). For his piece *Midlands* (2019), Salem made field recordings whilst walking 120km of the River Derwent. These materials were used to synthesise new sounds with Wavenet, one of the earlier deep-learning algorithms for audio generation, but the speed of the workflow made it

difficult to explore the full possibilities of the technique (documented in the PRiSM blog post *A Psychogeography of Latent Space*).

An alternative, SampleRNN, represented an opportunity to generate sound more quickly but the code was broken. Dr Christopher Melen, PRiSM Research Software Engineer (2019-2023), undertook a complete reimplementation of the original SampleRNN code[1], fixing broken dependencies and upgrading it to work with the latest versions of Python and Tensorflow. It constitutes a completely new implementation of the SampleRNN architecture (documented in the PRiSM blog post *A Short History of Neural Synthesis*).

4 First Training, June 2020

The first training data set I created for SampleRNN was based on equal proportions of animal and human voice examples. I composed a set of music studies in response to a selection of animals, for one voice. These studies responded to whales, wolves, crickets and bats, and took motifs and rhythms within the voices as their basis. I worked with soprano Peyee Chen to record these studies. The data set was compiled and sent to Dr. Melen at PRiSM for training.

SampleRNN's responses from this first training session were febrile and kaleidoscopic; like a high energy collision of species, heard through radio static. It was possible though to determine the beginnings of a type of voice hybridity, most often at the beginning or ending of a sonic gesture. For example: the in breath of a human followed by the decay tail of an orca or indri. The density and pace of the material felt challenging to penetrate, but the glimpses of hybridity motivated Dr Melen and I to prepare a second training phase.

5 Second Training, December 2020

I began this phase with long and repeated listening to SampleRNN's first training responses. I made edits of all the instances I thought I could identify SampleRNN creating hybridity. I then took this audio material into a second recording session with singers Peyee Chen and Héloïse Werner, which we did in two stages; firstly by singing back as accurately as possible SampleRNN's audio fragments in a call and response mode, and secondly by developing the fragment into a new vocal idea, through improvisation. In this way, we employed the processes of listening, repeating and elaborating to make a second human response to SampleRNN's initial proposal.

We applied this same iterative process to singing a replication of and answer to the same set of animal voices. This time though, I edited the collection to focus on recordings of animals that minimised background noise, so that the information SampleRNN was learning was as cleanly vocal as possible. I became aware of how this process separated the voice of the animal from its ecosystem context. I thought about how this separation perhaps perpetrated the very dislocation that the project seeks to question, and began to think about voice as a composite force: a mechanical process, an expression of identity and a manifestation of the systems of life.

Dr Melen trained SampleRNN on this second set of data, adjusting training parameters by point two percentage points. The results were identifiably more vocal and more hybridised in expression. What emerged now was a response that sounded more like animals attempting human speech; an eerie, anxious, inter-species puppeteering. Put another way, the neural net response breaks down the categories of human and non human, and reveals the human as another example of vocalising animal.

6 Creative development of machine generated material

I took the creative conclusions drawn from SampleRNN's second responses and the ethical questions surrounding the processes into the development of a new artwork. I also took the 'story' of the iterative learning process I had worked through, and the broader concepts of listening, repetition and memory.

My goal was to realise multiple creative outcomes from my research process, including a live music performance and a digital online work. I began with the digital work as a way to prototype a concise sonic and visual identity for the project as a whole. I wanted the initial results of my

research to reach broad audiences globally, and to use the online form of it to experiment with an individualised experience of listening, which could pre-figure a collective experience of listening in the later live work.

The creation of the digital piece was funded by Sound and Music and Arts Council England.

7 Project Outcome

Dreaming Species functions as a triptych, in which each episode shares a thematic cosmology, but is not narratively linked. It is designed to be listened to on headphones.

7.1 The Animals Are Leaving

Episode one takes its cue from John Berger's poem *They Are the Last,* featuring the thought 'the animals are leaving… only pets and carcasses remain'. The spine of the piece is a stream of consciousness voice, colliding encounters between humans and animals in everyday life. This vocal form is a creative translation of the neural net aesthetic of collision. The voice (recorded binaurally), is tightly woven together with SampleRNN's material and animal voice, creating an intricate, ambivalent experience of the conflicted terrain between humans and other intelligences.

The video element of this episode is composed of abstract colour fields that move and respond to the currents of the music. The aesthetic is deliberately restrained to foreground the listening experience, and to hold space for the listener to interpret the density of sound images for themselves.

7.2 Memory Palace

Episode two memorialises the many species facing extinction, introduced by the word cloud of animal species that begin. This episode fictionalises SampleRNN's recursive learning process, finding equivalence between its layers of learning, memorising and passing on of data with human traditions of oral storytelling and song sharing. In this episode, a group of humans become custodians of the voices of extinct animals, singing them in order to keep them alive. The music and the visual worlds interpret the iterative journey of creation with machine learning as a mythic space, and auditions its potential potency as an advocate for solidarity with other species.

7.3 Ritual

Episode Three proposes the idea of animal and human experience mapped across time. The attempt here is to bring attention to the reality of animal structures and bodies existing along side our own. The video uses algorithmic processes to interpose two differing habitats from opposite hemispheres, reflecting the story of duality between a human and orangutan's co-existance. This episode also engages with the aesthetics and ethics of extractive field recording practises. It inverts the microphone to scrutinise the human presence usually erased from field recording, and instead listens for the animal on the other side of the planet through the imaginative labour of song.

Citation

prism-samplernn
Led by Sam Salem and Christopher Melen
Initiated by Sam Salem
prism-samplernn code by Christopher Melen

A PRiSM Collaboration also involving David De Roure, Marcus du Sautoy, and Emily Howard.
The RNCM Centre for Practice & Research in Science & Music (PRiSM) is funded by the Research England fund Expanding Excellence in England (E3).

References

1 The original SampleRNN architecture was described in the paper SampleRNN: An Unconditional End-to-End Neural Audio Generation Model (2017). This version was based on Python 2 and Theano, a library for performing fast computations involving matrices. This version formed the basis for the Dadabots' famous Relentless Doppelganger livestream on Youtube (https://www.youtube.com/watch?v=MwtVkPKx3RA). Since both Python 2 and Theano are officially deprecated it was decided that PRiSM would offer its own implementation, based on Google's popular Machine Learning library TensorFlow 2.

Resources

prism-samplernn code Github
prism-samplernn Google Colab Notebook

Artist biography

Melanie Wilson is a U.K. based inter-disciplinary composer, writer and performance maker. Her work is founded on the dialogue between sound, experimental forms of composition, language, technology and live performance. Wilson's projects create spaces for audiences to feel and think deeply through listening, and are shown in theatre, new music, digital sound and new opera contexts.

Recent work includes 'Dreaming Species', an online binaural audio visual work for voices and AI, 'glass human', a chamber opera with electronics for Glyndebourne and 'Current, Rising', a hyper-reality opera for Royal Opera House.

Wilson also collaborates with artists and companies in the U.K and internationally, including long term sound and composer collaboration on the work of theatre director Katie Mitchell. She was Chair of the Association for Sound Design & Production 2020 - 23.