

# Neural Synthesis as a Methodology for Art-Anthropology in Contemporary Music

MARK DYER

Royal Holloway, University of London  
Email: [mark.dyer@rhul.ac.uk](mailto:mark.dyer@rhul.ac.uk)

**This article investigates the use of machine learning within contemporary experimental music as a methodology for anthropology, as a transformational engagement that might shape knowing and feeling. In *Midlands* (2019), Sam Salem presents an (auto)ethnographical account of his relationship to the city of Derby, UK. By deriving musical materials from audio generated by the deep neural network WaveNet, Salem creates an uncanny, not-quite-right representation of his childhood hometown. Similarly, in her album *A Late Anthology of Early Music Vol. 1: Ancient to Renaissance* (2020), Jennifer Walshe uses the neural network SampleRNN to create a simulated narrative of Western art music. By mapping her own voice onto selected canonical works, Walshe presents both an autoethnographic and anthropological reimagining of a musical past and questions practices of historiography. These works are contextualised within the practice and theory of filmmaker-ethnographer Trinh T. Minh-ha and her notion of ‘speaking nearby’. In extension of Tim Ingold’s conception of anthropology, it is shown that both works make collaborative human and non-human inquiries into the possibilities of human (and non-human) life.**

## 1. INTRODUCTION

A circular machine, of new design/In conic shape: it draws and spins a thread/Without the tedious toil of need-less hands. (Dyer 1757: book III, lines 292–4, 99, emphasis added)

This article explores the application of machine learning, specifically deep learning artificial neural networks, within contemporary practices of experimental music as a methodology for anthropology; namely, as a transformational engagement between people and machines that might ‘open a space for generous, open-ended, comparative yet critical inquiry’ (Ingold 2013: 4). It presents compositions by composers Sam Salem and Jennifer Walshe before contextualising moments of the uncanny in these works within both the practice and the theory of filmmaker-ethnographer Trinh T. Minh-ha. By examining the collaborative and cultural role that algorithms (and by extension, artificial neural networks) can play in art-documentary, this article offers new possibilities for Tim Ingold’s conception of anthropology (2017: 22). Ultimately, it is shown how practices of

algorithmic art-anthropology, using machine logic and *threads* of code, can tell a story (or spin a *yarn*) and constitute an attitude towards learning and life itself.

Experimental music, both historic and contemporary, provides numerous examples of the use of both artificial intelligence and, separately, ethnographic approaches. Examples of the former are multifarious and include composer–performer–programmer George Lewis’s composition *Rainbow Family* (1984) and subsequent virtual orchestra software *Voyager*. Developed at IRCAM, Lewis’s explorations into improvising human–computer interactions utilised early AI and cybernetic technologies and placed an emphasis on hybridity and ‘sociomusical networks’ (Lewis 2020). Similarly, sound artist and performer Laetitia Sonami uses machine learning to map interactive, digital instruments, including the *lady’s glove* (1991) and *Spring Spyre* (2012). In constructing the latter, Sonami collaborated closely with software developer Rebecca Fiebrink and her machine learning tool *WeKinator*. Like *Rainbow Family*, *Spring Spyre* provides an example of the back-and-forth ‘exchange’ (Fiebrink and Sonami 2020) between instrument and performer and, furthermore, between artist, programmer and algorithm, as discussed later. Finally, though not a *sonic* artwork, composer Alexander Schubert’s ongoing project *CRAWLERS* (published 2021) provides a compelling comparison to the works discussed later. Developed in conjunction with the ZKM Center for Art and Media, *CRAWLERS* comprises a collective of AI bots that crawl online user data and generate fake social media posts, creating a ‘parallel social network of warped truths’ (Schubert 2021). This emphasis on the uncanny duplicate and reimagining of cultural artefacts, rather than on live, interactive systems, is pertinent to the following discussion. Such examples, whilst not intended to be exhaustive, illustrate the varied applications of artificial intelligence both in the creative process and as part of live-interactive works by composers and artists. Two general trends relevant to the following discussion might be identified within such practices. These include a growing concern for machine learning – specifically, content generated using artificial neural networks – and, crucially, a transition from projects based within and

funded by large research labs (such as IRCAM) to informal, collaborative practices. These trends correlate with the development of relevant software and computing power, and the growing accessibility and affordability of such technologies and big data (Goodfellow, Bengio and Courville 2016: 19–36; Steels 2020: vi; Briot, Hadjeres and Pachet 2020: 1).

Separately, there are numerous examples of experimental music in which composers and sound artists adopt ethnographic methods. Composer and sound ecologist Hildegard Westerkamp self-scrutinises her listening experience of environmental ambient spaces in the autoethnographic *Kits Beach Soundwalk* (1989). Similarly self-reflexive, composer Trond Reinholdstein explores and critiques the cultural boundaries, perceptions and customs within Western art music in the Dadaist lecture-piece *Theory of the Subject* (2016). By contrast, composer Joanna Bailie explores a more archival approach by deriving quasi-fictionalised biographies from old photos in audiovisual works such as *The Grand Tour* (2015) and *Roll Call* (2018). Likewise, Cassandra Miller uses audio recordings as historic documents and transcribes the idiosyncrasies of particular performers in works such as *For Mira* (2012) and *Guide* (2013). Once again, such examples are not exhaustive, but rather demonstrate the myriad practices incorporating ethnographic methods and concerns within contemporary composition and sound art. Perhaps a salient, but by no means essential, feature of such practice is the utilisation of multimedia, including field recordings, spoken word and image, to present explicit and reflexive interpretations of their subject matter. The convergence between soundscape composition and ethnography is well documented (Drever 2002; Rennie 2014; Iscen 2014; Anderson and Rennie 2016; Martin 2017). Whilst John Drever (2002: 24) places an unconditional emphasis on ‘open-air-research’ and ‘fieldwork’, Tullis Rennie (2014) instead propagates a more rigorous socio-sonic methodology. This array of artistic examples demonstrate a more liberal approach to fieldwork, but perhaps find consensus in their presentation of ‘self-reflexive narratives’ in sound (Anderson and Rennie 2016: 226).

The two works explored in the following sections represent novel explorations in combining these two subdisciplines, in which artificial intelligence, namely deep learning, is used as a methodology for anthropology.

## 2. SAM SALEM: *MIDLANDS*

In *Midlands* (2019),<sup>1</sup> for chamber ensemble, performative electronics, tape and dual video projection,

<sup>1</sup>Commissioned by Bludenzener Tage Zeitgemäßer Musik and Distractfold, funded by the Ernst Von Siemens Music Foundation.

British-Jordanian composer Sam Salem explores various forms of ethnography. Written for the ensemble Distractfold with guest performers, the large-scale work offers a biomythographic<sup>2</sup> account of Salem’s often problematic relationship to his childhood hometown – the city of Derby, UK (Salem 2020d).

Like much of the Salem’s work, *Midlands* is concerned with space and place (Salem 2020a). In composing the piece, Salem walked the 120km of the River Derwent from source to mouth and ‘recorded, filmed and experienced the changing geographical and social landscape’ (Salem 2020d), drawing upon practices of psychogeography and ‘ambulatory divination’ (Salem 2020a). The materials gathered on this excursion form the compositional and conceptual basis for *Midlands*. Audio and video footage describing river valleys, cotton industry, suburban architecture and St George’s flags are used as raw and processed materials (detailed later), whilst objects that hint towards Derby’s industrial past, such as metal springs, a thunder sheet and transducers, are utilised as performative instruments. Extracts of John Dyer’s 1757 poem *The Fleece*, referenced in the Introduction, are also borrowed. In his assemblage and traversal of found recordings, objects and their connotations – a form of ‘compositional “hyperspace”’ (Salem 2020a) – Salem presents a sociohistorical and autobiographical narrative of industrialisation, post-industrialism, racial othering and belonging. *Midlands* is in effect (auto)ethnographic.

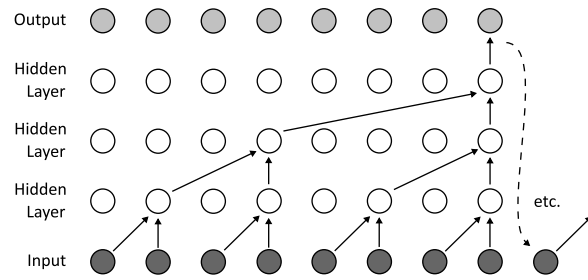
In the fourth section of *Midlands*, aptly named ‘How to Build A Machine’, Salem employs a model based on the WaveNet neural network architecture to generate musical material. Developed by Google’s DeepMind project, WaveNet is a deep generative model for raw audio waveforms. Salem trained the network to learn and generalise a dataset consisting of thousands of field recordings from his personal database (Salem 2020a). This dataset is fed to the network, which consists of multiple *convolutional* layers that filter and summarise their input, the output of one layer being fed as the input to the next. Convolutional networks are more commonly used to process image datasets, and involve a filter (or ‘kernel’) that passes over the 2D pixel surface to extract specific features. In contrast, audio sample data consist of 1D sequences. To counter this, WaveNet employs *causal convolution*, in which the filter has an inbuilt directional bias so that timesteps ahead of the current sample are not processed by the kernel until subsequent steps. In other words, the network effectively ‘listens’ to and generalises the audio from start to finish. The repeated application of these filters

<sup>2</sup>A term coined by feminist writer Audrey Lorde in reference to her novel *Zami: A New Spelling of My Name* (1982) to describe a narrative form in which biography, myth and history overlap.

generates, at each step, a probability distribution (i.e., the expected outcome) for the next sample, conditioned by all the previous samples (Melen 2020). Put more simply, the algorithm asks, having analysed an input sample  $x_1$ , what is the most likely output for the next sample  $x_t$ , given its knowledge of the entire sequence so far,  $p(x)$ ?

In addition to this novel form of convolution, each consecutive layer in the WaveNet model expands the *receptive field* (the size of the filter) of the previous layer by ‘poking holes’ in the kernel, effectively sub-sampling its input, a technique known as *dilated convolution*. As shown in Figure 1, through the process of dilation, the input is refined over successive layers and fed back into the prediction for the next step, creating a non-linear progression (Oord and Dieleman 2016; Melen 2020). This process of *amplification* is not dissimilar to the iterative evolutions in Alvin Lucier’s *I am Sitting in a Room* (1969). However, through the process of dilation, the model’s probability set – its understanding of the input so far and ability to predict new outputs – develops exponentially and circuitously. Depending on the variance of the training dataset and, more importantly, the width of the receptive field (i.e., the focus with which the network ‘listens’ to the input), audio outputs can become highly homogeneous on the one hand and wildly entropic on the other. Indeed, the reader need not know exactly *how* an audio output is generated but that such outputs can easily become structurally incoherent (Melen 2020) with ‘second-to-second variations in genre, instrumentation, volume and sound quality’ (Oord et al. 2016). That outputs are simultaneously derived from, and yet might be incoherent with, a given source is crucial in understanding the uncanny properties of the audio generated in *Midlands*.

Salem then sampled the network’s exponential learning and seeded it with a recording of the source of the River Derwent at Bleaklow moor in the Peak District, UK. With a widened receptive field, and therefore given an expanded range within which to make predictions, the trained model subsequently synthesised new audio (Salem 2020b). We first hear the unprocessed river source – two seconds of limp dribbling water. This is quickly subsumed by a more subterranean sound – the grinding of bedrock and crushing of fossilised matter – in turn frequently interrupted by incongruous explosions and clicks, an uncanny nod towards the industrial processes historically fuelled by the river. The generated audio is the algorithm’s sonic modelling of the river source, predicted through the distillation of sounds from both along the length of the river and Salem’s back catalogue of field recordings. This audio thus encapsulates and marries together the river and Salem himself – his personal sonic interests, experiences and identities.



**Figure 1.** Representation of WaveNet structure and convolution layers (amended from Oord and Dieleman 2016; used with permission).

In bars 53–90 of ‘How to Build A Machine’, Salem uses the generated audio as the basis for diffused sound design as well as for deriving performed timbres and textures. The voice, emulating the audio live via an earpiece, replies using unspecified and improvised extended techniques. In the 2019 premiere performance at Bludenzer Tage zeitgemäßer Musik, vocalist Ute Wassermann, with whom Salem developed the part, incorporates sucking in-breaths, croaks, hisses, tongue clicks and kisses. These noises marry with the tape part to create something utterly primordial; the synthesised river, generated by an algorithmic logic, is lent a quasi-human biology. Meanwhile, a quartet of violin, viola, bass clarinet and accordion provide an aeolian backdrop. From bar 58, the clarinet and accordion are asked to improvise using unpitched air sounds, joined by the viola and violin, whose players gently blow into their instrument microphones, at bars 64 and 69, respectively. During this terracing of mimicked environmental sounds, Salem invites the players onto his tour of the Derwent, asking the quartet to ‘Imagine wild cotton gently moving in the breeze’ (Salem 2019) while playing. The recorded signals of the voice and string players are subsequently transformed via a sound-on-sound loop through the thunder sheet and metal springs, respectively, whose resulting resonance also contributes to the ensemble texture (Salem 2020c).

Thus, Salem navigates the machine-generated audio through his compositional hyperspace. The resulting sound world, ASMR-inducing in its churning and tingling, is a composite of the human, the geologic and the machine. Derived from the Derwent, ‘whose flow shaped our world’ (Salem 2020d), and dilated through the composer’s sonic idiosyncrasies, the musical texture in performance follows and precedes video footage of the river, cotton plants, mill looms, post-war infrastructure, home video cricket and a musical quotation of *Acid Test* by Derby-based progressive metal band Gorilla – the eclectic human and non-human lives

and stories that, for Salem, spring from the same source.

'How to Build A Machine' is a not-quite-right representation of the river; sounds of the world that are not of this world. Salem describes machine learning as having the capacity to 'manufacture uncanniness' (2020a), the simultaneously strange and familiar, and as a way of bypassing personal taste and extending the compositional hyperspace established (2020a). We might understand his working through the outputs generated – his doubling of sounds using live performers and pairing with found audio and video materials to create an implicit and personalised collage – as 'a walk in a strange place, where objects do not fully make sense and are also not fully nonsensical' (2020a). We might also, however, equate this 'strange place', where familiar sights and sounds are now estranged from Salem, with his experience of 'growing up in Derby as other', subjected to racial discrimination and struggling with notions of home (Salem 2020d). This reference is allusive and made solely through abstract montage, but it is critical to our understanding of the work as autobiographical. *Midlands* is Salem's attempt to traverse these liminal spaces, both the manufactured and the past experiences recalled. His re-rendering of the River Derwent (and by extension his childhood hometown and questions of belonging) through the use of machine learning recalls an auto-sociogeologic narrative that does not exist and yet uniquely captures his impression of the world and his place within it.

### 3. JENNIFER WALSH: A LATE ANTHOLOGY OF EARLY MUSIC VOL. 1: ANCIENT TO RENAISSANCE

Irish composer and vocalist Jennifer Walshe similarly uses machine learning in her album *A Late Anthology of Early Music Vol. 1: Ancient to Renaissance* (Walshe 2020b). Featuring seventeen synthesised 'covers' of early notated music, ranging from the Seikilos epitaph of Ancient Greece to the florid polyphony of Giovanni de Palestrina, the album presents a simulated narrative of Western art music.

In creating *A Late Anthology*, Walshe collaborated with audio software engineers Dadabots (duo CJ Carr and Zack Zukowski) and their modified version of the neural network SampleRNN. SampleRNN also processes and generates raw audio but, unlike WaveNet, does not use convolution layers. Instead, SampleRNN processes sequential data, such as time series, using a Recurrent Neural Network that can retain an internal memory of previous states (Melen 2020). SampleRNN is similarly based on a probability distribution at each sample but adopts a hierarchical architecture whereby sample frames are consumed by

subsequent tiers of widening temporal resolution, as opposed to the concurrent structures of WaveNet (Melen 2020). Consequently, SampleRNN is more efficient than the latter in terms of both training time and processing power, being able to generate a greater number of outputs. This efficiency is exemplified by Dadabots's *Relentless Doppelganger* (2018), a continuous YouTube livestream of AI-generated technical death metal (Carr and Zukowski n.d.).

Dadabots trained SampleRNN on hours of a cappella recordings of Walshe's voice, producing 841 audio files made over 40 generations of learning (Walshe 2020b), in other words 40 tiers of developed probability distribution and increasing sample resolution. Walshe then mapped this raw audio and learning onto MIDI files of selected canonical works, obtained from open-source repositories such as ChoralWiki and the International Music Score Library Project (IMSLP)/Petrucci Music Library (Walshe 2020a; Haggett 2021: 113). The album begins with the Seikilos epitaph, one of the earliest notated pieces of music, performed with the network's most primitive training and basic ability to predict the next time sample. The result is a series of short shrieks and pops – vocal artefacts of Walshe's voice. As the training develops, the network is able to better predict the ensuing sample and generate more sustained sounds, the ensuing monophonic plainchants reconstituted as noisy drones. Throughout the album, progressing chronologically through the works covered, these sounds gradually evolve with greater dynamism and vocal fidelity into glitchy ghosts of the originals – there is even a clarion-like resemblance to the ground bass in the cover of John Dowland's 'Flow, My Tears'.<sup>3</sup>

In one sense, *A Late Anthology* is autoethnographic. As we hear the algorithm's progressive learning and growing ability to predict and imitate recordings of Walshe's voice, 'we hear and feel the neural network's comprehension of Walshe's singing change' (Poscic 2020), synthesising vocal habits and ticks she has developed over her performing career as a vocalist and improviser (Walshe 2018). Walshe describes the hearing-of-self through machine learning in *A Late Anthology* and other projects such as *ULTRACHUNK* (2018) as 'both uncanny and completely natural' (Walshe 2018). The generated audio perhaps tells us, if indirectly, the kinds of extended vocal sounds Walshe specialises in and favours, or sheds light upon the various aliases she adopts as an artist under the umbrella project Grúpat,<sup>4</sup> producing exponential doppelgangers. What does this neural learning tell us about Walshe's

<sup>3</sup>Available at: <https://jenniferwalshe.bandcamp.com/track/johndowland-flow-my-tears-air>.

<sup>4</sup>An introduction to Grúpat is available on the composer's website: <http://milker.org/anintroductiontogrupat>.



experiences as a musician, as a person, in a post-digital age? Consider Walshe's self-professed tendency to perform with a microphone and hear her own voice through various forms of signal processing and coding (Walshe 2018), or our more generalised collision with machine learning algorithms through customer profiling and marketing (Striphas 2015: 396), for instance. Perhaps the uncanniness of *A Late Anthology* is so natural to hear owing to our complete dependence upon, entanglement with and resemblance to such technologies (Walshe 2018).

In another sense, the album is anthropological. Walshe uses machine learning as a filter to listen to Western art music and how we document it. As a post-graduate teaching assistant at Northwestern University, Evanston, Illinois, Walshe taught the history of music in which two millennia were compressed into three terms, with selected works demonstrating a simplified narrative of linear and logical progression (Walshe 2020c; Haggett 2021: 113). The evolving sonic profile in *A Late Anthology*, with its own successive trajectory from recurrent squawks and murmurs to varied textures of vocal continuity, provides a satirical representation of the reductive narratives relayed in many anthologies and curricula. The album not only portrays a developing musical culture, but also 'is a way to think about why we write history, how we write history, who we choose to represent, who we choose not to represent' (Walshe 2020a). The choice of covers critiques, if only implicitly (Finan 2020), the gatekeeping of this history (e.g., by curricula, performance and recording programming and musical editions) in which works by primarily white male composers working within the Catholic Church are employed to neatly surmise a millennia of multifaceted artistic practices. *A Late Anthology* therefore not only reimagines a musical past, but also highlights and questions current practices of historiography in both formal and informal settings, and perhaps offers an alternative notion of an anthology and the canon it recognises going forward. It is a means of making sense of a culture that is 'utterly strange, utterly more bizarre ... than any science fiction' (Walshe 2018).

#### 4. SPEAKING NEARBY

In their use of neural synthesis – a process that relies on algorithms to learn a dataset – to generate sonic material within quasi-documentary formats, both works might be described as *algorithmic art-anthropology*. Both Salem and Walshe utilise machine learning to explore, refract and reimagine the pasts and presents of the personal contexts within which their works reside. Explore: as each composer must first gather the corpus of data from the world to train the algorithm, a process of finding, compiling and

sifting. Refract: as the logic of the network (unknown to the artists) decides what data to prioritise, what to disregard and reformulate accordingly, generating new sounds and sonic relationships. Reimagine: as both composers then employ the generated outputs to represent the worlds from which the data were gathered, turning the world back on itself through the lens of the machine. This turning offers new possibilities as to what might have been and what could be now. But what does the distance between the world and its representation by both artists achieve?

Both Salem's uncanny not-quite-river and the utterly strange Walshe-cum-Dowland-cum-machine, with their enigmatic logic and dislocation, are, for me, reminiscent of the holes, question marks and nonsense in the early documentaries of Vietnamese filmmaker-ethnographer Trinh T. Minh-ha. In works such as *Reassemblage* (Minh-ha 1982), *Naked Spaces: Living is Round* (Minh-ha 1985) and *Surname Viet Given Name Nam* (Minh-ha 1989), Minh-ha uses non-traditional methods of documentary filmmaking – including unpairing audio and video, arranging shots in disjointed montages, and voice-overs that consist mostly of non sequiturs – to 'speak nearby' her subjects (Balsom 2018). By purposefully suspending meaning and foregoing a position of authority, such an approach seeks 'to acknowledge the gap between [Minh-ha] and those who populate [her] film ... to leave the space of representation open so that, although [she's] very close to [her] subject, [she's] also committed to not speaking on their behalf' (Minh-ha, cited in Balsom 2018).

Minh-ha's concern for representation stems from her feminist and postcolonial ethics. For Minh-ha, identity is formed of 'infinite layers' (Minh-ha 1989: 94). The differences between and within these layers, between '*I* and *Not-I*, *us* and *them*, or *him* and *her*' might be understood as 'multiple presence' (Minh-ha 1989: 94). If words of difference between entities (i.e., between filmmaker/ethnographer and their subjects) serve to authenticate a discourse, narrative or ideology, rather than defer to this multiplicity and infinity, they are "noteworthy only as *decorations*" (Lorde 1979, cited in Minh-ha 1989: 101). Such decorations speak *for* and *about* a subject in finite and fixed terms in order to objectify (Minh-ha 1989: 101, 106). Ultimately, the only ideology such objectifications and their representations serve is the categorisation, discrimination and oppression of the 'other'. Instead, to speak *nearby* or *together* is to maintain a critical distance and allow another to fill the gap of representation, if they wish (Minh-ha, cited in Balsom 2018).

*Midlands* and *A Late Anthology* display a similar disjointedness to Minh-ha's ethnographically inclined films, a similar off-ness where fragments do not quite add up nor give us the whole story. Yet, the transformed sounds that populate both works – characterised by

uncanny glitches and audio artefacts – are not merely aesthetic decorations. Rather, in their displacement from the original sources, they open an interpretive space for both the composer and a listener to fill. Thus, Salem and Walshe invite us as an audience to reassess how we feel running water or Gregorian chant should sound; how nostalgia might be portrayed and problematised; or how Western art music could be canonised, for instance. Additionally, both Salem and Walshe demonstrate an affinity with Minh-ha's notions of identity and ethics. Salem, in his multimedia and collaborative traversal of a multilayered 'hyperspace' (Salem 2020a), and Walshe, speaking with her 'many, many voices' (Walshe 2018), both suggest a multiple presence similar to that described by Minh-ha. Salem here refers to the compositional process of navigating themes, objects and his own experiences and tastes but also suggests that both computer-assisted and collaborative methods of composition are effective ways of extending, augmenting and generating other hyperspaces (Salem 2020a). Such augmentation displaces personal taste and allows Salem to assume a critical distance from his subjects, in the instance of *Midlands*, the River Derwent, Derby and his own adolescent memories. Similarly, Walshe frames her voice as the 'staging area' for her experiences, memories and tastes, which might be drawn upon in 'infinitely different ways' (Walshe 2018). Walshe's interest in this 'polyphony, this confusion' (Walshe 2018), typified by the reconstitution of her voice using machine learning in *A Late Anthology*, allows her to maintain a critical distance from her voice and the music it is mapped onto as subjects. Rather than speaking *about* the world, Salem and Walshe, in their use of machine learning, in adopting another, unknowable logic and reconstituting the subjects that populate their work through the lens of the uncanny, speak *nearby*. In doing so, they acknowledge, similar to Minh-ha, the gap between representation and actuality, fact and fantasy, truth and meaning (Balsom 2018). The world, as described by each composer, is literally impossible and yet fittingly relays their experiences and interpretations.

## 5. ALGORITHMIC ART-ANTHROPOLOGY

Rather than using neural synthesis as a mere tool to navigate or even speak nearby culture, such practices of algorithmic art-anthropology are, first, collaborative and, second, recognise the cultural status and agency of algorithms (and thus neural networks) themselves.

In his proposition of ethnographic methods for defining algorithmic systems, technological anthropologist Nick Seaver suggests that algorithms (including the training algorithms used in neural networks) are 'not singular technical objects ... but are rather unstable objects ... composed of collective human

practices' (Seaver 2017: 5). The purpose, functionality and influence of any algorithmic system are not fixed entities, but infinitely variable depending on the changing habits of programmers, providers and users within a multidirectional exchange of cultural enactments. If we take this definition with regard to the artistic uses of machine learning algorithms outlined previously, both *Midlands* and *A Late Anthology* are assemblages of (at least) the artistic agencies of their respective composers (which, as earlier, are themselves accumulations of multiple presences), the algorithmic agency (the successive layers of refining and derived probability distribution sets) of their neural networks, and the creative-computational agencies of the networks' engineers – whether multinational companies or indie programmers such as Dadabots. Such an assemblage of human and non-human entities involves a shifting network of collaborative relationships and influence upon the generated audio and resulting artwork.

Furthermore, Seaver proposes that, owing to their enactment by collective practices, algorithms are 'not technical rocks in a cultural stream, but are rather just more water' (Seaver 2017: 5). In other words, algorithms *are* culture. This position differs from the conception of 'algorithmic culture' offered by technological historian Ted Striphas, whereby algorithms alter how culture is 'practiced, experienced and understood' (Striphas 2015: 396). The transformative and entropic force of algorithms outlined by Striphas does help to account for the more obscure agentic properties of both the WaveNet and SampleRNN algorithms in *Midlands* and *A Late Anthology*. However, Striphas's argument pertains specifically to those algorithms used to process big data for the purpose of online marketing and is less useful when analysing the small-scale collaborative acts undertaken by Salem and Walshe. If we take algorithms – like those used within these neural networks to learn datasets – *as* culture, as 'enacted by the practices used to engage with them' (Seaver 2017: 5), as actors within a multi-directional network, we account for the social aspect of these assemblages in which multiple agents exert varying influence on each other party. Thus, *Midlands* and *A Late Anthology*, as works of anthropology, should not be viewed as art using technology to describe culture, but rather as a back-and-forth, continuous flow of cultural exchanges between cultural agents. Such exchanges point towards a flattening of hierarchies between entities, forcing us to reformulate the autonomy of both composers, in terms of both the processes and the outcomes of such anthropological work.

Accordingly, we might situate such practices within Tim Ingold's conception of anthropology. For Ingold, anthropology is a process of 'learning to learn' that 'aims not so much to provide us with facts *about* the

world as to enable us to be taught *by it*' (Ingold 2013: 1–2). Ingold here describes an attitude (rather than a method, an openness to finding and learning that is not driven by outcomes or preconceived expectations but actively follows new experiences and instruction. The world becomes an opportunity for learning from and with its inhabitants. Thus, akin to Seaver's collective human practices, learning itself becomes 'an *outcome* of actions' (Seaver 2017: 4–5) and, crucially, interactions. Both Salem and Walshe adopt such an attitude in their decision to work with machine learning, whose unknown logic has the potential to show them new and unpredictable versions of the world it is trained upon, namely, the River Derwent and Salem's database of audio recordings, and Walshe's voice.

Subsequently, both composers engage in what Ingold refers to as the 'art of inquiry', allowing 'knowledge to grow from the crucible of [their] practical and observational engagements with the beings and things around [them]' (Dormer 1994 and Adamson 2007, cited in Ingold 2013: 6). Ingold here describes the approach of a craftsperson, one of learning through doing where thinking and knowing are in direct dialogue with the materials or entities with which they work. Salem and Walshe embark upon such an approach, working alongside programmers, seeding the algorithms and reviewing the numerous audio outputs generated. This back-and-forth exchange between cultural agents, as described earlier, is an opportunity for both composers to think through and know the world: for Salem to understand his relationship to Derby or his own field recording tastes in a new light; or for Walshe to reacquaint herself with her voice. Each composer will now know the world and perhaps (the multiple versions of) themselves a little differently, not simply through the lens of machine learning, but through this cultural exchange.

Finally, both artists make editorial decisions based on this learning. These decisions are multifarious, but might include what outputs to keep, how to employ them, or how to accompany them. For some, such as music critic Brendan Finan, this editorial role is the salient process of working with machine learning (Finan 2020). For Walshe, these decisions allow composers to be 'ethnomusicologists romping through the Wild West section of the Uncanny Valley' (Walshe 2018), to rewrite (musical) history and explore what could have been and what could come to be. Such decisions – opportunities to reframe the world anew – might be compared to Ingold's proposition of 'correspondence' whereby, having opened ourselves to what the world might teach us, we in turn 'respond' rather than describe (Ingold 2013: 7). Ingold's 'response' shares affinities with Minh-ha's endeavour to 'speak nearby' in her films, as described earlier. While the latter attempts to manifest the gaps or

miscommunications that arise in correspondence as a structural, and perhaps aesthetic, device, both allow and require interpretation; the meaning of one's learning is never fixed, never absolute. The editorial decisions made by Salem and Walshe in *Midlands* and *A Late Anthology* are made with a perception of the world that is opened anew by the neural networks they employ and by the network of cultural exchanges established. The decision to pair the machines' generated outputs with, say, video footage of cotton looms or midi files of renaissance polyphony, to speak nearby the worlds in which the composers, their work and the algorithms of the neural network inhabit, is to respond to the world and its teaching, and in turn invite further questioning from a listener. In extension to Ingold's conception of anthropology (2017: 22), both *Midlands* and *A Late Anthology* make human and *non-human* inquiries into the possibilities of human and *non-human* life and how we relate to it.

In *Midlands* and *A Late Anthology*, Salem and Walshe respectively employ varied formats, work with different neural networks and speak nearby very different subjects. Regardless, both works exhibit an attitude and approach to making art in relation to the world in the twenty-first century, one that is open, questioning and enlightening. Such work, or algorithmic art-anthropology, is literally 'composing (on) life in living it or making it' (Minh-ha 1990: 89).

## REFERENCES

- Anderson, I. and Rennie, T. 2016. Thoughts in the Field: 'Self-reflexive narrative' in field recording. *Organised Sound* 21(3): 222–32.
- Balsom, E. 2018. 'There is No Such Thing as Documentary': An Interview with Trinh T. Minh ha. *Frieze*. [www.frieze.com/article/there-no-such-thing-documentary-interview-trinh-t-minh-ha](http://www.frieze.com/article/there-no-such-thing-documentary-interview-trinh-t-minh-ha) (accessed 3 August 2021).
- Briot, J., Hadjeres, G. and Pachet, F. (eds.) 2020. *Deep Learning Techniques for Music Generation*. Cham, Switzerland: Springer Nature.
- Drever, J. L. 2002. Soundscape Composition: The Convergence of Ethnography and Acousmatic Music. *Organised Sound* 7(1): 21–7.
- Dyer, J. 1757. The Fleece: A Poem. In *Four Books*. London: R. and J. Dodsley.
- Fiebrink, R. and Sonami, L. 2020. Reflections on Eight Years of Instrument Creation with Machine Learning. *International Conference on New Interfaces for Musical Expression (NIME)*. Birmingham: Royal Birmingham Conservatoire, Birmingham City University.
- Finan, B. 2020. Against the Party Line of Western Music. *Journal of Music*. <https://journalofmusic.com/criticism/against-party-line-western-music> (accessed 16 July 2021).
- Goodfellow, I., Bengio, Y. and Courville, A. 2016. *Deep Learning*. Cambridge, MA: MIT Press.



- Haggett, G. K. 2021. Jennifer Walshe, *A Late Anthology of Early Music Vol. 1: Ancient to Renaissance*, Bandcamp (Record review). *Tempo* 75(295): 112–15.
- Ingold, T. 2013. *Making: Anthropology, Archaeology, Art and Architecture*. Abingdon: Routledge.
- Ingold, T. 2017. Anthropology Contra Ethnography. *Hau: Journal of Ethnographic Theory* 7(1): 21–6.
- Iscen, O. E. 2014. In-Between Soundscapes of Vancouver: The Newcomer's Acoustic Experience of a City with a Sensory Repertoire of Another Place. *Organised Sound* 19(2): 125–35.
- Lewis, G. 2020. Rainbow Family. *Bandcamp, Carrier Records*. <https://carrierrecords.com/album/rainbow-family> (accessed 4 January 2022).
- Lorde, A. 1982. *Zami: A New Spelling of My Name*. Watertown, MA: Persephone Press.
- Martin, B. 2017. Soundscape Composition: Enhancing Our Understanding of Changing Landscapes. *Organised Sound* 23(1): 20–8.
- Melen, C. 2020. A Short History of Neural Synthesis. *PRiSM Blog, RNCM*. [www.rncm.ac.uk/research/research-centres-rncm/prism/prism-blog/a-short-history-of-neural-synthesis/](http://www.rncm.ac.uk/research/research-centres-rncm/prism/prism-blog/a-short-history-of-neural-synthesis/) (accessed 31 August 2021).
- Minh-ha, T. T. 1989. *Woman, Native, Other: Writing Postcoloniality and Feminism*. Bloomington: Indiana University Press.
- Minh-ha, T. T. 1990. Documentary Is/Not a Name. *October*, 52: 76–98.
- Oord, A. V. D. and Dieleman, S. 2016. WaveNet: A Generative Model for Raw Audio. *DeepMind*. <https://deepmind.com/blog/article/wavenet-generative-model-raw-audio> (accessed 31 August 2021).
- Oord, A. V. D., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A., Kavukcuoglu, K. 2016. WaveNet: A Generative Model for Raw Audio. *Cornell University*. <https://arxiv.org/pdf/1609.03499.pdf> (accessed 5 January 2022).
- Posic, A. 2020. Jennifer Walshe: A Late Anthology of Early Music Vol. 1: Ancient To Renaissance. *The Quietus*. <https://thequietus.com/articles/28117-jennifer-walshe-a-late-anthology-of-early-music-vol-1-ancient-to-renaissance-review> (accessed 16 July 2021).
- Rennie, T. 2014. Socio-Sonic: An Ethnographic Methodology for Electroacoustic composition. *Organised Sound* 19(2): 117–24.
- Salem, S. 2019. *Midlands*. Musical score.
- Salem, S. 2020a. A Psychogeography of Latent Space. *PRiSM Blog, RNCM*. [www.rncm.ac.uk/research/research-centres-rncm/prism/prism-blog/a-psychogeography-of-latent-space/](http://www.rncm.ac.uk/research/research-centres-rncm/prism/prism-blog/a-psychogeography-of-latent-space/) (accessed 16 July 2021).
- Salem, S. 2020b. Model – 99100 River6 – Long. *Soundcloud, RNCM*. <https://soundcloud.com/user-922563269/4-model-99100-river6-long> (accessed 5 October 2021).
- Schubert, A. 2021. CRAWLERS. *Alexander Schubert*. <http://www.alexanderschubert.net/works/Crawlers.php> (accessed 4 January 2022).
- Seaver, N. 2017. Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems. *Big Data & Society* 4(2): 1–12.
- Steels, L. 2020. Foreword: From Audio Signals to Musical Meaning. In E. R. Miranda (ed.) *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity*. Cham, Switzerland: Springer Nature, v–xviii.
- Striphas, T. 2015. Algorithmic culture. *European Journal of Cultural Studies* 18(4–5): 395–412.
- Walshe, J. 2018. Ghosts of the Hidden Layer. *Milker Corporation*. <http://milker.org/ghosts-of-the-hidden-layer> (accessed 19 July 2021).
- Walshe, J. 2020b. Jennifer Walshe: A Late Anthology Of Early Music, Vol. 1: Ancient To Renaissance. *Milker Corporation*. <http://milker.org/a-late-anthology> (accessed 16 July 2021).
- Walshe, J. 2020c. A Late Anthology of Early Music Vol. 1: Ancient to Renaissance. *Bandcamp*. <https://jenniferwalshe.bandcamp.com/album/a-late-anthology-of-early-music-vol-1-ancient-to-renaissance> (accessed 16 July 2021).

## FILMOGRAPHY

- Minh-ha, T. T., director. 1982. *Reassemblage*. Women Make Movies, Inc.
- Minh-ha, T. T., director. 1985. *Naked Spaces: Living is Round*. Women Make Movies, Inc.
- Minh-ha, T. T., director. 1989. *Surname Viet Given Name Nam*. Women Make Movies, Inc.

## VIDEOGRAPHY

- Carr, C. J. and Zukowski, Z. n.d. *Relentless Doppelgänger*. *YouTube*. [www.youtube.com/watch?v=MwtVkPKx3RA&ab\\_channel=DADABOTS](https://www.youtube.com/watch?v=MwtVkPKx3RA&ab_channel=DADABOTS) (accessed 5 October 2021).
- Salem, S. 2020c. *Midlands by Sam Salem* (excerpt). *YouTube, PriSM*. [www.youtube.com/watch?v=RNoz23t9PA0&ab\\_channel=PRiSM](https://www.youtube.com/watch?v=RNoz23t9PA0&ab_channel=PRiSM) (accessed 5 October 2021).
- Salem, S. 2020d. *RNCM Research Forum – Dr Sam Salem*. *YouTube*. [www.youtube.com/watch?v=d-TzJCeQh8&ab\\_channel=rncmlive](https://www.youtube.com/watch?v=d-TzJCeQh8&ab_channel=rncmlive) (accessed 16 July 2021).
- Walshe, J. 2020a. Jennifer Walshe 'A Late Anthology of Early Music' (excerpts). *SoundCloud, eavesdropping.london*. [https://soundcloud.com/eavesdropping\\_london/jennifer-2?in=eavesdropping\\_london/sets/eavesdropping-archive-jennifer-walshe](https://soundcloud.com/eavesdropping_london/jennifer-2?in=eavesdropping_london/sets/eavesdropping-archive-jennifer-walshe) (accessed 19 July 2021).