The Roughness of Neural Networks. Jimi Hendrix, Holly Herndon, GPT-3, Timbre Transfer and the Promising Failure Aesthetics of Musical AIs

Jan Torge Claussen

Introduction

Writing a contribution after a talk should not be that difficult, because the core ideas have already been presented and sharpened semantically or artistically in previous works.[1] However, in the meantime, different theoretical and sonic inputs, for instance, the contributions of all the other participants in the symposium, have inspired and motivated me to access this new *data*. The processing of this data by a humanoid author, one might assume, is not totally different from that by an artificial intelligence, apart from the very high processing speed of the machine. But in contrast the perspective of the human author is shaped by highly individual experiences, evaluations and emotions. Moreover, despite the machine's pursuit of optimisation and perfection, I am more fascinated by the moments when it fails regarding to human expectations, and acts unexpectedly, in contrast to my own creations, which I always hope will turn out well.

During my research for this article, I began a conversation with ChatGPT (Generative Pre-trained Transformer), the chatbot from the Open AI research lab. Some answers have been thought-provoking and are documented and reflected at designated points in this article. Our conversation, that is accessible online at full length, starts with me prompting:

> *JTC*: Explain Sound in simple terms!
> *GPT-3*: Sound is a type of energy that travels through the air (or any other medium, such as water) as a wave. It is produced when an object vibrates, causing the air around it to vibrate as well.
> The vibrations cause pressure waves in the air, which can be detected by our ears. […] Sound is an important part of our everyday lives. It allows

1 Parts of this Article have been published in German language: "Risse in der Schallmauer. Musikalische KI und phonographische Stimme" (Claussen 2024).

> us to communicate with each other, to enjoy music, and to be aware of our surroundings.[2]

Even though the AI only computes the most statistically likely sequences of words, it is worth noting that by using the words "we", "our", or "each other", the AI seems to imagine itself as part of humanity, from which it is still very different, despite its fascinating successes: artificial intelligences are able to compose music (MuseGAN, Aiva), write texts (GPT-3), translate (DeepL) and produce voices (Holly +, Tone Transfer). Are these AIs, which all play a role in the following, creative or not, and what does this mean for art and artists? The roles of authors of art seem to be redefining themselves once again in the face of new media. Roland Barthes (1967) declared the author dead more than half a century ago to make clear, that the author is not the sole ruler of his work; rather, it is the readers, who constantly reinterpret the story, place it in new contexts, or connect it to their own experiences. But with the rise of more and more accessible media technologies different voices, sounds, images or texts could also be combined physically and not only in the mind of readers or listeners. However, in the digital realm materials are translated in numbers and codes, which lead to even further possibilities for creative manipulations with a huge variety of media interfaces. In the postdigital age the borders between authors and readers, of producers and users, artists and audiences combining different voices, sounds, images or texts are widely open. Even if living in the postdigital age means, that men are no longer fascinated by digitality per se and have recognised that all areas of life are under its influence (Bishop et al. 2016; Cascone 2000; Cramer 2014), AI as such exerts great fascination.

At the same time, the progress of AI seems to be a lot faster, than one imagines. In 2021, I tried to play the trumpet through my voice using AI more specifically so-called Timbre or Tone Transfer / DDSP (Differentiable Digital Signal Processing) by Googles open source research project *Magenta* (Carney et al. 2021). The sound was indeed amazing or as some musicians put it, who have been involved in the project:

> I think the thing that stands out the most to me is all of the nuances of the sound that make it so much more real than working with a sample. […] It gives you a lot of surprises.
> I am excited, by not knowing what could come out of it. (Google 2020)

2 ChatGPT is based on the GPT-3 language model of Open AI and can be accessed via https://chat.openai.com/chat. The entire conversation between me (JTC) and the ChatBot (GPT-3) from 20 December 2022 is documented at the following link: https://gegenwaerts.com/conversation-between-me-and-gpt-3/.

At that time timbre transfer still required the diversions of a recording of my own voice, which could then subsequently and monophonically sound as a trumpet, for example (Audio 1a und 1b[3]). Today, this process is instantaneous, as Matt Dryhurst, Holly Herndon's partner, impressively demonstrated when he sang with her voice at the Sonar Festival 2022[4].

In the following, I will present facts about GANs, a machine learning model that has been used for timbre transfer a lot, and what they might have in common with human learning methods. After that, I will elaborate on some works that appear cleaner and more abstract in contrast to the ones operating in the concrete realm of sound and I will go into more detail about further experiments and sound studies. These will be connected to media theoretical perspectives and artists featuring digital failure aesthetics before I present some compact experiments of my own involving machine learning of Hendrix's guitar and Herndon's voice to end with a discussion of my findings regarding media education and collaboration.

Creativity in Generative Neural Networks

My focus initially has fallen on so-called Generative Adversarial Networks (GAN), as these have been considered particularly innovative in computer music research and actually have a promising architecture concerning a form of artificial creativity. This is because within this specific method of machine learning, two neural networks are developed that work against each other. The so-called generator creates new data from randomly generated data, while the other neural network, the discriminator, checks whether the newly generated data could also come from the original data set. In other words, it checks whether the images are imitations or could be originals. The calculations are complete when enough imitations are thought to be originals.

This approach is not so different from human learning methods. A big part of learning a musical instrument is to copy the playing styles of existing artists and examine how good the self-made imitation is. A rock guitarist for example might learn to play a preferred guitar riff of Jimi Hendrix playing it over and over again, each time going through a little sometimes more or less unconscious self-assessment, asking her / himself, if this sounds like the original and optimizing his playing subsequently. While the guitarist becomes more and more satisfied with his adaption,

3 https://gegenwaerts.com/wp-content/uploads/2023/01/Audio_1a_Tone-Transfer-Web-trumpet.mp3 https://gegenwaerts.com/wp-content/uploads/2023/01/Audio_1b_Tone-Transfer-Web-original_Recording.mp3
All Audio Examples are available at this website: https://gegenwaerts.com/aix/ (21 January, 2024).

4 "AI and Music – Holly Herndon presents Holly+ feat. Maria Arnal, Tarta Relena and Matthew Dryhurst" (30.03.2022) https://youtu.be/Wk6T2WmhuJw (30 March, 2022).

he might integrate what he learned in his own playing style and sometimes create new riffs that sound partly like Hendrix, but they aren't his.[5] In the logic of a GAN, the musical sequences could be considered to be originals and with that, the machine learning process has come to an end. However, in machine learning, GANs have first been explored in the visual domain.

The machine learning model, developed by computer scientist Ian Goodfellow (2014), has been applied to images with astonishing results; it can, for example, produce images of people that do not even exist and still look absolutely convincing[6]. By learning the underlying distribution of data masses, GANs can generate new realistic sample data that have the same statistical properties as the training data. However, the resulting images are not simple replications or the result of averaging. As Rashid (2020) summarizes: "GANs learn to create images at a level far above simply replicating or averaging training data". The algorithms have effectively learned what is possible with a particular image in this case of a human face, but without actually knowing that it is a face. Unlike rule-based systems, no rules were given either, so the algorithm would determine where the nose or eyes were, for example, based on calculations.

As Cádiz et al. (2021) point out, this method of machine learning could even be called creative: "The whole idea of this approach is to make the generative model G (Generator) so good that eventually D (Discriminator) might be fooled by a false input. If this happens it means that G is generating false data that is indistinguishable from real data, also a possible indication of creativity". (Cádiz et al. 2021: 3) In their article, the researchers examine creativity in terms of two aspects within a case study based on Margaret Boden's (2009) definition. According to the cognitive scientist, the creation of something new, unexpected and surprising is essential for creativity. However, this alone is only an indication of possible creativity; what is also important is a form of evaluation, an appreciation, something meaningful. For this, the human being would at least be a central instance. The computer is capable of generating something that, according to its calculations, corresponds to general human listening habits, but it cannot value its compositions.

The strength of artists like Jimi Hendrix lay in defying listening habits and creating something new or surprising in collaboration with his instrument (Millard 2004; Trampert 1998). GPT-3 also seems to confirm this:

> Jimi Hendrix was widely regarded as one of the most creative and innovative guitarists in the history of rock music. He was known for

5 Listen for example to the guitar played in the verse of the song "Under the bridge" by *Red Hot Chili Peppers*' (1992) former guitarist John Frusciante, who has made a lot of guitar-playing styles from Hendrix his own.

6 https://thispersondoesnotexist.com/

> his unique and groundbreaking playing style, which incorporated a wide range of sounds and effects, as well as for his ability to improvise and create new musical ideas on the spot. […] He was known for writing complex and sophisticated musical pieces that incorporated a wide range of musical elements, including intricate guitar work, dynamic changes, and unexpected musical twists.[…]

So it seems as if the chatbot is able to value Hendrix's creativity, but only in terms of all the things other people have written about him on the web. Obviously, the AI is not able to listen and to develop its own opinion about Hendrix's creativity.

Musical AI doesn't listen

Unlike static images and texts, music has other properties that make its computations more complex: Music is time-based; it has multiple tracks or instruments that change dynamically but are still related to each other; and music is grouped both horizontally and vertically by chords, polyphonic melodies, arpeggios, etc. The MuseGAN (Multi-track Sequential Generative Adversarial Networks) project attempts to account for these interdependencies (Dong et al. 2017). To this end, the neural network was applied to a data set of over 100,000 bars of rock music to generate new piano rolls or images of the MIDI notations of five tracks corresponding to the instruments guitar, bass, drums, piano and strings. The aim was to generate short (4-bar) coherent pieces of music "right from scratch […]without human inputs", i.e. without human pre-selection, while the addressed coherence of the polyphonic music was guaranteed on the basis of the following parameters: "1) harmonic and rhythmic structure, 2) multitrack interdependency, and 3) temporal structure (Dong et al. 2017). For this purpose, the scientists put together different architectures of the GANs, which are also clearly noticeable in the sound of the results. Listening to the so-called composer model, jammer model and a hybrid model on the website[7], it becomes clear, apart from the differences in general, that the presented models make human beings believe that they learn something about music. This is confirmed by a survey in which various listeners positively evaluate the music pieces resulting from the process according to selected criteria (Dong et al. 2017). The authors' study recorded whether the varying four bars of music had a pleasant harmony, a consistent rhythm, a clear musical structure and references or "coherence". Significantly, however, the listeners were not asked to focus on interesting, unexpected or creative outcomes, which I think would be much more

7 Website with sound examples of MuseGAN: https://salu133445.github.io/musegan/results (21 January, 2024).

thought-provoking. After all, it can only be considered creative to a very limited extent if the machine only produces imitations (see above), i.e. credible variations of something that has already been there.

Many AI projects use methods that transform the music on a higher level of representation and not directly in the realm of the physical audio signal. So, as in the previous example, MIDI sequences or chords or text-based forms of representation. In the project "The lost tapes of the 27th club", song imitations of Kurt Cobain and Jimi Hendrix[8] among others were even produced, though it has been very hard to translate the unorthodox guitar playing into MIDI. However, the MIDI data used on the basis of about thirty songs each was then recorded by professional cover musicians (Grow 2021). Media music scholar Rolf Grossmann (2022) sees these methods as a step backwards, as the music, similar to what was practised in classical music history, would be thrown back on the reading of its score, while cultural and historical context, sound and performance would be neglected. What remains is a "pattern generator" which provides symbol structures ready for a machine or human to perform.

Despite the limitations mentioned above, piano rolls may be well suited to capture essential parameters of instrumental playing (pitch, duration, timing and velocity) in relation to the piano. However, other playing techniques, such as vibrato on the guitar, sliding over the fingerboard, plucking or bending are neglected in favour of better predictability:

> The physical process through which sound is produced is abstracted away. This dramatically reduces the amount of information the models are required to produce, making the modelling problem more tractable and allowing for lower-capacity models to be used effectively. (Dieleman 2019)

This is partly comparable to the gamification strategies used in digital applications for learning musical instruments, where the tracking of this type of playing technique on the guitar is also not reliable (Claussen 2019, 2021). But just as it is incomparably more complex to capture and translate the sounds of a conventional electric guitar in its entire sound spectrum than the pressed keys of the Guitar Hero controller, it is also much more complicated and involves more computational wiring to train neural networks phonographically directly based on sounds. In return, however, it frees one from being tied to a few symbolic parameters and encompasses timbre, space, amplification and mixing of the entire recording, as in the GAN Synth

8 Jimi Hendrix' AI Song "You're Going To Kill Me": https://youtu.be/6Ohf97p7u1w (21 January, 2024). The website of the project is dedicated to drawing attention to musicians struggling with mental health. https://losttapesofthe27club.com/ (21 January, 2024).

project, for example, where waveforms or spectrograms are used rather than piano rolls (Engel et al. 2019). In all cases, however, images of sounds are the basis of the calculations, even if representations of the entire audio signal are more comprehensive than midi scores, which hide sound and performance. It remains to be said: The machine does not listen, it reads preferably clean symbolic representations.

The Rasping of Neurons

What is striking about various academic research projects (Cádiz et al. 2021; Dhariwal et al. 2020; Dong et al. 2017; Engel et al. 2019) is that they largely focus on optimizing and perfecting sounds, timbres or musical sequences that are more familiar to us. This is all the more true for commercial applications such as AIVA[9] and Co., which produce soundtracks that are by no means aimed at developing a unique or creative AI voice, but rather at reproducing proven voices in an application-, a cost- and a licence-optimised way for video games, films or advertising clips. In the search for a peculiar, previously unknown voice, on the other hand, the focus is precisely on those imperfect places where the limits of the medium are exhausted, so that errors, bugs or glitches come to light. This is also a phenomenon that was particularly expressed in the playing of Jimi Hendrix (1967) on the electric guitar, as he made musical use of feedback, noises from amplifiers and electronic components of the instrument. So listeners as well as the transcriptionists who convert his songs into machine-readable MIDI notes could not always be sure whether it was a disturbance or part of his music.

The history of media formats is determined by two opposing tendencies. On the one hand, the focus is on optimization processes to create an almost perfect sound. So the most important characteristic of the digital CD was that it is not even perceptible as a sound carrier. On the other hand, artists in particular were interested in artefacts and errors of new technologies. In relation to digital media technology or the compact disc, for example, they created a post-digital error aesthetic at the end of the last millennium (Cascone 2000; Claussen 2020; Großmann 2003). Exemplary of this are the compilations entitled *Clicks & Cuts* (2000–2003) by the German label 'Mille Plateaux' as well as various works by artists such as Markus Popp aka Oval or the media artist Yasunao Tone, who, among other things, changed the CD surface by pasting and scratching it to create new and altered sounds that no longer correspond to the original recording (Stuart 2003). In this way, they elevated the jumps, skip noises and loops, which are otherwise perceived as a disturbance, to the aesthetic material of their music.

9 https://www.aiva.ai/ (21 January, 2024).

As philosopher Sybille Krämer (1998) describes referring to the theory of McLuhan (1964), media represent the blind spot of media use. Humans only recognize them by their malfunctions, the stuttering compact disc, the feedbacking guitar, the coarse pixels in the video stream when the internet connection breaks and precisely also by the noisy artefacts generated by the algorithm of an artificial intelligence when generating a sound. For example, such striking artefacts are generated by the artist Holly Herndon with her AI "Spawn". This instrument, or musicking thing (Ahlers et al. 2022; Ismaiel-Wendt 2016), was trained for two years based on data from her own voice, her compositions, her cooking and her "just living alongside it" to dynamically generate a unique repertoire of sounds. Traceable and audible, Spawn or Holly Herndon's hybrid voice on the album *Proto* (2019) is particularly evident in the first track "Birth". Herndon describes that based on her artistic approach, the neural network becomes audible, which more adequately represent the current state of AI research and its artistic application than would otherwise often be the case:

> A lot of the press releases present AI in this very glossy way that erases all of the human labor that went into training whatever the AI is doing. It also creates this illusion of it being more developed than it is. We wanted to be more honest about it. When you're dealing with automated composing, you get a MIDI score at the end, and when you push that through a digital instrument, it sounds really clean. We're dealing with sound as material, and by using audio material, you can really hear the roughness of the neural network trying to figure out what to do next. That was something that we chose, because we wanted to make that clear: [that] the current state of the technology is still developing. That's what's exciting about it: we still can have a say in which direction it goes. (Friedlander, 2019)

For Holly Herndon and other scientists, it is clear that AI alone will never be creative, but that creativity always comes from collaborations between humans and AI. She is already offering a tool for that.

Further Experiments from Holly+ to Jimi Hendrix and Tone Transfer

With *Holly+*, the artist has made a tool available that enables users to sing with her voice. As is the case with other AI tools, new sounds can be created directly in the prompt. This so-called "promtism" (Hayward 2022), as artists have already named this phenomenon, is easily accessible but comes with other challenges: "[S]ome of my AI images took 30+ tries to match closely to my end goal. I had a clear picture of what I wanted in my head, but it's a matter of articulating that in a way that's

friendly to the machine". (Hayward 2022) In the case of Holly+, all that needs to be done is to upload an audio file to be downloaded in the sound colour. The results of this process reveal the roughness described by Herndon (Audio 2)[10]. Even if the process is not yet live for every user, it is already possible (Herndon 2022). If it actually becomes more common to sing and produce music through other voices, this naturally raises questions about identity, authorship, exploitation or possible forms of cultural appropriation – even if these questions are less the focus of my contribution. It will be at least as challenging a task as in the cultural practice of sampling to decide how to deal with the diverse identities, artistic freedoms and ethical and legal issues in each individual case. The musical spectrum has expanded.

After experimenting with my voice or guitar as input for Holly + or Magenta's Tone Transfer, I wondered if, instead of mimicking guitar riffs to learn how Hendrix played, it would be possible to preserve his timbre. Of course, Chat GPT could provide a convincing answer to the question of what the guitar of Jimi Hendrix sounds like,[11] and the specialized literature (Clague 2014; Trampert 1998; Waksman 2010: 166–206) as well as the original recordings provide reliable, but also more varied and complex answers, which are, however, beyond the scope of this experimental study. Therefore, I first decided to use similar equipment that Hendrix used in most of his recordings. Then I created training data for machine learning by recording my own guitar playing on a Fender Stratocaster using a Marshall amp, and various effects devices such as wah-wah and fuzz distortion. And in doing so, I tried to imitate some significant playing techniques of Hendrix and replay song sequences. The training data generated could then be used within the environment of a so-called "Colab notebook" provided for the Tone Transfer project, which contains the Python code for the machine learning model including step-by-step instructions (Carney et al. 2021). After these technical hurdles, about 12 minutes of audio were used for training a neural network with 30 000 steps for about 3 hours. The resulting files could finally be played in the music software environment of Ableton Live with the help of the Tone Transfer VST plug-in. At this point, it is necessary to experiment again and this time with the input, which produces different fascinating results in interaction with the timbre (Audio 3–7)[12]. Even after several attempts with different inputs of one's own voice (Audio 3), a piano (Audio 4), a field recording of birds (Audio 5) or a guitar (Audio 6), the result does not seem to correspond particularly clearly to a Hendrix-like electric guitar, but unique references to the recorded training data (Audio 8) remain. The results clearly differ from the existing presets of other timbres (Audio 7) and produce a lot of aesthetic failures or "rasping of the

10 https://gegenwaerts.com/wp-content/uploads/2023/01/holly_plus-1_english-sentence.mp3 (21 January, 2024).

11 https://gegenwaerts.com/conversation-between-me-and-gpt-3/ (21 January, 2024).

12 Website with all audio examples: https://gegenwaerts.com/aix/ (21 January, 2024).

neurons". They also have an emotional effect that comes with having designed one's own timbre and making it usable in other contexts, similar to how musicians emphasise the value of specially recorded or discovered sounds and music sequences when sampling or djing. Otherwise, "spawning", as Herndon calls the methods of timbre transfer, clearly differs from digital sampling due to the specific media-technical and cultural collaboration: "So, for example, with sampling, usually you copy and remix a recording by someone else to create something new. But with spawning, you can perform as someone else based on trained information about them". (Herndon 2022 min 3:21)

Conclusion: Human Learning through Machine Learning

In the age of timbre transfer and after the death of the author (Barthes 1967), two things can be said. Thanks to spawning as a form of machine learning, it will be possible in the future to speak with the voices of dead authors without sampling them directly and thus reproducing something they have already said. And as this is at least unrecognisable to the amateur, it bears numerous dangers and uncertainties. But it also bears many creative potentials that can be found especially at the edges of various AI music productions. As the previous experiment showed, the result is not the trained Hendrix sound, but an in-between that is tempting and more than the reproduced imitation of the original. Such approaches are appealing not least because we never listen exclusively with our ears but also with our other senses, our memories and depending on the most diverse contexts (Schulze 2018; Sterne 2003). In this way, the reference to Hendrix becomes culturally significant without being recognized as a sound event.

The neural networks presented are characterized by their type of so-called unsupervised deep learning. Once the process is set in motion, it cannot be continuously monitored and interrupted by a human. In this sense, the machine acts autonomously at times. Nevertheless, the human determines the parameters of the process, as I have shown with some examples, and ultimately also orders the results with a view to an aesthetically meaningful outcome. Following Marshall McLuhan (1964), the question is: What is the message of AI? In other words: What influence does it have on music and its protagonists, on pieces, compositions and performances? For researching the messages of AI, and for recognizing the limitations during the use of the media that determine our situation, application-oriented approaches are crucial, such as those provided by the Magenta project from the environment of Google's AI Tensor Flow or artists like Holly Herndon.

Collaborating with the tools thus means relinquishing part of the control, as was already practised a decade ago in particular in the aleatoric and open works of John Cage, Pierre Boulez or Karl-Heinz Stockhausen in relation to the classical score or

the concert hall: "Those involved with the composition of experimental music find ways and means to remove themselves from the activities of the sounds they make. Some employ chance operations" (Cage 1961: 10). In principle, however, a certain degree of loss of control is present in every musicking thing. Game theory, for example, makes clear what may also apply to musicians and their instruments, namely that players not only play a game but are also played by the game (Claussen 2021; Huizinga 1998). And Pierre Boulez states for the aleatoric piece of music, that it is like a labyrinth offering a number of possible paths, that have been precisely designed by the artist, while chance plays the role of setting the course (Boulez et al. 1964).

In the face of machine learning, however, this collaboration seems to take on a new dimension. The playing field of predictable chance, the labyrinth with its diverse possible branches, becomes much faster, more intricate and more mobile and can persist in statistically similar things – or inspire something uniquely new. All the more important is to dedicate oneself to this field, of course without falling in love with the "gadgets" and losing oneself in it in the sense of McLuhan (1964: 22).

A good strategy is to approach the boundaries, those of predictability but also those of one's own listening habits. Otherwise, likely, users will only ever produce what is already present in the respective musicking thing (Ahlers et al. 2022; Ismaiel-Wendt 2016) and thereby correspond to widespread listening habits. Musical tools, games and instruments are part of social contexts and imply specific ways of dealing with them. In the best case, this empowerment of things leads to human-machine hybrids rich in variation; in the worst case, to the one-dimensional repetition of what the respective musicking thing demands at the most obvious level. Users, artists and producers are always in the process of negotiating the boundaries of this influence. Media education takes place within this process. For in musicking things, there is the chance to learn about music production cultures, heterochronous (Pelleter 2020: 25) music history or the power of well-tempered mood and timbre, in short, to practise what Kodwo Eshun (1999: 22) calls "beat education". Concerning machinic learning, this raises the obvious question of intelligence. What about the musical ghost in the machine? Eduardo Miranda (2021: 20) emphasizes that AI is a great tool to study musical intelligence. Where intelligence includes human abilities such as creativity, subjectivity and emotion, interaction and embodiment. At the same time, these properties inherent in music make it an interesting field for research. Both in AI and concerning human educational processes.

However, as machine learning could be used to create a future of endless repetitions of things that have been heard before, a future where no one can ever be sure if someone is embodying the original or only speaking in the voice of the original, it also offers the chance to value the moment, the live experience full of traces, raw materials, failures, experiments and collaborative performances.

Bibliography

Ahlers, Michael / Benjamin Jörissen / Martin Donner / Carsten Wernicke (2022): *MusikmachDinge im Kontext: Forschungszugänge zur Soziomaterialität von Musiktechnologie*. Hildesheim, New York: Georg Olms Verlag.

Barthes, Roland (1967): "Der Tod des Autors". *Texte zur Theorie der Autorschaft*: 185–197.

Bishop, Ryan / Kristoffer Gansing / Jussi Parikka / Elvia Wilk (2016): *Across and beyond: A transmediale reader on post-digital practices, concepts and institutions.* Berlin: Sternberg Press and Transmediale e.V. https://eprints.soton.ac.uk/410908/ (31 May 2023).

Boden, Margaret A. (2009): "Computer Models of Creativity". *AI Magazine* 30(3): Art. 3. https://doi.org/10.1609/aimag.v30i3.2254

Boulez, Pierre / David Noakes / Paul Jacobs (1964): "Alea". *Perspectives of New Music* 3(1), 42. https://doi.org/10.2307/832236

Cádiz, Rodrigo F. / Agustín Macaya / Manuel Cartagena / Denis Parra (2021): Creativity in Generative Musical Networks: "Evidence From Two Case Studies". *Frontiers in Robotics and AI* 8, 680586. https://doi.org/10.3389/frobt.2021.680586

Cage, John (1961): *Silence: Lectures and writings.* Middletown, Conn.: Wesleyan University Press. http://archive.org/details/silencelectureswi961cage (31 May 2023).

Carney, Michelle / Chong Li / Edwin Toh / Nida Zada / Ping Yu / Jesse Engel (2021): "Tone Transfer: In-Browser Interactive Neural Audio Synthesis". *Joint Proceedings of the ACM IUI 2021 Workshops,* April 13–17, 2021, College Station USA.

Cascone, Kim (2000): "The Aesthetics of Failure: 'Post-Digital' Tendencies in Contemporary Computer Music". *Computer Music Journal* 24(4): 12–18. https://doi.org/10.1162/014892600559489

Clague, Mark (2014): "'This Is America': Jimi Hendrix's Star Spangled Banner Journey as Psychedelic Citizenship". *Journal of the Society for American Music* 8(04): 435–478. https://doi.org/10.1017/S1752196314000364

Claussen, Jan T. (2019): "Gaming Musical Instruments: Music has to be Hard Work!" *Digital Culture & Society* 5(2): 121–130. https://doi.org/10.14361/dcs-2019-0208

Claussen, Jan T. (2020): "Audio-Hacks: Apparative Klangästhetik medientechnischer Störungen". S. Diekmann / V. Wortmann (eds.) *Die Attraktion des Apparativen.* Leiden, Niederlande: Brill Fink, 203–218. https://doi.org/10.30965/9783846765098_015

Claussen, Jan T. (2021): *Musik als Videospiel. Guitar Games in der digitalen Musikvermittlung*. Hildesheim, New York: Georg Olms Verlag. https://dx.doi.org/10.18442/167

Claussen, Jan T. (2024): "Risse in der Schallmauer – Musikalische KI und phonographische Stimme". S. Hegenbart / M. Kersten / M. Meuer (eds.) *Grenzen der Künste im Zeitalter der Digitalisierung.* Berlin: De Gruyter.

Cramer, Florian (2014): "What is 'Post-digital'?". *A peer-reviewed journal about_Postdigital Research* 3(1). http://www.aprja.net/what-is-post-digital/ (31 May, 2023).

Dhariwal, Prafulla / Heewoo Jun / Christine Payne / Jong W. Kim / Alec Radford / Ilya Sutskever (2020). "Jukebox: A Generative Model for Music". *ArXiv. 2005.00341 [Cs, Eess, Stat]*. http://arxiv.org/abs/2005.00341 (31 May, 2023).

Dieleman, Sander (2019): "Generating music in the waveform domain". Sander Dieleman. https://benanne.github.io/2020/03/24/audio-generation.html (31 May, 2023).

Dong, Hao-Wen / Wen-Yi Hsiao / Li-Chia Yang / Yi-Hsuan Yang (2017): "MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment". *ArXiv. 1709.06298 [cs, eess, stat]*. http://arxiv.org/abs/1709.06298 (31 May, 2023).

Engel, Jesse / Kumar K. Agrawal / Shuo Chen / Ishaan Gulrajani / Chris Donahue / Adam Roberts (2019): "GANSynth: Adversarial neural audio synthesis". *ICLR* 2019

Eshun, Kodwo (1999): *More Brilliant Than the Sun: Adventures in Sonic Fiction*. London: Quartet Books.

Friedlander, Emilie (2019): "How Holly Herndon and her AI baby spawned a new kind of folk music". *The FADER*. https://www.thefader.com/2019/05/21/holly-herndon-proto-ai-spawn-interview (31 May, 2023).

Goodfellow, Ian J. / Jean Pouget-Abadie / Mehdi Mirza / Bing Xu / David Warde-Farley / Sherjil Ozair / Aaron Courville / Yoshua Bengio (2014): "Generative Adversarial Networks". *ArXiv. 1406.2661 [Cs, Stat]*. http://arxiv.org/abs/1406.2661 (31 May, 2023).

Google. (2020): "Turning everyday sounds into music with Tone Transfer". https://www.youtube.com/watch?v=bXBliLjImio

Großmann, Rolf (2003): "Spiegelbild, Spiegel, leerer Spiegel. Zur Mediensituation der Clicks & Cuts". Marcus S. Kleiner / Achim Szepanski (eds.) *Soundcultures: Über elektronische und digitale Musik* Vol. 1. Frankfurt/Main: Suhrkamp, 52–68.

Großmann, Rolf (2022): "Hören, was die Maschine hört". *Acoustic Intelligence.* Düsseldorf: University Press, 83–94. https://doi.org/10.1515/9783110730791-005

Grow, Kory (2021): "In Computero: Hear How AI Software Wrote a 'New' Nirvana Song". *Rolling Stone*. https://www.rollingstone.com/music/music-features/nirvana-kurt-cobain-ai-song-1146444/ (31 May, 2023).

Hayward, Jeff (2022): "The Growing Art Movement of 'Promptism'". *Counter Arts*. https://medium.com/counterarts/the-growing-art-movement-of-promptism-9ec956d82a61 (31 May, 2023).

Herndon, Holly (2022): "Holly Herndon: What if you could sing in your favorite musician's voice?". *TED Talk*. https://www.ted.com/talks/holly_herndon_what_if_you_could_sing_in_your_favorite_musician_s_voice (31 May, 2023).

Huizinga, Johan (1998). *Homo ludens a study of the play-element in culture*. London: Routledge and Kegan Paul.

Ismaiel-Wendt, Johannes (2016): *post_PRESETS: Kultur, Wissen und populäre MusikmachDinge*. Hildesheim, New York: Georg Olms Verlag.

Krämer, Sybille (1998): "Das Medium als Spur und als Apparat". *Medien – Computer – Realität: Wirklichkeitsvorstellungen und Neue Medien*. Frankfurt/Main: Suhrkamp Verlag.

McLuhan, Marshall (1964): *Understanding media: The extensions of man*. London: Routledge.

Millard, André (2004). *The Electric Guitar: A History of an American Icon*. Baltimore: Johns Hopkins University Press.

Miranda, Eduardo R. (ed.) (2021): *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-72116-9

Pelleter, Malte (2020): *Futurhythmaschinen – Drum-Machines und die Zukünfte auditiver Kulturen*. Hildesheim, New York: Georg Olms Verlag. https://doi.org/10.18442/166

Rashid, Tariq (2020): *Make Your First GAN With PyTorch*. Independently published.

Schulze, Holger (2018): *The Sonic Persona: An Anthropology of Sound*. New York: Bloomsbury Academic.

Sterne, Jonathan (2003): *Audible Past: Cultural Origins of Sound Reproduction*. Durham: Duke University Press.

Stuart, Caleb (2003): "Damaged sound: Glitching and skipping compact discs in the audio of Yasunao Tone, Nicolas Collins and Oval". *Leonardo Music Journal* 13, 47–52.

Trampert, Lothar (1998): *Elektrisch!: Jimi Hendrix – Der Musiker hinter dem Mythos*. Augsburg: Sonnentanz.

Waksman, Steve (2010): *Instruments of desire: The electric guitar and the shaping of musical experience*. Cambridge, Mass.: Harvard University Press.

Discography

Hendrix, Jimi (1967, 2012): *Are You Experienced*. Sony Music (Sony Music).

Herndon, Holly (2019): *Proto*. 4ad/Beggars Group / Indigo.

Red Hot Chili Peppers, The (1992): *Blood,Sugar,Sex,Magik*. Warner Bros. Records (Warner).

Various Artists (2000): Clicks & Cuts. Mille Plateaux.