### Esmeralda Conde Ruiz: sensitive ears and insensitive infrastructures. Part II

### Interview by Michael Klipphahn-Karge | Section: Interviews

Abstract: From April to October 2022, the internationally renowned composer, conductor and artist Esmeralda Conde Ruiz (\*1980/Spain, lives in the UK) was a guest at the Schaufler Lab@TU Dresden. Whilst on residency she worked with TU Dresden's (TUD) scholars and scientists as well as experts from other institutions and researched ideas related to the Lab's main topic of enquiry which is Artificial Intelligence as a Factor and Consequence of Social and Cultural Change. Our editor Michael Klipphahn-Karge, currently a fellow of the Schaufler Kolleg@TU Dresden himself and an art scholar, interviewed Conde Ruiz about her stay and plans during the residency in Dresden.

# When composing, can AI be used not only - as is often the case - as a tool, but rather as an instrument? In this case, how can a human influence the AI and how does - for example - the use of AI systems have an effect back?

Finding out if I could use AI as an instrument was one of my hopes with this residency. All the AI sounds we hear are still labelled by humans or lifted from somewhere else and this for me was not a unique or musically innovative idea to compose with.

What I do find interesting about your question though is to consider the difference between a tool and an instrument. For example, in my artistic practice I play the piano, but I do not consider myself a pianist. It is a tool for me but the voice I actually consider as my instrument. I have spent years perfecting the voice, understanding it, researching it, analysing it and I have mastered this instrument. In addition, I use computers and technology as tools to create new effects, or provide me with more options, however the important part here is that I create the sounds, not the technology.

If AI was to be an instrument, I would like to take time to learn how to play it, in the same fashion as learning any other musical instrument. So, more time would be needed to master it and to learn its rules and then how to break those rules, how to create something unique with it. And here I found the biggest problem: time.

Junior Professor Matthew McGinity (*Immersive Media*, Institute of Software and Multimedia Technology, TUD) said something that really resonated with me: nothing meaningful with depth can currently be created as the industry is moving so fast that one can't keep up with the updates. It's like playing a new instrument every day and one needs to relearn everything again rather than mastering and exploring it.

If the instrument keeps being updated constantly, I can't really create anything meaningful other than showing how it works and that is what I currently see in most AI work. Obviously, someone who comes from an algorithmic music making and live coding background might respond very differently to this question, and I am sure someone else can create an instrument but for me it just didn't work. The nature of an instrument is accessibility and exploration within the parameters. None of this was possible with the idea for my model in such a short time.

Is it possible to give AI its own artistic voice and thus find or develop an individual sound? What role do humans play in this - on the one hand as the person who operates and programmed the AI and on the other hand as the person whose voice can be replaced by AI? It is hard to say what Al's own artistic voice might sound like. I have only come across voices *in the style of* humans' or newly created synthetic voices that are based on natural voices. Does copying a human voice count as giving Al its own artistic voice? A step into this direction of using synthetic voices as an instrument is *Holly*+. It is created by Holly Herndon, a Berlin based sound artist and sound designer. Anyone can upload polyphonic audio to the website and custom voice instrument by *Never Before Heard Sounds* to get a download of the song spoken back in Herndon's voice.

But the role that humans play is vital here: as synthetic voices are getting more and more human sounding there is also currently the discussion around if it should be signified that the voice you are hearing is not human. On the other hand, when we have recorded phone calls this recording will be fed into another AI model to learn to sound more human. Without the human there is no synthetic voice.

Another fascinating voice is currently the sound of AI translators. The models are incredible but equally show the differences of voice qualities depending on the country. In Google's AI translator specific countries such as Armenia don't have a voice, whilst Bosnia's voice sounds very robotic, and Catalunya sounds incredibly human and accurate. Does the wealth and political situation of a country determine the quality and accuracy of its artificial voice then?

There is so much energy and money being pooled into AI and generating human sounding voices. Gender neutral voice assistants. Commercial voices. Singing voices. A choir of artificial voices. As there is no AI voice without humans the human element is much stronger to be considered here.

### What role do authorship, artistic autonomy and originality play in this respect?

Authorship, artistic autonomy and originality have always been topics of dispute in the arts. There are endless court cases before the use of AI that have already divided the industry. An example: *Stable Diffusion 1*, a machine learning-based text-to-image model by *Stability AI* generated images in the style of artist Greg Rutkowski, to the shock of Greg Rutkowski. The model was clearly trained with his images, without him knowing. The harm this has caused to the artist and others is immense and unrepairable. Because that model has had very precise training to reproduce such an image, a copyright free image. A recent update to *Stable Diffusion 2* makes copying artists harder and apparently users are furious. This update has probably not been implemented to protect artists but to avoid legal challenges, which raises the question, who are these models really for.

The problem gets more complicated as many of these models are subscription based, *Stable Diffusion* is for the open-source community, it can be further developed to individual ideas through one's own adaptations and code. Which equally means it is harder to control and to understand what happened to that training data set, what has been created with it and what has been added to it. Unlike humans, those models can't forget or unlearn.

Another interesting case is the one of Jason M. Allen, who won first place in the digital art category at the Colorado State Fair Fine Arts Competition this year. His image was made with *Midjourney* which sparked a debate about what it is exactly to be an artist. The judges didn't understand how the image was created by AI.

What that means for sound-based work is still unclear. Sound is more complex due to the length of longer pieces, and it is much more challenging to compress sound as opposed to the pixels in an image. The current text to sound models aren't nearly as impressive as their visual cousins, but you can sense that they will get better. And what do we do then? Copyright law is one of the most crucial laws to protect an artist's creations.

How do you judge an artwork created with similar models? Do you request to see the code, a score, the

training data? How do we protect artists' ideas and often lifetimes of work?

# What role does the idea of networking, human to human, acting in collectives and so to say *actual contact* - for example in choirs - play when you measure these concepts of *making music together* against digital possibilities?

Art critic Lucy Lippard says in Sweeping Exchanges: The contribution of Feminism to the Art of the 1970s :

"We take for granted that making art is not simply expressing oneself but is a far broader and more important task: expressing oneself as a member of a larger unity, or community"

... and I couldn't agree more. I think that is what draws me back to the material of working with groups and choirs. Without them I can't create my work and through them the work finds meaning.

Digital possibilities are therefore exciting in terms of finding collectives and new more inclusive ways of collective creation. For example, recently kids on TikTok have made a completely player-generated, decentralised game, playing, and creating it together as they go, world-wide. What an incredible collective achievement and it shows that technology can be a tool of human connectivity. My artwork *Cabin Fever* for example is also fully created digitally, an audio-visual artwork created on Zoom with professional and community singers from all over the world.



Esmeralda Conde Ruiz: Cabin Fever (2022). Photo: Ros Kavanagh.

Many formed incredible bonds online and have tried to actually meet in real life. The main focus for all performers was how to stay connected and how to feel connected even though technology would often make their voices inaudible. I have great hopes for the future that we will find ways of inclusive world-wide music making.

In our research and development phase during *Cabin Fever* we spoke to many choirs around the world, choirs who did not have reliable access to the internet or who for other reasons found it difficult to participate. Globally the excitement of connecting and performing together was mutual felt though.

Through Johanna Elisabeth Moeller and Katrin Etzrodt at the Institute of Media and Communication, TUD, an extract of *Cabin Fever* was shown at their *Exploring Socio-Technical Research* Workshop, a trans-

disciplinary and multi-method workshop. Sharing the work with researchers was very inspiring for them and also for me and it sparked many conversations on digital communal music and how it can be used to make and create new communities.

# What role does the idea of an audience play for you in general, for example in terms of the experimental participation of others, and how important is the performative aspect intertwined with it?

In my work I like to blur the line between audiences and performers. To me they are all part of the piece. Without the audience, the observer, the thinker, it is just a rehearsal. Also visually, audiences bringing their bodies into the space. Their ears, their hearts. My work is audio-visual, both equal parts but my work does not exist without an audience that listens.

Sonically my sound installations and performances are always mixed with an audience in mind, an audience that is walking freely around the space. Spaces sound different with bodies occupying them, the resonance changes. We always do tests before an actual audience gets access to the space. So, in a way the audience is already there in the first sketches. Visually and sonically.

During this residency I envisioned a piece without humans, but I learned that there are a hell lot of humans involved in AI systems. Therefore, my visual ideas also changed. I suddenly saw people standing in a square room drenched in light, listening. Experimental involvement didn't feel right here. I tested it and it didn't work. What worked though was movement, from A to B. Movement as an interpretation of the system architecture of an AI model perhaps.

# For w/k, it is important to determine as precisely as possible the interplay between art and science that is taking place here. Therefore, I would like to work out with you the individual components of the project and thus the concrete connections between art and science.

Time with researchers and these valuable new conversations have really shaped the direction of the work. Below a few examples:

TUD's Junior Professor in Empirical Musicology Miriam Akkermann spent a period of months engaging with me in conversations about my initial AI Choir idea. She especially liked the framework of not copying the human sound and explored ways of how to build such a model in theory, what to measure, talking through ideas and inspiring new angles for me to think about.

Sebastian Merchel, Chair of Acoustic and Haptics (Institute of Acoustics and Speech Communication, TUD) ran tests with me with sample sounds taken from different machines that sounded similar to server farm sounds that I had heard on my first site visit. We examined together how these sounds differentiate in frequency, physical vibration, and listener experience. We tried different floor vibrations for different sounds and examined our experiences together. As our methods of using test audiences are very similar, we hope to work in the future together, for me to create an artwork that involves vibration and for him to be able to run experiments and further research through the artwork.

Orit Halpern, professor and chair of Digital Cultures and Societal Change (Institute of German Studies and Media Cultures, TUD) was generous in her time and introduced me to her deeper knowledge of digital infrastructure and the digital waste and additional researchers outside TUD such as Nanna Thylstrup. Nelly Yaa Pinkrah and Michelle Pfeifer from Halpern's team have joined the conversation to share their latest research on voice, decolonizing AI, and voice recognition algorithms that we could discuss during the exhibition time in form of public events.

Another fantastic opportunity was being invited by Matthew McGinity, who I have already mentioned, to collaborate on his two weeklong *IXLAB* workshop for informatics students. Together with his team we

wanted to explore the creation of spatial sonic experiences by constructing real-time systems to transform sound and music into structures and environments that can be explored in virtual reality. The students examined the use of generative and emergent systems and how, when coupled with musical instruments, sounds and the voice, the systems might be used as interfaces for sculpting virtual worlds. They studied concepts related to emergence, complexity, and real-time systems, as well as cross-sensory perception and binding and ideas of musical, spatial and a modal perception and cognition. In particular the students studied sonic experience through the lenses of *4E cognition* as an activity that is fundamentally embodied, enactive, and embedded and extended through space. We moved the entire workshop to *COSMO* at *Kulturpalast* Dresden and used the foyer as our augmented reality space. As the Philharmonic Choirs are based in the same building, we visited rehearsals and started initial conversations with the Dresden Philharmonic to find ways to collaborate.



IXLAB Workshop (2022). Photo: Schaufler Lab@TU Dresden.

What are the advantages of working as an artist in the field of science, how can science benefit from your work, what can scientists learn from a decidedly artistic approach to research material, technologies and so on?

Advantages of multidisciplinary work have been researched in depth. If the groundwork is the right setting, these collaborations can be very inspiring, for both sides. For true collaboration one needs to leave the comfort zone and then new ideas can be explored together in a more fruitful way. Listening skills, patience, taking time to do something different, being curious and open, creating space for such an encounter are the essential ingredients.

And in the end then both sides can benefit from the exploration and new ideas.

I have completed many interdisciplinary projects. The difference when collaborating with a researcher though is to have in depth conversations that might challenge and inspire one's own thinking in a new way. Insightful new knowledge, different and new ways of thinking that you would never be able to achieve in any other way.

There can be an entire new world explored together.

### Links:

▷ <u>www.econderuiz.com</u>

Symposium Zukunftsmusik. What the future might sound like: https://www.youtube.com/watch?v=UXx3w9pM898

https://www.hellerau.org/de/event/24-7/

In spring 2023, Conde Ruiz will present a new exhibition at *HELLERAU* – European Centre for the Arts, one of the most important centres for contemporary fine art, dance, and music in Germany. Conde Ruiz will be one of the main artists at the international festival 31. Dresdner Tage der Zeitgenössischen Musik. She will present a site-specific artwork which is the culmination of the ideas explored whilst on residency: <u>https://www.hellerau.org/en/festival/dtzm/</u>

Details of the cover photo: Symposium at TUD: *Zukunftsmusik – What the future might sound like* (2022). Photo: André Wirsig/© Schaufler Lab@TU Dresden.

#### Tags

- 1. Artificial Intelligence
- 2. Esmeralda Conde Ruiz
- 3. Interviews
- 4. Michael Klipphahn-Karge