

In Conversation with Jennifer Walshe: Performing with Intelligent Machines

Mark Dyer¹, Zubin Kanga¹ and Jennifer Walshe²

¹*Department of Music Royal Holloway, University of London, Egham, United Kingdom*

²*Faculty of Music, University of Oxford, Oxford, United Kingdom*

Abstract

In this interview, composer, performer and improviser Jennifer Walshe talks with Mark Dyer and Zubin Kanga about her work performing with machine learning algorithms and other intelligent machines. In contrast to the many articles and interviews that have largely focused on Walshe's practice as a composer, this interview focuses on her practice as a performer integrating these technologies into her work, centering around *ULTRACHUNK* (2018) and *IS IT COOL TO TRY HARD NOW?* (2017). As well as discussing the effect machine learning technologies might have on the experimental music industry, Walshe reflects on the role of the performing human body in an era of technological proliferation.

In Conversation with Jennifer Walshe: Performing with Intelligent Machines

The music of Irish composer Jennifer Walshe has been commissioned, broadcast and performed across the world. In addition, Walshe frequently performs as an experimental vocalist, exploring a wide gamut of vocalities. Many of her works are commissioned specifically for her voice, either as a soloist or with other instruments. Walshe is also an active improviser, performing

regularly with musicians in Europe and the US, and in her duos Ma La Pert with Tony Conrad, PUTIF with Tomomi Adachi, Ghikas & Walshe with Panos Ghikas, and together with Wobbly (Jon Leidecker). She has been the recipient of fellowships and prizes from the Foundation for Contemporary Arts, New York, the DAAD Berliner Künstlerprogramm, the Internationales Musikinstitut, Darmstadt and Akademie Schloss Solitude, among others. Walshe is currently Professor of Composition at the University of Oxford.

A great deal has already been written about Jennifer Walshe's practice as a composer. In this interview, we instead focus on her role as a performer working with machine learning algorithms and other intelligent technologies. After a general discussion about her relationship to technology as a performer, we discuss two key works written by Walshe for herself as vocalist, *ULTRACHUNK* (2018) and *IS IT COOL TO TRY HARD NOW?* (2017), and then conclude with a conversation regarding future projects and possibilities. The interview provides a window into Walshe's perspective on the performing body, on agency and collaboration in an era of intelligent technological proliferation. Ultimately, Walshe perceives her voice and body to be inextricably entangled with such systems in a symbiotic, creative companionship.

Your Relationship with Technology as a Performer

Mark Dyer and Zubin Kanga: You've previously stated that your voice operates closely in collaboration with technology and, in particular, with the microphone. Could you elaborate on this collaboration and how it has evolved throughout your career?

Jennifer Walshe: I regard my voice as an amplified voice, first and foremost, not an *a cappella* voice. I think of my voice as happening in tandem with the microphone, the PA, the space and

the acoustic. It's an amplified voice, designed to have a wide dynamic range and have everything picked up.

In *Performing Rights*, Simon Frith (1998, 187–189) describes Frank Sinatra as a master of the microphone, using it to make people feel intimately connected to him. Since the advent of the microphone, there are two types of performers: those who always perform with the microphone and use it as an instrument in their proximity, including how or whether they hold it, where it attaches to their instrument and how it relates to their instrument; and those who operate independently from that.

When we talk about intimacy in terms of sound, there's a huge difference between tiny vocal sounds—those heard in proximity to a human—and tiny sounds you might make on a violin that you would only hear in proximity to a violinist. That's where gender, sexuality and the body come flooding in. I'm interested in an array of sounds on all instruments, at all dynamic levels. The microphone is handy with the voice because you can catch little whispers and glottal clicks that would be impossible to hear otherwise.

MD and ZK: In your recent work in which you're also a performer, you use machine learning and neural synthesis. How does performing with and alongside these technologies differ from experiences of performing with other forms of technology, e.g. live electronics, MIDI synthesisers, samplers, etc.?

JW: I love performing using different technologies. Machine learning (ML)¹ is different for two reasons. Firstly, with the systems I've worked with, you have to create a huge amount of data on which to train the network. You have a different level of investment because it forces you to

¹ The term machine learning (ML) is used here to broadly denote a field of artificial intelligence (AI) that uses deep learning techniques to generate statistical or sample outputs based on a training data set.

think about datasets and training cycles, and to come to grips with the ML technology itself. What's interesting about these networks is that they are trained, they're not symbolic AI where somebody has coded every possible permutation. It's important to be aware of the limits of the dataset that you train these networks on, and that the network only understands this learning. Your own biases within that dataset, in bringing it together, are all fed in.

In *Cybernetic Forests*, Eryk Salvaggio (2022) claims that an image generated by a machine learning network is an 'infographic' of the dataset it was trained on. We should be thinking in this way, in that pieces of music produced by machine learning networks are simply musical infographics of the training datasets, rather than an absolute form of music that has somehow sprung forth.

Secondly, you don't know exactly what the system you're performing with is going to do. You have a ballpark idea because you know the dataset and it's not going to go completely off-piste (though these systems are extremely sophisticated, they're also very dumb and limited in what they can do). Having to carry that is an interesting performative challenge. You have to take responsibility for long-term dramaturgical thinking, structure and callbacks. The AI isn't planning a minute in advance or remembering what happened half an hour ago, it's generating the next grain, the next sample.

MD and ZK: Do these two criteria—the curation of training datasets on the one hand and developing the macro structures within which to situate the network's outputs on the other—change your creative role as a composer or performer?

JW: For a lot of composers, working with AI is a sophisticated form of the *I Ching*. They're interested in new forms of randomness or juxtapositions they might not have thought of before. Many composers are currently training AI systems on their own output. If you're a film

composer, ML tools are useful because you can use them to come up with 30 epic themes for a *Game of Thrones*-like TV show. You can use them to iterate music and generate ideas quickly. I imagine they'll be integrated into a workflow for a lot of people.

In chess, the best unit is a human-computer pair called 'centaurs.' They are good at beating either a computer or a human, because you have assistance, something which helps you. I think there'll be more of this within music, especially in pop music. For me, as a performer, I'm curious as to where it will go.

But, also, these systems change very quickly and a lot of money goes into pieces with expensive tech that will be obsolete—it'll have the wrong OS or the company will have been bought out—5 or 10 years from now and it won't function anymore. We don't talk about that planned obsolescence.

MD and ZK: What hasn't changed, when you're performing with machine learning systems, from working with older technologies?

JW: I think of it in terms of what hasn't changed from working with *people*. When I perform with *ULTRACHUNK*, I'm performing with a network that is giving me things in real time, but it can't think long-form. It's not aware of who's in the room or the energy in the room, for instance. I feel like I'm dealing with an intelligence that isn't on a par with a sober, adult human. I have to manage and massage that situation, to make it work, style it out and make it coherent, or embrace the incoherence in such a way that's interesting.

MD and ZK: Your use of humour reoccurs in a number of your performances. But it's not always directly a result of the composition; it's often about allowing the audience to find humour in the relationships you've created with the technology or the strange outputs of the machine

learning. Is this something you are aiming for as a performer? Do you think about that aspect of audience reaction?

JW: Of course, I'm aware of comedy and its broad human practice—I love reading technical analyses of stand-up routines.

But what I'm doing is the equivalent of somebody messaging you saying, "My God, did you see this? This is completely fucked up!" I feel I'm presenting the audience with the absolute absurdity of life as it is lived now. I'm not trying to write a joke or comedy routine. It's the equivalent of me sending a screenshot of a Reddit post with no context.

It's liberating that people feel happy and able to laugh. There are times when people laugh at things that you couldn't have predicted. So, I don't try to write gags because you don't know how anything will land. I suppose my approach is very internet-like, in that there's this flood of material often without context.

ULTRACHUNK (2018)

Written in collaboration with computational artist Memo Akten, *ULTRACHUNK* is a live, improvisational duet between Walshe and her AI doppelgänger. Every day over the course of a year, Walshe recorded solo improvisations in front of her webcam. Akten trained a machine learning system on this footage in order to mimic the key components of Walshe's voice and face. During the performance, the video and audio outputs are generated live in real time by the neural network in reaction to Walshe's similarly responsive improvisations.

MD and ZK: Do you view *ULTRACHUNK* as a duet between you and your AI doppelgänger, between you and Memo, between the neural network and the dataset, memories of your past self, something else or a conglomeration of these? In a sense, what's the performing 'ensemble' here?

JW: During performance, I'm not thinking about Memo nor the dataset. Of course, Memo is at the side making sure the whole thing doesn't melt, because it's doing so many computations per second. But onstage, I'm just thinking about the sound. Being a free improviser, I'm thinking, what sound is happening and how do I react to it? I'm listening and I'm in that moment.

But offstage, I would compare *ULTRACHUNK* to Donna Haraway's concept of the 'companion species' (2003), in that you have different species with varying levels of intelligence working together, trying to find a meeting ground. Still, I would argue that any dog Haraway has done her competitions with is far more intelligent than any neural network we could work with.

MD and ZK: In this meeting ground, are there elements of friction, of working in spite of or against one another, as well as collaboration?

JW: You're always doing that with machine learning. In *ULTRACHUNK*, for example, Memo had to put a resonator on the sounds to try to wrangle them a little bit. That then locks the sounds in, in a certain way. As well, Memo really wanted to work with video as well as audio. I'm not so comfortable seeing videos or images of myself. That was quite bizarre.

In fact, the thing that caused the most friction for me—not between myself and Memo, we had a very nice working relationship—was an installation version of *ULTRACHUNK*. It was in a gallery, running in real time. People could sing into a microphone and *ULTRACHUNK* would generate live videos of me and sounds using my material, because that's what it's trained on. I would walk by a gallery and see a stranger sock-puppet this weird, machine learning version of me. I couldn't be in the same room! I think Holly Herndon's *Holly+* (2021) is a

brilliant project² but that's not my path. There's a lot of intimate information locked into the voice, that you're allowing other people to access. That, for me, was a bit too much.

MD and ZK: Do you think a performance of *ULTRACHUNK* says something about you as a performer? About the sounds you favour, your facial expressions, the sounds these expressions entail or about you as a person?

JW: It's horrific, in that it shows all the ticks and little habits. I kept a journal when I was making the recordings and that's what I kept writing. You notice things about yourself.

But, it's also limited. It demonstrates things that can be gleaned from a limited dataset of a limited set of my experiences. If I can quote Salvaggio (2022), *ULTRACHUNK* is a musical infographic of the dataset, and the dataset was Jenny. It's a very strange snapshot of where I lived, where I travelled to, what haircuts I got and what clothes I wore. Doing these multiple 5-minute improvisations in front of a webcam in different settings, often when I was exhausted or in the dressing room just before I went on stage, it's a strange dataset.

MD and ZK: When you're performing *ULTRACHUNK*, do you have any preconceived structure or trajectory?

JW: No, it's purely free improv. I performed it at Rewire Festival (2019) and it was 40 minutes long. That was a real challenge because I had to hold it for the duration and didn't know if the system was going to break down. It's listening at different degrees of closeness and distance, which again feels like a classic free improv setup, quite natural and organic. It's a different thing to do a long structure—I was completely wrecked by the end, psychologically and physically.

² A publicly-available audio model of the artist's voice, available at <https://holly.plus>.

MD and ZK: Does working with these kinds of technologies alter how you incorporate the body and physical gesture in performance, and, flipping that around, do your physical demands instead inform your uses of the technology?

JW: I like to move around on stage, look into the audience and point to people. I perhaps have a different physicality to many new music performers—I use my body more like someone in free improv or pop.

With technology, specifically, what's great about a microphone is that you can hold it in your hand; I hate headsets and I never use them. I like tech that I can get my hands on, that I can be organic and tactile with. *ULTRACHUNK* is a challenge because, whilst I'm singing into a microphone, a lot that's happening is disembodied and that's strange.

MD and ZK: Has performing with these technologies altered your sense of what liveness and presence mean?

JW: Yes, definitely. One of the systems that I've performed with a lot is Jon Leidecker's (aka Wobbly) Monitress. Monitress is a coalition of pitch-tracking codes, each installed on numerous broken or early generation iPads and iPhones, running obsolete OSs. It's not a sleek Max patch on a state-of-the-art M1 computer. Even though it's tech and code, they're running on physical devices that you can hold in your hands. Because they're running on different OSs, they're all running at different speeds and making slightly different decisions.

When I'm singing with Monitress, it feels like I have this flock of birds that are swarming around me. It's very organic. When I work with Jon, he's making changes the whole time and responding to what I'm doing. So, it's two people completely embedded. It's different from *ULTRACHUNK*, because it's me, Jon and Monitress, all intertwined in that moment. When I'm

performing with Monitress, I know roughly what's going to happen because I've sung with it many times, but I don't know exactly and I'm happy with that.

ULTRACHUNK is more on the edge because it's just me with a version of myself. This makes me feel very exposed as a performer because of the personal aspects of vocal sounds, as I mentioned. It's challenging because I've got to hold and carry it. When it's a video of yourself, you also feel a little self-conscious. It was a fascinating project to do for that reason, because it was something different.

At the premier, I remember saying to Memo, we need to explain to the audience what's happening because they can't possibly understand how cutting edge it is, they'll just think this is all pre-recorded video. At the first performance, people were freaked out! Their minds were melted by the GAN-ish looking Jenny with two eyes in each socket and multiple sets of teeth, and the fact that it was being generated live.³ It felt like a demonstration of where Memo could get the tech at that point in time—how fast we could get a PC working using the code that existed. I think of 'technology' as a type of technology that is produced by certain people or stakeholders at specific points in time. If we were doing *ULTRACHUNK* now, we would have a different suite of options available.

MD and ZK: Are there future plans to do an evolution of *ULTRACHUNK*?

JW: No, I don't need to do that. That technology has become par for the course now. So the pieces I want to make are different because those systems have been integrated into people's lives. It's about what's possible right now or what one version of the future could look like,

³ GAN (Generational Adversarial Network) is a class of machine-learning frameworks, commonly known for its application using image data to generate realistic-looking animals and human faces.

trying to think that millisecond ahead. Because what AI is and the way we relate to it is constantly changing. I find that constant change interesting.

IS IT COOL TO TRY HARD NOW? (2017)

IS IT COOL TO TRY HARD NOW? is a concert work written and performed by Walshe for voice, video and electronics. During composition, Walshe trained the deep neural network WaveNet on a varied dataset, including her diary entries and text, images, audio and video from wildly divergent sources. She uses the outputs generated by the network to derive materials for performance.

MD and ZK: In *IS IT COOL TO TRY HARD NOW?*, you trained WaveNet on your voice and bodily gestures. But there's a lovely symmetry, or feedback, where you're then training yourself as a performer to mimic the network's outputs. Does this mimicry, this symmetry, represent an implicit treatment of your voice and body as technologies or, even, the neural network as pseudo-biological? You've previously referred to the inhuman qualities of the vocal sounds such systems can produce.

JW: I'm interested in the weird, messed up sounds WaveNet, or other programs, can generate through vocal synthesis. They're interesting to mimic as a person. It involves me at times learning to mimic files that people are not designed to perform, and it makes my voice function in a different way. It's similar to Oneohtrix Point Never's use of vocal samples at speeds and in registers that nobody can sing in.⁴ You have this human voice but it's used in a machinic way.

⁴ For example, in 'Child Soldier' from *Replica*, released by Mexican Summer Studios, 2011.

In *IS IT COOL TO TRY HARD NOW?*, I was interested in trying to embody music or sounds that we don't normally hear, sounds that result from different things being mashed up together in a neural network. The limits of this are interesting, because your body's trying to do something in which the aesthetic comes from machines rather than bodies.

I'm currently writing a book about this blurry line between beings and non-beings. I'm interested in why we think certain things are alive and certain things aren't, and how you can trace these blurry categories right back to the beginning of people. If you go to a Shinto temple in Japan, there'll be a stone in which it's said a spirit lives. In the same way, it's hilarious to watch people pretend to be robots in movies, like what Michael Fassbender does to his voice in *Alien: Covenant* (2017). Or, the idea that, for a lot of people, the internet is something that is haunted or has an occult. I'm interested in flipping these things back and forth, turning them on their heads.

MD and ZK: Do you think this 'haunting' is exclusive to technology here, or is there also a haunted aspect to your voice, a performance of memories?

JW: Oh, yes, definitely. That's how Irish people use their voices, they mimic people and do accents when they retell stories. In shows like *Father Ted* there's a huge range of voice types and each priest has a different vocal tone. That's really important, and makes it ten times more funny. If you're a natural mimic, you notice those things. Then, of course, that goes into the work, for better or for worse.

MD and ZK: *IS IT COOL TO TRY HARD NOW?* engages directly with your love of collecting and, in particular, your love of language. Has performing this work made you think differently about language and your relationship to it?

JW: No, because I'm always collecting stuff and using it in all of my work. It's part of a larger interest. There were simple technological tools that made writing *IS IT COOL TO TRY HARD NOW?* easier, like Evernote, content syncing across devices and being able to search by a word, rather than flicking through a journal (though I still work that way!). Certainly, with all the things I collect, I could have a machine learning system that would make suggestions or connections, but I want to conceptualise that more.

What Now? Looking Ahead

MD and ZK: Have you seen any work recently that's changed the way you think about how the body, specifically, can interact with technology?

JW: I recently watched Future Islands performing *Seasons (Waiting on You)* on the David Letterman show in 2014.⁵ The lead singer, Samuel T. Herring, has this very strange way of moving on the stage. He does these physical gestures where he's showing you where the music is going with his body. He's just got an SM58 and he knows exactly when to bring it in. You keep thinking he's too late and he's going to miss his line, but he knows exactly where that mic is and when it's going to come up to his face, as sure as he knows that he's breathing and he'll be able to do the next note.

I thought to myself, for all the tech in the world I can watch somebody just kill it on stage with an SM58 in their hand and it can still make my day as a human being. I see people do incredibly complex and interesting things, technologically. But, I also see people just commit to using their bodies, where the technology becomes second nature.

⁵ Available online here: https://www.youtube.com/watch?v=upP19mZW_zw&ab_channel=Letterman.

Sometimes, work using high-tech can feel more like a trade show demo. From that angle, the Museum of Emerging Science and Innovation in Tokyo really impacted me. I spent the whole day interacting with the ASIMO robot and all the Otonaroid robots. As a result, I wrote a short story, *Waiting for Asimo* (2022). So, these things don't even necessarily turn out as musical pieces.

MD and ZK: If performance-with-technology is a discipline, broadly speaking, do you feel it's at a crossroads or precipice? Related to that, what do you think or hope the coming years might bring and what might your future plans be?

JW: I don't know if it's at a precipice. I'm keeping an eye on what's happening in the art world right now. Art websites are being flooded with images that were made by people with no visual art skills other than pumping the text prompt of machine learning networks. Some communities don't want these images and others want them labelled as such.

But, systems like DALL·E, MidJourney and Stable Diffusion are all predicting text-to-audio and text-to-video arriving very soon. If they are correct and it's good, that will change things. It's certainly going to shift things for people who write in the commercial space. For some, it will be a helpful tool to iterate things. For others, it might make it difficult to ever get their foot in the door because the market will be so flooded. It's different from the invention of the drum machine or the home synthesiser, simply because of the speed and scale at which these systems can be deployed.

MD and ZK: And our ubiquitous complicity with these technologies, whether we like it or not or whether we're aware of it.

JW: Yes, they're everywhere. I'm curious how things will pan out. A lot of these networks are little toys that people use to make NFTs, which means that the use of AI to make art is inadvertently feeding the crypto market, which inadvertently will or will not affect banking structures in general. It's complicated!

Funding

This research was undertaken as part of Cyborg Soloists, supported by a UKRI Future Leaders Fellowship and Royal Holloway, University of London (Grant number MR/T043059/1).

Research Data Access

Full documentation of the interview cannot be made publicly available as participants have not consented to public sharing of these materials. Selected documentation materials cited in the article can be made available on request to researchers.

Reference List

Frith, Simon. 1998. *Performing Rights: On the Value of Popular Music*. Cambridge, MA:

Harvard University Press.

Haraway, Donna J. 2003. *The Companion Species Manifesto: Dogs, People, and Significant*

Otherness. Chicago: Prickly Paradigm Press.

Salvaggio, Eryk. 2022. 'How to Read an AI Image: The Datafication of a Kiss.' *Cybernetic*

Forests.Substack. October 2. <https://cyberneticforests.substack.com/p/how-to-read-an-ai-image>

Walshe, Jennifer. 2022. 'Waiting for Asimo.' *The Stinging Fly* 47 (2).