Four surprises of electronic music

Miller Puckette[1]

Abstract: The evolution of electronic music practice over the past half century has not gone precisely as we expected. In this paper, we describe four respects in which the development of electronic music was either overtaken by unexpected developments, or, itself, headed in new and unexpected directions.

Keywords: Electronic music. Pure data. Spatialization. Sound art.

Quatro surpresas da música eletrônica

Miller Puckette[2]

Resumo: A evolução da prática da música eletrônica ao longo do último meio século não ocorreu precisamente como esperávamos. Neste artigo, descrevemos quatro aspectos nos quais o desenvolvimento da música eletrônica foi ultrapassado por desenvolvimentos inesperados ou, por si só dirigido em novas e inesperadas direções.

Palavras-chave: Música eletrônica. Dados Puros. Espacialização. Arte Sonora.

[1] PhD in Mathematics (Harvard), Prof. Department of Music University of California, San Diego. Wrote Max, a computer music software environment and developed Pure Data ("Pd"), an open-source real-time multimedia arts programming environment. E-mail: msp@ucsd.edu

2 PhD em Matemática pela Harvard University. Professor do Departamento de Música da University of California, em San Diego. Criou um software de música para computador, o Max, e desenvolveu o "Pure Data" (Pd), um ambiente de programação de artes multimídia em tempo real, com código aberto. E-mail: msp@ucsd.edu

Introduction

Things often go according to plan, but sometimes they don't. There is value in a thing going according to plan---if there weren't, we wouldn't bother to make plans at all. But there is a different value to things that don't: they are opportunities to learn something, or at least invitations to think more deeplyor in a different way about something.

Over a career several decades long in electronic music and art, I've had the pleasure of being surprised often. Looking back, I find myself able to group or conglomerate them into a few rather large, slowly-emerging surprises which I'll take this opportunity to examine in their own right. I'll arbitrarily choose a point 50 years ago as a reference point.

One very interesting thing happened that, nonetheless, was no surprise to any one who watched Star Trek in the 1960s: we all now have telephones and computers in our pockets. (There were minor variations between predictions. Dick Tracy's phone was in his wristwatch, but Maxwell Smart's was, more magically, in his shoe). What I describe here is not a list of cool gadgets, but instead a meditation on changes in our relationship with them and its implications onartistic practices. The next four sections will each take on one major surprise I've encountered. In a last section, I'll try to give at least one good reason that there might be more surprises in store.

Computers became primarily communication devices

In the 1960s, computers were used in banks and other businesses that handled large numbers of repetitive financial transactions, and in research laboratories, mostly but not exclusively military-funded. Communications, in contrast, were mostly mediated via paper or telephones, or sometimes even direct human-to-human contact. One did not use a computer and communicate at the same time; these were quite separate activities and if any project involved both computing and communicating, the two belonged to separate phases of the project.

Today the main difference between a telephone and a computer is physical size (the telephone fits into a pocket or purse; the computer into a briefcase or handbag). Because of the size difference, the computer can dissipate about 10 times the wattage that the phone can, and can include a reasonable keyboard and display a page or two of text. One uses the computer sitting down, and the telephone standing up (or while driving). The computer is a better platform for reading and writing and the telephone better for oral communication, but each is perfectly able to stand in for the other for short periods of time.

The driver for this change has been that the primary mode of interfacing between computer and human is no longer paper as it once was, but is now sight and sound. It was rare outside of a research laboratory in, say, the mid 1960s, to see a camera or microphone or loudspeaker connected to a computer. (Various CRT-based displays were starting to appear however.) In some research settings it was starting to become desirable to connect environmental sensors directly to computers, but computers were not often used directly as actuators, as in a robot or computer-controlled tool. This may have partly been because computers were not very reliable, and thus were ill-suited for real-time applications.

A computer program's output might be more voluminous than its input, and anyway was mostly numeric, so it is natural that visual output was considered more useful than audible output. Early CRT displays gradually were standardized into 70s-era CRT terminals, primarily alphanumeric, but some with primitive graphics output capabilities as well. Graphics-capable printers were starting to supplant alphanumeric ones.

Even though Max Mathews's MUSIC program appeared in 1957, the practice of connecting DACs and ADCs to a computer to enable audio output and (somewhat later) input was still exotic. The MIT Experimental Music Studio, where Barry Vercoe wrote what would become csound, had audio output capability in 1979 when I started working there, but no ADC until 1981. The device cost about $20,000 in 1981 dollars.

It was hard to imagine then that computers would later come standardly equipped with audio input and output. The first computer I saw that to offer that was

the NeXT computer in 1988; at $10,000 or so for a reasonably-equipped machine, this was still a bit of an experiment; but in less than 10 years PCs were being shipped with stereo "CD quality" input and output built into the motherboard.

This innovation marked the transition of computers from computing devices to communications devices. At first the most common use for audio I/O was to watch movies and play music and games: not "work" but "entertainment", rising, occasionally, to the level of art. Images were fine and good as long as it was work that was being done; but when art was being created or even just consumed, audio became essential.

The addition of audio I/O as a standard computer feature brought about the erasure of the distinction between computer and telephones, although that did not become apparent until cellphones (which had processors from the outset) started to hold enough non-volatile storage to host what were by then called "applications".

The convergence of the computer and the telephone was part of the same gesture that established the computer primarily as a media tool suitable for creating artworks, both fixed-media and interactive.

The instrument became an app

Computer Music in the 1960s was a batch affair, although apparently not by Max Mathews's choice; it was a limitation in the technology of the day. Back in the analog world, patchable synthesizers appeared in the mid 1960s, which, unlike computers of the era, ran in real time and could be used (with difficulty) on stage, as a musical instrument. Many people worked toward, and looked forward to, the day a computer, too, could act as a musical instrument.

Such an instrument would have advantages over the patchable analog synthesizer: lower noise floor and distortion; recallable presets without the need to manually move patch cords; and, in time, greater sound-making power as measured, say, in the size of the patch one could run.

Researchers and artists knew approximately when that day would come by simply extrapolating computer power. Personally, I put the date at 1997, when off-

the-shelf CPUs broke the 100 megahertz mark. (It is no coincidence that the first releases of Pure Data came out that year.) At that point artists were relieved of the need to program their sounds; they could just perform them computer-instrumentally.

The surprise was that, far from renouncing programmability as a no-longer-useful relic of an earlier time, if anything, the complexity and autonomy of the computer programs that we used in music performance increased. It was both easy and fun to make patches that simply ran by themselves, or else that generated large numbers of musical events semi-autonomously under the control of higher-level inputs supplied by the "performer".

It was easy to imagine two separate strands of development continuing into the real-time era: the programmable, automatic computer music orchestra that could "render" music in the way of the MUSIC N programs (descended from Mathew's original MUSIC program, the main instance now being csound) and the computer-as-musical-instrument that would be essentially a digital patchable synthesizer. The surprise was that the distinction between the two ways of approaching computer music simply disappeared, and we now have something more general: the smart musical instrument.

Perhaps related or even because of this, we are now seeing a gradual dissolution of the distinction between composer, performer, and programmer. This has freed many an artist to forge his or her own role with respect to the tools of the trade.

Space became a manipulable musical quantity

Spatialization has been an important element in classical electronic music, showing up in work by Stockhausen, Berio, Nono, Chowning, and many others. A mainstream practice emerged in which spatialization consisted of the simulation of static and/or moving sound sources. One places the audience at the interior of a ring or a shell of loudspeakers, at least four of them but possibly hundreds, and then by adjusting the levels at which a given sound appears in each speaker, and perhaps

using other cues such as amount and type of reverberation, one can simulate sounds coming from different locations. This is clearly something that can be done much more flexibly electronically than one could do physically by, say, asking instrument players to run from here to there as they play.

This was catnip to formalist twentieth-century composers (such as those named above), in that any out-of-time musical structure, that might be manifested as a set of pitches or time durations, could also be manifested as a spatial location. A textbook example of this is Pierre Boulez's piece, *Dialogue de l'Ombre Double*, in which sounds jump among the six loudspeakers (twelve divided by two) in ways that make spatially defined musical phrases and shapes.

However, this approach gives rise to a serious theatrical problem. One is simply not used to sitting at a concert pointed in a direction different from that of the sound source. Admittedly, when working to the sound of a transistor radio we don't insist on facing the radio; but when we listen attentively, as to a fireside chat or a Beethoven symphony, we want to face the loudspeakers. I think this is for the same reason that we face a human speaker to whom we are paying attention. Things that have narrative import, we face; things that are environmental cues we accept when coming from oblique directions, and only when such a sound becomes either interesting or threatening do we turn toward it. In this understanding of the listening process, it makes no sense to have a musical narrative that darts from place to place as it unfolds.

Related to this problem is one of balance between live players and electronic sounds, particularly if the electronic sounds are derived directly from the instrumental ones. To see a sound made in one place but to hear it from another gives rise to a kind of cognitive dissonance that can easily cloud the musical message. (In *Dialogue*, Boulez brilliantly sidesteps this problem by darkening the hall when sounds come from the loudspeakers so that we no longer see the live clarinet player; the live sounds and the electronic ones take turns. But this is not a thing one could do habitually; it would get tiresome).

A further, well-known problem with the moving-source model of music spatialization is that of the "sweet spot", the tendency for sound localization effects to

fail for listeners who aren't placed nearly equidistantly from all the loudspeakers. If the loudspeakers (at least, the ones that are in the musical foreground) are all frontal this can be achieved over a reasonably large area; but when we add speakers off to the sides the "sweet spot" shrinks to a small area, a substantial fraction of which is often occupied by the composer, sound crew, and gear.

The surprise here was the emergence of a broad spectrum of new possibilities brought about by a focus on the space itself, as opposed to its earlier, rather dismissive treatment as a mere blank canvas parametrized by three coordinates. This change certainly has roots in the distant past, but perhaps first came into sharp focus in 1969 with Alvin Lucier's piece, *I Am Sitting in a Room*. In this piece the air and walls of the room are the musical instrument itself. The performed-in space is absolutely central; it is not performed in, but is itself performed.

Today we do not much go to concerts in concert halls with their rings of loudspeakers, nor sit down after dinner to give our close attention to the radio or victrola as Enrico Caruso sings an aria. Listening happens in very non-neutral spaces: in galleries; between us and our car dashboards; between earbuds and eardrums. Sound on its way from the actuator to our ears may bounce off odd surfaces; travel through media other than air; bend around corners or obstacles; mix with sounds from other sources in the environment.

The composer Pauline Oliveros has greatly contributed to our awareness of the importance of space itself, not least in her Deep Listening practice which included performances and recordings in spaces so extremely reverberant that any idea of a localized sound source is rendered absurd. Maryanne Amacher's performances explored the behavior of sounds in architectural spaces, to the point of treating the building as a co-performer. The spatial aspect of music has been freed from the limiting, Cartesian framework bestowed on us by the pioneers of postwar electronic music.

Sound art

It is only a slight oversimplification to say that, fifty years ago, an artwork whose medium was sound was likely either a radio play or a piece of music. There was a narrative structure spanning a fixed length of time between a beginning and an end. The seated audience, if more than one person, all faced the same direction. There were a writer/composer, performers, and audience, each playing a well-understood role.

Continuing to oversimplify somewhat, one can now apply the name "sound art" to a vast new field of artistic endeavor that uses sound as a medium but does not fit the narrative-performative model. It should not be any surprise that there is a great deal more space outside any given set of strictures than within it. The surprise has been that, rather than being a barren void, the space beyond the pale proves to be littered with all sorts of interesting possibilities.

Sound art is closely tied to electronic music practice because it requires a level of adaptability and malleability that is not offered by a traditional musical instrument. It is hard to think of many ways one could incorporate a violin, for example, into a piece of sound art. But there are many ways in which electronic transduction of sound can enter into an artwork. The groundwork laid by electronic music researchers, especially those of us who write computer environments for realizing electronic music, turns out to be ideally suited for use in sound art as well.

The sound artist must take into consideration the fact that the viewer may come and go at will, and unless one places rigid controls, there will often be more than one viewer at a time, each entering and exiting independently. The work should somehow reward a longer-staying visitor without denying itself to a more casual onlooker (and moreover, should somehow invite the passing visitor to linger). An example of a work in which this is managed extremely well is Ben Rubin and Mark Hansen's *Moveable Type* (1986), commissioned for the lobby of the New York Times building, which shows decontextualized snippets of text from the massive back catalog of Times newspapers while making quiet, unobtrusive typewriter-like sounds whose ebb and flow punctuate the updating of the texts. Over a longer period of time

the viewer perceives stylistic changes reflective of the different sections of the paper (editorials; obituaries; etc) that are being mined. But a frequent visitor, an employee for example, might merely pass thorough the lobby each day; for her the piece serves as an atmosphere setter, even as, over the years, it perpetually rewards an occasional, momentary glance and listen.

A sound artist has great freedom in deciding whether, and how, the artwork is to react to visitors. This has been apparent since the earliest days of the form, for instance in sound sculptures made by Liz Phillips before the term "Sound Art" was coined. Phillips used capacitive sensing (as in a Theremin) to allow the position of one or more visitors to alter the pitch and timbre of a synthetic drone sound. The sculpture itself was the antenna that responded to the capacitive coupling, so that the shape of the sculpture determined a pickup pattern, like that of a non-omnidirectional microphone. Later artists have used more sophisticated proximity sensors such as sonar or video cameras, but the robustness and simplicity of Phillips's approach still stands out. (Full disclosure: I'm lucky to count Phillips among my teachers.) A viewer could come away with the feeling that she is somehow emitting invisible energy that the sculpture is making audible.

In contrast to proximity sensing, which does not require the viewer's conscious engagement, one can choose to invite the viewer to interact with an artwork on purpose, as in, for instance, Katharina Rosenberger's *Viva Voce* (2013). In this work, video footage of three singers (of very specific and contrasting styles and approaches) appears on three larger-than-life screens, each with a touch-sensitive device mounted at a close but reasonable viewing position. The piece must be housed in a large enough room that a viewer can watch one or another singer while hearing what the other two are doing at a somewhat lower volume. The visitor can either "play" one or another singer in turn via the touchpads, or just wander through the space. Up to three visitors may be at the touchpads at a time. The result is a bit like a cross between a farmer's market and a video arcade. In sharp contrast to Phillips's work, here the visitor is empowered as a sort of computer user, and is called upon to make executive decisions---a stance that sets up a fundamentally different mindset in which to receive the work.

For some reason that I can't explain, these and many other of the best works of sound art that I could cite play with issues of presence and absence. The biggest splash of any piece of sound art recently was made by Janet Cardiff's *Forty Part Motet*, in which forty loudspeakers, mounted on stands at about the height of a human head, each sing one voice in an old and very beautiful piece of vocal music. Besides being beautiful, the result makes the visitor feel, almost urgently, the absence of the singers themselves. Are we hearing the voices of the men and women who once sang this motet 550 years ago? To visit this piece is to be made aware of unbridgeable spans of time and space, and yet to experience a sweet and intimate closeness with the unseen singers.

Why this might be interesting

The preceding paragraphs are concerned with the work of the last fifty years in electronic music and sound art. It would be interesting to see what any of these developments might imply, or incite, by way of new work in electronicmusic. My own way of approaching this question is to consider what new possibilities might have arisen for designers and users of electronic musical instruments.

The idea of the electronic musical instrument seems important to me because of my belief that music is primarily a performed medium (I haven't been much attracted to musical automata, except, as described above, when they are used in some way as a musical instrument). When we look closely at what goes on when a human plays an instrument, the process starts to appear mysterious. The closer we look at it, the deeper the mysteries appear. Thus:

Over the years researchers have made many attempts to capture the stream of control that flows from the human body into a musical instrument, for instance by full-body-tracking a percussionist in an attempt to capture the relationships between legs, torso, arms, and mallets as they move through space. The result is always a large collection of data. I have yet to see anyone manage to glean any real insight from such data, and I'm ready to conclude that nobody is likely to anytime soon, because it's impossible to capture the gestures that the performer for

whatever reason decided not to make. In other words, the gesture exists inside a space of possibilities but we can only observe the gesture itself, not the space it lives in. If we ask the performer to repeat the gesture, perhaps varying some aspect such as tempo, we get one other point, or perhaps 50, but such a collection of data points is still dwarfed by the huge potentiality of gestures not made.

Further, the human gestures of a musician are only made meaningful by their context, that is, the musical practice of our day that works like a language in which new things can be said using pre-existing words. It is the web of pre-existing meanings between the words that make it possible to make meaning by stringing them together. To make any sense of recorded gestural data from a musician, you would need somehow to codify how it relates to that context, so instead of studying a specimen you must now studying a whole world, or else learn nothing meaningful at all.

Further, the human doesn't just output a stream of data that the instrument ingests, but there is a tight feedback loop between the two. It is quite possible that, for many wind and string instruments and especially for the human voice, the same muscular gestures repeated exactly would result in audibly different results because of the chaotic or ill-posed nature of the evolution in time of the instrument. Imagine trying, for instance, to ride a bicycle without being able to sense your and its motion and adjust accordingly; it and you will soon fall over. So any understanding of what a human is doing as she plays the instrument would need to take into account not just her muscular output, but how that output would have changed had the instrument's physical motion evolved differently, or, to put it differently, her output described as a function of all possible ways the instrument's sound and feel might come back to her.

How all this happens in time is a mystery as well: the timing accuracy of an instrumental gesture is far greater than the speed at which one can make the gesture. It is not meaningful to treat human input and output as a sampled function; it is something else in which we can control aggregate facts (the moment of an attack for instance) without having moment-to-moment control on anything like the same time scale. It is also mysterious that we can control the pitch of a violin, a trumpet, or

most of all a voice, down to a part in a thousand or so. We certainly can't hold a fingertip at a particular spot in the air with anything like that accuracy.

Even all these pessimistic considerations don't fully account for the difficulty of understanding what is going on, because the whole is also somehow guided by a computing device, the brain, which will probably prove impossible ever to analyze in any meaningful way. It is a complicated device, sporting 100 billion neurons, each with on the order of a thousand connections to other neurons, each via electrochemical reactions which time constants are on the order of a millisecond, and each of which we might conservatively estimate would take a hundred or so computer operations to simulate. That adds up to ten quintillion operations per second, or approximately a billion times the power of a CPU. (It's a marvelous fact that the brain does all this computation on a mere 100 watts of power or so, perhaps a hundred millionth of the power of a hypothetical ten-quintillion-operations-per-second server farm.)

Of course we don't aim to replace a brain here; I was only proposing to seek to understand what it is doing. However, that would be even harder than building a brain-like, intelligent computing device would be. Suppose one is given a circuit (an audio dynamics processor, for example) into which we can feed inputs and whose outputs we can measure. The difficulty of analyzing and understanding the behavior of the box is perhaps not great if there are only one or two degrees of freedom (capacitors, say, each of whose charge is a time-dependent state variable). Depending on the degree of nonlinearity of the components, it may even prove possible to analyze a circuit with four or five state variables. Once there are ten, the multiplicity of possible interactions between the state variables is so great as to resist any attempt at measuring them from the outside. The difficulty grows exponentially. The brain, with its 100 trillion synapses each of which can perhaps be reductively modeled as having ten or a hundred state variables, will almost certainly be beyond our understanding forever.

Setting a monkey at a typewriter would be a more efficient way of designing a new musical instrument than modeling any system that incorporates humans. But if this makes a scientist (or a strong-AI-hyping technologist) cry, it

should make an artist start to think. Art, and especially music, is always a dance with the ineffable. Anything that is possible to explain completely is not likely to succeed as more than pedantry, and certainly seems unlikely to rise to the level of art. The complete unknowability of the human/instrument system is an invitation to try out new ideas, and new musical instruments.

These may take the form of further software developments, new algorithms, and/or new musical interfaces. I'm experimenting with two possible examples of interfaces these days: hand-gesture control using the Leap game controller, which gives millimeter-resolution tracking of all the moving parts of one or both hands at about a 100-Hz. resolution, and control of synthesis algorithms by analyzing the moment-to-moment timbre of a human voice.

As a further possibility, perhaps someone, maybe a sound artist, will find a way to turn the paradigm of the musical instrument on its head in some way. Can instruments play each other, or people act as instruments? Are there sources of possible feedback to human players that go beyond the sound output, and/or the tactility, of the instrument? Can other media than electricity be used to carry musical signals, perhaps biochemical or photonic? Can the semantic boundaries between what is the music, what is the instrument, what is the musician, and who is the audience be transformed or made more interpermeable?

* * * * *

I can offer no summarizing conclusion, but will take this moment to once again register my surprise. As a youngster, one knows intellectually that many surprises might be in store. I'm lucky that the ones I describe here can be seen (as I see them) as openings, as new fields of experimentation that have yet to be codified and rectified; as opportunities to make new and different art. Speaking for myself, I'm in a hurry to try to make good on them.