Interview with Prof Atau Tanaka at Goldsmith University, London
30 March 2022

Atau Tanaka conducts research in embodied musical interaction. This work takes place at the intersection of human computer interaction and gestural computer music performance.

He studies our encounters with sound, be they in music or in the everyday, as a form of phenomenological experience. This includes the use of physiological sensing technologies, notably muscle tension in the electromyogram signal, and machine learning analysis of this complex, organic data.

Tanaka's compositions feature no external sensor object, but draw on internal physiological muscle-sensing data. By tensing his arms and making concentrated gestures, he sculpts sound coming from the computer by shaping and modulating parameters of sound synthesis and sampling. The similarity with gestural sensor systems is that the muscle-sensing data also reacts in relation to the body and not to three-dimensional space as in a motion detection system. All these musicians have a deeply rooted physical practice with gestural systems. Their insights are of great value to situate and contextualise various perspectives at the intersection of human–computer interaction and gestural computer music performance.
https://www.gold.ac.uk/computing/people/tanaka-atau/

The Biomuse Instrument

Franziska Baumann: Your Myo wristband uses the EMG signal for musical performances in which the human body becomes a musical instrument through muscle tension recognition. The choice of interface determines the type of interaction. In dance, the motivation is often to create different ways of experiencing the body and space through movement. Here, video or body tracking with markers may be preferred because it converts absolute spatial parameters into data. In contrast, sensors built into a wearable interface act in relation to the body rather than space. That is also the case with your Myo interface.
Why did you choose an interface that is controlled by your muscles?

Atau Tanaka: Yes, you cite the Myo, but that's just the most recent device that allows this. Using muscle in music is a practice. I've been doing it for much longer than the Myo itself exists. And so, for me, it's a question of not what is the interface but what is the instrument.

You are right, if the muscle of the performer's body is what is being captured, this is fundamentally different from motion capture systems because motion capture will detect the body's movement in XYZ space. So, mocap is Cartesian. It catches the movement of the body's limbs, the arms, and the legs as movement is being done. So, motion capture systems report on the result of movement. Muscle sensing is interesting because you need to activate the muscles to make movement. It's not the result of musical action, it's the intention of musical expression. We pick up musical gesture with electromyogram sensing before movement happens.

Did I know all this before choosing this modality? No, this goes back over 30 years to 1990 when I was a PhD student at Stanford University at the Center for Computer Research in Music and Acoustics, CCRMA. In the lab, there was a biomedical researcher and electrical engineer, Hugh Lusted and Ben Knapp, who had created a device called the BioMuse. This interface I came in contact with gave the possibility of biosignal interaction or interfacing the body's physiology with music. This was the early days. Also, the beginning of what today we know is brain-computer interaction because these bioelectrical signals are electrical impulses of the body from the central nervous system neurons in the brain.

Similarly, the muscle cells respond to neuron impulses firing electricity to make muscle tension. The BioMuse allowed reading brainwaves and allowed reading muscles. In the history of

contemporary and modern music, we have Music for Solo Performer by Alvin Lucier from 1965 . He reads brain waves to trigger a series of percussion instruments in the room.

By the time I was thinking about this in 1990, this was the beginning of the early sort of alternate MIDI instruments, various interfaces for computer-human interaction and the first generation of virtual reality. So in my thinking, as a young musicians in those days, given the proposition to make music with the body signals from the body, brain or muscle, I was less interested in the brain and more interested in the muscle. Alvin Lucier's piece is a kind of anti-performance performance, whereas I was a pro performance performer. If you were watching a brain wave piece, as audience you just have to believe what's going on. Whereas if you perform with a muscle system, it is voluntary control - the body makes intentional gesture, and you see what you get. And for me, this was important as a form of the beginning of thinking about embodied interaction in the visceral relationship that we could have with digital computer music systems.

FB: Have you been an instrumentalist?

AT: Yes. All my initial musical training was in classical music. I played the piano throughout my childhood, the classical, romantic repertoire, and post-romantic repertoire. In University, I tried to switch over from classical piano to jazz piano. I was having some trouble switching from the harmonic systems of classical music to the "wrong" Blue Notes of jazz. Then I played electric guitar in a very physical way in the 80s in Boston.

I followed the scene with musicians like Fred Frith and John Zorn, being a fan of that downtown New York scene. I was playing Free Improvised Music. I was coming from classical music, composed and structured and organized, and then I was going towards Free Jazz kind of direct expression. The body is vital on both sides. And so, for me to go into electronic and computer music, the body had to have its place.

For example, I was an assistant for Fred in a gallery installation piece back in the early 90s at the festival Banlieues Bleues in Paris and Musique Action in Nancy.

I remember seeing the Golden Palominos, Naked City, and those bands. The energy and the freedom inspired me. Fred Frith was the first example for me of someone who undid the guitar, and this was very interesting for me.

FB: Do you get any haptic feedback from your Myo?

AT: The question of having feedback is an important one. My first encounters with muscle interaction EMG interaction for music were happening in 1990 at the first wave of virtual reality when Jaron Lanier famously had the company, VPRL. We had the first VR systems when I was in California during my PhD. There was a magazine called Mondo 2000. Wired Magazine was still about to arrive on the scene. So the virtual was something interesting for us. In the commercial music field, we had MIDI controllers, alternate MIDI controllers, MIDI saxophones, and MIDI guitars. And I thought, okay, if there's a VR virtual reality, we will need music for Virtual Worlds and virtual musical instruments. That seemed an interesting concept to me. By having the electromyogram muscle interface, we don't have an object that gives the haptic feedback. We are playing in thin air, like a Theremin. I thought, okay, with this, we can create a virtual musical instrument because it's an invisible musical instrument. I realised that you need some resistance to make muscle tension.

It's different from playing the theremin, or maybe not because we see theremin players are pretty tense. They try to hold their arm in midair to find the correct intonation, the fine-tuning.

It is tricky if you don't have a string to push your finger on. We need haptic feedback. A concept that I developed at that time was called the "boundary object." We need to push against the wall to make muscle tension. In some of my early BioMuse performances, even though I was waving my arms in midair, I used a rubber ball in my hands to give resistance, so my muscle had

something to act on. It's not haptic feedback in a technological sense. But it's haptic feedback in the proprioceptive sense where the object is a boundary. The object gives a passive resistance and provides the body with something to act on.

FB: You could create an instrument with a boundary having various resistance on which the muscles can become reactive.

AT: A couple of years later, I'd gone to Paris to do the last year of my PhD at IRCAM. And after that, I was involved in the multimedia scene, making CD-ROMs and early web pages mid-90s. We started to work with students from the design school, ENSAD in Paris. There was the student Remy Bourganel., who now is quite a well-known designer. During his internship with me, he created some objects and designed objects with which you could play. The object gives the thing upon which we act, and then the body makes the muscle tension play this part in reality.
     Or air guitar, which became one of the popular culture early demos of the system.

FB: A more metaphorical approach can be applied to sensor interfaces because technology as a necessary and creative mediation can be considered a source of ideas rather than merely a means for their transmission. As a result, the mediating artefact or prop comes to the fore. With each specific prop/interface, one chooses an idiosyncratic embodiment, appearance and affordance with a particular type of body image and gestural inscription. The chosen prop and software define the instrument character and the symbolic control as a source of embodied activity.
     A SensorGlove, invisible wireless Wave Rings, Alex Nowitz's strophonion, Pamela Z's sensor board worn directly on the wrist, and your Myo are interfaces that offer the performer a very different character of visual affordance. My SensorGlove, with its cables that reach the fingers to the end of the arm, has almost a vintage feeling. We all know electricity, and that's why the glove is a prop reminiscent of electrical devices. It attracts attention, and people want to know how it works.
     Is the Myo an instrument, a body extension, or a prop? Are these phenomenological characterisations an issue in your work?

The Myo is an object that allows the body to articulate energy.

AT: Well, there are several things in your question: the question of extension, the affordance, the idiosyncratic, and question of the prop. All are interesting to me except the "prop: because it is a word from the theatre.
     I remain a musician. Even if what I do, people see it and say: oh, could you imagine doing a dance performance? But I'm not a dancer. I'm not a choreographer. I'm presenting musical gestures. Prop is a funny word because prop also implies its non-functional. It's a theatre kind of thing, and it stands in for reality sometimes.

FB: I refer to prop because all the acoustic instruments are based on movement energy. Digital Interfaces don't sound themself, and they need to be programmed and mapped to sound and sound modification. Therefore, the way they look says nothing about their sound behaviour.
     Your Myo is not making the sound, and it's not translating the movement energy. However, it is an object that has a design.

AT: The Myo is an object that allows the body to articulate energy. The armband is a very successful piece of industrial design. It's very convenient to wear, but I'm not particularly interested in how it looks. It looks too cool. It looks too easy, but I don't want to make something look difficult either. This is always the question in musical performance. If someone plays something incredibly difficult and at the same time makes it look easy, perhaps we are impressed.

However, if we cause the performance to look complicated than it actually is, we are not impressed.

We build the tools, and the tools build us.

I think technology has a lot of similar problems. There is a lot of imagination about what technology can do. And the companies that make the technology try to sell you the technology by making it look magical and easy. We know that working with technology often comes with crashes. It doesn't do what we want it to do. We struggle with it, and also, it's challenging to do anything creative and fluid with it. So why lie about it? I'm not interested in looks. I don't care if I look like Frankenstein in the biomedical case or a cyborg in the Myo case. The essential thing for me is the notion of affordance. What affordance does the instrument have? What does it offer me musically in terms of gestures?

When we come to this question of extension, then to the idiosyncratic. These instruments are, in a sense, the most natural extension of the body. Both in a phenomenological sense and a media sense, it seems. In Merleau-Ponty's phenomenology, the blind man's cane is an extension, and at the tip of the stick, the blind man sees. Then in Marshall McLuhan's media theory, we build technologies that are the "extensions of man", "We build the tools, and the tools build us." This sometimes brings us to the question you might ask: Is it just an extension? Are they just allowing us to do what we would do anyway? Or do we do something else because we have made this thing. And so Marshall McLuhan's idea that technology as an extension is quite interesting to counterbalance the phenomenological extension of perception.

Music is created because we try to work with the imperfect.

And then comes the question about the idiosyncratic that you asked. Is it just an invisible extension of what we do anyway, or does the instrument take on a character or personality of its own? All musical instruments are idiosyncratic, aren't they? If we were to make a perfect violin, it wouldn't have the shape that it has. It would be made of a different material than the old wood used hundreds of years ago in Cremona, Italy. But a scientifically made violin made of carbon fibre, optimised in terms of technology, does not have an exciting sound. So there is something peculiar about a Stradivarius violin. It is a deep, very magical sound. In addition to this unique character, we compose for these instruments as musicians and composers. It's about understanding the character of the device and composing for it. I think that is a "respect" for the object. We are in technology and engineering. We usually try to create an ideal situation that doesn't usually exist, but we try to solve all the problems and create something perfect. On the other hand, there is no such thing as "perfect" in music. Music is created because we try to work with the imperfect.

I learned characterisation in composition and orchestration classes at the conservatory. We learn that the violin and the flute have a very similar tessitura in orchestration classes. The range is almost the same, but we would never compose for the flute the way we compose for the violin. Because of the arrangement of the fingers and the holes in the flute, a particular melodic passage is very different from the four strings and intervals of a violin. This is what we call idiomatic. To write idiomatic music for each instrument. There is one of my book chapters in the Oxford handbook [Tanaka, A. (2009) "Sensor-Based Musical Instruments and Interactive Music", in Dean, R. (ed.) The Oxford Handbook of Computer Music, Oxford University Press, pp.233-257.] about this. This is quite important for me to go from the extension idea to idiomatic writing.

FB: The Myo Armband is not an instrument. If I would compose mapping strategies for the same interface, it may sound completely different.

AT: Hmm. It's true because it's not by itself. An instrument is it, and that's why to come back to your first question yet. The fact that you use the word interface was then perhaps appropriate, and

it's in that same chapter in the Oxford series where I write about the idiomatic. I write about what is the "instrument"? The interface by itself is not the instrument because in this case, it requires the body. Or it requires the sound synthesis that you're doing.

FB You compose an instrument with its specific ergodynamics in the way you want to play it.
That's the difference with a violin, where the principle stays the same.

The concept of affordance

FB: With an interface, I have to design an intention and how it is played in the sense that there's organology. Sensor interfaces are entirely free from traditional instrument because the physical laws that force a particular gesture to produce a specific sound no longer exist.
Nevertheless, each sensor has a specific characteristic, a way of sensing the real world and recognising physical or chemical properties. For example, with a bend sensor, I can control continuous musical parameters very precisely in contrast the accelerometer has another characteristic and can hardly be assigned to continuous controlling.
Each sensor asks for the corresponding movement to utilise the sensor qualities optimally. To what extent does your sensory system influence the type of gestures?

AT: Here, with this question, we get to the concept of affordance because, as you said, the bender sensor and the accelerometer ask you to make different gestures. So they are inviting a different action. And this is indeed the original definition of affordance in the field of environmental psychology from JJ. Gibson: "What is in the environment that invites the subject to act?" This has been taken up in design practice by people like Don Norman to consciously create affordances.
As musicians, we are somewhere in between that, aren't we? We take technology modules, like different kinds of sensors, and we start to put them together to build instruments. The sensors themselves may have affordances. The bricolage and the assemblages of sensors that we build as instruments will have their meta-affordances, and our bodies perform them and are invited to create different kinds of gestures that are appropriate for the sensors that allow each sensor to operate in a way that best articulates it. I'm avoiding saying "it is designed to do so" because that's engineering. Often as artists, we use the technology in unintended ways.
With electromyogram muscle interaction and its affordance, the body or the human performer is no longer separate from the thing that affords. If we take the original case of Gibson's case of affordance, there is the environment on the one side and the subject on the other. And so we're in a classic kind of duality, whereas in my embodied musical performance, I'm interested in phenomenological experience that challenges dualities. There may not be a subject-object relationship in the traditional structural way.
The way of picking the interface itself may or may not have an affordance. The Myo fits on the arm. But it doesn't fit somewhere else on the body. Yes, it has a certain constraint or limitation that could be seen as an affordance, but the Myo is not representative of all muscle interaction. If we had an electromyogram system with more free cables that allow you to put electrodes on the legs, we could have a different affordance. We are picking up the electricity of the body itself. So it's not the interface that has its affordance in that sense. We can ask: Does can the body has an affordance for itself? This is an interesting open question.
With this question of affordance, I have developed the concept of sonic affordance where we ask, "Can the sound have an affordance?" [Altavilla, A., Caramiaux, B., Tanaka, A. "Towards Gestural Sonic Affordances." In Proc New Interfaces for Musical Expression (NIME). 2013.]

How sound invites a specific gesture

So if we take the Gibsonian approach, sound in the environment can invite the body to have some relationship with that sound. As composers and musicians, can we then take that concept to start to compose sounds that themselves have certain gestural affordances?

FB: So you mean that the sound invites a specific gesture?

AT: Absolutely. That is what I'm working on. The fact that I'm working with muscle, the interaction does change the kinds of sounds I compose as a musician because there are certain kinds of sounds that I find appropriate or that I feel viscerally coherent or somehow meaningful.
    And these sounds are sounds that can follow a particular morphology - to recall Dennis Smalley's notion of spectromorphology. They are continuous sounds that can be sculpted because the kinds of musical gestures that I make almost shape physical gestures.
    In this sense, they get me into a mindset that is going away from the keyboard, whether it's a computer keyboard or the piano keyboard, or going away from button culture, the "trigger culture" of electronic music. Rather than triggering events, I'm interested in shaping sounds.

Mapping Strategies

FB: What qualities or what characteristics have your mapping strategies? You say you shape the sound according to the musical gesture? How can we imagine your gesture findings? Are gestures borrowed from instruments, or do they have so-called meta meanings? For example, for Alex Nowitz, the design of his strophonion has a strong relation to piano playing. For him, the strophonion has an instrument character, and he assigns always the same gesture to the same function. As if in my case, I create pieces, and the assignment of gestures and sound may be different from one piece to another.

AT: It depends on the piece, it depends on the instrument. It depends on the artistic intent.

FB: It depends on how the music invites action.

AT: Yes.

FB: What are the characteristics or qualities of your mapping strategies. Do you also play metaphorical interactions in the sense that you add additional gestures to the technically necessary ones? Or is the gesture always a purely technically necessary gesture?

AT: So the mapping strategies and what gestures are essential are crucial questions because I am originally a traditional instrumentalist. Those references are important to me, but not as references to copy or duplicate but as references that are experientially meaningful. I had enormous satisfaction playing the piano. I had enormous physical satisfaction playing the electric guitar. And what was important for me, working with embodied interfaces for computer music, was to get that similar visceral satisfaction playing digital technologies. And so, that's the gold standard.
    Some instruments may provide a reference. The Theremin historically is the first electronic instrument of embodied interaction. And the fact that one antenna does pitch in the other does amplitude gives a beautiful example.
    I have such a piece for EMG. It's not a simple oscillator because some subtle things are going on. But essentially, it can be understood as "one arm is pitch and the arm is amplitude." And if it sounds simple, I find that the simple case can be communicative and quite powerful, both as a performer experience and, I hope, as a listener experience.
    From there, I get into more complex mappings that may have to do with actual musical material I'm interested in. So, working with granular synthesizers has been a fascinating hybrid for me of working with recorded sounds but not triggering samples. Working with recordings from the

real world to stretch them and to transpose and pitch-shift them where pitch and time are independent, one of the other, and to be able to freeze sound into a single grain and then shape that frozen sound is quite emblematic of what we can do with muscle tension.

The human in the loop of machine learning

In mapping: as we try to get more sophisticated in mapping, we map muscle tension, one source of muscle to many musical parameters, to create a more sophisticated mapping. But it quickly becomes incomprehensible. So the whole system needs to "speak", or we may start to take more channels of muscle input to control a few synthesis parameters. These are the classic modes of mapping that Marcelo Wanderley, and Andrew Hunt laid out at the very beginning of the field of digital musical instruments: The idea of "one-to-many" or "many-to-one" and "many-to-many".

Then, in a multi-dimensional space, no matter whether we're going convergent, divergent, or parallel, the more dimensions of information you have, the more difficult it is to do this by hand. More recently, I've been beginning to use technologies of machine learning. Neural networks that you can teach a mapping system by giving examples. And then to provide the machine learning algorithm some examples: "Here is a position of the body, and here is the sound that I would like to be associated with it." With AI, I can design a sound and design a gesture. By providing a sequence of such examples, we train a neural network, and the neural network based on these examples will then create what we call a regression model. That is a mapping matrix that a human could never create. To start to play these mappings proposed by the machine, to explore them to see if they're musically interesting or not, can become part of the compositional process or, in my case, the performance process on stage. If muscle performance is compelling because it is visible and comprehensible on stage, I would like machine learning too to be the case. Instead of going on stage with an already trained data set to say: "Oh, here's magic. This is AI". In that case people have to believe it. Instead, I go on stage with an empty data set, and the composition is the training of the neural network.

It takes the live music performance case to make an interesting example or exciting case for the science of machine learning. It has to be live and performative. We may provide examples to the algorithm. The algorithm may create a mapping and make it somehow, but we may or may not like it. So we need to be able to interact with the training set to delete, add, edit and have the human in the loop of machine learning training. This is the Interactive Machine Learning paradigm that Rebecca Fiebrink brought to music performance with her Wekinator software.

With Federico Visi we have worked on Reinforcement Learning tools where an artificial agent proposes mappings. If the system proposes a mapping that we don't like, we can give it a reward: yes, or no, "go more like this" or "less like these propositions".

Federico worked here in our lab when he was a researcher. The examples that he put online are a kind of an extesion to the interactive machine learning paradigm that we call Assisted Interactive Machine Learning.

Intensity of Meaning-Making

FB: Finally, I would like to draw attention to the question of the intensity of meaning-making. You can create an instrument and simultaneously decide how it is played. I see it as creating relational intensity in meaning-making human-computer interaction. How can we think about meaning-making in gestural systems?

AT: So this is the question of organology. Thor Magnuson has written a nice book called Sonic Writing. There is an instrument builder called Bert Bongers who used to work at STEIM and is now in Sydney, Australia. He is a key figure because he's not a performer. He's not a musician. He is an interactive instrument builder [https://profiles.uts.edu.au/Bert.Bongers]. There's Norbert Schnell, who used to work at IRCAM and is now in the southwest of Germany near Freiburg, and

he had the concept of "Composed instruments."The question of Intensity is hugely important for me. It is a key word for me in visceral practice. It's more than physical, it's visceral. Intensity, yes, it just can be loud sounds, and we've played with my trio, Sensorband, back in the day on massive sound systems. Maybe it was a kind of rock and roll spirit of the day, but I think it was more than that. It had to do with sound intensity and power as a physical medium of air pressure waves. And this gives the haptic feedback that you asked about. That physicality of sound can then become the feedback channel through the concert space and the public space, which gives haptic feedback. Such intensity can create meaning.

In more recent American phenomenology, philosopher Mark Johnson has written a book, The Meaning of The Body. He thinks about how meaning is made through embodied experience and memory. He gets into body schemata and how our life experiences and bodies start to create meaning based on past experience. In that sense, meaning is not an abstract semantic and semiotic mental construct, but it comes directly out of corporal experience. If we transpose this idea to music-making, I think this might be one answer to your question about the intensity of meaning-making.

Re-connect to the memory with situated action

And I think these experiences are meaningful for us because of the intensity and because they are powerful because they are new.

But not only because they are new but because there is memory somewhere I think. To find that balance is important because intensity by itself will in some moment just be overpowering or if you only have power we get numb eventually or if it's only that, people might become afraid. So getting beyond the numbness of too much power or the fear of too much unknown, somewhere there is the life experience. And this is what Mark Johnson gets into in the body and how it connects. What we encounter is important and I think this then comes back to all music-making, all composition somehow. There is new music, there are new melodies, there are new harmonies or anti harmonies. But if they don't reconnect, within one piece, or within the history of music to our memory of where we're coming from, it's difficult to situate meaning .

FB: This probably also relates to the aspects of clarity and coherence.

AT: In human-computer, interaction research we have the idea of situated action. This is Lucy Suchman. [https://books.google.fr/books/about/Plans_and_Situated_Actions.html]

FB: Thank you so much for your time!