Exploring Musical Agents with Embodied Perspectives

Çağrı Erdem[1,2,3]([✉])

[1] RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion,
University of Oslo, Oslo, Norway
cagrie@uio.no
[2] Department of Informatics, University of Oslo, Oslo, Norway
[3] Department of Musicology, University of Oslo, Oslo, Norway

Abstract. This chapter presents a retrospective of five interactive systems I have developed focusing on how machines can respond to body movement in music performance. In particular, I have been interested in understanding more about how humans and non-human entities can share musical control and agency. First, I give an overview of my musical and aesthetic background in experimental music practice and a less conventional approach to sound and music control. Then follows a presentation of embodiment and music cognition theories that informed the techniques and methods I employed while developing these systems. Then comes the retrospective section structured around five projects. *Biostomp* explores the unintentionality of body signals when used for music interaction. *Vrengt* demonstrates musical possibilities of sonic microinteraction and shared control. *RAW* seeks unconventional control through chaos and automation. *Playing in the "air"* employs deep learning to map muscle exertions to the sound of an "air" instrument. The audiovisual instrument *CAVI* uses generative modeling to automate live sound processing and investigates the varying sense of agency. These projects show how an artistic–scientific approach can diversify artistic repertoires of musical artificial intelligence through embodied cognition.

Keywords: Musical Artificial Intelligence · Multi-Agent Systems · Embodied Cognition · Human-Computer Interaction

1 Introduction

Artificial intelligence (AI) and multi-agent systems (MAS) can already accomplish highly complex musical tasks, such as modeling instrumental acoustics (Damskägg et al., 2019), synthesizing raw audio (Caillon and Esling, 2021), symbolic music generation (Briot et al., 2020), and generating music from text prompts (Agostinelli et al., 2023). However, real-time musical interaction with AI and MAS is still in its infancy. Music performance is a highly embodied phenomenon, and less is known about how machines can perceive humans as

embodied entities and how humans can communicate with machines with multiple modalities. This chapter presents a retrospective of five interactive systems I have developed with these questions in mind and focuses on how machines can respond to body movement. The chapter provides an overview of a multi-year artistic–scientific exploration, its iterative methodology, and how theories and methods from the performing arts, computer science, and music cognition informed each other.

I have been particularly interested in exploring human and non-human entities controlling sound and music together, which I call *shared control*. What are the benefits of shared performance control? Following brief introductions of the key terms, I will begin with an overview of my musical and aesthetic background in experimental music practice. This is important to understand where these projects come from. Next is a presentation of embodiment and music cognition theories that informed the techniques and methods I employed while developing the systems, clarifying the emphasis on "embodied perspectives" and reflecting on the interdisciplinarity of my entwined artistic–scientific research model. The retrospective and discussion of the interactive systems I developed based on five shared control strategies will follow: *Biostomp*, *Vrengt*, *RAW*, *Playing in the "air"*, and *CAVI*. Together, these projects show how applying embodied cognition theories can help diversify artistic repertoires of musical AI and MAS.

1.1 Musical Agents

In the field of New interfaces for musical expression (NIME), it has been common to use a variety of machine learning (ML) techniques for action–sound mappings since the early 1990s (Lee et al., 2021, Jensenius and Lyons, 2017). Over the last decades, there has been a growing interest in researching musical agents within the broader field of artificial intelligence (AI) and music (Miranda, 2021). *Agent* comes from the Latin *agere*, meaning "to do" (Russell, 2010). Essentially, anyone or anything that can act with a purpose can be seen as an agent. For example, an agent's sole task might be to recognize the music's particular rhythm while others track simple musical patterns, such as repeating pitch intervals (Minsky, 1981). Such artificial agents are concerned with tackling musical tasks and are what I call *musical agents*. They are artificial entities that can perceive a human performer through sensors, process that information, and act upon their environment by generating sounds and visuals.

1.2 Embodied Perspective

Musical embodiment is concerned with how the body shapes human musical experiences. For example, the effort a musician and a listener exert often depends on the uncertainty of some musical situations, such as technically challenging tasks. Then, one can use the body to communicate, such as nodding to signal their bandmate to return to the tune's main melody. From an enactive perspective, human perception is shaped by our actions (Schiavio, 2015). The enactivist approach asserts the living body as the cognitive system. In other words, the

regulation and control of cognition as a homeostatic system are determined by its biological structure (Schiavio and Jaegher, 2017). Thus, cognition can be seen as the action Varela et al. (1991, p. 172):

> By using the term *action* we mean to emphasize once again that sensory and motor processes, perception and action, are fundamentally inseparable in lived cognition. Indeed, the two are not merely contingently linked in individuals; they have also evolved together.

Since cognition emerges not just through information processing but mainly from the dynamic interaction between the agent and the environment, the embodied perspective is concerned with an agent's percept of receiving input and processing abilities. More concretely, it questions an agent's ability to perceive the human body and map percept sequences to actions. Although numerous examples of interactive AI and MAS exist in the literature, only a few have dealt with such embodied perspectives.

1.3   Musical Control

In my work, I question the sound and music control—or the lack thereof—in many interactive music systems. As a noise music artist and improviser, my practice focuses on techniques and approaches that foster unconventional expression in music performance. In particular, I have been inspired by John Cage's (1991) exploration of nonintention, which led me to ask how machines could be given more initiative. How can I share the performance control with another musical agent? An analogy can be two persons playing the same guitar, one exciting the string while the other modifying the pitch on the fretboard. Technically, these two entities are agents, regardless of whether they are human or not. If they practice, they can have reasonable control over the system, which, however, can be possible if they lower their expectations of what to expect from their actions. The outcome will always be contingent on the other entity's influx. One may not even be able to make a sound if the other does not allow it. That is inherently different than two agents improvising on their instruments.

2   Artistic Foundation

It is common for experimental musicians to use electronic hardware in unusual ways. Some tutorials, such as that of Collins and Lonergan (2020), teach, for example, how to hack household electrical appliances. Still, shorting a handheld radio's circuit board to make wizard sounds can be considered "wrong" by many people. One such "wrong" instrument that could spark off a niche performance tradition within the experimental music scene is the "no-input mixing board". The principle is the same as creating loops between a speaker and a microphone. It does not require specialized equipment, and any mixing board can be used. Albeit rare examples of meticulously controlled performances with elaborate rigs,

such as Marko Ciciliani's composition *Mask* (2001), no-input mixing is known for its emergent peculiarities (Charrieras and Hochherz, 2016). Performing on a mixer involves sharing musical initiatives with the tool, hence waving the control and being dependent on it. According to Locke (1959), actions are performed in a two-stage temporal sequence. First, possibilities randomly blossom. Then, we choose one action possibility in the next phase: de-liberation. When we act, what was previously out of control is now a determined action. In playing instruments like a no-input mixer, the thought and action processes, hence the decision-making, are distributed between the player and the tool's internal dynamics. Toshimaru Nakamura, states (Paul, 2009):

> You shape the feedback into music. It's very hard to control it. The slightest thing can change the sound. It's unpredictable and uncontrollable, which makes it challenging. It's because of the challenges that I play it. I'm not interested in playing music that has no risk.

The "risk" that Nakamura remarks here implies a preferred uncertainty rooted in a lack of control. That is unconventional in most traditions of playing a musical instrument. Artistically, however, it enables new approaches to performance techniques and music technology innovation.

2.1   Feedback

To better understand the concept of feedback, we can develop an analogy between playing music and driving a boat. The helm of a ship can be seen as analogous to the control interface of a musical instrument, such as a no-input mixer mentioned above. The sea is the electrical current circulating in the components and becoming sound waves through the speakers. As the captain, you shift the steering according to the feedback from the environment concerning waves, winds, and so on. In other words, you continuously evaluate the possibilities, introduce a move, and validate the result before restarting the "loop".

We see such information-feedback paths in all living systems adapting to their environment (Kline, 2015), which can be described as an *autopoietic* organization Maturana and Varela (1980). Poiesis is Greek for "creation" while auto denotes "self". Thus, autopoietic systems consist of self-creating processes (Straussfogel and von Schilling, 2009), which refers to the recursive interactions between the components of living organisms, such as proteins, nucleic acids, lipids, etc. That is a basic understanding of *cybernetics* (Wiener, 1948), which comes from the Greek word *kubernetes*, meaning the helmsman.

The idea of feedback can be traced at least as far back as the beginning of humankind's written record. The first premise of today's rule-based systems is based on the *if...then* condition, which can be found in *modus ponens* of antiquity. Ctesibius' water clock (*clepsydra* c. 250 BC) is considered the first machine to operate under its control. Fast-forward to the 20th century, Nicolas Schöffer created *CYSP 0 & 1* in 1956, human-scale robotic sculptures responsive to changing sound, light, and movement, premiered in a performance with

the Maurice Bejart dance company (Shanken et al., 2012). "We are no longer creating a work; we are creating creation," remarked Schöffer (Whitelaw, 2004), signaling the artistic paradigm shift. John Cage, Eliane Radigue, Steve Reich, and David Rosenboom were some of the composers who incorporated feedback into their music. David Tudor's *Bandoneon! (A Combine)* was one of the first pieces that transformed an entire physical space into a self-oscillating instrument via acoustic feedback loops (Goldman, 2012). A milestone was the *Cybernetic Serendipity* exhibition (1968), which happened with 130 contributors, from composers, artists, and poets, to engineers, scientists, and philosophers (Reichardt, 1968).

2.2 Biofeedback

In cybernetics, a particular topic called *biofeedback* emerged as a medical technique that uses electronic devices to measure the physiological processes (Moss, 1999) in the form of visualization or sonification. In the arts, Alvin Lucier's 1965 piece, *Music for Solo Performer, for enormously amplified brain waves and percussion*, was the first to use electroencephalography (EEG) electrodes on a performer's scalp to capture the alpha rhythm of the brain (typically 8–12 Hz). Following an amplification apparatus created by Edmond Dewan, the amplified alpha rhythms excited the sounding body of percussion instruments (Straebel and Thoben, 2014).

In the following years, several other pieces employed biofeedback techniques, such as John Cage's *Variations V* (1965) (Miller, 2001), David Rosenboom's *Ecology of The Skin* (1970) (Rosenboom, 1972), and Stelarc's *Third Hand* (Dixon, 2019). Eventually, the biofeedback paradigm shifted into a new paradigm of *biocontrol* in the 1990s (Tanaka and Donnarumma, 2018). One of the first pieces here was Atau Tanaka's *Kagami*, featuring *The BioMuse* (Lusted and Knapp, 1988), a "biocontroller" that monitors the electrical activity in the body in the form of both EEG and electromyography (EMG) (Tanaka, 1993). The main difference between biofeedback and biocontrol is that the former focuses on measuring bodily processes regardless of the level of intention or willfulness. At the same time, the latter aims at deliberate control.

2.3 Biocontrol

Easier access to fast computers allowed a widespread interest in using the human body as part of musical instruments at the turn of the 21st century. The Myo sensor was particularly important in making bio signals available to larger groups of people through its wireless 8-channel EMG armband with a built-in inertial measurement unit (IMU). Ata Tanaka's *Myogram* (2015) is a piece composed for two Myo armbands and an octophonic sound system, described as "spatial sound trajectories of neuron spikes projected in the height and depth of the space, with lateral space divided in the symmetry of the body" (Tanaka and Donnarumma, 2018, p. 13).

In addition to bioelectric signals, muscle contractions also produce mechanical vibration, which can be captured as acoustic signals through *mechanomyograms* (MMG) (Caramiaux et al., 2015). Donnarumma (2011) pioneered "biophysical music" using his custom device Xth Sense, which uses an electret microphone-based armband to capture "muscle sounds". Donnarumma describes his experience using such bio-interface as "a relationship of configuration, where specific properties of the performer's body and the instrument are interlaced, reciprocally affecting one another" (Tanaka and Donnarumma, 2018, p. 15).

2.4   Coadaptation

Artist-scholars, such as David Borgo David Borgo Borgo and Kaiser (2010) and Marco Donnarumma (2016) suggest a mutual configuration with the (technological) practice and the environment. The latter actively co-constitutes music with living bodies and their activities. If your microphone faces the speaker too closely on a concert stage, creating audible acoustic feedback, you will most likely be triggered to change the microphone direction spontaneously. This could be seen as similar to reaching out the hands while falling. According to Chi et al. (2000), we execute several physiological and biological processes for a single, deliberate task, most of which are often not deliberate or intentional. In that regard, the biological signals produced by muscles reflect the in-betweenness of the human body's voluntary and autonomic functions.

Over the years, I have performed with several different muscle interfaces. This includes the MMG- and EMG-based devices I have developed myself, as well as various commercial products, such as the consumer-grade Myo armband and the medical-grade Delsys Trigno system (some of these works will be introduced in later sections). My experience is that using muscle signals for precise control is challenging. I agree with Tanaka (2000) describing biosignals as "truly living signals," which reflect the in-betweenness of the human body's voluntary and autonomic functions. The causality flows in one direction when we move toward a specific goal. Simultaneously, the dynamic interaction with the environment bestowing the body can flow back via the body's autonomic responses. In other words, the bodily experience of the environment feeds back into one's actions. Starting from these perspectives, I wanted to explore embodied strategies and approaches for interacting with non-human musical agents in artistic settings.

2.5   Musical AI and MAS

Embodied perspectives are scarce in the literature on (musical) human–computer interaction. Literature reviews of artificial intelligence and multi-agent systems for music, such as those made by Collins (2006) and, more recently, Tatar and Pasquier (2019), highlight that musical AI & MAS prioritize interaction based on symbolic audio (e.g., *M & Jam Factory* by Joel Chadabe and David Zicarelli (Zicarelli, 1987), *Cypher* by Rowe (1992), or *Band-out-of-a-Box* by Thom (2000)); audio (e.g., *Voyager* of Lewis (2000), and (*FILTER*) system of Nort et al. (2013)); or cognitive/affective systems (e.g., *OMax* by Dubnov

and Assayag (2005), or *MASOM* by Tatar and Pasquier (2017)). However, body movement is also integral to musical interaction and a focal point in developing and performing with new interfaces for musical expression. What is relatively underexplored is how musical agents can interact with embodied entities, e.g., humans, other than merely listening to the sounds of their actions. Rare examples include *Robotic Drumming Prosthesis* by Bretan et al. (2016), *RoboJam* by (Martin and Torresen, 2018), the multimodal agent architecture proposed by Camurri and Coglio (1998), and the musical robot swarm of Krzyżaniak (2021).

3   Embodiment

Embodiment in music interaction essentially refers to actions originating in the body (Leman et al., 2018). As such, the body is the prime medium for interaction. *Gesture* is a commonly used term to describe meaning-bearing human actions and has attracted growing attention in music research (Gritten and King, 2006 Godøy and Leman, 2010, Gritten and King, 2011), spanning new musical interactions (Cadoz and Wanderley, 2000, Jensenius et al., 2010, Tanaka, 2011). However, the term gesture is overwhelmingly multifaceted and differently used in the literature (Jensenius, 2014). In the following, I will clarify the term by dividing it into different levels of body movement, for which using a single term—gesture—is confusing (see Jensenius and Erdem (2022) for more details).

3.1   Low Level

Using a bottom-up approach, I start with low-level body movement, which refers to physical phenomena. Such as *force*, a biomechanical phenomenon that sets the object in *motion*, which refers to the physical displacement of the object. Humans and animals generate voluntary and passive muscular forces to process energy while interacting with the environment (Uliam et al., 2012). When playing a musical instrument, all its different parts transmit forces, motion, and energy from one to another. The experience of playing a musical instrument originates in the sum of the material properties of the instrument and the features of interactive human motion. See, for instance, how the upper harmonics vary by alternating the bow pressure (Motl, 2013), or the amplitude modulation (AM) in a vibrato effect (Dromey et al., 2009). Physical phenomena like force and motion and their variations' influence on the resultant sound can be objectively measured via several motion capture technologies (Jensenius, 2018).

3.2   Middle Level

Differently from force and motion, (embodied) *actions* denote intentionally executed motion fragments, which are subjective phenomena. Godøy and Leman (2010) refer to "cognitive units" to describe such *chunking* of continuous motion and force. Thus, one can think of the action as mental imagery (Godøy, 2009a). As long as an action is not communicated intentionally, it does not necessarily
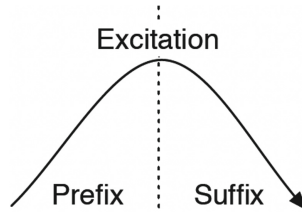
Fig. 1. An action, such as hitting a guitar string, is realized through an excitation phase, which incorporates a prefix and a suffix (Jensenius, 2007, p. 24).

bear a meaning. Hence, I place it in the middle level, between low-level physical signals and high-level communicative actions. Since this middle level is subjective, it is impossible to precisely define, for example, the start and endpoints of an action. Consider the case of hitting a guitar string once. As Godøy (2009b) suggests, the attack has an *excitation phase* having a *prefix* (lifting the arm) and *suffix* (moving down) as illustrated in Fig. 1. *Fidgeting* are the motion parts not directed by a goal nor intentional or conscious.

Since motion and sound are temporal phenomena, we perceive different features in different timescales (Godøy, 2009a). That is a necessity of our cognitive apparatus, for example, in chunking the action segments. Godøy suggests a three-level grouping:

- *Sub-chunk level*: The *micro* timescale for pitch, loudness and timbral features (<0.5 s)
- *Chunk level*: The *meso* timescale as well as the timescale for sound-producing actions (0.5–5 s)—short-term memory
- *Supra-chunk level*: The *macro* timescale for longer contexts (>5 s)—long-term memory

There are many types of music-related body motion (see Jensenius et al. (2010), for an overview), but in the following, I will primarily focus on *sound-producing* actions. Cadoz (1988) suggested that these can be subdivided into *excitation* actions, such as right-hand guitar fingering, and *modification* actions, such as left-hand pitch modifications. As depicted in Fig. 2, excitation actions can be divided further into the three main categories proposed by Schaeffer (1966) and presented by Godøy (2006):

- *Impulsive*: A fast attack resulting from a discontinuous energy transfer (e.g., percussion or plucked instruments).
- *Sustained*: A more gradual onset and continuously evolving sound due to a continuous energy transfer (e.g., bowed instruments).
- *Iterative*: Successive attacks resulting from a series of discontinuous energy transfers.

Identifying the excitation phase can be relatively straightforward when dealing with a single impulsive action but becomes highly complex when combining
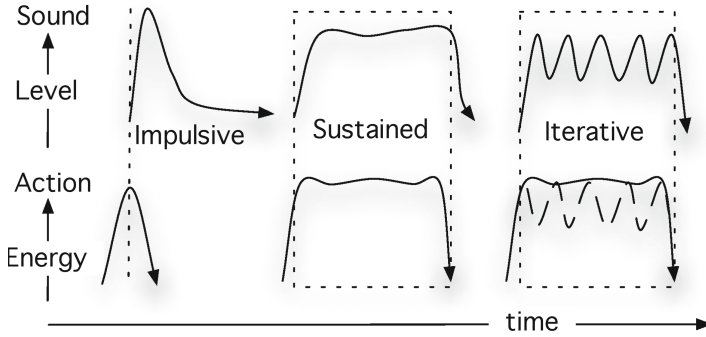
Fig. 2. An illustration of three categories for the main action and sound energy envelopes resulting from different sound-producing action types (Jensenius, 2007, p. 26). The dotted lines correspond to the duration of contact during the excitation phase.

multiple actions. Such action series can be seen as a form of *coarticulation*, the merging of individual actions into larger shapes (Godøy, 2013). Analyzing such action shapes can be challenging from an empirical point of view, particularly segmentation of motion capture recordings for motion–sound analysis.

3.3   High Level

Gestures are actions with an associated *high-level* communicative meaning. The meaning-bearing aspect of gestures has been studied in linguistics: "Gestures exhibit images that cannot always be expressed in speech [...] With these kinds of gestures, people unwittingly display their inner thoughts" according to McNeill (1992, p. 12), emphasizing that bodily gestures are essential to communication.

In music, the word gesture is often used synonymously with both motion and action. However, the challenge is to define the *musical gesture* in a way that covers both motion-related definitions and sonic properties, such as the sound shapes presented by Smalley (1997). The threefold grouping presented in this section provides an embodied perspective on such different levels and definitions of musical gesture.

4   Retrospective

In this section, I present an overview of some of my interactive music systems:

1. *Biostomp*: a muscle-based motorized audio effects controller that explores the boundaries between control and the lack thereof (Erdem et al., 2017)
2. *Vrengt*: an interactive dance piece in which two performers share the control of the system (Erdem et al., 2019)

3. *RAW*: a muscle-based instrument exploring a chaotic behavior in control and automatized ensemble interaction (Erdem and Jensenius, 2020)
4. *Playing in the "air"*: a predictive action–sound model using deep learning based on a custom dataset collected throughout a series of laboratory experiments (Erdem et al., 2020)
5. *CAVI*: an agent-based interactive system using a generative model trained on the data collected in the previous study (Erdem et al., 2022)

Since each system has been described elsewhere, I will breeze through their implementations and focus on details about control structures and sonic design.

4.1   Biostomp

Biostomp is an interface that lets the performer use muscle contractions to control audio effects parameters in live performance situations (a video playlist is available at https://youtu.be/cgnns9z-Nl4). Unlike wearable integrated motion units (IMUs) that measure three-dimensional motion, muscle contractions do not always happen intentionally, which is typical of most biological processes. That can be challenging when using muscles for control. On the other hand, biological idiosyncrasies can also be used creatively in music, similar to how musicians benefit from nature's indeterminacy (Borgo, 2005, Cantrell, 2007).

Biostomp relies on the mechanomyogram (MMG), which denotes low-frequency mechano-acoustic signals generated by contractions in muscle fibers (Watakabe et al., 2001). MMG is the signal resulting from contracting a muscle and can be captured via electret condenser microphones worn on the body part, such as limbs, in the case of Biostomp. When recording audio signals from "inside" of the body, these recordings include multiple bodily "sounds," such as blood flow and heart rate.

Direct transmission of biologically-occurring muscle signals was the primary design consideration for *Biostomp*. It was designed as a self-contained system and avoided any complex mapping and sound design. Instead, it is based on a one-to-one mapping between the MMG amplitude and a motorized headpiece designed to be hooked on potentiometers. The performer then decides which audio effects to control.

The variety of playing modes of *Biostomp* depends on the effects type and the variable signal intensity ("predictability"). In the user study, I observed how different users reacted to different combinations of control and effects. For example, there is a drastic difference in controllability between dynamic (e.g., overdrive) and time-based (e.g., delay) effects. Several users were positive about the system's surprising and less controllable aspects. Nevertheless, most reported that it became more predictable after practicing for some time, which may or may not be favorable. I will return to this aspect later since predictability and user reactions are fundamental commonalities among the five interactive systems being presented.

4.2   Vrengt

Vrengt (Norwegian for "inside-out") is an interactive system that allows a dancer and a musician to control the same sound and music parameters in the interactive system (a video teaser is available at https://youtu.be/vXJ0l9Q68nc). It was designed through a recursive process: capturing and sonifying the dancer's (micro)motion and the shared control of the sonification parameters, which, in turn, affected the dancer's motion. The idea was to work on *sonic microinteraction*, an interaction mode common in acoustic instruments but rarely found in interactive systems (Jensenius, 2017).

In *Vrengt*, I used muscle sensing through electromyograms (EMG), the signal that puts the muscles in motion. Tanaka (2015) describes EMG as capturing the intention to move. It is a bioelectric signal that captures human micromotion indirectly as this level of interaction does not always result in overt body movements (Tanaka, 2015, Jensenius et al., 2017). EMG often reports small or non-visible motion akin to consciously executed actions and automatic body processes (Ortiz et al., 2011). As for the specific sensor device, we chose to work with the (at the time) commercially available Myo armband.

The second interaction method employed in *Vrengt* was capturing the dancer's breathing through a wireless audio signal. Breathing is fascinating in that it is mostly involuntary and unconscious but can also be voluntary and conscious. We preferred using audio over a wireless headset microphone so that the dancer could create acoustic feedback loops by changing her proximity to the speakers on the stage. In doing so, breathing was also used as an aesthetic element. Since the dancer's position on stage influenced the produced sound, the physical space became an integral part of the performance. This was particularly effective in the piece's opening when the dancer was blindfolded for artistic purposes. Then, she had to rely on the auditory feedback from the system to orient herself.

Sonification was a core method used in the sound design of *Vrengt*, which gave the dancer a direct and immediate sonic response. Sonification is often seen as an objective approach to representing data through sound (Hermann and Hunt, 2011). However, in our context, sonification was not the end goal. Instead, we used sonification as part of the creative process.

Drawing on our perceptual and cognitive capacity regarding the link between sounds and sound resources, what Godøy (2001) describe as *mental imagery*, we focused on two techniques in the sound design: (1) Physics-based synthesis of everyday sounds and (2) abstract techniques. In doing so, we could also explore the dancer's sensations concerning the sound synthesis techniques' sonic imagery and mappings. As for abstract techniques, we explored waveshape distortion, ring modulation, and exponential frequency modulation. According to the dancer, while physics-based sounds evoked more straightforward imagery, abstract techniques for sound synthesis resembled shapes that she could "fill with any image you want".

We decided to work with fixed mappings in Vrengt. This was decided early on to accommodate that two human performers would share the control of the

system. The dancer's incoming sensor and audio data were processed and interpreted in real-time by the musician, who used knobs and faders on a MIDI controller (Fig. 3). This way, both performers could experience the other's agency. Both performers perceived this as inspiring and fuelled further implementation of artificial agents.
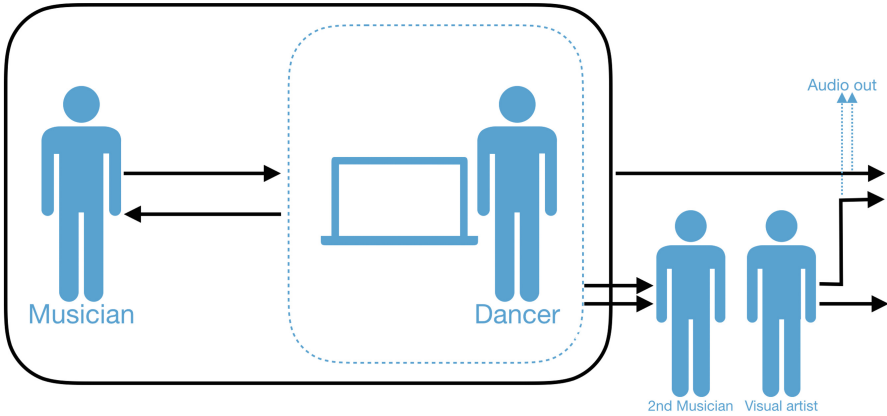

Fig. 3. The setup for the final collaborative performance, showing the levels of connection between performers and instruments (Erdem et al., 2019).

4.3   RAW

The name of *RAW* comes from the system's primary distinctive property, using raw bioelectric muscle signals (EMG) at the audio rate (a video teaser is available at https://youtu.be/_--dzA5pl9k). This was inspired by *Myogram* by Tanaka (Tanaka and Donnarumma, 2018), which uses a direct audification of EMG signals. *RAW* uses two Myo armbands, one on each forearm. Four EMG channels (two per forearm) are buffered every quarter of a second, which is then converted to an audible level by increasing the frequency via a time-scaled sawtooth signal. In doing so, the inherent noise of the raw signal is also frequency-shifted, thus creating a quite noisy high-frequency layer in the audible spectra, requiring filtering. This is where the performer can start being creative as a composer. For example, speeding up the signal to extreme values introduces glitches reminding of well-known electronic music textures, similar to those of Ryoji Ikeda (Emmerson and Landy, 2016).

Two channels of EMG per forearm are sonified, corresponding to extensor and flexor muscle groups. This provides four drone sound channels, controlled by each wrist's extension and flexion. Other poses, such as ulnar or radial deviation, open or closed hands, and neutral poses, create different combinations. One can imagine such a scenario as mixing four audio channels using faders on a mixing

board. This approach can be awe-inspiring, but requiring a multi-channel sound system limits its applicability in different ensemble settings. Therefore, I explored several algorithmic approaches for generating control signals.

In the control part of *RAW*, I used multiple feature extractors simultaneously. First, amplitude envelopes were extracted as the continuous EMG signal's root mean square (RMS). For more precise actions, such as event triggering, I used the IMU data, particularly the *jerk*, the rate of change of the acceleration. In air performance, where the performer can move in any direction, the relativity of jerk-based excitation may not always be favorable. Therefore, I trained a support vector machine (SVM) classifier to recognize pinch grips, which I use for triggering purposes. Such gesture recognition helps when performing based on muscle signals for more precision-requiring actions.

A second control part was based on chaotic attractors, such as Hénon-and-Heiles or Lorenz systems, to create melodic motives. The EMG was pitch-shifted at the audio sample rate using additional oscillators. When using a pinch grip, the SVM model can recognize and draw a new set of points on the orbit, where each point refers to a frequency. Although the new frequency may sound random compared to the previous one, it converges into a melodic line. In practice, that does not always work as expected. For example, if the interval between two points is too long, it never converges to a globally familiar pattern. However, the interval can become too repetitive if it is too short.

A third control part was based on two multi-layer perceptron (MLP) artificial neural networks (ANNs). They can be used both in pre-trained mode or in online training mode. The networks were used with a simple gamification strategy. Each ANN mapped eight EMG channels of one armband to a point in an XY plane, of which both axes were mapped into an oscillator parameter. The goal of the "game" is to make two points meet so that a new random event is triggered. As a performer, this is one of the fascinating features of the system.

RAW is based on real-time audio analyses for automated ensemble interaction. Real-time audio analysis is challenging at many levels, particularly in free improvisation settings. The solution was to use an adaptive algorithm and limit the system's scope to rhythm-related tracking using mainly spectral flux and dynamics-tracking using envelope-following. The system also incorporates an effects outboard with a selection of time-based processing modules. These can be employed for live sound processing, producing highly efficient duo performance results. However, in bigger ensembles, such processing can introduce too much ambiguity.

4.4   Playing in the "Air"

Later versions of *RAW* inspired a new project on guitar *ergomimesis*. Magnusson (2019, p. 36) suggests this term for mimicking the *ergon*, Greek for work or function. Thus, ergomimesis denotes carrying out the function and the incorporated working memory, *ergogenetic* memory, from one context or domain into another. I began from an "air guitar" perspective, although the aim was never to mimic the guitar in the air. Instead, I wanted to employ the embodied knowledge of

playing the guitar and use these possibilities and constraints in constructing a new instrument.

The first part of the project involved a controlled experiment in a laboratory context. A total of 36 participants performed tasks based on guitar-like versions of each of the three basic sound-producing action types proposed by Schaeffer (1966): impulsive, iterative, and sustained. Analyses of the motion capture, EMG, and sound data from the experiment showed explicit action–sound correspondences compatible with theories of embodied music cognition (Erdem et al., 2020, p. 15).

Following the empirical exploration of how biomechanical energy transforms into sound, we used these transformations as part of a machine learning framework based on Long Short-Term Memory (LSTM) networks and compared nine model configurations. The aim was to determine how much latency these models would be subject to when used as a musical instrument (Erdem et al., 2020, p. 30). Our results showed that the models could predict audio energy features of free improvisations on the guitar, relying on an EMG dataset of three distinct motion types (a video is available at http://bit.ly/air_guitar_smc). Our modeling approach provided empirical support for the embodied music cognition theory.

4.5   CAVI

The inspiration for *CAVI* came from the concepts of emergent coordination (Knoblich et al., 2011), collaborative emergence (Sawyer and DeZutter, 2009), and temporal (un)predictability (Haggard et al., 2002). Following the considerable latency of the trained models, I focused on generative modeling. Instead of a discriminative supervised model, I used a recurrent neural network (RNN) combined with a mixture density network (MDN) layer (Bishop, 1994). This MDRNN model continuously tracked the data streamed from a Myo armband worn on the right forearm of the performer and generated new electromyogram (EMG) and acceleration data.

One interesting question is whether coordination or joint action can emerge between a performer and a musical agent that somewhat simulates the performer's likely actions using generative predictions. To explore that, *CAVI* continuously tracks the performer's motion input, consisting of 4-channel EMG and 3-channel ACC signals, and generates what will likely come next. In brief, *CAVI* generates control signals solely based on the performer's excitation actions. The generated data were used as control signals mapped to digital audio effects module parameters. This could be seen as playing the electric guitar through some effects pedals while someone else is tweaking the knobs of the devices.

CAVI's effects modules rely on time-based sound manipulation, such as delay, time-stretch, and stutter. The jerk of the generated acceleration data triggers the sequencer steps, functioning as a matrix that routes the effects and sends & returns. The generated EMG data (corresponding to the same flexion and extension muscle groups similar to previous projects) is mapped to effects parameters. The real-time analysis modules track the musician's dry audio input and adjust

the parameters according to pre-defined thresholds. These machine listening agents include trackers of onsets and spectral flux. For example, if the performer plays impulsive notes, *CAVI* increases the reverb time drastically, becoming a drone-like continuous sound. If the performer plays loudly, the system decides about its dynamics based on the particular action type of the performer (A video is available at https://youtu.be/kmYEEEnjm0s).

CAVI is an audiovisual instrument not only for aesthetic reasons but also to avoid potential causality ambiguities. The design presents CAVI as an uncompleted, creepy but cute creature with only legs that are too small for its body, no arms, a tiny mouth, and a big eye (Fig. 4). In real-time animation, the body contracts but does not make full-body gestures. Instead, the eye blinks from time to time when *CAVI* triggers a new event, opens wide when the density of low frequencies increases or stays calm according to the overall energy levels of sound.

Fig. 4. A still image from the performance piece "Me & My Musical AI 'Toddler'", recorded for the online NIME 2022 conference. The performance setup comprised the author, *CAVI*, and, in addition, six self-playing guitars (Photo: Adrian Axel).

5   Discussion

From playing acoustic instruments to performing with computers, my journey illuminated a gap: the intimate, embodied experience of the former seemed absent in the latter. The intrigue around translating the sensation of effort—an inherent yet elusive aspect of human experience—to computational systems drove me to explore embodied music cognition theories. Rolf Inge Godøy's

decades-long work on shape cognition (Godøy, 2019) grounded my approach, enabling systematic analysis and fostering innovation in music technologies.

Employing muscle sensing as a motion capture method revealed the intriguing complement that motion-based interfaces could bring to existing interaction paradigms. While biological processes might be challenging for direct control due to their involuntary nature, their unpredictability can be harnessed for improvisational musicking.

My work then expanded on the concept of "air performance", where, unlike acoustic instruments, there is no tangible feedback. Explorations into Godøy's gestural-sonic objects and his idea of chunking on varying timescales informed my work's evolution from biofeedback to biocontrol. These ideas and conceptions inspired me to think and design in terms of dynamic sound shapes. For example, *RAW* is heavily based on responding to a sustained chunk with an impulsive action. Similarly, mental imagery became instrumental in *Vrengt*, where sonic design and dance interplayed through metaphoric mappings. Mental imagery can serve as a shared language, bridging the communication gap between musicians and dancers.

The culmination of these investigations led to the development of systems for coadaptation. By embracing biological unpredictability, I aimed for shared control structures rooted in the embodied human experience. This was not about using machines as tools but promoting more initiative in musical interactions, adapting mutually, and shifting the narrative.

While much has been achieved, the journey is ongoing. As an artist–researcher, I stand at the confluence of embodiment, artificial intelligence, and multi-agent systems. The challenge ahead is not merely about integrating human complexity with machines but envisioning a harmonious coexistence and diversifying the known ways of musicking. As we continue to develop human-in-the-loop technologies, there are many unanswered questions: How do we strike a balance between the urge to take over musical control and the serendipity in waving it? How can we employ our communicative skills and human understanding in musical human–machine interactions? How do we ensure that as we innovate, we foster creativity and expression? I will aim to answer some of these in the years to come.

References

Agostinelli, A., et al.: MusicLM: Generating music from text. arXiv:2301.11325 [cs, eess] (2023)

Bishop, C.M.: Mixture density networks. Technical report NCRG/97/004, Neural Computing Research Group, Aston University, Birmingham. Aston University (1994)

Borgo, D.: Rivers of consciousness: the nonlinear dynamics of free jazz. In: Jazz Research Proceedings Yearbook, pp. 46–58 (2005)

Borgo, D., Kaiser, J.: Configurin(g) KaiBorg: interactivity, ideology, and agency in electro-acoustic improvised music (2010)

Bretan, M., Gopinath, D., Mullins, P., Weinberg, G.: A robotic prosthesis for an amputee drummer. arXiv:1612.04391 (2016)

Briot, J.-P., Hadjeres, G., Pachet, F.-D.: Deep Learning Techniques for Music Generation. Computational Synthesis and Creative Systems, Springer, Cham (2020). https://doi.org/10.1007/978-3-319-70163-9

Cadoz, C.: Instrumental gesture and musical composition. In: ICMC 1988 - International Computer Music Conference, Cologne, Germany, pp. 1–12 (1988)

Cadoz, C., Wanderley, M.M.: Gesture - music. In: Wanderley, M., Marc Battier, I.-C.P. (eds.) Trends in Gestural Control of Music (2000)

Cage, J.: An autobiographical statement. Southwest Rev. **76**(1), 59–76 (1991)

Caillon, A., Esling, P.: RAVE: a variational autoencoder for fast and high-quality neural audio synthesis. arXiv:2111.05011 [cs, eess] (2021)

Camurri, A., Coglio, A.: An architecture for emotional agents. IEEE Multimedia **5**(4), 24–33 (1998)

Cantrell, M.: Enactive reading: John Cage, chance, and poethical experience. Genre **40**(1–2), 131–156 (2007)

Caramiaux, B., Donnarumma, M., Tanaka, A.: Understanding gesture expressivity through muscle sensing. ACM Trans. Comput.-Hum. Interact. **21**(6), 1–26 (2015)

Charrieras, D., Hochherz, O.: Chasing after the mixer, pp. 253–254 (2016)

Chi, D., Costa, M., Zhao, L., Badler, N.: The EMOTE model for effort and shape. In: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 2000, pp. 173–182. ACM Press/Addison-Wesley Publishing Co., USA (2000)

Collins, N., Lonergan, S. (eds.): Handmade Electronic Music: The Art of Hardware Hacking, 3rd edn. Routledge, New York (2020)

Collins, N.M.: Towards autonomous agents for live computer music: realtime machine listening and interactive music systems. Ph.D. thesis, University of Cambridge (2006)

Damskägg, E., Juvela, L., Thuillier, E., Välimäki, V.: Deep learning for tube amplifier emulation. In: ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 471–475 (2019)

Dixon, S.: Cybernetic-existentialism in performance art. Leonardo **52**(3), 247–254 (2019)

Donnarumma, M.: Xth sense: researching muscle sounds for an experimental paradigm of musical performance. p. 9 (2011)

Donnarumma, M.: Configuring corporeality: performing bodies, vibrations and new musical instruments. Ph.D. thesis, Goldsmiths, University of London, London, UK (2016)

Dromey, C., Reese, L., Hopkin, J.A.: Laryngeal-level amplitude modulation in vibrato. J. Voice **23**(2), 156–163 (2009)

Dubnov, S., Assayag, G.: Improvisation planning and jam session design using concepts of sequence variation and flow experience. Zenodo, Salerno (2005)

Emmerson, S., Landy, L.: The analysis of electroacoustic music: the differing needs of its genres and categories. In: Landy, L., Emmerson, S. (eds.) Expanding the Horizon of Electroacoustic Music Analysis, pp. 8–28. Cambridge University Press, Cambridge (2016)

Erdem, C., Camci, A., Forbes, A.: Biostomp: a biocontrol system for embodied performance using mechanomyography. In: Proceedings of the International Conference on New Interfaces for Musical Expression, pp. 65–70. Zenodo, Copenhagen (2017)

Erdem, C., Jensenius, A.R.: RAW: exploring control structures for muscle-based interaction in collective improvisation. In: Proceedings of the International Conference on New Interfaces for Musical Expression, pp. 477–482. Zenodo, Birmingham (2020)

Erdem, C., Lan, Q., Jensenius, A.R.: Exploring relationships between effort, motion, and sound in new musical instruments. Hum. Technol.: Interdisc. J. Hum. ICT Environ. **16**(3), 310–347 (2020)

Erdem, C., Schia, K.H., Jensenius, A.R.: Vrengt: a shared body-machine instrument for music-dance performance. In: Proceedings of the International Conference on New Interfaces for Musical Expression, pp. 186–191. Zenodo, Porto Alegre (2019)

Erdem, C., Wallace, B., Jensenius, A.R.: CAVI: a coadaptive audiovisual instrument-composition. PubPub (2022)

Godøy, R.I.: Imagined action, excitation, and resonance. In: Musical Imagery, vol. 13, pp. 237–250. Swets & Zeitlinger Publ., Lisse (2001)

Godøy, R.I.: Gestural-sonorous objects: embodied extensions of Schaeffer's conceptual apparatus. Organ. Sound **11**(2), 149–157 (2006)

Godøy, R.I.: Chunking sound for musical analysis. In: Ystad, S., Kronland-Martinet, R., Jensen, K. (eds.) CMMR 2008. LNCS, vol. 5493, pp. 67–80. Springer, Heidelberg (2009a). https://doi.org/10.1007/978-3-642-02518-1_4

Godøy, R.I.: Gestural affordances of musical sound. In: Musical Gestures, pp. 1–23. Routledge (2009b)

Godøy, R.I.: Thinking sound and body-motion shapes in music: public peer review of "gesture and the sonic event in Karnatak music" by Lara Pearson. Empir. Musicol. Rev. **8**(1), 15–18 (2013)

Godøy, R.I.: Musical shape cognition. In: Grimshaw-Aagaard, M., Walther-Hansen, M., Knakkergaard, M. (eds.) The Oxford Handbook of Sound and Imagination, vol. 2. Oxford University Press, Oxford (2019)

Godøy, R.I., Leman, M.: Musical Gestures: Sound, Movement, and Meaning. Routledge, London (2010)

Goldman, J.: The buttons on Pandora's box: David Tudor and the bandoneon. Am. Music. **30**(1), 30–60 (2012)

Gritten, A., King, E.: Music and Gesture. Ashgate, Aldershot (2006)

Gritten, A., King, E. (eds.): New Perspectives on Music and Gesture. SEMPRE Studies in The Psychology of Music. Ashgate, Routledge, Farnham (2011)

Haggard, P., Clark, S., Kalogeras, J.: Voluntary action and conscious awareness. Nat. Neurosci. **5**(4), 382–385 (2002)

Hermann, T., Hunt, A.: Interactive sonification. In: Hermann, T., Hunt, A., Neuhoff, J.G. (eds.) The Sonification Handbook, pp. 273–298. Logos Verlag, Berlin (2011). OCLC: ocn771999159

Jensenius, A.R.: Action-sound: developing methods and tools to study music-related body movement (2007)

Jensenius, A.R.: To gesture or not? An analysis of terminology in NIME proceedings 2001–2013. In: Proceedings of the International Conference on New Interfaces for Musical Expression. Zenodo, London (2014)

Jensenius, A.R.: Sonic Microinteraction in "the air". In: Lesaffre, M., Maes, P.-J., Leman, M. (eds.) The Routledge Companion to Embodied Music Interaction, 1st edn., pp. 429–437. Routledge, New York, London (2017)

Jensenius, A.R.: Methods for studying music-related body motion. In: Bader, R. (ed.) Springer Handbook of Systematic Musicology. SH, pp. 805–818. Springer, Heidelberg (2018). https://doi.org/10.1007/978-3-662-55004-5_38

Jensenius, A.R., Erdem, C.: Gestures in ensemble performance. In: Timmers, R., Bailes, F., Daffern, H. (eds.) Together in Music: Coordination, Expression, Participation, pp. 109–118. Oxford University Press, Oxford (2022)

Jensenius, A.R., Lyons, M.J. (eds.): A NIME Reader: Fifteen Years of New Interfaces for Musical Expression. Current Research in Systematic Musicology, Springer, Cham (2017). https://doi.org/10.1007/978-3-319-47214-0

Jensenius, A.R., Sanchez, V.G., Zelechowska, A., Bjerkestrand, K.A.V.: Exploring the Myo controller for sonic microinteraction. In: Proceedings of the International Conference on New Interfaces for Musical Expression New Interfaces for Musical Expression, pp. 442–445. Zenodo, Porto Alegre (2019)

Jensenius, A.R., Wanderley, M.M., Godøy, R.I., Leman, M.: Musical Gestures: concepts and methods in research, pp. 12–35 (2010). 978-0-415-99887-1

Kline, R.R.: The cybernetics moment: or why we call our age the information age (2015). OCLC: 890127838

Knoblich, G., Butterfill, S., Sebanz, N.: Psychological research on joint action. In: Psychology of Learning and Motivation, vol. 54, pp. 59–101. Elsevier (2011)

Krzyżaniak, M.: Musical robot swarms, timing, and equilibria. J. New Music Res. **50**(3), 279–297 (2021). https://doi.org/10.1080/09298215.2021.1910313

Lee, M., Freed, A., Wessel, D.: Real-time neural network processing of gestural and acoustic signals, pp. 277–280. International Computer Music Association, Montreal (1991)

Leman, M., Maes, P.-J., Nijs, L., Van Dyck, E.: What is embodied music cognition? In: Bader, R. (ed.) Springer Handbook of Systematic Musicology. SH, pp. 747–760. Springer, Heidelberg (2018). https://doi.org/10.1007/978-3-662-55004-5_34

Lewis, G.E.: Too many notes: computers, complexity and culture in voyager. Leonardo Music J. **10**(1), 33–39 (2000)

Locke, J.: An essay concerning human understanding (1959). OCLC: 371280

Lusted, H.S., Knapp, R.B.: Biomuse: musical performance generated by human bioelectric signals. J. Acoust. Soc. Am. **84**(S1), S179–S179 (1988)

Magnusson, T.: Sonic Writing: Technologies of Material, Symbolic, and Signal Inscriptions (2019). OCLC: 1023599945

Martin, C.P., Torresen, J.: RoboJam: a musical mixture density network for collaborative touchscreen interaction. In: Liapis, A., Romero Cardalda, J.J., Ekárt, A. (eds.) EvoMUSART 2018. LNCS, vol. 10783, pp. 161–176. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-77583-8_11

Maturana, H.R., Varela, F.J.: Cognitive function in general. In: Maturana, H.R., Varela, F.J. (eds.) Autopoiesis and Cognition: The Realization of the Living. Boston Studies in the Philosophy and History of Science, vol. 42, pp. 8–14. Springer, Dordrecht (1980)

McNeill, D.: Hand and Mind: What Gestures Reveal About Thought. University of Chicago Press, Chicago (1992)

Miller, L.E.: Cage, Cunningham, and collaborators: the odyssey of variations V. Musical Q. **85**(3), 545–567 (2001)

Minsky, M.: Music, mind, and meaning. Comput. Music. J. **5**(3), 28–44 (1981)

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

Moss, D. (ed.): Humanistic and Transpersonal Psychology: A Historical and Biographical Sourcebook. Humanistic and Transpersonal Psychology: A Historical and Biographical Sourcebook, pp. pp. xxi, 457. Greenwood Press/Greenwood Publishing Group, Westport (1999)

Motl, K.R.: Multiphonics on the double bass: an investigation on the development and use of multiphonics on the double bass in contemporary music. Ph.D. thesis, University of California San Diego, San Diego (2013)

Nort, D.V., Oliveros, P., Braasch, J.: Electro/acoustic improvisation and deeply listening machines. J. New Music Res. **42**(4), 303–324 (2013)

Ortiz, M., Coghlan, N., Jaimovich, J., Knapp, R.B.: Biosignal-Driven Art: Beyond Biofeedback. CMMAS (2011). Accepted 2017-11-29T13:51:16Z

Paul, E.: Subsonics - Episode 4 (2009). https://vimeo.com/3799720

Reichardt, J.: Cybernetic serendipity-getting rid of preconceptions. Studio Int. **176**(905), 176–77 (1968)

Rosenboom, D.: Method for producing sounds or light flashes with alpha brain waves for artistic purposes. Leonardo **5**(2), 141–145 (1972)

Rowe, R.: Machine listening and composing with cypher. Comput. Music. J. **16**(1), 43–63 (1992)

Russell, S.J.S.J.: Artificial Intelligence: A Modern Approach, 3rd edn. Prentice Hall, Upper Saddle River (2010)

Sawyer, R.K., DeZutter, S.: Distributed creativity: how collective creations emerge from collaboration. Psychol. Aesthet. Creat. Arts **3**(2), 81–92 (2009)

Schaeffer, P.: Traite des objets musicaux: essai interdisciplines. Éditions du Seuil, Paris (1966). OCLC: 301664906

Schiavio, A.: Action, enaction, inter(en)action. Empir. Musicol. Rev. **9**(3–4), 254–262 (2015)

Schiavio, A., Jaegher, H.D.: Participatory sense-making in joint musical practice. In: The Routledge Companion to Embodied Music Interaction, pp. 1–9. Routledge (2017)

Shanken, E., Clarke, B., Henderson, L.: Cybernetics and Art: Cultural Convergence in the 1960s (2012)

Smalley, D.: Spectromorphology: explaining sound-shapes. Organ. Sound **2**(2), 107–126 (1997)

Straebel, V., Thoben, W.: Alvin Lucier's music for solo performer: experimental music beyond sonification. Organ. Sound **19**(1), 17–29 (2014)

Straussfogel, D., von Schilling, C.: Systems theory. In: Kitchin, R., Thrift, N. (eds.) International Encyclopedia of Human Geography, pp. 151–158. Elsevier, Oxford (2009)

Tanaka, A.: Musical technical issues in using interactive instrument technology with application to the BioMuse. In: Proceedings of the International Computer Music Conference (1993)

Tanaka, A.: Musical performance practice on sensor-based instruments. Trends Gest. Control Music **13**(389–405), 284 (2000)

Tanaka, A.: Sensor-based musical instruments and interactive music. In: The Oxford Handbook of Computer Music, roger t. dean edn. Oxford University Press (2011)

Tanaka, A.: Intention, effort, and restraint: the EMG in musical performance. Leonardo **48**(3), 298–299 (2015)

Tanaka, A., Donnarumma, M.: The body as musical instrument. In: The Oxford Handbook of Music and the Body (2018)

Tatar, K., Pasquier, P.: MASOM: a musical agent architecture based on self-organizing maps, affective computing, and variable markov models, p. 8. MuMe, Atlanta (2017)

Tatar, K., Pasquier, P.: Musical agents: a typology and state of the art towards musical Metacreation. J. New Music Res. **48**(1), 56–105 (2019)

Thom, B.: BoB: an interactive improvisational music companion. In: Proceedings of the Fourth International Conference on Autonomous Agents - AGENTS 2000, pp. 309–316. ACM Press, Barcelona (2000)

Uliam, H., de Azevedo, F.M., Ota Takahashi, L.S., Moraes, E., Negrao Filho, R.D.F., Alves, N.: The relationship between electromyography and muscle force. In: Schwartz, M. (ed.) EMG Methods for Evaluating Muscle and Nerve Function. InTech (2012)

Varela, F.J., Thompson, E., Rosch, E.: The Embodied Mind: Cognitive Science and Human Experience. MIT Press, Cambridge (1991)

Watakabe, M., Mita, K., Akataki, K., Itoh, Y.: Mechanical behaviour of condenser microphone in mechanomyography. Med. Biol. Eng. Comput. **39**(2), 195–201 (2001)

Whitelaw, M.: Metacreation: Art and Artificial Life. MIT Press, Cambridge (2004)

Wiener, N.: Cybernetics; or control and communication in the animal and the machine. In: Cybernetics; or Control and Communication in the Animal and the Machine, pp. 1–194. Wiley, Oxford (1948)

Zicarelli, D.: M and Jam factory. Comput. Music. J. **11**(4), 13–29 (1987)