ASPECTS OF GESTURE IN DIGITAL MUSICAL INSTRUMENT DESIGN

William Brent

American University, Washington DC
Audio Technology
`w@williambrent.com`

ABSTRACT

The flexibility of current hardware and software has made the mapping of relationships between a sound's parameters and physical source of control a relatively trivial task. Consequently, the endeavor of sophisticated digital instrument design has been accessible to the creative community for several years, which has resulted in a host of new instruments that explore a variety of physical mappings. The emphasis on physicality exhibited by so-called "gestural controllers" stands in contrast to the practice of conventional laptop performance. While the laptop computer is certainly a digital musical instrument, its associated performance practice is often criticized based on a perceived lack of physicality. This paper examines motivations behind the foregrounding of gesture in computer-based performance. Critical theory and neuroscience research are drawn upon in order to consider ways in which the desire for connections between sound and motion amount to more than mere fascination with virtuosity.

1. INTRODUCTION

For centuries, the pipe organ was the most sophisticated synthesizer employed in western music. The sounds it produced could be specifically sculpted in novel and artificial ways, and were called upon from a keyboard that allowed for rapid and accurate control. Today, these same tasks are often accomplished with a general purpose tool: the laptop computer. The two technologies bring many patently obvious differences into relief, but they also share a significant similarity. Namely, both possess the ability to decouple a source of sound from its means of control. In general, this is a feature regarded as exclusive to electronic or computer-based instruments. For example, [19, p. xvi] asserts that "perhaps uniquely in the history of the performance of music, we are able to separate entirely the production of sound from the means used to control it."

With respect to the majority of instruments, this is a fundamentally important point to make. In the woodwind family, for instance, design of the mechanical means of pitch control is strictly constrained according to the physical properties of an instrument's standing wave. Digital musical instruments offer an unprecedented degree of freedom in relaxing such constraints. However, the organ's control interface has varied considerably over the course of its extraordinary lifetime, often in ways that were completely independent from the development of its sound generation unit [6]. Theoretically, the freedom offered by 19th century electric actions that output simple control voltages allowed for any imaginable type of playing interface; yet alternate controllers were not explored extensively in the years that followed. Thus, we cannot attribute the staying power of the keyboard to a lack of sufficiently adaptable technology.

Strictly speaking, the keyboards of both organs and laptop computers offer rows of buttons that are either active or not. Yet the pipe organ's status as a musical instrument was never challenged to the same degree as the laptop's. Most attempts to explain negative views of laptop performance focus on an absence of physical performance gesture [22][28][32][23].

> The crucial missing element is the body . . . With the emergence of the laptop as instrument, the physical aspect of the performance has been further reduced to sitting on stage and moving a cursor by dragging one's finger across a track pad in millimeter increments [22, p. 12].
>
> Whilst we hear great leaps in sound in these performances, from screeching slabs of noise to fragile and tinkly droplets of high-end tones, the performer sits behind their screens with little or no perceivable movement, lost in thought as they manipulate files and patches. [28, p. 59]

There are of course several valid points of view on this subject. Laptop performance has carved a permanent space for itself within the concert tradition, and some audiences simply do not expect strong action-sound relationships. To give but one example, computer musician Pamela Z has said that she is perfectly content watching a performer "standing on stage just flipping knobs," despite the fact that her own performance practice using the BodySynth controller is very consciously gestural [1, p. 3].

Whether one chooses to highlight, de-emphasize, or ignore it, the issue of physicality is of central importance to the performance of computer-based music. And, regarding the design of digital musical instruments that make use of gestural controllers, the degree of perceived physicality associated with performance can now be chosen relatively freely. The purpose of this paper is to consider the motivations behind such a choice in gradual stages.

First, by drawing on Miranda and Wanderley's model of the Digital Musical Instrument (DMI), we can re-examine qualities assumed to be unique to the laptop computer and other types of DMIs with unfamiliar control interfaces. This will help us move away from the notion that they represent an irreconcilable break from the history of western musical instruments, allowing us to place them along a continuum that draws relationships to other types of instruments while still acknowledging what may very well be different. To discuss the way that DMI performers relate to their instrument through physical gesture in performance, we will look to frameworks for describing musical performance gesture, or as it is termed in [15], "music-related movement". We can then attempt to locate the music-related movements of DMI performance within the concert tradition, and consider how this practice challenges assumptions.

One common assumption is particularly relevant: that there should be a perceivable correlation between the movements of a performing musician and the resulting sound. This is closely tied to virtuosity—an aspect of music that from the most cynical perspective is viewed as empty exhibitionism. Thus, we must ask: are desires for clear action-sound relationships merely historical remnants linked to a perverse obsession with technical prowess? The mimetic hypothesis [8] and supporting research related to the mirror neuron system can be seen to suggest otherwise, potentially establishing a physiological basis for such expectations. With this in mind, we can re-evaluate references to action-sound relationships that dismiss the phenomenon as extra-musical, or mere spectacle.

2. DIGITAL MUSICAL INSTRUMENTS

[19] presents a very thorough catalogue of efforts in interface design for the control of computer-realized sound. Before covering specific instrument systems in depth, the authors put forward a simple but crucial model of DMIs in order to solidify concepts and terminology. The model is comprised of two primary units: a gestural controller and a sound production unit. As stated above, a significant feature of DMIs is the artificially established mapping between these modules. It is fundamentally important to acknowledge that neither a gestural controller nor a sound production unit constitute a musical instrument. Both components and their network of connections must be considered as a whole.

[19] also introduces a gestural controller taxonomy that does not insist on discrete categorization. The four stable points of reference along its spectrum are: augmented musical instruments, instrument-like gestural controllers, instrument-inspired gestural controllers, and alternate gestural controllers. Here, we focus on alternate gestural controllers, exemplified by the *Hands* [31], the *T-Stick* [18], the *Silent Drum* [21], and the *Peacock* [20]. To this group we can add hardware not originally intended for DMI use, such as the Wii Remote, touch screen devices, and the laptop keyboard interface itself.

Alternate controllers do not clearly resemble any traditional acoustic instruments. In this sense, the choice of an alternate controller for DMI design opens many creative avenues—the controller can be designed from the ground up in order to facilitate arbitrary types of performance actions, and the process of composition is not influenced by long-standing idiomatic practices. But alternate controllers are also subject to difficulties that typically surround the unfamiliar. Whether or not audiences will be able to make any sense of the unconventional gestural vocabularies that emerge from performing with such DMIs is a question that must be considered.

2.1. Alternate Controller Digital Musical Instruments

In general, what sets alternate controller DMIs (AC-DMIs) apart from acoustic and analog instruments in the context of performance? One point has already been raised—that the means of controlling sound has been decoupled from the physical source of vibration. Two other unique characteristics are the ability to reference an extremely diverse range of sound sources, and the ability to control musical processes on many different time scales.

With the exclusion of systems that exploit robotic control of acoustic instruments, the sound production units of AC-DMIs rely on loudspeakers to articulate sound. Consequently, the range of sound sources that can be convincingly evoked is enormous. It is worth considering whether or not such broad referentiality is one thread in the complex of factors contributing to criticism of AC-DMI performance. Do the action-sound relationships of such performers present a unique challenge in this regard? Michel Chion's discussion of the concept *synchresis* suggests that there is no implicit difficulty surrounding the use of sounds that (in terms of physical source) are incongruent with their perceived causes.

Synchresis—a conflation of the words synchronism and synthesis—is "the spontaneous and irresistible weld produced between a particular auditory phenomenon and visual phenomenon when they occur at the same time. This join results independently of any rational logic." [7, p. 63] Chion's context for the term is cinema, where "causal listening is constantly manipulated," and "most of the time we are dealing not with the real initial causes of the sounds, but causes that the film makes us believe in." [7, p. 28] Likewise, performers can make us "believe in" the causal relationships they present in spite of any logical incongruities.

Though it is obvious, Chion specifies a condition to the potential relationship between an auditory and visual event: they must occur at the same time. Yet, in the context of music, a "weld" between action and sound does not necessarily require synchrony. In terms of the sonic time scale proposed in [25], the most straightforward action-sound relationships exist at the level of a *sound object*. Here, a single sound event with a clear onset and unambiguous decay simply needs to begin at an appropriate point in relation to a well defined action. We can imagine a percussionist striking a surface with a mallet and hearing a piano note. The relationship might be unexpected (unless the surface *is* a piano), but through repeated instances we gain trust in the established cause and effect.

A different set of considerations arise when AC-DMI performers control events at the next larger (*meso*) time scale, measured in seconds and associated with the traditional idea of a musical phrase. This would correspond to the mallet strike initiating a multi-onset melody that might unfold over twenty seconds. Continuing this line of thought to single actions that trigger events on the *macro* time scale (large-scale formal elements measured in minutes), we can begin to see that the strength of synchresis is inversely proportional to the time scale of musical sound events. When trying to establish a desired level of action-sound association, the extreme flexibility offered by AC-DMIs can become a disadvantage.

3. ACTION-SOUND RELATIONSHIPS

The differences between digital and acoustic instruments raised in the previous section point to the importance of examining physical relationships between performers and instruments, and the way that performed actions correspond to sound. The discussion thus far has already made reference to Jensenius' concept of action-sound, introduced in [15]. Here, we can consider some of the related details. The most basic distinction exists between action-sound relationships and action-sound couplings. Couplings represent correlations between action and sound that are bound by laws of nature. Jensenius gives the example of a glass falling toward the ground. As the glass falls, we anticipate a range of possible sonic outcomes: either the sound of shattering glass, or a simple "clink" (if we are lucky). The result of an infant crying would defy our understanding of reality, leading us to formulate other possible explanations. These tendencies are strongly linked to survival instincts, and it is difficult to imagine functioning in everyday life without the support of reliable action-sound couplings. All couplings are also action-sound relationships, but the relationship category is open to cases that are inconsistent and entirely fabricated. For instance, Chion's examples of synchresis must be classified as relationships, not couplings.

In the gestural control of AC-DMIs, action-sound couplings are essentially impossible, which impacts the expected set of movements that accompany such instruments. What are the physical gestures associated with performing AC-DMIs? Imagining a familiar acoustic instrument (a violin, cello, or contrabass) it is easy to visualize the vocabulary of gestures that go along with it. Likewise, little effort is needed to imagine the movements required to make specific types of sound on these instruments. As a cellist leans forward to reach the end of the fingerboard, we expect to hear sounds in a higher register. In contrast, it may seem that there are no gestures that an AC-DMI performer can make in relation to their instrument that will necessarily produce the expectation of a specific sound. However, if predictable consequences are established in the course of a performance through temporarily consistent action-sound relationships, there is no reason that AC-DMI gestures cannot generate expectation. From a perceptual standpoint, a child's first experience witnessing a violinist in performance is a very similar scenario.

A more accessible analogy is that of multi-percussion setups, where many of the instruments may be found objects that audiences have never heard in a musical context. Such diversity is encouraged by mainstays of the percussion repertoire like Cage's *27'10.554"* and Ferneyhough's *Bone Alphabet*, which call for fundamental instrumentation choices to be made at the discretion of the performer. Even for seasoned audiences, each performance of these pieces can require adjustment to new action-sound relationships. Members of the audience may have vague notions of the sounds that a particular set of objects will create (what Jensenius calls an action-sound palette), but accurate gesture-based expectations on par with those accompanying familiar acoustic instruments are not in place until action-sound couplings are made manifest by the performer. Furthermore, through extended techniques and mallet changes, percussionists often subvert the expectations that begin to form in response to coupling perception.

In other words, there is no reason to question the effectiveness of establishing action-sound relationships in the course of a performance. Modern percussion performance practice has complicated our conception of musical instruments, and DMI design will continue to make its own contribution in this regard through creative design of action-sound relationships.

3.1. Music-related Movement

In [5], Cadoz and Wanderley draw together and develop information from an impressive range of gesture-related theory and research, including work by Choi, Delalande, and Ramstein. Building upon this and other sources, Jensenius forms four areas of music-related movement in [15]: sound-producing actions, ancillary movements, sound-accompanying movements, and communicative movements. In all, these frameworks offer sophisticated tools for gesture-based performance analysis.

Both [5] and [15] describe two main types of sound producing actions: excitation and modification. Jensenius further specifies the area with a varying scale of direct to indirect excitations, classified according to the level of remove between a player and the source of vibration. For instance, a pizzicato string articulation is purely direct, while a vibrating string that results from the mechanized activation of a piano key is more indirect. Rather than the instantaneous vs. continuous distinctions used in [5], Jensenius describes excitation actions in terms borrowed from Godøy's discussion on the gestural-sonorous-object [12]—i.e., impulsive, sustained, or iterative. In the case of vibraphone performance, these descriptors would be appropriate for a mallet strike, a bowed bar, and a multiple-bounce roll, respectively. Both models include parametric and structural modification actions, exemplified by the changing of pitch on a string instrument's fingerboard and the insertion of a mute into the bell of a trumpet, respectively.

The third type of instrumental gesture in [5]—selection gesture—is not included in Jensenius's area of sound-producing actions. Selection gestures relate only to choosing between different areas of an instrument, and hence do not produce sound. The appropriate place for such motions is the area Jensenius calls ancillary movement. Support movements also fall in this area, and both play an important role in the formation of short-term expectation. In the case of a pianist, the support movement of a raised arm for an excitatory sound-producing action telegraphs the sonic result of that action before it is heard. If complex aesthetic experience can be partly explained in terms of expectations, violations, and explanations [4], it is clear that supportive ancillary movements make a significant contribution to the process.

Alongside supportive movements, Jensenius proposes phrasing and entrainment movements as distinct areas within the ancillary category. Functionally, phrasing movements are considered unnecessary for sound-production, but Jensenius notes their consistency in connection with specific musical phrases. In music that emphasizes formal considerations, the reappearance of phrases accompanied by phrase-specific ancillary movements arguably impacts the perception of form. From this point of view, form is articulated both sonically and visually. The other type of ancillary movement—entrainment—synchronizes with a "continuous underlying feature of the music," [15, p. 50] and is most easily exemplified by pulse-oriented foot tapping.

While there are a great number of further distinctions to be made within these frameworks, we can introduce and apply some of them in context by considering specific examples of AC-DMI systems.

3.2. Action Vocabularies in AC-DMI Classification

In addition to using points along the spectrum of gestural controllers presented in [19], (i.e., augmented instrument, instrument-like, instrument-inspired, and alternate controllers), we can describe the gestural vocabulary associated with individual instruments in order to make finer distinctions that are often quite necessary. For instance, referring to laptop musicians as AC-DMI performers with no further specification ignores creative design choices that distinguish individual laptop-based instruments. Though the hardware is similar, the performance gestures associated with laptop instruments described in [9] vary considerably.

Even in the case of AC-DMIs whose custom hardware makes them easily distinguishable by sight, reference to gestural vocabulary is needed in order to arrive at characteristics that make each instrument unique. Taking the *Silent Drum* [21] and *Peacock* [20] into account, the former involves haptic feedback from a flexible membrane, while the latter makes use of proximity sensors and involves no physical contact whatsoever. Relative to a pizzicato string pluck, the excitation gestures performed on these instruments are both indirect, but they are significantly more so for the *Peacock*. Drawing on Cadoz' notion of the gestural channel [5], we can also say that performance of the *Silent Drum* makes use of the ergotic function—associated with forces applied to an object—while that of the *Peacock* does not.

As a result, the nature of excitation gestures themselves are drastically different for these two instruments. On the *Silent Drum*, contact with the membrane is in most cases the understood cause of event onsets[1]. Excitation gestures are best classified as impulsive in this case, though subsequent parametric modification gestures that exploit the instrument's flexible surface are continuous. In contrast, excitation on the *Peacock* is executed via entrance of the hand into a specific performance zone above the instrument (i.e., within range of the grid of sensors). Continuous modification gestures are then performed by moving the hands higher or lower in relation to the instrument's surface, and additional excitations can be carried out by crossing the boundaries of each sensor in the grid.

In both cases, once the consequences of an excitation gesture are understood, the support movements leading up to excitation significantly contribute to a dimension of musical anticipation. That is, our perception of the performer's approach to the playing surface (or area) as either fast, slow, aggressive, or tender is meaningfully connected to the sounds we hear before we hear them.

For AC-DMIs, the nature of a selection gesture may be quite different from that of an acoustic instrument. Rather than movement toward or away from specific areas of the instrument, a selection gesture might be tied to the simple press of a button that initiates a pre-established remapping between control data and synthesis parameters. In such a case, the different "area" of the instrument is virtual, not physical. The performance of *Silent Construction I* and *Black Vox* (for the *Silent Drum* and *Peacock*, respectively), both involve such selection gestures, resulting in a series of unique gestural environments as the pieces unfold[2]. Of course, more conventional selection gestures exist in these cases as well. In *Black Vox*, the performer reaches toward specific rows of sensors in order to access desired sounds and processes.

4. THE MIMETIC HYPOTHESIS

The detailed treatment that performance gesture is given in the frameworks described above speaks to its importance as a musical dimension. Clearly, music-related movement analysis based on these models can provide meaningful insight for instances of AC-DMI performance. However, in order to consider the possibility that action-sound relationships constitute a fundamental aspect of music per-

---

[1] Among the exceptions in *Silent Construction I* is a moment where the initiation of a vibraphone chord is connected with release of the membrane.

[2] Depending on the scope of the remapping, it may be more accurate to consider such modifications a change of instrument entirely.

ception, we will require more than well constructed terminology-based frameworks. To that effect, a compelling intersection of abstract concepts and concrete findings can be seen between Godøy's notion of the gestural-sonorous-object, Cox's mimetic hypothesis, and the human audio-visual mirror neuron system.

In [12], Godøy draws on a classical concept from electronic music—Pierre Schaeffer's objet sonore—to assert that all sounds have implicitly gestural associations. This idea seems to contradict the standard notion of a sonorous object as a purely abstract sonic entity arrived at through a process of reduced listening that actively attempts to nullify the cultural and causal significations of a sound. In response to potential skepticism, Godøy is quick to provide Schaeffer's own words, which acknowledge the fact that it is impossible to rigidly adhere to a single mode of listening, and that the focus of listening attention will inevitably pass unconsciously from one system to another. In other words, there is room to consider the sonorous object in conjunction with other modes of perception. In fact, Godøy claims that Schaeffer's methods for excising a sonorous object from a recording—and the very descriptors he applied to his classifications in [27]—are themselves evidence of the connections between movement and sound. In Godøy's words, "there is a gesture component embedded in Schaeffer's conceptual apparatus." [12, p. 154]

Building on the relatively easy to accept notion that the perception of sound might generate images of related movements, Godøy reasserts his previously expressed belief in a motor-mimetic component of music perception:

> There is an incessant simulation and reenactment in our minds of what we perceive and a constant formation of hypotheses as to the causes of what we perceive and the appropriate actions we should take in the face of what we perceive. I believe this points in the direction of what I would like to call a motor-mimetic element in music perception and cognition, meaning that we mentally imitate sound-producing actions when we listen attentively to music. [11, p. 318]

We are still a long way from establishing that music perception literally involves "incessant simulation" and a "constant formation of hypotheses as to the causes of what we perceive." But the instinctual need to account for the sounds in our environment does seem to extend to benign as well as life-threatening situations. With regard to the implications of motor-mimesis playing an active role in music perception, it would go far in explaining the frustrated reactions to laptop performance referenced above. It is not as if the process of relating the minimal actions of a performer to a dense stream of musical events is an outrageously difficult task for contemporary audiences—it is that such a relationship between performer movement and sound may not clearly invite processes of reenactment or mimesis. If human beings are unconsciously accustomed to following a musical performance through vicarious participation or imagined action, it is plausible that efforts to understand and relate to music largely depend on clear action-sound relationships.

Arnie Cox begins his presentation of a similar theory—the mimetic hypothesis—with an invitation:

> Recall the Beethovenian theme of the last movement of Brahms's 4th Symphony. As you recall, ask whether your voice is involved or activated in any way, whether imagining singing, or singing along, or feeling only the impulse to sing along. [8, p. 195]

In what follows, Cox introduces the idea of subvocalization, which in its most subtle form is the mere impulse to vocalize. The article suggests that subvocalized mimetic response to music is responsible for both the culturally ingrained metaphor of "greater is higher" in relation to pitch, and the unidirectional application of vocally-inspired terminology to instrumental sound (e.g., cantabile, sotto voce, etc.). A fundamental assumption of the argument is that while children take part in overt mimesis, the mimetic participation of adults gradually becomes covert, yet never disappears completely. In the case of music perception, this would mean precisely what Godøy implies above—that we are covertly but incessantly imitating the sounds that we hear.

Subvocalization is the type of motor-mimesis that Cox sees as most important, but in the case of musical instruments, he suggests that a range of motoric responses exist. According to Cox, upon hearing sounds we first react through motor-mimesis, then compare that motor response to motor patterns associated with our previous experience of creating the sound (or a similar sound) ourselves. This is a direct comparison if we actually have experience making the sound in question, and indirect if not. Thus, when we hear an instrument—say, the Brazilian cuíca—mimesis enables both a direct comparison to our own experience of making such sounds, as well as an indirect comparison to making similar sounds via other types of action (perhaps vocalizing).

The cuíca is an especially apt example because its vocal character makes the prospect of imitating instrumental sounds with the voice very approachable[3]. From there, one can imagine imitating the sound of non-voice-like instruments and even completely synthetic sounds. Suspending, for the moment, any disbelief that we automatically engage in such a process, we can consider a few of its implications as explained by Cox. The most radical suggestion is that "speech imagery and musical imagery are actually special cases of motor imagery in general." [8, p. 201] That is, our systems for imagining the sounds of speech and music could be subsystems of imagining movement. If true, this would have further implications in the area of musical affect—a search for understanding

---

[3] Alternatively, it is known as the "laughing gourd", and the name of a similar instrument—the "lion's roar"—also seems significant in relation to vocalization.

why it is that music should elicit such strong emotional responses. In Cox's words, "the hypothesis suggests that muscular-emotional response to music is . . . integral to how we normally perceive and understand music, because we normally imagine (most often unconsciously) what it is like to make the sounds we are hearing." [8, p. 205] Regarding semiotics in music, the embodied approach to music perception that the mimetic hypothesis encourages would not completely disambiguate signification, but at least constrain it within a definable range of options as dictated by the human body.

The notions of musical mimetic response put forward by Godøy and Cox, compelling as they may be, cannot be taken at face value; however, their significance to discourse on DMI design and performance is obvious. The most extreme consequence of these theories is that music is understood only through movement.

5. THE MIRROR NEURON SYSTEM

Neuroscience research from the early 1990s identifying so-called mirror neurons in the macaque monkey premotor cortex has garnered a great deal of attention in recent years. Here, it provides support for theories of embodied music cognition. Mirror neurons fire during both the execution and mere observation of specific actions [24]. By recording the discharge patterns of mirror neurons, it has been established that this area of the brain reacts similarly whether a monkey picks up a piece of food or observes another monkey picking up a piece of food. This phenomenon occurs at various distances, with various objects, and regardless of whether the observed action is carried out by a monkey or a human. A hypothesized function of these neurons is that they play an important role in processes of imitation and action understanding.

Rizzolatti and Craighero cite two experiments that test this hypothesis [30][16]. In the latter study, Kohler et al. exploited actions with prominent sonic by-products to show that movement can be understood without visual information. Thus, in the absence of visual cues, it is suggested that mirror neurons mediate understanding of actions that are taking place. The response of mirror neurons was recorded in relation to the sight and sound (V+S) and sound alone (S) of two actions: paper ripping and the dropping of a stick. In both the (V+S) and (S) cases, increased activity of these neurons clearly aligned with the sound event, with a slightly weaker response in the case of (S). It should be noted that the authors recorded responses to non-action-related sounds as well. Neither white noise nor synthesized clicks elicited an increase in activity whatsoever, indicating the specific significance of established action-sound relationships. Kohler et al. named the subset of mirror neurons that responded strongly in both the (V+S) and (S) cases *audiovisual mirror neurons*.

Audiovisual mirror neuron research has established that for macaque monkeys, sound perception involves the firing of cells that are associated with the actual execution of movements known to make the perceived sound in the first place. According to these results, sound literally elicits an appropriate motor response. With regard to whether or not a similar system exists in human beings, Rizzolatti and Craighero's references and subsequent studies carried out by others [2][10] point to the affirmative. The relevance of these findings to the mimetic hypothesis is that sounds alone may indeed trigger mimetic responses as we listen to music.

Building off of studies establishing human motor response to speech sounds, [2] sets out to investigate other action-related sounds. Paper tearing and typing sounds (action-sounds associated with the hand) and control sounds (footsteps and thunder) were played to participants while transcranial magnetic stimulation (TMS) was applied. During this process, motor evoked potentials (MEPs) were recorded from the hand muscle associated with the stimulated hemisphere. The study documented significantly larger MEP amplitudes when participants heard the hand-related sounds during stimulation of the left hemisphere, indicating that "in the left hemisphere, actions are coded through auditory, visual and motor components." In short, simply hearing hand-related sounds resulted in motor response, while footsteps and thunder clearly did not. This is in spite of the fact that participants did not report mentally imagining the actions associated with sounds presented to them during trials.

In [10], the authors investigated audiovisual mirror neuron properties from both directions, recording activity while participants listened to sounds and when they later performed the actions associated with these sounds. Monitoring was carried out using functional magnetic resonance imaging (fMRI). The study demonstrated further selective characteristics of the mirror system, identifying regions that were activated by either the sound or performance of hand-related actions but not mouth-related actions, and vice-versa.

Music-specific studies making use of fMRI have shown that levels of musical training have a strong effect on response patterns. Thus, we must be careful not to make generalizations concerning perception with respect to these issues. In [13], Haslinger et al. presented pianists and a non-musician control group with video of piano playing actions without sound. Pianists exhibited activation in auditory areas, while non-musicians did not. [3] identifies consistent brain activity unique to professional pianists vs. non-musicians when either listening to piano tones or playing a mute piano keyboard.

Changes in motor response have also been tracked relative to short-term musical training. Focusing on a specific action-sound relationship—right-handed piano playing—the experiment described in [17] required non-musicians to learn 24-note, 15-second melodies by ear, and later monitored frontoparietal motor-related brain regions via fMRI during a variety of listening situations. Participants were instructed to remain completely still while listening to eliminate the possibility that recorded motor activity could be due to actual movement. The authors identified motor areas that were active only when participants heard the melodies they had practiced. Unfamiliar melodies composed of different pitches failed to cause any such response; however, unfamiliar sequences made up of the same pitches used in a practiced melody triggered a degree of motor activity as well, despite the fact that "subjects were completely unaware . . . that the piece was composed of the same notes as the trained-music," and that unfamiliar sequences using non-practiced pitches triggered no activity. Musical novices thereby exhibited appropriate motor response based on absolute pitch.

If the mere sound of a piano can evoke motor imagery of piano performance movements and vice-versa, could listening to a recording of an AC-DMI performance by artists that we have witnessed in action elicit appropriate motor mimesis? This would require that connections between action and sound are able to form within a very short time span. A statement from David Rokeby, creator of the *Very Nervous System*, which tracks body motion to control sound, seems to confirm this possibility anecdotally:

> An hour of the continuous, direct feedback in this system strongly reinforces a sense of connection with the surrounding environment. Walking down the street afterwards, I feel connected to all things. The sound of a passing car splashing through a puddle seems to be directly related to my movements. [26, p. 28].

In Rokeby's case, the consistent practice of artificially triggering sound via body motion produced a state of mind in which he perceived his body to be controlling independent sound events outside of the video tracking environment. This perceived control could actually be the result of a learned mimetic response.

The full implications of Rizzolatti et al.'s theories of action understanding and reproduction via the human mirror neuron system are contentious [29][14]. However, the notion that viewing actions or simply hearing sounds associated with certain actions triggers activity in the motor systems of our brains appears to be firmly established. At the very least, conclusions from the research reviewed in this section are relevant to both Godøy and Cox's propositions. To a certain extent, their intuitions are supported by empirically determined evidence.

6. CONCLUSION

Indications that music perception involves covert motor mimesis help us view appreciation of instrumental virtuosity in a more favorable light, and provide meaningful insight for gestural considerations in DMI design. At the same time, acknowledging the musical role of action-sound relationships does not imply that practices that make no attempt to invite mimetic response are somehow less justified. In fact, based on the line of thought presented here, one could easily argue for the importance of establishing modes of musical perception that deviate from such well-worn paths.

More broadly, the ideas presented here pose a host of questions relevant to computer-generated music. For instance, can abstract sounds be categorized according to the types of mimetic responses they trigger—revisiting Schaeffer's project in [27] from a new angle? Are interests in generating music based on biological processes or the activity of artificial life forms partly motivated by our inclination to understand music through movement? What kinds of motor responses are triggered by observing the performance of a DMI that generates sound via robotic control of acoustic instruments? Would the performer's control actions (action-sound relationships) or the machine's consequent actions (action-sound couplings) dominate mimetic processes? Could hearing computer-based music with features that are currently beyond the limits of human execution ability result in motor coding that allows us to push these very limits? In other words, do various types of human-machine interaction establish new human performance abilities through motor mimesis?

To this collection, we can add the question that AC-DMI performers pose: what are the consequences—both musically and in general—of continually redefining action-sound relationships?

7. REFERENCES

[1] M. Alburger, "Z'wonderful, Z'marvelous: An interview with Pamela Z," *20^{th} Century Music*, vol. 6, no. 3, pp. 1–16, 1999.

[2] L. Aziz-Zadeh, M. Iacoboni, E. Zaidel, S. Wilson, and J. Mazziotta, "Left hemisphere motor facilitation in response to manual action sounds," *European Journal of Neuroscience*, vol. 19, pp. 2609–2612, 2004.

[3] M. Bangert, T. Peschel, G. Schlaug, M. Rotte, D. Drescher, H. Hinrichs, H. Heinze, and E. Altenmüller, "Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction," *NeuroImage*, vol. 30, pp. 917–926, 2006.

[4] K. Burns, "Bayesian beauty: On the ART of EVE and the act of enjoyment," in *Proceedings of the AAAI06 Workshop on Computational Aesthetics*, 2006.

[5] P. Cadoz and M. Wanderley, "Gesture-music," in *Trends in Gestural Control of Music*. IRCAM—Centre Pompidou, 2000, pp. 71–94.

[6] M. Campbell, C. Greated, and A. Meyers, *Musical Instruments: History, Technology, and Performance of Instruments of Western Music*. New York: Oxford University Press, 2004.

[7] M. Chion, *Audio-Vision: Sound on Screen*. New York: Columbia University Press, 1994.

[8] A. Cox, “The mimetic hypothesis and embodied musical meaning,” *Musicae Scientiae*, vol. 5, no. 2, pp. 195–212, 2001.

[9] R. Fiebrink, G. Wang, and P. Cook, “Dont forget the laptop: Using native input capabilities for expressive musical control,” in *Proceedings of the2007 Conference on New Interfaces for Musical Expression*, New York, 2007.

[10] V. Gazzola, L. Aziz-Zadeh, and C. Keysers, “Empathy and the somatotopic auditory mirror system in humans,” *Current Biology*, vol. 16, pp. 1824–1829, 2006.

[11] R. Godøy, “Motor-mimetic music cognition,” *Leonardo*, vol. 36, no. 4, pp. 317–319, 2003.

[12] ——, “Gestural-sonorous objects: embodied extensions of schaeffers conceptual apparatus,” *Organised Sound*, vol. 11, no. 2, pp. 149–157, 2006.

[13] B. Haslinger, P. Ehard, E. Altenmüller, U. Schroeder, H. Boecker, and A. Ceballos-Baumann, “Transmodal sensorimotor networks during action observation in professional pianists,” *Journal of Cognitive Neuroscience*, vol. 17, no. 2, pp. 282–293, 2005.

[14] G. Hickok, “Eight problems for the mirror neuron theory of action understanding in monkeys and humans,” *Journal of Cognitive Neuroscience*, vol. 21, no. 7, pp. 1229–1243, 2009.

[15] A. Jensenius, “Action-sound: Developing methods and tools to study music-related body movement,” Ph.D. dissertation, University of Oslo, 2007.

[16] E. Kohler, C. Keysers, M. Umiltà, L. Fogassi, V. Gallese, and G. Rizzolatti, “Hearing sounds, understanding actions: Action representation in mirror neurons,” *Science*, vol. 297, pp. 846–848, 2002.

[17] A. Lahav, E. Saltzman, and G. Schlaug, “Action representation of sound: Audiomotor recognition network while listening to newly acquired actions,” *Journal ofNeuroscience*, vol. 27, no. 2, pp. 308–314, 2007.

[18] J. Malloch and M. Wanderley, “The T-Stick: From musical interface to musical instrument,” in *Proceedings of the 7th International Conference on New Interfaces for Musical Expression*, 2007.

[19] E. Miranda and M. Wanderley, *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard*. Middleton, Wisconsin: A-R Editions, Ltd., 2006.

[20] C. Miyama, “Peacock: A non-haptic 3D performance interface,” in *Proceedings of the International Computer Music Conference*, 2010, pp. 443–445.

[21] J. Oliver and M. Jenkins, “The silent drum controller: A new percussive gestural interface,” in *Proceedings of the International Computer Music Conference*, 2008.

[22] B. Ostertag, “Human bodies, computer music,” *Leonardo Music Journal*, vol. 12, pp. 11–14, 2002.

[23] G. Paine, “Interfacing for dynamic morphology in computer music performance,” in *Proceedings of the International Conference on Music Communication Science*, Sydney, Australia, 2007.

[24] G. Rizzolatti and L. Craighero, “The mirror-neuron system,” *Annual Review of Neuroscience*, vol. 27, pp. 169–192, 2004.

[25] C. Roads, *Microsound*. Cambridge: MIT Press, 2001.

[26] D. Rokeby, “The construction of experience: Interface as context,” in *Digital Illusion: Entertaining the Future with High Technology*. ACM Press, 1998, pp. 27–47.

[27] P. Schaeffer, *Traité des objets musicaux*. Paris: Éditions du Seuil, 1966.

[28] C. Stuart, “The object of performance: Aural performativity in contemporary laptop music,” *Contemporary Music Review*, vol. 22, no. 4, pp. 59–65, 2003.

[29] L. Turella, A. Pierno, F. Tubaldi, and U. Castiello, “Mirror neurons in humans: Consisting or confounding evidence?” *Brain and Language*, vol. 108, pp. 10–21, 2009.

[30] M. Umiltà, E. Kohler, V. Gallese, L. Fogassi, L. Fadiga, C. Keysers, and G. Rizzolatti, ““I know what you are doing”: a neurophysiological study,” *Neuron*, vol. 32, pp. 91–101, 2001.

[31] M. Waisvisz, “THE HANDS: A set of remote MIDI controllers,” in *Proceedings of the International Computer Music Conference*, 1985.

[32] M. Zadel and G. Scavone, “Different strokes: a prototype software system for laptop performance and improvisation,” in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Paris, 2006.