

**This is an electronic reprint of the original article.
This reprint *may differ* from the original in pagination and typographic detail.**

Author(s): Mendoza Garay, Juan Ignacio; Thompson, Marc

Title: Gestural Agency in Human-Machine Musical Interaction

Year: 2017

Version:

Please cite the original version:

Mendoza Garay, J. I., & Thompson, M. (2017). Gestural Agency in Human-Machine Musical Interaction. In M. Lesaffre, P.-J. Maes, & M. Leman (Eds.), *The Routledge Companion to Embodied Music Interaction* (pp. 412-419). Routledge, Taylor & Francis Group. Routledge Companions. <https://doi.org/10.4324/9781315621364-45>

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

Gestural Agency in Human-Machine Musical Interaction

Juan Ignacio Mendoza and Marc Richard Thompson
Music Department, University of Jyväskylä

Abstract

Musical technologies are evolving in such a way that they start to resemble the people that use them. In this vision it seems pertinent to abandon the conception of musicians as users of musical instruments in favor of machines and humans interacting to make music. Both the machine and the human can be modeled as embodied cognitive agents that comprise a network of musical gestures. These gestures are multi-modal signals that allow the agents to exert influence upon each other towards making music. This standpoint integrates traditional and novel music-making technologies, towards a better understanding of musical interaction.

Introduction

In this article we discuss the interaction between humans and machines with the purpose of making music. We first present a generalized model of a musical instrument, to which we simply refer as a *machine*. Then we make an analogy between the model of the musical machine and a model of the human with whom it interacts. Finally, we examine the human and machine models acting within a system of musical interaction.

The discussion is formulated in light of recent scholarly literature mostly related to the development of electronic musical instruments, as they have seen great development and are a logical step after non-electronic musical instruments. Take for instance the guitar followed by the electric guitar and then the electric guitar as a controller of software. The design of electronic musical instruments has progressively diverted from the design of non-electronic musical instruments and the interaction of human beings with these musical machines has changed in

turn, however still retaining some of its fundamental characteristics.

We define the concept Human-Machine Musical Interaction as a holistic view of the interactions between humans and machines with the goal of making music. The concept applies to any human, not necessarily one that has musical training or one that practices music professionally. Likewise, we consider machines as a generalization of musical instruments, whether they have computational capabilities or not.

The dynamics of Human-Machine Musical Interaction described in the following sections are based on the assumption that musical machines can be considered agents, whose behavior and inner workings have evolved to the point of resembling their human users. In this conception a musical instrument becomes an entity coexisting with the human, instead of a tool used by the human. Machines and humans communicate via various signals produced by different modalities. With the most obvious mode being auditory, we will explore how interactive systems also integrate visual and haptic signals for communication. We also remark that the information that those signals carry can be encapsulated under the broad term *gesture* (Cadoz & Wanderley, 2000; Jensenius et al., 2009). The discussion describes the process in which signals emanating from humans and machines affect each other in ways to produce music. We call this process *gestural agency*, expanding upon previous uses of the term (Cumming, 2000; Hatten, 2012; Robb, 2015).

Generalized model of musical machines

We start by analyzing electronic musical instruments, the most recent evolutionary stage of musical instruments. They can be the most complex kind of musical instruments, offering possibilities unforeseen in non-electronic musical instruments. We assume that very sophisticated realizations of electronic musical instruments could offer the same musical interaction possibilities of non-electronic musical instruments. Therefore, a model of electronic musical instruments shall also serve for non-electronic musical instruments. This generalized model will allow us to analyze the interaction of musical machines with humans. Wanderley

(2001) proposed to study the interaction of musicians and electronic musical instruments by dividing the inquiry into four parts: 1) definition and typologies of gesture; 2) gesture acquisition and input device design; 3) mapping of gestural variables to synthesis variables; and 4) sound synthesis. Broadly speaking, musical gesture has been defined as “human body movement that goes along with music” (Jensenius et al., 2009, p. 13). In Wanderley’s proposal, gesture refers to the bodily movement that is communicated to the musical instrument to produce musical sound. Further in this text we elaborate on musical gesture, redefining it as a concept beyond bodily movement. Previous to that we build on the latter three parts of Wanderley's proposal, which have been often presented as an electronic musical instrument’s main constituting modules (Wessel & Wright, 2002; Hunt, Kirk & Neighbour, 2004; Armstrong, 2006; Magnusson, 2010; Fabiani, Friberg & Bresin, 2013; de Campo, 2014).

The first module (gesture acquisition and input device) is comprised of an interface called *controller*, which is a device or several devices, for example a keyboard, knobs, buttons and so forth. The second module (mapping of gestural variables to synthesis variables) involves the connections between the first and the last modules. The last module (sound synthesis) is a device or system to produce a result from the instrument. This result is at least sound, but sometimes might also be a visual display, for example to inform the user about the state of the instrument. The human user (e.g. a musician) is connected to the first module by means of the controller. Using this interface the human transmits an action, also called gesture, to the instrument. This action is a signal, which is transduced from its original form into electrical signals, analog or digital, flowing towards the final module.

Caution is needed when interpreting the human-machine musical interaction as being modular. Modularity implies that the steps involved in the interaction are categorical and discrete. However, the true nature being represented is continuous and often the boundaries of the modules are not evident. This continuous nature can be explained by looking at some non-electronic instruments. In a piano, organ, or harpsichord the model is quite explicit as the first module represents the keyboard and additional commands to modify the sound. The module in the middle represents the connections between keys and other commands (e.g., pedals, levers, and knobs) to the mechanism that produces sound, which are represented by the final module. In

these instruments each module is associated with a physical part of the instrument. Now, how can we describe the violin with this modular approach? It can be argued that the first module corresponds to the bow and strings, the final module to the body, and the module in the middle to the bridge as it connects the strings to the body (Chadabe, 2002). However, all the parts of a violin play a role in each of the three modules. For example, the bow and strings are not only related to the communication with the musician's body, but also with sound production. In this case the categorical modular approach becomes problematic. In the clarinet, the control module can be thought of as the mouthpiece and keys, mapping given by the key mechanism and sound production at the body. Again, a modular approach poses important overlaps as, for example, the mouthpiece, apart from controlling the sound, is also producing the sound. These two instruments, as do probably many other non-electronic instruments, require a model that overcomes the rigid association of the model's modules with a physical part of the instrument.

The problem vanishes when the modules are understood as processes rather than physical parts of instruments. Therefore, in what follows we will call the first module simply Input and the last module Output, the module in between remaining as Mapping. For example, the Input for the violin is the bow and strings; while Mapping is strings, fretboard, bridge, and body; and Output is bow, strings, fretboard, and body. This model, now more general, also serves for electronic musical instruments in which the mapping can be designed to from simple to quite complex. One connection might be directly made from the controller to the synthesizer, whereas other connections might undergo some further processes in between, for example learning the connection that the user prefers (Caramiaux & Tanaka, 2013). To this extent, the incorporation of machine learning into electronic musical instruments has been a crucial evolution. It allows a user to build instruments quickly, for example just by mapping parameters between controller and synthesis, a paradigm that has been called *composed instrument* (Schnell & Battier, 2002; Fiebrink et al., 2010).

In this scenario, the signals are not always flowing in the direction from the controller to the sound production. In a musical instrument, most possibly an electronic one, there might be internal feedback signals going from the Output module back to the Input module or into the Mapping module without leaving the Output module, enabling the musical machine to monitor

its own internal behavior. Theoretically an unlimited amount of internal feedback signals can be mapped in any direction. Likewise, the signals entering the machine, the signals exiting the machine and all internal feedback signals can be of any kind. Obviously, the one kind of signal that is required to be output from a musical instrument is an auditory signal. Returning to non-electronic musical instruments we can observe that visual and haptic signals aid in the playing. For example, looking at the piano keyboard and feeling the keys or visually finding the approximate location of a position when playing the double bass and feeling the weight of the string at that point. Even olfactory signals are present in the interaction with musical instruments, although they are often less influential in making music. In the case of electronic musical instruments, output other than audio has been used more or less in the same fashion as in non-electronic musical instruments, as an aid to the production of sound. An exception is the case of electronic musical instruments that incorporate a visual component that is not an aid to playing but a signal with the same valuation as the auditory output signal. In the same way, the electronic musical instrument could produce haptic and olfactory signals.

Clearly the generalized model of musical machines turns aside from a more conservative conception in which the design and use of an electronic musical instrument is more or less bounded to the same rules as for non-electronic musical instruments.

Human and machine embodied music cognition

The human being that interacts with a machine can be represented using the same logic as for the machine. The Input, Mapping and Output modules that constitute a musical instrument can be seen as parallel to Perception, Cognition, and Action in the human, from a Cartesian point of view (Armstrong, 2006). Likewise, if we consider a machine in which the Mapping module has a component that can learn, we can easily notice that it parallels the human module between Perception and Action: Cognition. In this view, the intermediate modules Cognition and Mapping are the ones actively modifying the signals flowing in an enactive direction: towards Action in the human and towards Output in the machine.

Although it might seem useful to represent musical interactions between humans and machines, this approach holds a caveat. It places cognition between perception and action, like a sandwich (Hurley, 2002, p. 401), as if cognition was a process concerned solely with the brain. This idea has been rendered obsolete in general by the viewpoint of Embodied Cognition (Anderson, 2003) and in particular by the viewpoint of Embodied Music Cognition (Leman, 2008) which, generally speaking, consider cognition as a process involving not only the brain but also the body and its environment.

The embodied quality of a model of a musical human being can be illustrated in the same way as we described a model of musical machines in the previous section: in terms of processes rather than physical units. This means that its modules do not necessarily represent areas of the brain, limbs, or sensory organs. In this way each module can represent functions performed by more than one organ. It can also represent one organ associated with more than one module. It becomes clear that the module between Perception and Action, if we think of it as the human brain, cannot account for the cognitive processes that occur throughout the whole body (Wheeler, 2005). Therefore we prefer to call the human's intermediate module Connections, allowing the term cognition to refer to a process occurring throughout the three modules. Also we take into account that there are internal feedback signals connecting the modules without being sent outside. These signals represent, for example, the modification of a perception organ's behavior triggered by an action, with or without the mediation of thinking.

On the side of the machine we can state similarly that *machine cognition* takes place. The concept of machine cognition, as in the human counterpart, comprises perception, understanding and learning by a machine (Haikonen, 2007). In the model of a musical machine that we have described so far, these areas are more or less represented by the modules Input and Mapping. Yet, in the design of machines that interact with their environment, it has been necessary to consider cognition as an enactive process that leads to consciousness, being embodied as the process involves the awareness of the machine's body and its surroundings (Holland, 2007). The awareness of the machine is related to its internal feedback signals, as described in the previous section. Akin to the musical human model, we shall consider that the cognitive capacity of a musical machine takes part not only in the Mapping module but also at the Input and Output

modules.

The generalized model of musical machines and the model of the musical human being now appear highly resemblant of each other. When human and machine interact, the Action module of the human is connected to the Input module of the machine and the Output module of the machine is connected to the Perception module of the human. In the resulting loop, the signals are continuously updated by the human and the machine. Moreover, these connections are not exclusive. Human and machine “catch” signals with their first modules, whether these signals are their own (e.g., to listen to their own sound) or coming from elsewhere (see FIGURE 1).

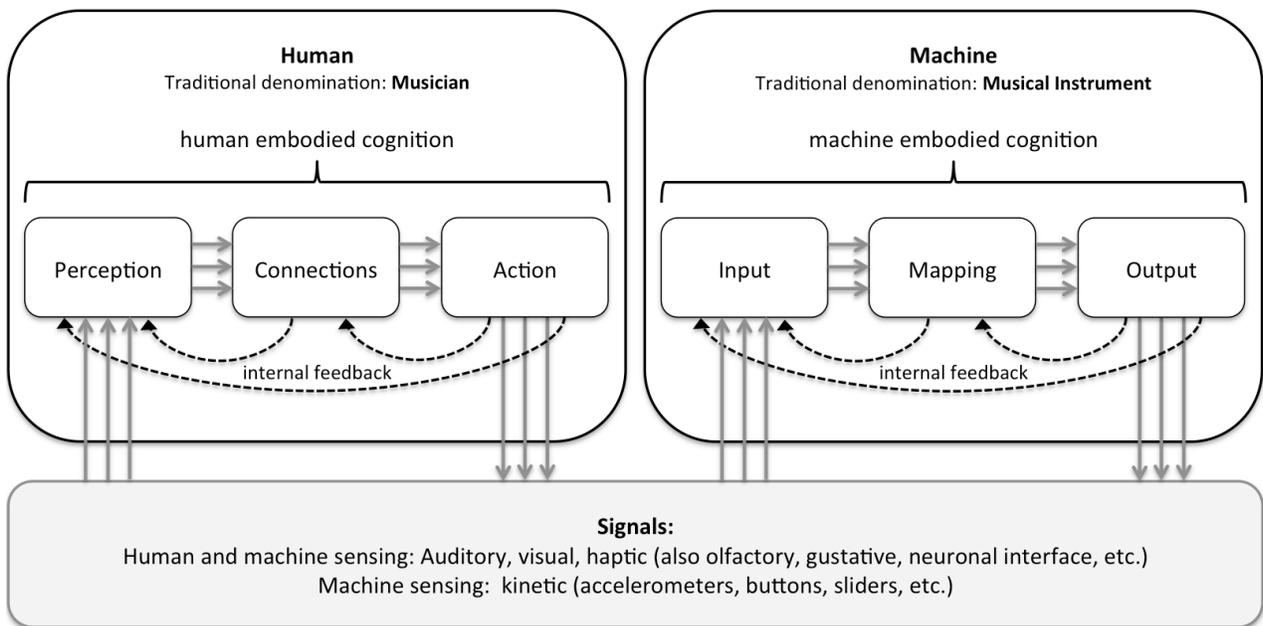


FIGURE 1: Human and machine are agents connected by signals. The model allows for the incorporation of more agents into a network.

Gestural Agency

Now we examine the interactions between the human and machine models that we have described. In this interaction system the flow of signals is enactive, in other words, they are

directed towards Action in the Human and Output in the Machine. This implies that there is a target receiver of an outgoing signal. Thus, both human and machine have the potential to become agents exerting equal or unequal influence upon each other to produce music. These agents' behavior is affected by the conditions imposed, demanded or proposed by other agents (Gurevich & Treviño, 2007; Bown, Eldridge & McCormack, 2009). The process starts with exploration and discovery of the intentions of the other agents, gradually turning into an objective-based task as a musical aesthetic emerges (Caramiaux et al., 2014). In this way the machine resembles an entity, more a musician than a musical instrument (Van Nort, 2011).

The evolution of these agents requires some internal adaptation. In the model of the human that we described, the Connections module is dynamically affected by reasoning and experience, to satisfy demands from the musical environment or to accomplish individual musical goals. The machine can also go through such a process, as the mapping module can be affected by generative algorithms and machine learning.

This idea seems to challenge a more traditional conception of a musical instrument as an extension of the musician's body (Nijs, Lesaffre & Leman, 2009; Simoens & Tervaniemi, 2013). In that conception the flow of signals is bi-directional, as the instrument can affect the way the musician plays as much as the musician affects the sound of the instrument. We integrate this notion with the idea of musical machines being agents, regarding the relationship of a musical machine and a human being as two-fold. While the musical machine can be seen as external to the human body, it is still connected to the human body by the gestural signals, both comprising a greater organism. This organism is the interaction system that we have been describing, where not only the musical machine is an extension of the human body but also the human body is an extension of the musical machine. Rather than bi-directional, the signals in this system are multi-directional, as the output or action of one agent can be "caught" by the same agent or by other agents. Following this reasoning, musical interaction becomes a network of musical agents in a musical ecosystem.

For example, one agent proposes a musical idea like a rhythmical pattern or a tonic note, which is acted upon by the rest of the agents in the system. This could be the case of musicians playing together, a single musician interacting with a musical instrument, or intelligent machines

playing together. Agents have limitations and thus impose constraints, to which the interacting agent has to respond. A violin cannot play a very low drone of Tuvan chant as would a group of vocal performers or an electronic musical instrument. The one who plays the violin adapts to take most of the instrument. In the same way it would be possible for a musician to learn how to interact with a musical instrument that allows to shape complex drones but does not allow for such a detailed command of pitch and dynamics as the violin does.

As suggested in the previous section, the agents in the musical ecosystem are connected by *multimodal* signals (Françoise, Schnell & Bevilacqua, 2013). These signals act as interfaces between agents (Di Scipio, 2003) carrying information that we call *gesture*, regardless of the nature of the signal. In this context, the concept of gesture is not bounded to movement of the human body. Gesture can be visual, kinetic, auditory, or of any other kind, as long as it is information that could be perceived by an agent as meaningful and thus having musical influence. In this sense we conform to the idea that gestures are information that can carry more information within them (Godøy, Jensenius & Nymoen, 2010). The signals carrying gesture are the pathways that conform the network of agents, which can be abstracted as a topological space (Mazzola & Andreatta, 2007; Van Nort, Wanderley & Depalle, 2014). This space can be called *gestural space*, as the substance of the topology is gesture. Gesture can be observed at any point of the gestural space. In less abstract terms, there is always gesture, and therefore agency, in the various communication channels between those who are making music.

The dimensions of the gestural space can be multiple as they are given by the parameters that agents can modify in the signals they process. For example, parameters can be pitch, level, filter cutoff and so forth. The production of musical sound can be seen as navigating through this space (Choi, Bargar & Goudeseune, 1995; Chadabe, 2002; Schwarz, 2012). The amount of dimensions in gestural space could be large, providing the musician with a rich set of possibilities. This could be the case of a symphonic orchestra or an electronic musical system. However, humans can only handle a limited amount of dimensions in real-time and thus the rich set of possibilities remains feasible only in non-real-time music making such as composing or rehearsing (Tubb & Dixon, 2015).

In greater or lesser amount there are correspondences between the signals of the system,

whether they are auditory, visual, or of any other kind. Several investigations have pointed out the resemblances between bodily movement and music (Godøy, Jensenius & Nymoen, 2010; Toiviainen, Luck & Thompson, 2010; Cox, 2011). For instance, the movement of the human body that produces a sound has remarkable similarities to bodily movement as a response to listening to that sound (Altavilla, Caramiaux & Tanaka, 2013). In the same way the bodily action of playing an instrument has similarities with the sound produced by the instrument (Haga, 2008). Another example is movement helping to visually convey the meaning of the music (MacRitchie, Buck & Bailey, 2013). As phenomena correlates among the gesture space, it is possible to reduce dimensionality to bring the topology closer to a human experience of real-time performing (Arfib et al., 2002; Zappi & McPherson, 2014). An electronic musical instrument that offers a highly dimensional gestural space will implement dimensionality reduction to respond to the ability of a human to handle a limited amount of dimensions in real-time music making. This is a demand imposed by the human agent upon the machine agent.

As a consequence of the exposed analysis, we call *gestural agency* the influence that an agent exerts over other agents within a musical ecosystem. The extent of this influence is a means of power that an agent has on shaping its musical environment, including the behavior of other agents. A conservative view on this has the human in possession of most of the power but we can see that in this system the human is more a participant than a user (Kaipainen et al., 2011). In this way the whole musical ecosystem is enactive as the production of signals is linked to a function, a role of each agent in relation to the other agents (Matyja & Schiavo, 2013).

As the interaction within a gestural agency system demands action to be taken in order to reach a musical goal, agents are necessarily presented with challenges. It has been argued that a musical instrument needs to pose a balanced challenge that is neither too much to be frustrating nor too little to be unappealing in order to be interesting (Wanderley & Orio, 2002; Wessel & Wright, 2002; Levitin, McAdams & Adams, 2002; McDermott et al., 2013; Fabiani, Friberg & Bresin, 2013). This notion conforms to the concept of *flow*, defined as an optimal state of well-being achieved by an activity that provides challenges or opportunities for action accommodating the skill level of an individual (Nakamura & Csikszentmihalyi, 2014). As skills grow, *flow* is maintained by newer and slightly bigger challenges.

In sum, a musical machine demands a certain gesture from the human, to which the human has to respond, thereby changing internal connections to adapt to this requirement. This can take the form of learning to use an instrument, but can also be to learn how to interact with a more independent musical agent. In the same way the machine can be challenged by the human, leading to the reconfiguration of the machine. The *flow* state can be achieved by the regulation of the expectations and tasks by the human. Also this regulation could be carried out on the side of the machine, especially in the case of electronic musical instruments incorporating learning algorithms based on reward (Smith & Garnett, 2012).

Conclusion

We have presented a generalized model for a musical machine that resembles a model for human embodied cognition. Both models depict enactive agents interacting in a musical ecosystem. We have observed that these agents are connected by signals that carry gesture. Gesture is the means that an agent has to exert influence over other agents to produce a musical result. This understanding of musical interaction between humans and machines fits well to traditional and newer technologies for making music. In practice this serves as a framework to analyze musical interactions that integrate humans, traditional musical instruments, and newer electronic musical instruments.

References

- Altavilla, A., Caramiaux, B., & Tanaka, A. (2013). Towards gestural sonic affordances. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (pp. 61–64). Daejeon, Republic of Korea: Graduate School of Culture Technology, KAIST.
- Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial intelligence*, *149*(1), 91-130.

- Arfib, D., Couturier, J. M., Kessous, L., & Verfaillie, V. (2002). Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces. *Organised Sound*, 7(02), 127–144.
- Armstrong, N. (2006). *An enactive approach to digital musical instrument design* (Doctoral dissertation). Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.5347&rep=rep1&type=pdf>
- Bown, O., Eldridge, A., & McCormack, J. (2009). Understanding interaction in contemporary digital music: from instruments to behavioural objects. *Organised Sound*, 14(02), 188–196.
- Cadoz, C., & Wanderley, M. M. (2000). Gesture-music. In M. M. Wanderley & M. Battier (Eds.), *Trends in gestural control of music* (pp. 71–94). Paris, France: IRCAM - Centre Pompidou.
- Caramiaux, B., & Tanaka, A. (2013). Machine learning of musical gestures. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (pp. 513–518). Daejeon, Republic of Korea: Graduate School of Culture Technology, KAIST.
- Caramiaux, B., Françoise, J., Schnell, N., & Bevilacqua, F. (2014). Mapping through listening. *Computer Music Journal*, 38(3), 34–48.
- Chadabe, J. (2002). The limitations of mapping as a structural descriptive in electronic instruments. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (pp. 38–42). Dublin, Ireland.
- Choi, I., Bargar, R., & Goudeseune, C. (1995). A manifold interface for a high dimensional control space. In *Proceedings of the 1995 International Computer Music Conference*, Banff

Centr for the Arts, Canada (pp. 385–392).

Cox, A. (2011). Embodying music: principles of the mimetic hypothesis. *Music Theory Online*, 17(2).

Cumming, N. (2000). *The sonic self: Musical subjectivity and signification*. Bloomington: Indiana University Press.

de Campo, A. (2014). Lose control, gain influence – Concepts for Metacontrol. In *Proceedings of the 2014 International Computer Music Conference*, Athens, Greece (pp. 217–222).

Di Scipio, A. (2003). ‘Sound is the interface’: from interactive to ecosystemic signal processing. *Organised Sound*, 8(03), 269–277.

Fabiani, M., Friberg, A., & Bresin, R. (2013). Systems for interactive control of computer generated music performance. In A. Kirke & E. Miranda (Eds.), *Guide to Computing for Expressive Music Performance* (pp. 49–73). London, United Kingdom: Springer-Verlag.

Fiebrink, R., Trueman, D., Britt, C., Nagai, M., Kaczmarek, K., Early, M., ... Cook, P. (2010). Toward understanding human-computer interaction in composing the instrument. In *Proceedings of the 2010 International Computer Music Conference*, New York, NY, USA (pp. 135–142).

Françoise, J., Schnell, N., & Bevilacqua, F. (2013). A multimodal probabilistic model for gesture-based control of sound synthesis. In *Proceedings of the 21st ACM International Conference on Multimedia*, Barcelona, Spain. (pp. 705–708). New York, NY, USA: ACM.

Godøy, R. I., Jensenius, A. R., & Nymoen, K. (2010). Chunking in music by coarticulation. *Acta Acustica united with Acustica*, 96(4), 690–700.

- Gurevich, M., & Treviño, J. (2007). Expression and its discontents: Toward an ecology of musical creation. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (pp. 106–111). New York City, NY, United States.
- Haga, E. (2008). *Correspondences between music and body movement* (Doctoral dissertation). Retrieved from <https://www.duo.uio.no/handle/10852/26916>
- Haikonen, P. O. (2007). *Robot brains: circuits and systems for conscious machines*. Chichester: John Wiley & Sons.
- Hatten, R. (2012). Musical Forces and Agential Energies: An Expansion of Steve Larson's Model. *Music Theory Online*, 18(3).
- Holland, O. (2007). A strongly embodied approach to machine consciousness. *Journal of Consciousness Studies*, 14(7), 97–110.
- Hunt, A., Kirk, R., & Neighbour, M. (2004). Multiple media interfaces for music therapy. *IEEE MultiMedia*, 11(3), 50–58.
- Hurley, S. L. (2002). *Consciousness in action*. Cambridge, MA, USA: Harvard University Press.
- Jensenius, A. R., Wanderley, M. M., Godøy, R. I., & Leman, M. (2009). Musical gestures. In R. I. Godøy & M. Leman (Eds.), *Musical gestures: Sound, movement, and meaning* (pp. 12–35). New York, NY, USA: Routledge.
- Kaipainen, M., Ravaja, N., Tikka, P., Vuori, R., Pugliese, R., Rapino, M., & Takala, T. (2011). Enactive systems and enactive media: embodied human-machine coupling beyond interfaces. *Leonardo*, 44(5), 433–438.

- Leman, M. (2008). *Embodied music cognition and mediation technology*. Cambridge, MA, USA: MIT Press.
- Levitin, D. J., McAdams, S., & Adams, R. L. (2002). Control parameters for musical instruments: a foundation for new mappings of gesture to sound. *Organised Sound*, 7(02), 171–189.
- MacRitchie, J., Buck, B., & Bailey, N. J. (2013). Inferring musical structure through bodily gestures. *Musicae Scientiae*, 17(1), 86–108.
- Magnusson, T. (2010). Designing constraints: Composing and performing with digital musical systems. *Computer Music Journal*, 34(4), 62–73.
- Matyja, J. R., & Schiavio, A. (2013). Enactive music cognition: Background and research themes. *Constructivist Foundations*, 8(3), 351–357.
- Mazzola, G., & Andreatta, M. (2007). Diagrams, gestures and formulae in music. *Journal of Mathematics and Music*, 1(1), 23–46.
- McDermott, J., Gifford, T., Bouwer, A., & Wagdy, M. (2013). Should music interaction be easy?. In *Music and Human-Computer Interaction* (pp. 29–47). London, United Kingdom: Springer-Verlag.
- Nakamura, J., & Csikszentmihalyi, M. (2014). The concept of flow. In *Flow and the Foundations of Positive Psychology: the collected works of Mihaly Csikszentmihalyi* (pp. 239-263). Dordrecht, Netherlands: Springer.
- Nijs, L., Lesaffre, M., & Leman, M. (2009). The musical instrument as a natural extension of the

musician. Retrieved from <https://biblio.ugent.be/publication/844863/file/944424>

Robb, H. J. (2015). Imagined, Supplemental Sound in Nineteenth-Century Piano Music: Towards a Fuller Understanding of Musical Embodiment. *Music Theory Online*, 21(3).

Schnell, N., & Battier, M. (2002). Introducing composed instruments, technical and musicological implications. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (pp. 156–160). Dublin, Ireland.

Schwarz, D. (2012). The sound space as musical instrument: Playing corpus-based concatenative synthesis. *Proceedings of the International Conference on New Interfaces for Musical Expression* (pp. 250-253). Ann Arbor, Michigan, USA: University of Michigan.

Simoens, V. & Tervaniemi, M. (2013) Musician–Instrument Relationship as a Candidate Index for Professional Well-Being in Musicians. In *Psychology of Aesthetics, Creativity, and the Arts*, 7(2), 171-180.

Smith, B. D., & Garnett, G. E. (2012). Reinforcement learning and the creative, automated music improviser. In *Evolutionary and Biologically Inspired Music, Sound, Art and Design* (pp. 223–234). Berlin; Heidelberg, Germany: Springer-Verlag.

Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied meter: hierarchical eigenmodes in music-induced movement. *Music Perception*, 28(1), 59–70.

Tubb, R., & Dixon, S. (2015). An evaluation of multidimensional controllers for sound design tasks. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, Seoul, Republic of Korea (pp. 47–56). New York, NY, USA: ACM.

Van Nort, D. (2011). Human: Machine: Human: Gesture, sound and embodiment. *Kybernetes*,

40(7/8), 1179–1188.

Van Nort, D., Wanderley, M. M., & Depalle, P. (2014). Mapping control structures for sound synthesis: Functional and topological perspectives. *Computer Music Journal*, 38(3), 6–22.

Wanderley, M. M. (2001). Gestural control of music. In *International Workshop Human Supervision and Control in Engineering and Music*, Kassel, Germany. Retrieved from <http://recherche.ircam.fr/equipes/analyse-synthese/wanderle/pub/kassel/>

Wanderley, M. M., & Orio, N. (2002). Evaluation of input devices for musical expression: Borrowing tools from HCI. *Computer Music Journal*, 26(3), 62–76.

Wessel, D., & Wright, M. (2002). Problems and prospects for intimate musical control of computers. *Computer Music Journal*, 26(3), 11–22.

Wheeler, M. (2005). *Reconstructing the cognitive world: The next step*. Cambridge, MA: MIT press.

Zappi, V., & McPherson, A. (2014). Dimensionality and appropriation in digital musical instrument design. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (pp. 455–460). London, United Kingdom: Goldsmiths, University of London.