

Interactive Music: Social Considerations of Gesture & Vocality

ARTAUD Forum 2: Konnecting Gestures, International Conference-Workshop on Performance and Sound Technologies, 30 March 2012

Antonin Artaud Performance Centre, Brunel University, Uxbridge, West London (UK)

Dr. Pieter Verstraete

In this presentation, I would like to critically reflect on what interactive music is about and how it has developed in recent years. Among interactive music, I include what is often called Interactive (Computer) Music, Interactive Gesture Music (IGM) or ‘gesture controlled’ music, ‘multimodal’ (interaction) environments (MEs) and live electronics.

Central are two small case studies of interactive music performances that share some features in gestural control and composition through, but are also very different in many ways: Swiss interactive vocal performer, singer and sound artist Franziska Baumann’s *Electric Renaissance* and German singer and composer Alex Nowitz’s *Studies for Selfportrait*. Both are (or were) artists in residence at STEIM in Amsterdam, which is an independent live electronic music centre for the performing arts based in Amsterdam. Both are classically schooled as singers (Baumann at Winterthur Conservatory and Nowitz in Berlin, Munich and the University of Potsdam). Both use live electronics particularly based on live controlled audio feedback, segmented into sequences, which allows them to fully explore the sonorous qualities, particularly the gestures, in their own voices. Baumann uses a Sensorglove (or cyberglove), an interface custom-designed at STEIM, as well as a sound dress with inbuilt speakers to manipulate the interactive system. Nowitz uses the *Stimmflieger* and the LiSa live sampling software (also produced by STEIM), which he manipulates by means of two Wiimotes (the low-cost game controller of the Nintendo Wii). (Let’s have a quick look at both performances first so that we know what we are talking about.)

Rather than discussing these clips at great length, I would rather focus on the theoretical underpinnings of my own considerations regarding these performances, which leaves more room for discussion. The second clip does bring up an interesting development, which developers envisioned decades ago but which only slowly finds resonance in the commercial industries with corporate interests: the accessibility of such interactive media, interfaces and tools for public use and musical training. This has particularly social consequences for how we think in terms of bodily and musical gestures and how we appreciate or depreciate such interactive music performances. For this presentation, I would therefore like to critically focus on related issues of interactive music technology, which have an inherent social interest, particularly:

- the computer layer vs. the cultural layer;
- the nature of gestures (instrumentalisation & overdetermination);
- how the interplay of sonic and bodily gestures can induce a sense of vocal bodies.

In previous research, I discussed the oral and literate modes of listening that these performances may elicit in the listener's search for meaning. I also studied the difference in embodied experiences (in particular the contrast between Cartesian eye and embodied ear) from the perspective of performers and spectators in the related art form of interactive dance. But I thereby ignored the specific nature of visible and audible gestures in performance and the possibility of a combination of first and third-person perspective by the audience (as suggested by Marc Leman of the IPEM Institute in Ghent). I will focus on these issues in more detail later. Although all of these issues are interconnected, they do not immediately imply a coherent argument for which I apologize. I do invite you to think along with me and I hope that the issues I will raise here will stimulate a critical debate on where interactive music stands today and which directions it could take further.

One note, though, on methodology. Because of the complex nature of the computer-assisted coupling of bodily and musical movement into gestures, the study of interactive music needs an interdisciplinary approach, encompassing approach:

- Musicology and theories of music perception;
- (Neuro-)cognitive science and psychology;
- Ergonomics and the study of body movements (human movement science) which are mainly preoccupied with how to make human movement better;
- Dance with Laban's theory of effort and his notion of the kinesphere as most central (i.e. an imaginary box surrounding a person that defines the maximum movement possibilities from a certain point in space, a mental construct of which one is always aware when interacting with the environment and others);
- Media archeology and the history of modern media cultures;
- Computer programming, computer science; to name just a few.

I am in no way an expert on either field, nor would I be claiming so. My interest lies purely in the social and cultural aspects to understand and describe the phenomenon of interactive music. But I would not dismiss interdisciplinary aspects and concepts from the mentioned fields of study and I welcome the discussion from multiple perspectives.

Starting with Some Critical Remarks

Let's start with an initially impulsive gesture: our critical considerations with what

interactive technology promises to do. Here is an abstract of a paper in the *Proceedings of the 1995 International Computer Music Conference* by Todd Winkler from Brown University:

The increase in sophistication of new devices that allow gesture and movement to be translated into computer data holds great promise for interactive composition, dance, and creating responsive music in virtual reality systems. Data describing human motion can produce musically satisfying results by their impact on sound and musical processes.

The abstract has clearly been written in vein of Alan Kay's 1989 optimism and Marshall McLuhan's aphorism that computer technology would not only be a tool but an extension of man, a vehicle for creative and artistic expression. Indeed, interactive systems are improving in capturing, recognizing, analyzing and mapping motion data in much more refined ways for creative communication. As we just saw in the performances by Baumann and Nowitz, there are indeed unexplored musical qualities that can influence compositional as well as listening structures. However, what have been more artistically interesting are the restraints of these systems. The restraints are the result of the predetermined range of the system, which provides a matrix and a space to explore the boundaries of where the system fails but likewise, of the capacities of the body and particularly the vocal apparatus in response to its immediate environment through the system's capacities and constraints. So the great promise of 'musically satisfying results' is often of a minor concern in these boisterous compositional experiments, challenging thereby the borders of the listener and her modes of listening.

Moreover, this begs the question of what precisely a 'satisfying result' is in an artistic context that is still little known to the wider public. There is a common belief among performers that interactive music succeeds when it encourages spontaneity while residing within the boundaries of a dynamic artistic context that is whole and engaging. However, as is a rule of thumb in postmodern performance, the peripheries of those boundaries are equally interesting. This necessitates that both performers and developers need to study the nature of human gesture and its laws in order to tamper with it in those virtual environments for provocative and artistically intriguing results. And this is also the thrill of interactive media: there are at least as many different sensors on the market now as there are distinguishable gestures which produced rich and intricate typologies (roughly body-attached sensors, spatial sensors among which those indirectly in touch with the body such as pressure sensor and detached sensors such as ultrasound sensors, infrared sensors, camera-based sensors; some might be supporting MIDI). Yet it is usually the unnatural, unexpected correlation between gesture and sound that offers the most interesting and meaningful experiences.

There is also the question of ‘sophistication’ in the above abstract. Indeed, the world only slowly caught up with the first interactive instruments, the Theremin developed by Léon Theremin in 1919 followed by the Ondes-Martenot developed by Maurice Martenot around 1928, both allowing by electrical means to produce sound without physical touch. These low-tech predecessors of interactive systems today gave rise to experimentation and a whole spectrum of new radar-like instruments in the 1960s and 70s when several composers rediscovered movement as a basis for exploring electronic music, thus turning space in a virtual musical instrument. But it was only from the 1980s with the rise of the (personal) computer as a multimedia instrument and the accuracy of digital technology that new motion-based interfaces started to be developed. However, the so-called ‘sophistication’ of these interfaces did not always match the interests of the experimental base, such as the Nintendo version of the sensorglove, namely the PowerGlove (produced by Mattel), demonstrates. (Show clip) Although a symbol for the potentially unrestrained capacities of modern, computer-based technology and the concomitant thrill of the boy in us that says ‘it’s so bad’, today we look with nostalgia and a healthy dose of skepticism to the grasp of the PowerGlove into its immediate future.

Although such new interactive instruments are around for almost a century, the sophistication comes along with a great deal of specialization and inaccessibility due to the high expense of these still developing technologies. Corporate interests such as Nintendo’s with the recent Wii game console and the Wiimote control (which Nowitz uses) are pushing their own research agenda’s, which are not immediately following the interests of experimental performance but which might offer a solution to democratize and thereby consolidate the art form in terms of technical schooling of specific skills and musical appreciation by the wider public. Yet, at the moment interactive music as an art form is still living of the mystery of interactivity that is the privilege of the have’s over the have-not’s. Performers need to be engineers, computer programmers or they are heavily dependent on specialized centres and institutes for the production and maintenance of custom-designed software and interfaces such as STEIM in Amsterdam, Logos in Ghent, IRCAM in Paris, MIT Media Lab in Boston. For instance, IRCAM has an extensive online database of sensor classifications (including cost, suitability, availability) but is only accessible to IRCAM members. Equally, the underlying studies of gesture controlled environments is restricted to highly specialized academic departments and research institutes such as IPEM in Ghent University (Marc Leman), ICSRiM at University of Leeds (Kia Ng), the dance department at UIUC in Illinois (John Toenjes), MEME at Brown University (Tod Winkler), the InfoMus Lab at DIST-University in Genova (Antonio Camurri, who is co-author of the famous EyesWeb), most of which have a predominant interest in empirical research to make better systems. The

concurrent technical language often heightens the sense of inaccessibility of these technologies to the layman musician.

This should not be so much of a problem as any kind of professional music implies a level of sophistication and even ‘fetishization’. Yet it seems that the sophistication of such ‘invisible’ technology also takes away some of the labour and zeal of the traditionally trained musician, who in many cases, has spent hours of practice time, developing body alterations such as calluses “ – as a sort of ‘filter’ – in order to tune the bandwidth of the tactile channel” so to speak (Rovan & Hayward 6). As with every new media something gets lost, it is certainly the physical, predominantly tactile labour that seems to have been replaced by technological sophistication. However, the constant remapping and tuning of body and machine – such as in the case of Baumann’s sensorglove – increasingly searches for more tactility in the experience and responsivity to human touch, thereby highlighting the aspects of embodiment that traditional musical instruments seem to have lost in our perception and Western musical culture. As such, if new interactive instruments are re-inventing the wheel, there seems to be a socially and cultural necessity for it.

The Computer Layer vs. the Cultural Layer

This brings us to a well-known and perhaps now outmoded opposition, formulated by Lev Manovich in *The Language of New Media* (2001) as a contrast between the computer layer and the cultural layer which seem to be also at odds in our perception of interactive music environments.

Motion capture by sensors and interfaces treats human gestures in the first place as patterns. Interactive systems consists generally of a gesture recognition engine (a gestural interface) and a multimodal mapping engine (usually a sound/feedback engine) supported by ideally real-time, multilayered modular software to control and generate audio (and sometimes also visual) feedback. Well-known examples are Camurri’s EyesWeb, STEIM’s BigEye, Rokeby’s Very Nervous System.¹ Such modular systems ‘transcode’ the movements into streams of data and, from there, into other formats. Most of them do not necessarily incorporate an understanding of human movement, but rather depict information through camera-based motion capture at the low level of pixel variations. It is only the human agent in these environments who re-interprets the output in terms of gestures as part of her cultural understanding and making sense of the world.

¹ Other systems include Big Briar’s Theremin, Buchla’s Lightning II, Interactive Light’s Dimension Beam, the Polhemus device, Palindrome Inter-media Performance Group’s Dynamic Fields, WaveRider, BodySynth and BioMuse (Rovan and Hayward 7).

The virtual space of the computer system is, however, purely biomechanics of algorithms, which constitutes the computer layer, according to Manovich. This layer is similar to Vivian Sobchak's notion of the electronic space as opposed to cinematic space, which must have also inspired Manovich: a space which constructs "objective and superficial equivalents to depth, texture, and invested bodily movement" (Sobchak 2000: 149). According to Manovich, the transcoding into the realm of the computer layer – with its own epistemology, ontology and pragmatics – is part of a global process of cultural reconceptualization (Manovich 47).

The new blend or 'composite' of cultural and computer layers in new media presupposes opposing – perhaps even mutually exclusive – spaces which is, however, flawed since both layers are discursively produced in relation to the same cultural repertoire of the perceiver. True, the binary can explain for the constant negotiation in our experiences of interactive music performances between our musical tastes, sense of structure and the impersonal computer protocols of which only the dancer seems to have ingrained knowledge after hours and hours of training but which remains hidden for most layman spectators/listeners. Yet the performer's knowledge of the system as well as the listener's experience of the system as a controllable instrument or 'hyper-instrument' in this case – and in most cases, as an adaptable agent – makes the computer layer inherently part of our discursive – and therefore, culturally based – search for meaning. As such, these interactive systems seem only to confirm and perhaps to makes us more aware of our human insistence on meaning, even when we flirt with the literally 'empty' signifying computational system based on 0s and 1s.

The Nature of the Gesture

This brings us to the most fundamental issue of interactive music performance: meaning and gesture. When we look at the performances of Baumann and Nowitz, despite their inherent differences in their use of interfaces and compositional ideas, one question begs asking: do you primarily look at what the gestures *mean* or what they *do* to you? Is there a separation between meaning and doing?

Theories of performativity would suggest – contrary to Wittgenstein's separation between music and language – that musical gestures are similar to the gestures of language in the way they make us do things. They make us respond in our ways and attitudes of looking and listening. In this sense, gestures are the quintessence of our positions and responses that inhere in our pursuit for meaningful relations. And this is what gestures in performances such as Baumann's and Nowitz's can make us aware of: a questioning of the nature of body and sonic gestures as meaningful to our experiences. It

is the embodied nature of the ‘gesture’ that makes the correspondence between physical motion and music possible. Gesture is then part of the discursive matrix that makes any social communication, be it through body or sonic gestures, possible. Be it sound-accompanying, sound-supporting or sound-producing, gesture always requires in some way a carrier of expression and meaning, something that forces us to understand the gesture as movement in a meaningful way. As a vehicle of meaning, gesture implies social interaction and its central role to the way we communicate – there can be no music without gesture (Cadoz 1988: 64), even in the most formalistic, computer-layered play with audio feedback. Through perceiving gestures, we impose culturally related actions to movement.

Moreover, it is commonly accepted that gesture is also used to denote our embodied engagement with our direct environments. The sensorglove and the Wiimote are different interfaces, which provide different affordance structures – to borrow Gibson’s ecological notion – to the performers to interact with their environments. The resulting gestures of their highly personal interactions also call for a different engagement from the audience. As I mentioned in the beginning, in previous research I questioned the possibility of a first-person, embodied experience of the audience, which rather dwells in a ‘floating ear’ and ‘mirrored eye’ “that specializes in the spatialisations of a ‘pure, image-system’” against the embodied ear of the dancer or performer, which “privileges corporeality, verticality of being, collective experience” (Kroker 1988: 93). However, such categorical distinctions fail to explain the precise mechanism of the body and sound gestures that elicit such specific modes of looking and listening. Because of its appeal on embodiment, interactive music performances demonstrate how gestures have the capacity to communicate or elicit a point of view.² The original meaning of Latin ‘gestus’ (from the verb ‘gesture’) is quite telling as it denotes posture or pose (indicating certain types of expressive behaviour), but also attitude. Another meaning of the verb ‘gesture’ also directs us to the illocutionary power of gestures: wishing, desiring, craving for. Gestures call for us to position ourselves.³

² This is the expressive power of gesture which comes close to, for instance, Kurt Weil’s reference to ‘gestische Musik’ in his collaboration with Bertolt Brecht in 1927, highlighting the political potential of musical gesture.

³ In some rare cases, the proprioceptive awareness can direct us to a sense of egolocation or even ‘cenaesthesia’. *Egolocation* refers to the awareness of one’s overall position within a defined space, or with respect to objects in that space. This primarily holds for the performer, who through proprioceptive feedback receives the ability to orient, critique and calibrate her performance gestures (Rovan & Hayward 4). But it could equally include the listener when, in most silent ways, the bodily and sonic gestures make her aware of her own position. *Cenaesthesia*, on the contrary, happens when the interactive music calls for a social “sensation of existence, caused by the functioning of the internal organs” (Leman 150). Related terms are ‘sensus communis’, ‘soatognosis’, ‘Gemeingefühl’.

In that sense, I am tempted to accept that a combination of first-person and third-person perspective is highly possible in interactive music performances, even fundamental if they mean to bridge the Cartesian gap between body and mind. By highlighting the gestures in a sense of overdetermination and instrumentalisation, they reveal their in-betweenness as both mental and corporeal. As such, both virtual and physical aspects of the concert space could appeal to a proprioceptive experience, a sense of one's sensing, that is not 'free floating' but anchored in our meaningful responses to the implied perspectives that the gestures call for.

The Imagination as Social Gesture: Voice-Bodies and Gender Critique

The question remains if in interactive music's call for one's positioning through action-oriented third-person and experience-based first-person perspectives, one could also conceptualize a second-person point of view. According to Leman, in this particular perspective "gesture appears as a mediator for music-driven social interaction or as the vehicle through which a 'me-to-you' relationship is established in space and time, through musical engagement" (Leman 142). It is the perspective par excellence that enables gesture as communicative act, social interaction, entrainment and intentionality. "(I)t is the mirroring through which 'my' perception of 'your' movement is grasped in terms of the (overt and covert) deployment of 'my own' body movement in the environment, so that 'your' movement is corporeally understood as an action" (Leman 143). It explains for the kind of mental dance that we experience through our bodies when reacting to other bodies through physical/visual and possible also sonic gestures.

Now, I would contend that it is in this empathic second-person relation of 'me-to-you' that gestures can contribute to the sense of vocalic or voice-bodies, a term coined by Steven Connor to explain how we imagine virtual bodies to the voices we hear in combination with physical gestures. I realize that this is the most subjective, hypothetical and thereby weakest point in my considerations of the social implications of interactive music. Yet I feel strengthened by the comparison between Baumann's and Nowitz's performances of its potential and leave this to your own good judgment.

Connor's notion of the voice-body is to describe the composite experience of voice and gesture. As a result of the voice's atypicality and 'disembodiment' once emitted into space, we lend it a body: "For voice is not simply an emission of the body; it is also the imaginary production of a secondary body, a body double: a 'voice-body'" (Connor 2004, 158; see also Connor 2000, 35-42). The disembodied voice invites the listener to conjure an imaginary body, as Connor suggests, that "may contradict, compete with, replace, or even reshape the actual, visible body" (Connor 2000, 36). This

imaginary body is an ‘invocation’, so to speak, which includes certain postures, gestures, facial expression, movements of the shoulders, hands, arms, such as the physical gestures in interactive music performances as well as the sonic gestures of the disembodied voice:

The leading characteristic of the voice-body is to be a body-in-invention, an impossible, imaginary body in the course of being found and formed. But it is possible to isolate some of the contours, functions, and postures by means of which vocalic bodies come into being (Connor 2000, 36).

Connor distinguishes the voice-body most significantly from the voice *in* the body, as the former incorporates all of the gestures that we ascribe to the utterance of the voice:

[T]he voice also induces and is taken up into the movements of the body. The face is part of the voice’s apparatus, as are the hands. The shaping of the air effected by the mouth, hands and shoulders marks out the lineaments of the voice-body (which is to be distinguished from the voice in the body). When one clicks one’s fingers for emphasis, claps one’s hands, or slaps one’s thigh, the work of gesture is being taken over into sound, and voice has migrated into the fingers (Connor 2004, 163).

The voice, in its simultaneous bodily and sonic gestures, may then produce a second body, which speaks to us, addresses us in our imaginations, as a ‘phantasm’. This voice-body restores the body to a voice in an urge to place it in relation to ourselves, to our own listening bodies. In my view, this idea corresponds to Leman’s second-person perspective which related my body to yours in the physical embodiment and reading of gestures. This new voice-body that we imagine is not only a mental construct, but also a result of the mirroring activity as part of our own embodiment in our response to the bodily and sonic gestures.

When we relate Connor’s principle of the voice-body to Baumann’s and Nowitz’s performances, we could describe how the bodily and sonic gestures give rise of imaginary bodies that surpass the body images and visual postures that define our relation to these performers. On a metaphorical level, one could see in Baumann’s performance a highly feminine, Gothic figure, heightened by the long black glove and the sound dress, which acts as a veil between the physical world and the imagination, between the present and the world of the baroque (supported by phrases by Haendel). This stands in contrast to Nowitz’s rather sluggish appearance as a man letting the boy within him come out whilst playing with the Wii controllers as if in the middle of a pinball game. I can hear him almost say: ‘I love the Wiimote. It’s so bad!’ The sonic gestures are also quite

crucial. Whereas Baumann accentuates and then gently deconstructs her soprano voice, Nowitz seems to immediately go for an exploration of extended vocal noises. Both performers seem to have developed their own personal palette of extended vocal techniques in relation to the respective interactive technologies. Through the interactive systems, they express their 'selves', their auditory fingerprints whilst controlling them and altering them through the audio feedback.

The aspect of gestural control is also significant in the imagination of voice-bodies. Baumann's control seems to master the audio feedback with a soft touch. The sensorglove gives her control over a wide spectrum of parameters, whereas some of the gestures seem to be added for dramatic effect. Nowitz's rapid movements of the Wiimotes also highlight the sense of control. But his movements are often almost robotic configurations of movement, which sometimes look like puppetry or juggling as if he is trying to keep the invisible sounds into a defined frontal space which he can manipulate. In a reciprocal way, Nowitz constantly tests the distribution and distortion of his disembodied vocal sounds as auditory self-images, while constantly overlaying them with new disembodied sounds.

But the voice-body is more than such semiotic and metaphorical reading of vocal gestures, postures and body images. The voice-bodies come into play when we relate our own bodies to the physical and virtual ones we perceive in these performances. As such, voice-bodies remind us how the imagination is always embodied, never the result of a mental coupling alone. It is therefore also always inherently social. It is as such that interactive music performances can have relevance for future audiences.

This play with instant sound through interface control produces an interactive space that is both fascinating to listen to and to watch, as we assess the sounds and the creator's gestures. It produces a highly focused space that calls attention to the performer's body as well as to our own bodies, as they 'feed back' to the performing bodies from which the sound and musical gestures once originated. It is in this both imagined and embodied space that interactivity can do something meaningful.

Ankara, 29 March 2012
Dr. Pieter Verstraete

References

- Claude Cadoz. "Instrumental gesture and musical composition". In: *Proceedings of the International Computer Music Conference*, 1988: 1-12
- Connor, Steven. *Dumbstruck: A Cultural History of Ventriloquism*. Oxford: Oxford University Press, 2000.
- . "The Strains of the Voice". *Phonorama: Eine Kulturgeschichte der STIMME als Medium*. Catalogue. Zentrum für Kunst und Medientechnologie Karlsruhe, Museum für Neue Kunst, 18 Sept. 2004 - 30. Jan. 2005. Ed. Brigitte Felderer. Berlin: MSB Matthes & Seitz, 2004: 158-172.
- Godøy, Rolf Inge & Leman, Marc, eds. *Musical Gestures: Sound, Movement, and Meaning*. New York, NY: Routledge, 2010.
- Kroker, Arthur & David Cook. *The Postmodern Scene: Excremental Culture and Hyper-Aesthetics*. London: MacMillan Education, 1988.
- Leman, Marc. "Music, Gesture, and the Formation of Embodied Meaning". Godøy & Leman 2010: 126-53.
- Manovich, Lev. *The Language of New Media*. Cambridge, Massachusetts: MIT Press, 2001. (= Leonardo Books)
- Rovan, Joseph & Hayward, Vincent. "Typology of Tactile Sounds and their Synthesis in Gesture-Driven Computer Music Performance". *Trends in Gestural Control of Music*, Wanderley, M. & Battier, M., eds. Paris: Editions IRCAM, 2000.
- Sobchack, Vivian. "The Scene of the Screen: Envisioning Cinematic and Electronic 'Presence'". *Theories of the New Media: A Historical Perspective*. Ed. J.T. Caldwell. London: Athlone Press, 2000.
- Verstraete, Pieter. "Interactive Dance: Looking at space through the eyes of the matrix". *Performing the Matrix: Mediating Cultural Performances*. Eds. Meike Wagner & Wolf-Dieter Ernst. Munich: Epodium Verlag, 2007: 157-73. (= Reihe Intervisionen, Vol. 8)
- Winkler, Todd. "Making Motion Musical: Gesture Mapping Strategies for Interactive Computer Music". In: *Proceedings of the 1995 International Computer Music Conference*. Ed. E. Michie. San Francisco: International Computer Music Association, 1995.