

# UC San Diego

## UC San Diego Electronic Theses and Dissertations

### Title

To un-button : strategies in computer music performance to incorporate the body as re-mediator of electronic sound

### Permalink

<https://escholarship.org/uc/item/9t03q218>

### Author

Oliver La Rosa, Jaime Eduardo

### Publication Date

2008

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

**To Un-button: Strategies in Computer Music Performance to Incorporate the  
Body as Re-Mediator of Electronic Sound**

A thesis submitted in partial satisfaction of the requirements for the degree  
Master of Arts

in

Music

by

Jaime Eduardo Oliver La Rosa

Committee in charge:

Miller Puckette, Chair  
Philippe Manoury  
Shlomo Dubnov

2008

Copyright

Jaime Eduardo Oliver La Rosa, 2008

All rights reserved.

The thesis of Jaime Eduardo Oliver La Rosa is approved and is acceptable in quality and form for publication on microfilm:

---

---

---

Chair

University of California, San Diego

2008

## DEDICATION

to my wife...

## EPIGRAPH

It is through my body that I understand other people; just as it is through my body that I perceive 'things'. The meaning of a gesture thus 'understood' is not behind it, it is intermingled with the structure of the world, outlined by gesture, and which I take up on my own account. It is arrayed all over the gesture itself (Merleau-Ponty 1962:186)

## TABLE OF CONTENTS

Signature Page . . . . .	iii
Dedication . . . . .	iv
Epigraph . . . . .	v
Table of Contents . . . . .	vi
List of Figures . . . . .	vii
Acknowledgements . . . . .	viii
Abstract . . . . .	ix
Chapter 1. Introduction . . . . .	1
Chapter 2. Instrument - Gesture - Sound . . . . .	4
Chapter 3. Gesture, the Studio and the Computer . . . . .	18
Chapter 4. 'Live' Performance, the Body and Concert in Computer Music . . . . .	29
Chapter 5. Conclusions . . . . .	54
References . . . . .	56

## LIST OF FIGURES

Figure 2.1. Model of a Digital Musical Instrument. . . . .	12
Figure 4.1. Acousmonium. . . . .	30
Figure 4.2. Circularity in Performance. . . . .	38



## ACKNOWLEDGEMENTS

I gratefully acknowledge all faculty and students in UCSD for their contribution to my knowledge.

ABSTRACT OF THE THESIS

**To Un-button: Strategies in Computer Music Performance to Incorporate the  
Body as Re-Mediator of Electronic Sound**

by

Jaime Eduardo Oliver La Rosa

Master of Arts in Music

University of California, San Diego, 2008

Miller Puckette, Chair

This thesis reviews the current literature for instrumental gesture and digital musical instruments. It then elaborates on the gestures of electronic and computer music in the studio and as a result of algorithmic procedures as well as briefly review the composer-performer-listener model of music production. This is followed by a brief history of performance and the concept of 'live' electronic music, from which a theory is derived whereby computer music is incorporating the body as its primary means of expression, elaborating on its perceptual impacts.

# Chapter 1

## Introduction

Before recording and sound synthesis devices, music performance consisted almost exclusively of instruments performed by people. With the advent of music technology, the practice of public audition of music started undergoing a set of radical changes, from having no people in stage to having no concert hall at all. Computers, machines that have been capable of real-time production of high quality sound for a few decades, have introduced radical changes in the performance of computer/electronic music that we are still exploring and coming to grasp.

The sound generating possibilities that music technology opened up in the middle of the 20th century changed the way we understand, produce and consume music and sound. Computer or electronic textures and timbres have had a tremendous impact on the way we understand and process gesture, gradually distancing from the culturally embedded to the unknown.

Tape music, either as a product of manipulation of sound recordings or through the synthesis of sounds through the electronic medium, opened up a world of sounds impossible to imagine in the strictly acoustic world. With the use of computers, the ability to represent gestures as functions of time or computer automations as well as the development of other sound generating techniques native to the new medium, these new gestures and organizational possibilities increased. This new sound world developed

its own gestures, shaped at first by the human gestures that produced the sounds in the studio, although usually unrecognizable due to the complex layering process and tape construction procedures, thus creating remote and surrogate gestures.

However, the fast emergence of the fields of controller design and HCI, act both as a need to feed computers with live human gestures that control sounds, but also to provide the computer with elements of decision to generate interaction. The concept of live performance of technologically created or mediated material has become an issue of interest not only in the music discipline, but also in other performative arts. New ways of presenting new media are continually emerging and evolving.

These new practices, namely, performance with gestural controllers and interactive computer systems will be analyzed as attempts by composers and performers to re-incorporate the body as principal medium of expression and as mediator between the known acoustic instrumental world and the unknown electronic sound. There is a substantial body of theoretical reflection outside the computer music discipline, which can help understand this and which will be reviewed.

This thesis is an exploration of the issues surrounding the performance of live computer music as it is evolving, evaluating the impact in the traditional composer-performer-listener model, from the point of view of humans who practice and listen to electronic and computer music.

The second chapter reviews some of the literature on the conceptualization of acoustic instruments, instrumental gesture and digital musical instruments. The third chapter explores the new gestures that the electronic music studio and the computer created as a result of the new musical practices that emerged from that technology mainly in acousmatic music. The fourth chapter, explores briefly the history of 'live electronics', exploring the concepts of interactionism in cognitive science, our understanding of the word 'live' and the impact of the body in the perception and performance of music.

This thesis, does not pretend to solve the problems of electronic music performance, but to explore them. As Heidegger once said, we can not know the forest, we can only walk through its paths. And it is in this spirit that certain theorization needs to

be achieved to understand why we are doing what we are doing.

## Chapter 2

# Instrument - Gesture - Sound

### *Acoustic Instruments*

Musical instrument is a self-explanatory term for an observer in his own society (Kartomi, 2008)

We understand acoustic musical instruments as organized physical objects that are played to produce musical sounds. From this definition, we can extract the idea that a gesture is performed on an object to obtain sound. We cannot imagine an acoustical musical instrument without a human performing it, that is, without human actions that bring out the potential sound that the instrument affords.

We can think of an instrument in terms of its material qualities and the sound it makes. When seeing a clarinet we imagine its sound, and when we hear a clarinet we imagine it being played.

This sound has a quality that varies with playing technique and material characteristics, but exists within a well-defined set of boundaries that makes it impossible, for example, for a clarinet to make the sound of a guitar and for us to confuse it with one.

We usually refer to the sound produced by the clarinet with common playing technique as ‘the timbre of the clarinet’ and to alternate ways of playing it as ‘extended techniques’. These gestures are a set of movements that include those of the respiratory system, mouth, arms and fingers. Specific configurations of these body parts produce a

particular sound.

The instrument is then defined by its material limits as well as by the human limits of the performer. The material limits are somehow clear, acoustic instruments are acoustic because the sounds they produce are a physical acoustic phenomenon. These organized physical objects have the function of: (1) transducing a gesture into vibrations; like a plucking a string, blowing into a reed or hitting a membrane; vibrations that become sound waves. (2) Other objects like keys and frets change the size of these waves for pitch control. Finally, some sort of (3) resonating body amplifies and models the sound.

In the words of Choi (2000) “Musical Instruments provide physical action-spaces with auditory affordances”. Instrumental performance is then a negotiation between what the instrument affords in terms of acoustic phenomena and the extent to which a performer controls/explores these affordances. In performing the instrument, the performer needs to exert physical energy upon it, resulting in various types of proprioceptive movement. We call this movement instrumental gesture.

Most occidental musical instruments have reached a level of standardization that allows performers to perform a wide repertoire of music that can be composed with certain knowledge of its performing boundaries. This standardization consists of a specific organization of its material elements (resulting for example in specific tunings, range, etc.); all of these have a reflection on technique.

The consequences of standardization have shifted our cultural understanding of instruments in several ways. Culturally, we understand instruments through their repertoire and it is this repertoire that seems to determine what is and what is not an instrument and how a particular instrument should be played. Contemporary music practice has significantly extended what we understand as music, and therefore, the practices that define what an instrument sounds like.

Standardization has also brought a view of instruments that are perfectly deterministic, ones that are defined by the high predictability of its output relative to a performer’s controls. (Chadabe, 2002)

Finally, the small variation from instrument to instrument has led us to define instruments as “singular physical objects” in Schick’s words.

Percussion practice extends and challenges some of our ideas of what instruments are, especially those ideas that stem out of standardization. “There is not a single instrument that defines percussion playing in the same way that the piano, for example the singular physical object of the piano, ubiquitous and universal defines piano playing” (Schick, 2006)

In Schick (2006), the author states the impossibility of defining percussion through a single physical object or through the reduced contemporary solo percussion repertoire and proposes the following: “The most succinct definition of percussion comes from the German, Schlagzeug; Schlag means ‘hit’ and Zeug means ‘stuff’ ”

This definition is particularly interesting for it has the duality we stated to begin with. An instrument is an object-gesture compound. It is interesting how sound is implicit when defining this sound producing activity. What we usually call musical instruments are object-gesture compounds that have been standardized. Marimbas and vibraphones, sets of bars of similar material and arranged in ascending size, are clear examples of standardized percussion instruments.

This brings us to the concept of percussion setup. This concept is very common amongst percussionists. It consists of arranging percussion instruments within the reach of the percussionist; a sort of modular construction of a supra-instrument; an instrument made of instruments.

Composers have gladly adopted this practice. When Stockhausen and Boulez talked of composers not only organizing sounds, but also making these sounds themselves, a new perspective on music making was being forged. The construction of the percussion setup seems to be in the same spirit, the construction of the percussion meta-instrument allows constructing an instrument for a specific piece; choosing the sounds an instrument will have is in a similar spirit as making the sounds.

This concept is however not new, the orchestra, string quartet, wind quintet are clear examples of standardized meta-instruments. These supra-instruments could how-



ever be considered simply an instrument; a violin is an element of the string quartet and a string an element of that violin. It is still an object-gesture compound and the art of writing string quartets, from Haydn to Lachennmann, provide us with the palette of gestures that are possible. In any event, the practice of choosing instruments for a piece grew very quickly in the 20th century in comparison to previous ones, where standardized formations were the rule.

### *Gesture*

In our definition of acoustic musical instruments as object-gesture compounds we made a brief definition of instrumental gesture as proprioceptive movement of a performer that brings out the sound potential in an object. In this way, gestures are dependent on the instrument that is being played and the context of a score or plan. In this section I will attempt to review some ideas and conceptualizations about gesture, emphasizing certain facts that pertain to percussion playing and computer tracking strategies.

The study of instrumental gesture in music brings out a consciousness of the body that isn't present in the traditional analysis of music. Gestures seem to be studied and classified in terms of bodily affordances and types of movement or in terms of the function they perform.

Cadoz and Wanderley (2000) reviewed the current literature on gesture, serving as an excellent guide for this section on gesture. Bonnet et al. (1994) recognize two kinds of motor units and corresponding movements. Slow motor units and movements are related to posture and fast motor units and movements, like fingers or hands, related to gestures such as strikes. Goldstein proposes a similar idea with the terms "Current control" and "Ballistic control":

The root of all gesture is muscular action, and neurophysiologists have recognized two types of muscle control [3]. "Current control" movements are sustained and can be changed while they are being performed. "Ballistic control" movements are short and energetic. They send a limb on a trajectory while the ballistic movement itself ends before the limb has completed its action. (Goldstein, 1998)

This brings us to the question of gesture and posture:

The coordination between posture and movement condition the efficiency of the gesture. In fact, postural muscular activities, static or dynamical, anticipate, accompany and follow the execution of the movement in an automatic manner. (Bonnet et al., 1994)

The question of posture and gesture, specifically in percussion performance is of importance in that it helps us determine what to track. There are many approaches to tracking the movement of a performer that include the tracking of several limbs. However, from this idea we can derive that the movement of a mallet head, hand or finger contains somehow all the other movements including posture; as if it was a derivation or final expression of the general movement of the body towards the gesture that actually produces the sound.

Viviani (1994) states a very interesting fact: the bigger the curvature of a movement, the slower its speed. This fact, especially in percussion will give us two basic types of gesture as extremes in the speed continuum, where slow movements produced curved gestures, fast ones linear ones and combinations of these in between.

Choi (2000) proposes the idea of gestural primitives as fundamental human movements that relate the human subject to dynamic responses in an environment, presented by the author as an attempt to bring a formalization of human motion in terms of performance gesture into the computable domain.

Choi recognizes three types of gestural primitives: (1) Trajectory-based primitives, (2) Force-based primitives and (3) Pattern-based primitives. These are related to rotation or changes of orientation, gradient or linear changes and period, respectively. Choi understands these primitives as a set of resources that can be used across instruments and as substructures of individual gestures or gesture sequences. He exemplifies it in the following way

... a trill may be performed as a gradient event by its crescendo/decrescendo property, a force-based primitive. However, a trill could also be performed emphasizing the rate of repetition of individual notes, a pattern-based primitive.

In a more functional approach, Delalande (1988) classifies gestures in three levels “from purely functional to purely symbolic”:

- Effective Gesture - necessary to mechanically produce the sound  
bow, blow, press a key, etc.
- Accompanist gesture - body movements associated to effective gestures - chest, elbow movements, mimics, breathing for a piano player, etc.
- Figurative Gesture - perceived by the audience, but without a clear correspondence to a physical movement - a melodic balance, etc.

Cadoz considers instrumental gesture as effective gesture, however, it must be quite clear that this separation is somehow artificial, which is of great use for academic analysis, but not directly applicable in describing real gestures; in a performance situation, these are inseparable. For example in percussion, only the down-stroke produces the actual sound, but it needs a previous lift and a movement that positions the mallet and performer in the space of the object that will be hit.

Cadoz proposes as well an Instrumental Gesture Typology based on its function:

1. Excitation Gesture:
  - Instantaneous - Sound starts when gesture finishes
  - Continuous - gesture and sound coexist.
2. Modification Gesture:
  - Parametric - continuous variation of a parameter. (e.g. vibrato)
  - Structural - when modification is related to categorical differences (e.g. inserting a mute in a trumpet)
3. Selection Gesture - Choice among similar elements in an instrument

Perhaps Cadoz’s most interesting contribution, besides his comprehensive view of the research literature, is the concept of the gestural channel.

The gestural channel is unique if compared to other human communication channels (Visual, auditory, and Vocal) in that it is both a means of action on the physical world and a means of communication of information. In this second role, the gestural channel has a double direction emission and reception of information. It is therefore impossible to dissociate action from perception. (Cadoz and Wanderley, 2000)

He recognizes three different functions in the gestural channel:

- Ergotic function - material action, modification and transformation of the environment.
  - Epistemic function - perception of the environment.
  - Semiotic function - communication of information towards the environment.
- (Cadoz and Wanderley, 2000)

These ideas bring us to the concept of feedback. Having defined an instrument as a gesture-object compound that results in musical sounds, gesture as an action performed on an object implies the response of that object in the form of physical feedback as well as sonic feedback. That impossibility of dissociating “action from perception” is key in our understanding of gesture.

Delalande brings forth two important facts. The first is that a performer perceives his performance with the whole body. The second is that this perception constantly shapes the gesture; that is, in the process of shaping a gesture in rehearsal or during the execution of a performance, his assimilation of perceptual information or feedback helps him accommodate his gesture. Or in Piaget's words “The hand adapts to the shape of the object”.

The semiotic function presented by Cadoz refers to the fact that gestures address an audience and is similar to Delalande's Figurative Gesture, which is described as purely symbolic. There is a general conception that gesture is emotional and communicative. From a linguistic point of view, we use gestures to complement speech as non-verbal communication, which gives us the ability to change or alter the meaning of the verbal. It is clear that an excess of gestures is read as exaggeration or dramaticness and a lack of gestures as inexpressiveness or mechanical, but there doesn't seem to be a convincing way of addressing expressivity and emotion in gesture.

For Choi, the ‘traditional’ relationship of gesture and expression is that “expressions are said to be the product of gestures”. He however proposes the following:

With respect to Gestural Primitives we propose the inverse of this tradition: an expression is not the child of a gesture, rather a gesture is the child of an expression, where the rehearsal and planning to perform an expression is defined by the performer's orientation to a gestural primitive. Gestural Primitives provide a movement substrate that defines expression resources. (Choi, 2000)

Finally, it is interesting to read Delalande's appreciation of Glenn Gould's performance. From this description it is clear that the measure of expressivity relies also in gestures that do not produce any sound, but that are read by the audience as expressive:

Pianist Glenn Gould progressively reduced the range of gestures he used to a certain number of types corresponding to what musical terminology would call 'expressive traits'. What is meant by an expressive trait? Merely a type of productive gesture that has become generalized. When Glenn Gould struck the keyboard either vigorously or lightly, the sound produced was not the only indicator - in addition, body position, facial expression -including movements of the eyebrows - expressed vigor or lightness. In a 'vigorous' body position, his shoulders would contract and his head pull in. A contraction of the whole upper body could be observed. On the other hand, to play light, successive notes, Gould's head would no longer be drawn back into his shoulders, and he adopted another typical posture, leaning over the keyboard, and wrinkling his forehead. Thus his whole body was involved. (Delalande, 2003)

### *Digital Instruments*

As opposed to acoustic musical instruments, electric and digital musical instruments are those in which the gestural input and the sound production are independent linked together by mappings. Sound production independent from acoustical means, that is recording, synthesis and transformation, has basically been practiced for artistic creation since the 1950's. We'll come back to this in the next chapter.

A digital musical instrument implies that the gestural input and sound production units involve digital technology and therefore electricity. Although historical instruments like the theremin didn't use a computer (for obvious reasons), computers have become the most flexible way of recording and producing sound. For that reason I'll refer indistinctly with the term digital musical instrument to both electric and digital, even

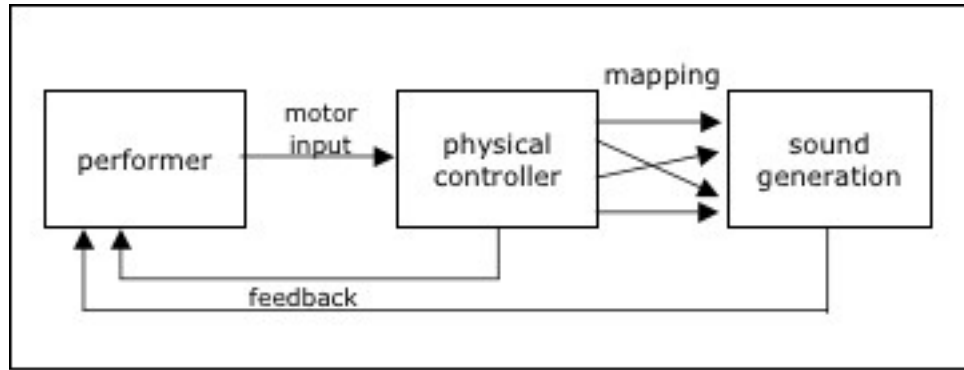


Figure 2.1: General Model for a Digital Musical Instrument

though the analog sound world produces sounds without one. A common representation of a musical instrument is the one seen in Figure 2.1.

The gestural controller is the unit in charge of transducing the continuous physical gesture of the real world into digital data through some sort of sensor technology. This process is commonly referred to as gesture acquisition, also called by Choi as “Gesture extensive research”.

According to Miranda and Wanderley (2006), gesture acquisition can be of three types: direct, indirect and physiological. In direct acquisition sensors are used to monitor the actions of the performer. A comprehensive description of available sensors for direct acquisition is found in chapter three of the same book. In indirect acquisition, gestures are monitored through the sounds they produce (eg. through a microphone). Physiological acquisition is concerned with the capture of bio-signals and not of physical gestures and is done with specialized equipment like ECGs, etc.

To this classification I would distinguish between sensors that are obtrusive and non-obtrusive. This extra category is important when designing controllers that don't interfere with the performer's gesture. In this sense microphones and video cameras aren't obtrusive.

Another important factor in the design of digital and electric musical instruments is that of feedback. Digital Musical instruments provide at least auditory, in most cases

visual and in some haptic feedback. This feedback allows the performer to evaluate his actions. “Auditory feedback is enabled by an auditory display mechanism, which offers a fine degree of resolution for the data field, and sensitive responsiveness to the observer’s performance with low lag time. Only then do we have an environment where the observer is able to construct auditory percepts and relate them to her own performance.” (Choi, 2000)

Miranda and Wanderley (2006) adopt a classification of new gestural controllers or digital musical instruments based on the degree of similarity to existing acoustic instruments in four categories:

1. Augmented musical instruments - acoustic instruments augmented by the use of various sensors
2. Instrument-like gestural controllers - gestural controllers that are modeled after the control surfaces of acoustic instruments, with the goal of completely reproducing their initial features.
3. Instrument-inspired gestural controllers - gestural controllers inspired by existing instruments or that intend to overcome some intrinsic limitation of the original models, but that do not attempt to reproduce them exactly
4. Alternate gestural controllers - that do not bear strong resemblance to existing instruments.

Tanaka (2000) proposes a classification into physical and non-physical, based on the mode of interaction with the controller and independently, mechanical or non-mechanical. In his examples, a potentiometer is a physical mechanical controller, while gestures that modulate light captured by a photocell is, from the point of view of the sensor, neither physical nor mechanical.

Choi proposes as well a “gesture intensive research” concerned with “the application to sound production of movement data retrieved from a measurement and storage system.” This is also referred to as mapping.

Iazzetta defines mapping in the following way: “While in traditional acoustic instruments the effects of the performer’s physical activity on an instrument are already established by the physical properties of the instrument, in electronic instruments this relation must be previously designed.” (Iazzetta, 2000)

While in acoustical instruments sound production is a direct result of the physical-mechanical application of energy to an object that creates sounds by acoustical means, in a digital instrument gestures are transformed directly into data-streams. The job of mapping is to assign these data streams or gesture parameters to parameters of sound production. Most authors on mapping issues claim that this is a critical feature of electronic instruments.

Rovan et al. (1997) classify mapping strategies into three categories: (1) One-to-One (2) Divergent, and (3) Convergent.

In One-to-One mappings each independent gestural output is assigned to one musical parameter. It is regarded as the least expressive. In divergent mapping, also known as One-to-Many, one gestural output is used to control more than one simultaneous musical parameter. Although it may initially provide a macro-level expressivity control, this approach nevertheless may prove limited when applied alone, as it does not allow access to internal (micro) features of the sound object. In convergent mapping, or Many-to-One, many gestures are coupled to produce one musical parameter. This scheme requires previous experience with the system in order to achieve effective control. Although harder to master, it proves far more expressive than the simpler unity mapping.

Hunt et al. (2000) explored mapping models for ‘expert interaction’, concluding that while complex mappings are harder to learn they allow control of complex parameter spaces.

Mappings can be implemented in two ways. Using generative mechanisms, such as neural networks or high dimensional interpolators or to define the mappings explicitly. (Hunt and Wanderley, 2003). In most musical situations controllers produce a different number of controls (usually less) than the number of parameters that need to



be used to control a sound, simply because a human being can only pay attention to a limited number of things at a time. Both gesture control and sound parameters can be represented as their derivative or their integral, increasing the ways of looking at these parameters, but also increasing the number of parameters to be mapped. The question is then how to match the number of controls to the number of mappings using complex mappings?

Goudeseune (2003) proposes a high dimensional interpolator that allows the user to specify pairs of groups of values; that is to specify what values we want in the sound variables when the control variables have another set of values, or, as Goudessene puts it: “when the controls have these values, make this sound.” This process continues by adding other points in the ‘map’ and then interpolating through them with a high dimensional ‘simplicial’ interpolator. With this method, the number of control parameters a performer needs to pay attention to is reduced, increasing his “intuitive understanding of the instrument.” In a similar spirit, Lee et al. (1991) proposed mappings with multi-layer neural networks trained by back propagation to control sound synthesis.

Besides the strategy of mapping that is used, there is the question of the kind of sound we are trying to control. Goldstein (1998) proposes observing the gestural affordance of real instruments to design our controllers or observing the sound’s structural qualities to choose or design the controller. This is done as an attempt to obtain what he calls gestural coherence, avoiding for example having violin sounds being controlled by keyboards.

Levitin et al. (2003) define the role of mapping as a means to “exploit some intrinsic property of the musicians cognitive map so that a gesture or movement in the physical domain is tightly coupled in a non-arbitrary way with the intention of the musician.” For him, our brains have evolved by incorporating certain specific physical principles of the world and developing cognitive maps that condition our perception. These cognitive maps include notions such as “ ‘harder means louder’ (for breathing or striking), ‘gestural wiggling yields pitch or timbre wiggling’ (such as in creating vibrato on a stringed instrument), and ‘tighter means higher in pitch’ (such as when stretching

a membrane on a drum, or tightening the embouchure on a wind instrument).

The problem with both Goldstein and Levitin's approach is that by focusing on the way we control the physical sounds of acoustic instruments we leave certain kinds of sounds aside. For example, audio feedback is a phenomenon exclusive of the electronic medium, impossible to model in terms of an acoustic instrument and its gestures. The kinds of gestures used to control feedback are well exemplified by Ostertag in his description of Jimi Hendrix's playing style: "Hendrix's crucial innovation was playing at high volume and standing close to the speaker to obtain feedback that he could control in an extremely nuanced way with the position and angle of the guitar, the weight and position of his fingers on the strings, even the exact position of his entire body." (Ostertag, 2002)

Evaluating a mapping strategy is an issue that has to be addressed in a case-by-case basis based on the sounds to which we are trying to map. The scope of this thesis can not include a review of all sound production techniques. However, we can still abstract ways of approaching sound production that mapping does not fully address.

Hunt et al. (2002) raise the issue of mapping as a specific feature of a composition or as an integral part of the instrument. They consider the second point of view as the best, giving the instrument consistency through time and offering the performer the possibility of learning these mappings at an expert level. It is clear that this issue raises all sorts of questions surrounding the concepts of instrument, which will be discussed later in more detail, but some questions can be posited now. Computers are machines capable of performing many tasks, many of which include to a certain extent, 'making decisions'. It is clear that in certain mappings-for example in controlling a granular texture-we are bound to map our gestures onto higher level parameters like density, envelope shape, delay times, random generators, etc. This is simply because if we were to map open air hand gestures to each of the grains, either the sounds would not be as interesting-or at least would not sound granular-or we would be forced to move hysterically and fail in the attempt. So for this matter, we entrust the computer with the task of generating the texture from our general indications.

Chadabe recognizes two types of instruments, deterministic and indeterministic ones. As mentioned previously, certain determinism characterizes standardized acoustic instruments in the hands of trained performers. However in digital musical instruments, the sounds generated by the gestures are not always entirely determined. For Chadabe an indeterministic instrument outputs “a substantial amount of unpredictable information relative to a performer’s controls. In working with such an instrument, a performer shares control of the music with algorithms as virtual co-performers...” (Chadabe, 2002)

In mapping strategies like Goudessene’s interpolation or Lee and Wessel’s cellular automata, the designer doesn’t fully describe the outcome desired, but only certain reference points. The computer then complements the final behavior of the controller and the performer is left to discover them in a learning process. But the overall behavior of the instrument is the same unless the reference points change.

But Chadabe goes further in assessing the limitations of mapping in determining the way an instrument works: “To the extent that an automatic mechanism generates information, even while remaining obedient to a performer’s commands, it becomes more difficult to conceptualize a performer’s control gestures as mapped onto an output the computer’s ability to expand simple but powerful instructions into coordinated controls for multitudes of variables, to redefine controls in different contexts, and to maintain goal-orientation while introducing enough unpredictability to keep the instrument interesting.” (Chadabe, 2002)

The question then starts to reveal itself. Can a button be considered an instrument? Can pressing a button be considered a performance?

## Chapter 3

# Gesture, the Studio and the Computer

The model Composer-Performer-Listener is the product of a long standing tradition in the history of music and indeed, when no means of generating sound independently from instruments existed, a concert/performance was the only way to deliver sounds to an audience. Although composer/performers existed, mainly due to the fact that traditional music education proposed the knowledge of the piano as an indispensable pre-requisite, the model persevered as a kind specialization system, where one was trained in composition and the other in performance. This also led to the concert ritual, the space and moment in which the performer delivered the music from the composer to the audience. It is necessary to state however that this model is to say the least simplistic, but it still reflects somehow the ideas of occidental culture towards music.

In his Introduction to the Sociology of Music, Adorno (1976) states that “works are objectively structured things and meaningful in themselves, things that invite analysis and can be perceived and experienced with different degrees of accuracy.” This point of view subscribes to the idea that the work is an abstract idea, a product of the mind, perfect in itself and independent from its sound realization. This is supported by a tradition of sound analysis that focuses exclusively on scores and particularly on a form that is determined by the organization of pitch as the main material. This leaves the performer as a body in charge of realizing this idea into sound, which can then be received

by an audience, the minds that can, at least in Adorno's view, perceive and experience it with certain degrees of accuracy.

The 20th century also saw an increased desire from composers to obtain new sounds. This was reflected in the increased use of percussion instruments, extended techniques, prepared pianos, timbral use of the orchestra, etc. There are many accounts of composers complaining about the inability of the instrumental medium to realize their ideas. In many ways the search for new ways of producing sound was a result of technology as much as music's use of technology was a search for new sounds, specially given that these technologies existed for several years and, in the case of recording, decades before.

In the 1950s two major research centers concentrated the research on electronic sounds. On the one hand the Radio Television Française (RTF) in Paris, directed by Pierre Schaeffer, associated with *Musique Concrète*, which focused in the manipulation of recorded sounds as the main means of sound production. On the other Northwestdeutscher Rundfunk in Cologne, directed by Herbert Eimert, associated with *Elektronische Musik*, which focused on the use of sound generators and modulators to generate sounds. Other studios and research centers gradually evolved and the electronic music studio gradually formed itself into a sort of standard with some form of synthesizer and tape recording device as central units.

The unifying element in both cases was the fact that the music produced, regardless of the sound sources, had an almost strict dependence on the use of magnetic tape as a medium to fix the sounds not only as a final product, but also throughout the composition process. Although we could argue that the recording process changes the sound, by fixing a sound in a digital or analog format, the concrete character of sound is preserved (as opposed to abstracted and notated) and acquires a material existence. This materiality not only allowed the possibility of sound reproduction but also of transformation and organization. In this sense, composition shifted to work directly with sounds and not with abstract representations or ideas of sounds, as in the case of instrumental musical scores.

In one of Schaeffers first definitions of *Musique Concrète* published in the magazine *Polyphonie* in December, 1948

We apply (...) the term abstract to conventional music, because it is first conceived by the spirit, then theoretically notated, finally realized in an instrumental performance. We have named our music "concrete", in opposition, because it is constructed from pre-existing elements, regardless of what sounding material caused it, be it noise or conventional music, then composed experimentally by direct construction. (quoted in Chion 1991)

When Schaeffer refers to a direct construction he not only refers to the physical cutting and pasting of the tape, but to the fact that musical notation and instrumental performance were not part of the composition process, the composer would work directly with the sound material that would constitute the piece; with the concrete material fixed in a recording format. *Musique Concrète* and music fixed in a recording format was later coined as acousmatic by composer Francois Bayle. This originally greek term refers to the phenomenon of hearing a sound which we can not see the source or cause. Various composers then adopted the term to describe music that comes out of loudspeakers.

So at the outset of acousmatic music the music produced by these composers challenged the composer-performer-audience model in many ways, but principally in the fact that there was no visible performer, and at a first glance, the performer was removed completely: the music that was produced in the studio was in practice the same to be heard by the audience, it would not be performed for them.

Several issues arise from these facts. Music, an art form characterized by the performance of music in a concert hall was not the same anymore. Furthermore, this performance practice, which traditionally featured a human being playing a musical instrument through musical gestures, didn't seem to have a place anymore. But what if we try to picture the electronic music studio as an instrument.

Luigi Nono considered the studio as "new instruments which need to be studied over time to learn and study again and again, to explore other possibilities, different from the ones usually chosen and given, other musical thoughts, other infinite spaces..."

(Nono, 1984)

In our initial definition of musical instrument, we talked about object-gesture compounds that produce sound. In our distinction of electric and digital musical instruments, we talked about the independence of gestural input and sound production units. Electronic music studios then, fit perfectly well in these definitions. Knobs, faders and switches constituted the gestural input unit. The mapping layer was the patching bay and the cables that connected the unit generators, modulators to speakers or to the magnetic tape. Magnetic tape manipulation on the other hand presented similar features with knobs for playback speed, direction and volume, providing even the possibility of pressing the tape on the reader with hands and fingers. Most studios quickly incorporated keyboards and Pierre Schaeffer even developed his own tape-controlling keyboard, the *phonogene*.

The studio provided as well many limitations. The first limitation was the size of these studios and the specificity of the equipment, factors that reduced the possibilities of moving it anywhere (and therefore onto a stage). Although the set of possible sounds that could be produced with this instrument increased dramatically, it was impossible to jump from one sound to the other, for it implied a change in the patching bay. On the other hand, the amount of sounds that could be produced at the same time was limited to the number of unit generators, modulators and tape decks available. For this reason the tape became the medium to store sounds temporarily, giving the composer time to produce another sound that could be mixed with this one, to change the mappings for the next sound, to change the tapes that were being manipulated, to store all the sounds that would then be assembled into the final piece and to fix that final result on tape. In a way, the composer performed the sounds into the tape, which recorded his original gestures.

All of these facts resulted in a kind of music that had a temporal and spatial dislocation, between the place and moments in time when the sounds were recorded to the place and moment in time in which the audience heard the sounds. While in the traditional composer-performer-audience model, the composer worked on a composition or a

score in a separate, previous time, the audience perceived the sounds in the performance immediately after they were produced by the performer in the concert. Now, the sound perceived by the audience in a concert was the result of sound produced in other spaces and times. So if we were to sketch the model again we would have to create a hybrid composer/performer, if our model reflected the time process, we would have to include performance-composition as much as composition-performance, or even more to establish a more dialectic way: composition-performance where each one renders the other in a slow process. This is precisely the process from which the audience is removed, left to hear the result through speakers.

Although human gestures are employed in the construction of electronic sounds in the studio, there are several points to consider. The first is that the manipulation of electronic sounds with machines often involves gestures that are different from our known instrumental gestures, first because the interface is different and second because the sound material is different. Even when using recorded physical sounds (and therefore physical gestures), the results can become so distant from the source sound that the original gesture is lost. The process of construction is usually so layered that the resulting sound has probably no direct relation to what we have learned to recognize as musical gestures in acoustic instruments. Composer Denis Smalley describes this process in the following way:

the ‘working gestures’ of the acousmatic compositional process do not carry perceptual information equivalent to an intuitive knowledge of the physical gestures of traditional sound-making. Therefore, while in traditional music, sound-making and the perception of sound are interwoven, in electroacoustic music they are often not connected. (Smalley, 2001)

Smalley describes the final sounds produced by acousmatic music as spectromorphologies referring “to the interaction between sound spectra (spectro-) and the ways they change and are shaped through time (-morphology). The spectro- cannot exist without the -morphology and vice versa: something has to be shaped, and a shape must have sonic content.”



Parallel with the growth of electronic music studios, several research centers in the USA were researching the production of sound with computers. Amongst those studios were Bell Telephone Laboratories in New Jersey. The first computers were as big as several rooms and rendering one sound could take hours or days. As early as 1963 Max Mathews proposed the idea of using the computer as a musical instrument (Mathews, 1963) and in 1970, along with F. Richard Moore published their GROOVE system: A Program to Compose, Store, and Edit Functions of Time (Mathews and Moore, 1970). This system had the virtue of being able to produce sounds by capturing gesture input by a computer that controlled analog devices for sound production. Again, the size of the machinery made it impossible to stage in a concert.

The first line in Mathews and Moore (1970) states: “Many tasks now done by people are best described simply by one or more functions of time.” This concept of functions of time is key to computer technology and especially to computer music. Although in their program, gestures inputted to the system were represented as functions of time to be used immediately, it also allowed to store, edit and compose them. That is, gestures could be manipulated and composed as much as sounds as functions of time and therefore the control of analog sound devices did not necessarily depend on humans at knobs, but could now be automated.

Due to the rapid improvement in computer technology, the electronic music studio has gradually become just a computer. The practice of acousmatic or tape music is now usually realized directly on the computer. (The value judgments of analog vs. digital technology are not part of this thesis.) This increase in computer speed also allows for freer and less mechanic ways of composing acousmatic/tape music, with no need of using magnetic tape for storage or assemblage.

This also means that gestures and sounds are now commonly automated not only in the sense that control functions are not performed by human performers, but also through the use of algorithms. In this way computers became the primary tool for music produced in the studio and then played back in a diffusion system or concert hall, but also a tool to produce sequences of instructions that could assist a composer in the

production of a composition, either as a score intended for instrumental performance or directly modeling the sounds.

The gestures performed by the computer or even the gestures we create as a result of the direct construction mentioned by Schaeffer are not the gestures we have trained ourselves to recognize as coming from humans playing instruments. So what are these gestures and how do we make sense out of them?

Dennis Smalley proposes a classification of gesture in terms of the perceptual distance it presents from the gestures we know from experience. For him listening is a sense making activity whereby we do not only think of the gesture process as cause-source-spectromorphology, but also as a reverse process spectromorphology-source-cause. We listen to sounds in a referral process by which we hear a spectromorphology, recognize the source and detect human activity behind it, building the gesture that caused it. An example of this is when we hear to recorded instrumental music. “The listeners experience of listening to instruments is a cultural conditioning process based on years of (unconscious) audiovisual training. A knowledge of sounding gesture is therefore culturally very strongly embedded.” (Smalley, 2001)

But when we hear acousmatic sounds, the causes can become remote or detached from known. Smalley calls this increasing remoteness continuum gestural surrogacy. Gestural surrogacy then has the following categories:

- First-order Surrogacy: The original, primal gesture, on which sounding gesture is based, occurs outside music in all proprioceptive perception and its allied psychology. Traditionally, this first level does not become music; it develops into second-order surrogacy. To consider them first-order we need to recognize source and gestural cause.
- Second-order surrogacy: Traditional instrumental gesture, a stage removed from the first order, where recognizable performance skill has been used to develop an extensive registral articulatory play. Much music which uses the simulation of instrumental sound can also be regarded as second-order since, although the

instrument may not be real, it is perceived as the equivalent of the real.

- Third-order surrogacy: where a gesture is inferred or imagined in the music. The nature of the spectromorphology makes us unsure about the reality of either the source or the cause, or both. In his examples we can infer a cause for an unknown source or we can only partially its behavior.
- Remote surrogacy: concerned with gestural vestiges. Source and cause become unknown and unknowable as any human action behind the sound disappears. The listener may instead be concerned with non-sounding extrinsic links, always, of course, based on perceived spectromorphological attributes. But some vestiges of gesture might still remain. To find them we must refer to tensile, proprioceptive properties, to those characteristics of effort and resistance perceived in the trajectory of gesture. Thus, remote surrogacy, while distanced from the basic, musical first order, can yet remain linked to the psychology of primal gesture. (Smalley, 2001)

While most of Smalley's ideas are based on experience and could be debatable in many ways, his intuition and ability to identify, categorize and name phenomena is one of the few that exist. Before leaving his ideas two more points need to be made.

In this sense-making activity and in relation to the extrinsic links he mentions in remote surrogacy, Smalley created the term *source bonding* to refer to connections between intrinsic qualities of the sound and extrinsic qualities of the world. In this way, the connections a listener makes between a granular texture and rocks falling are a form of source bonding.

The second point is that of texture. For him, "if textures are weak, if they become too stretched out in time, or if they become too slowly evolving, we lose the human physicality. We seem to cross a blurred border between events in the human scale and events on a more worldly, environmental scale" Smalley (2001) To this I would add that the conglomeration of small gestures in a small or repetitive way are also textures, although

not necessarily detached from the human scale. In this sense, long orchestral notes or even repetitive figures like the alberti bass act as a texture. However, we can deduce the many possibilities in which textures can be created in the acousmatic medium, either as granular processes created by a computer or through cutting and pasting magnetic tape like Xenakis' "Concret PH", Cage's "Radiomusic" or Tenney's "Blue Suede Shoes".

As stated previously, computers are machines able to receive and produce sequences of instructions that could assist a composer in the production of a composition, either as a score intended for instrumental performance or directly modeling the sounds.

The use of algorithms in composing is not entirely new in the sense that there have been innumerable attempts at formalizing compositional processes from D'Arezzo and Serialism to Xenakis as the most cited examples. Computers are machines perfectly fit for this task and therefore, since the advent of computers in music, some of the first tasks they were to perform were those of algorithmic organization of sounds.

The specific applications of algorithmic composition have included the generation of melodic, rhythmic and harmonic material created by algorithms that are then analyzed and modeled by the composer. Other approaches have attempted more radical positions, that is, where the composer doesn't modify the computer result.

The question of to what extent instrumental composition is determined by the physical gestures of humans playing acoustic instruments is one that needs to be assessed. The traditional disciplines of music education provide us with hints on this issue. Besides music language courses, the basic courses in traditional conservatory curriculums for music composition are music theory, harmony, counterpoint, analysis, instrumentation and orchestration. Instrumentation and orchestration are concerned with issues of performability in acoustic instruments, and furthermore, the issues of comfortable ranges, and dynamic and articulatory possibilities in those ranges. Gestures like trills, glisses, tremolos, chords are instrument dependent and have intrinsic limitations within each instrument, limitations that are determined both by the instrument and the corporeal limits of the player.

Iazzeta points out these issues in the following way:

Concerning music, one can say that physical gesture is directly related to music interpretation while composition is much closer to mental gesture: As Bernadete Zagonel stresses, if the composer goes from gesture to the composition, the performer goes the opposite way, that means, he goes from the score to the gesture. To this statement we can add that the listener completes this chain by mentally recreating the performer's physical gestures while listening to music. (Iazzetta, 2000)

This mental gesture produced by the composer before writing in the score is based on how a particular musical idea would sound on an instrument, but is also based on his cultural training on the way instruments work and sound, on the analysis, perceptual or formal of the way in which instruments work and the limits they present.

To write a crescendo going from ppp in the highest register of the flute to fff in the lowest is not adequate and won't work. These kinds of considerations form what has been called idiomatic composition, which we could define in very similar ways to instrumentation, but adding the consideration of "what kinds of phrases and passages are most appropriate for the instrument." Tanaka (2000) This approach not only defines the instrument through composition, but also composition through the instrument. In other words, mental gestures are those instrumental gestures we have processed and retained through past experience. In this way, could the embedding of instrumental gesture in our cognitive system be determining the gestures we make in electronic music? Could electronic music be hiding a body beneath it?

So what kinds of mental gestures can a computer have? The computer's unawareness of instrumentation rules (even of its own performative capabilities) renders it incapable of good acoustic instrumentation unless coded specifically for the task as a constraint to the composition algorithm. Such a task is better performed by the composer, who as a human, shares the body with the performer and can assess, by imagination, previous experience, education, score analysis and case-specific testing, the particularities and affordances each instrument presents.

So the composer can act on the computer in the following ways: either specifying parameters for the creation of material, encoding the boundaries and general concept of

his mental gesture or encoding a general idea that is then mediated by the composer to make it adequate for the instrument.

But, can we have non-instrumental mental gestures? The computer can easily perform certain things that are impossible for acoustic instruments and humans. Can we talk about gestures of the instrument-computer or even about a computer-composer?

I will leave these questions partially unanswered. For Smalley, acousmatic composition is a sort of negotiation between levels of surrogacy:

Acousmatic music, therefore, can stay close to traditional, gestural cause-source relations, but at its most adventurous extends into third-order ambiguity and beyond to a music which, although remote from traditional sound-making activity, can nevertheless maintain a humanity. I venture to suggest that an electroacoustic music which is confined to the second order does not really explore the potential of the medium, while a music which does not take some account of the cultural embedding of gesture will appear to most listeners a very cold, difficult, even sterile music. (Smalley, 2001)

## Chapter 4

# 'Live' Performance, the Body and Concert in Computer Music

### *History: Electronics and Performance*

The advances introduced by the computer and technology in general, namely the radical expansion in sound material and new spectromorphologies with the ability to create surrogate gestures independent from instrumental ones and their organization through algorithmic methods presented new problems: how to perform or present them to the public, the audience.

Due to convention, music produced in studios was presented in the traditional space for presenting music: the concert hall. The first concerts introduced the concept of a loudspeaker orchestra as opposed to traditional performance where a human was in charge of bringing the sounds out of the score. The concept of a loudspeaker orchestra had inevitable direct connotations with the idea of going to a regular orchestral performance. As we can see in Figure 4.1., the loudspeakers were placed in the space of the orchestra and the audience remained seated. This format of presentation, the concert, with an audience facing a stage in a darkened hall, and in the stage a dissimilar array of speakers, seemed to contrast with instead of assimilate the performing tradition.

Audiences were already familiar with radio and record playback. Since the ad-



Figure 4.1: François Bayle in the Acousmonium in 1974

vent of recording technology, music had acquired a dual nature, ‘recorded’ and ‘live’. Where ‘recorded’ music was the recording of a live performance and ‘live’ music was the actual ‘performance’. Now music presented new sounds, produced through a process unknown to the audience and whose sonic results were usually unknown as well. Moreover, the resulting pieces were fixed in a format associated with recorded music, they were actually records, the same the audience had at home, in their living room. So why go to a concert which is a sort of expanded living room? There seemed to be an association whereby the concert hall belonged to the ‘live’ sphere and the living room, where the radio and the record player lived, to the ‘recorded’ one.

Sound spatialization, the distribution of sound in a space, became key in the diffusion of acousmatic music and recorded sound in general. Initially, the works of composers like Schaeffer and Henry, who came from the radiophonic tradition were presented from the front, monophonic or at best multi-monophonic Bayle (2007). This situation set the sounds to come from one place, each one from one loudspeaker. The acousmonium, created in the 1970’s by François Bayle, was based on the concept of the



orchestra and was intended specifically for tape concerts. For him, the performance of acousmatic music, implied the adaptation of a work from an internal space designed in the studio and “formed within the work itself, made of reflections of the sonic contours, of the movement of entities, presenting itself to the hearing as a sensation of composed volume to the external space “no longer concerned with the work but with the configuration of the space wherein it is heard, with its particular peculiarities.” He called this the transposition of space.

The creation of stereo techniques allowed composers to position sounds in space. Instead of a sound coming from a speaker it came as a combination of two speakers with different relative amplitudes. In Bayle’s words: “We know that stereo functions through two channels but only carries a single message. My interest in stereo is that the object-message becomes alive when it is no longer dependent upon one or other of the projectors but goes on to establish itself in immaterial space between the two”. In the Acousmonium, Bayle presented the idea of spatialization as sound projection as a metaphor to the projection of images. For him loudspeakers are sound projectors, and the space, the projections are the screen.

The acousmonium consisted of 78 loudspeakers each with a specific sound quality and mixing consoles that controlled pairs of speakers or ‘stereo images’. Through time, these environments included speakers within the audience and surrounding them as opposed to a frontal arrangement in a stage. The loudspeakers are sometimes illuminated to emphasize the fact that they produce the sound. However in Martha Brecht’s view, “because of the strong visual aspect and a reception situation which does not separate stage and audience, it seems that an environment like this can hardly be called a concert anymore. It is rather a sound installation ... simply its visual aspect is stronger than Bayle had planned. (Brecht, 2002)

Contemporary sound spatialization or sound diffusion systems operate in multi-channel environments designed to surround the audience with speakers of the same kind, as opposed to arrangements of stereo pairs. This reinforces the idea of immateriality of the sound and puts the audience at the center as opposed to the speaker as element of

attention. This is now usually reinforced in concerts by completely darkening the room as an attempt to exclusively focus on the sounds.

Parallel to the development of loudspeaker orchestras and sound diffusion systems, attempts to deliver electronic sounds in a 'live performance setting were attempted both in Europe and in North America. For Martin Supper, there are several possible definitions of live electronic music.

[The German conception] considered as an extension of music made for loudspeakers. Extension in a double sense: in the first place, the electronically generated sonic material is not synthesized in the studio, but in real time on stage; in second place, the sound of the acoustic instruments or the human voice is transformed electronically also in real time. In North America, this concept is used in a wider sense, Live Electronic Music is also understood as the reproduction, simultaneous to the performance of one or more musicians, of a magnetic tape previously produced. (Supper, 2004)

Historically, the first piece to be considered live electronic music is John Cage's "Imaginary Landscape No. 1" for two variable speed turntables playing discs with sine tones, prepared piano and cymbal from 1939.

Studio productions in magnetic tape were quickly incorporated into instrumental performance. Some of the first examples include Varese's "Déserts (1949-1954), the collective compositions of Otto Luening and Vladimir Ussachevsky "Rhapsodic Variations" (1953-54) and "Poem in Cycles and Bells" (1954), Bruno Maderna's "Musica su due dimensioni" (1952) and Davidovski's "Synchronisms" (1963-70).

A practice called *performed tape* was concerned with the manipulation of magnetic tape during concert and in which Supper includes Cage's "Williams Mix" (1952) and "Fontana Mix" (1956), Robert Ashley's "The fourth of July" (1960) and Stockhausen's SOLO (1966) which included a melodic instrument and tape feedback. (Supper, 2004)

Other pieces did not use magnetic tape at all and were concerned with the live processing of instruments in real-time or their combination with live synthesis. Some

of the first pieces that attempted this were Stockhausen's "Mikrophonie I" (1964) and "Mantra" (1970) and Boulez's "...explosante-fixe..."

In terms of live performance, the tape and instrument pieces involved a composer that set the tape or tapes on motion and left the performer to synchronize with them, while in pieces like *Mikrophonie*, two of the four percussionists held a microphone which was moved throughout the tam-tam to capture the sounds being made with it at different distances. At the same time, two people transformed the sound of each microphone with a variable bandpass filter, transforming the spectral content of the sounds and spatializing it in the hall through faders.

While the tape and instrument pieces presented two separate worlds happening at the same time and sometimes synchronized, the resultant gestures tended to be separate. While in most pieces the gestures and surrogate gestures of the tape sounds remained intact, these were superimposed with the live acoustic instrument. On the other hand, in *Mikrophonie*, a sense that the electronics are being performed is developed, although the gestures of the faders are very simple, they are very effective and well coupled with the movement of the microphones.

Increases in computing power allowed for computers to process and synthesize sounds live. Although as early as the seventies composers like Chadabe were exploring the idea of interactive composing and interactive instruments, the eighties saw the flourishing of this activity.

IRCAM started functioning in the year 1977, and had as one of its central research themes the use of computers for real-time transformation and synthesis of sounds.

In many representative pieces, like Manoury's series *Jupiter* (1987), *Pluton* (1988), *Neptune* (1991) and *En Echo* (1993-4), the computer had the double task of following the score through microphones or MIDI input and generating the sound output. The sound output would consist of the live capture of samples for further transformation, transformation of live sound input, synthesis of sounds and spatialization. These transformations were events marked in the score and triggered by the score follower. In many cases the control of sounds was done through generative processes such as markov

chains and determined by the analysis of incoming sound and gesture.

In 1971, the studio for electro-instrumental music, STEIM, was created with the aim of creating new interfaces for musical interaction and control of live electronics. In 1986, Waiswicz composed and performed the “Archaic Symphony” with his controller “The Hands” which allowed him to track the movements of fingers, hands and arm, and a software called “Lick Machine”, in charge of re-interpreting and mapping all the information that came from “The Hands”.

George Lewis, composer and trombonist, wrote the software “Voyager” between 1985 and 1987. This software ‘listens’ to Lewis’ trombone improvisation and reacts as an improvisation partner.

During the 90’s and the first decade of the 21st century this trend has continued. The composition and performance of interactive pieces of music is increasing and the creation of controllers is growing at a very fast pace. Although it is hard to assess the degree to which it has been successful, it has certainly attracted the attention of many researchers and composers. This is reflected in the number of journal publications and conference papers on the subjects of interfaces and interactivity.

It is hard to establish a line of continuity between the first attempts at making live music with electronics with those done since the 80’s where the computer becomes a central element, mainly because of its ability to interact and produce any sound that can be programmed. Although most traditional electronic pieces are being transferred to modern computer technology, it is impossible to do otherwise, that is to use older technology to make the pieces done today with computers.

It is even harder to establish what this technological sense of going forward means. In what ways has the reductionist composer-performer-listener model changed? How is music incarnating its time?

In what follows I will try to characterize these new practices, namely, ‘live’ performance with gestural controllers and interactive computer systems as attempts by composers and performers to re-incorporate the body as principal medium of expression and as mediator between the known acoustic instrumental world and the remote gestures

of the electronic sounds created from the studio years to date.

*Embodiment in Philosophy and Cognitive Science*

The traditional view of the mind-body problem has its origins in Platonic ideas of immateriality of the soul and religious views of incarnation, such as in Christianity and Hinduism. This concept was further developed by Descartes, reason for which this division is today widely attributed to him:

What we have inherited from Descartes is a way of thinking about our relation to the world - in particular our epistemological relation to the world - which serves to support and strengthen this ontological stance. Descartes' famous ontological realization that he is a thinking thing is conditioned and tempered by the epistemological admission that all he had accepted as most true had come to him through the senses; yet it is precisely this intrusion of the body between knowledge and world which is in the end unacceptable. The body must be part of the causal order to allow for perceptual interaction, but is therefore both unreliable (a cause of the senses deception) and, as it were, too reliable (driven by physical forces, and so a potential source of unfreedom). Thus, the body is for Cartesian philosophy both necessary and unacceptable, and this ambivalence drives mind and body apart in ways Descartes himself may not have intended. (Anderson, 2003)

For the anthropologist John Blacking, the mind-body problem can also be explained as a cultural phenomenon:

The mind/body dichotomy can be partly explained ... as an artefact of certain modes of production and social formations, reflecting the division of manual and mental labour in production. This props up the bourgeois argument for class difference, namely, that one class must do the manual labour so that another can look after the mental needs of society. (Blacking, 1977)

In cognitive science, we can distinguish three approaches to a science of mind: *cognitivism*, *connectionism* and *interactionism*. These approaches have in common the idea that cognition is a kind of computation (Kim and Seifert, 2006). "*Cognitivism* uses 'the mind' as a main metaphor" and view it as "a sort of computational system in the

sense of a Turing machine” and *connectionism* uses the ‘brain metaphor’ where “artificial neural nets are used as operational computational models of cognition”. However, both of these approaches “neglect the body and its environment”.

Although with slightly differing approaches, *interactionism* is also called *ecological*, *enactive*, *embodied* and *systems view*, referring to cognition and perception. This set of theories propose a view of a system where “an entity interacts with its environment and other entities. ‘Embodiment’ in this sense stresses the importance of bodily mediated motor-behavior in an action-perception cycle in human and animal cognition as well as in social interaction.” (Kim and Seifert, 2006)

The mind/body dichotomy indirectly proposed that the sensory receptors were passive and objective transmitters of a reality that is ‘out there’, while the mind/brain was in charge of processing that perceived reality into information. Then the mind/brain uses that information to create an output in our sensory motor system.

A classic counter example of this is the fact that when we burn our hands, we move them away of the heat source faster than we realize we had been burnt, then we actually acknowledge the pain and inspect the source visually. That is, that our hands ‘decided’ to move before telling the mind/brain. In Lettvin et al. (1959) an analysis of the frog’s visual system showed that the eye of the frog provided its brain with highly processed information. In Pritchard et al. (1960) an image was stabilized in the retina of a human eye, and with this stabilization, the image disappeared, the eye doesn’t ‘see it’ anymore. The concept behind these facts and experiments is that the mind doesn’t receive the world as it is, but it receives information processed by the body, several senses at a time. In this way, the mind/brain concept can’t even be conceived as independent from the body, for all it knows comes from the body and the body tells it what the body believes; it is embodied.

Another classic example, this time from music, is commonly exemplified with the fact that musical performers develop an ability to learn to play music, either from a score or improvising which seems to be independent from the mind/brain. Performers refer to this as the memory of the fingers. Borgo (2005) makes an excellent analysis

of saxophonist Evan Parker and a thorough review of embodiment theory in the context of improvisation. Parker constantly refers to moments which he “can’t explain”, where his body takes over the performance and he thinks literally “about the laundry”. In an interview by Volker Krefled, Michael Waisvisz established a clear link to this idea in the following way: “I see the hand as a part of the brain, not as a lower instrument of the brain. Of course, you can see a hand as a transmitter and sensor, but in the consciousness of the performance, the hand is the brain. You can’t say that its precision is surpassed or even equaled by computers because we simply don’t know what we control in detail when we play an instrument. ” (Krefeld, 1990)

The idea of an action-perception loop in these enactive or interactionist approaches is worth noting. The idea that seeing is a purely perceptual phenomenon is confronted with the fact that it doesn’t just transmit the world as it is. Our senses don’t ‘just perceive’; they also perform an image of the world. Perceiving is an action. It is clear then that action is also a perception, that the hand that acts on a controller, like Waisvisz’s hands is anticipating a result and confronting it with perception.

In this spirit, Blacking proposes the following:

And just as the ultimate aim of dancing is to be able to move without thinking, to be danced, so the ultimate achievement in thinking is to be moved to think, to be thought. It is sometimes called inspiration, insight, genius, creativity, and so on. But essentially it is a form of unconscious cerebration, a movement of the body. We are moved into thinking. Body and mind are one. (Blacking, 1977)

#### *The body and musical performance*

We can easily connect musical performance and instrument with the interactionist or enactive approaches. Paraphrasing Merleau-Ponty, Stubley (1998) states the following about instrumental performance “as if musician and instrument are symbiotically tuned to one another, with each affecting the other as parts of a system that collectively demonstrates the property of ‘mind’.” This symbiosis to which she refers is at the heart of the action-perception loop, but it raises the issue of the instrument as an independent organism.

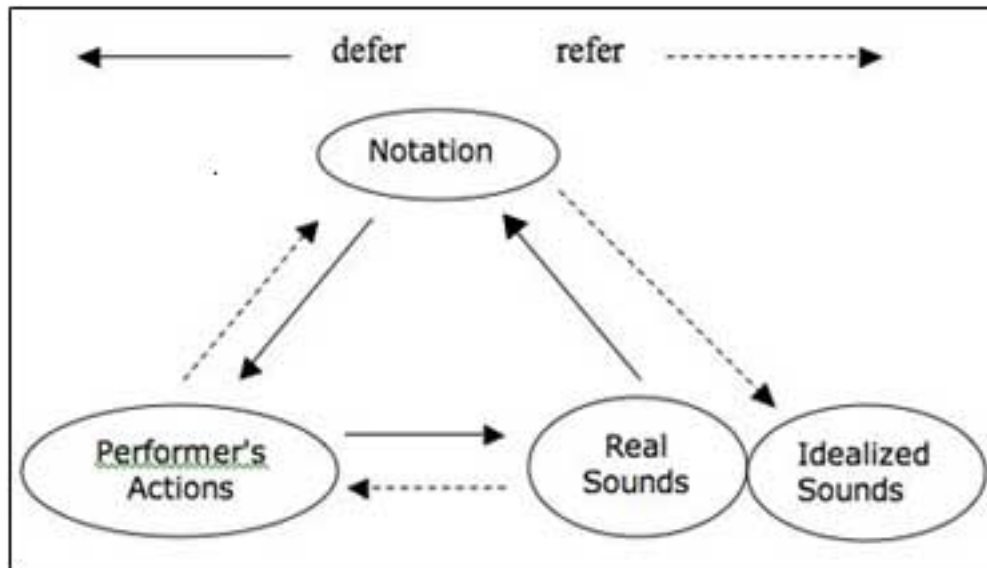


Figure 4.2: Choi's diagram of semiotic circularity in music notation and performance

Choi's model of semiotic circularity in music notation and performance, seen in Figure 4.2, implies a series of connections between gesture, sound and score that need to be analyzed. "By circularity, we mean the references are in a feedback loop for the production of sounds and movements." For Choi, musical gestures imply an intended auditory sequence, either following a score, plan or intention, that drive motion sequences as idealized gestures.

The musical score defers to the performer's movements while the performer's movements defer to the resulting auditory sequence. Ultimately, the musical score is an idealized reference system to the auditory sequences. The performer's task is to maintain these references in calibration with the semiotic circularity of sound production. Musical gestures arise in the semiotic circularity of the performer's physical calibration between their body and their instruments. (Choi, 2000)

From our previous analysis of the influence of instrumentation and orchestration in the compositional process we explored how human gestures determined a composition. However, the written score presents only a small part of what actually happens in a musical performance. Clarke (1993) explores ways in which performers create a



mental image of a work which is not contained in the score. This is reflected on the fact that different performers usually have varying interpretations of the same score, but also in the fact that repeated performances of the same piece show little variation over time. Cusick raises the role of the body in the realization of a score: “But the score is not the work to a performer; nor is the score-made-sound the work: the work includes the performer’s mobilizing of previously studied skills so as to embody, to make real, to make sounding, a set of relationships that are only partly relationships among sounds.”

Cusick goes further in her statements:

Music, an art which self-evidently does not exist until bodies make it and/or receive it, is thought about as if it were a mind-mind game. Thus, when we think analytically about music, what we ordinarily do is describe practices of the mind (the composer’s choices) for the sake of informing the practices of other minds (who will assign meaning to the resulting sounds). We locate musical meaning in the audible communication of one creating mind to a cocreator, one whose highly attentive listening is in effect a shared tenancy of the composer’s subject position. We end by ignoring the fact that these practices of the mind are non-practices without the bodily practices they call for - about which it has become unthinkable to think.

To deny musical meaning to purely physical, performative things is in effect to transform human performers into machines for the transmission of mind-mind messages between members of a metaphorically disembodied class, and, because disembodied, elite. (Cusick, 1994)

We can derive many interesting directions from this remarks. The first of course is the idea of a mind-mind game, between composer and listener to which we had referred to in Adorno’s characterization of works of art as ”objectively structured things and meaningful in themselves, things that invite analysis and can be perceived and experienced with different degrees of accuracy” as well as to Blacking’s cultural reference to the mind-body problem. As we explored the implications that studio electronic music had on musical performance we stated that the removal of the audience from the creative gestural process seemed to create a composer-listener model where the performer was in simple terms ’skipped’. In a way, this might seem like the perfect mind-mind model, exemplified in Jon Appleton’s words: “The debate over the role of the performer in

electroacoustic music goes back to the German theoreticians of the 1950's who claimed that both the complexity of 'new' music and the imprecision of human performers made tape music a necessity." (Appleton, 1984) Computers of course, reinforce this model in which algorithmic procedures and mathematically described gestures are perceived as precise machines. However, on one hand, the absence of a score made it impossible to carry on with traditional analysis based on it and on the other, the listeners increased difficulty in relating electronic sounds to known gestures created a vacuum in the model. In any case, the process of composing in the studio created a different kind of composer, a composer that performed and assembled his sounds into a composition, a composer-performer and a computer-composer whose 'working gestures' ended up unrecognizable in terms of traditional gesture. The studio became an action-perception loop in its own right, but failed to transpose it into the 'live' concert setting.

When asked about his musical 'role' Michael Waisvisz responded in the following way:

*Krefeld:* What do you consider yourself in the first place: an inventor of musical instruments, a composer, or a performer?

*Waisvisz:* A composer - a composer of timbres. Due to the state of technological developments in the current era, I'm a composer using electronic means because of their differentiated and refined control over timbre. 'The current era' has lasted 36 years... In my view, the term 'electronic music composer' implies being a performer as well; you cannot sit behind a desk and write electronic timbral music without hearing it. Aside from this, serialism has taken many of us away from composing by ear. I think that a composer has to be able to make immediate compositional decisions based on actual perception of sound rather than making decisions derived from a formal structure that - as happened in serialism - tends to drift away from our pure musical needs. Composers must go back on the stage and listen and think; they must work and perform where the music actually reigns.

With respect to the inventor role, I consider the creation of a specific electronic music instrument as being part of the compositional process ... The way a sound is created and controlled has such an influence on its musical character that one can say that the method of translating the performer's gesture into sound is part of the compositional method. Composing a piece implies building special instruments to perform it as

well. The inventor role is thus an integral part of composing. Your question suggests divisions that don't exist for me; I cannot see a personal involvement in the technical functionality of the instruments and performance as separate from the work of composing, so simply consider me a composer. (Krefeld, 1990)

Michael Waisvisz's words, clearly exemplified in his own work, propose not only a composer/performer, but a listener and instrument builder. In a way, one of the most important implications of his statements is the need to create alternatives to studio performance/automation/layering technique to control the spectromorphologies to which Smalley makes reference. The way sounds are controlled define the sounds, so building interfaces to perform these sounds become a natural need of composers.

But what are these instruments? Schnell and Battier (2002) introduced the term *composed instrument* to refer to the fact that "computer systems used in musical performance carry as much the notion of an instrument as that of a score, in the sense of determining various aspects of a musical work." In this sense talking about instruments and composition in Computer Music becomes impossible to separate, almost in the same way the mind cannot be dissociated from the body. Battier (1999) articulates the idea of a composed instrument in three categories: (1) Musical Instrument, (2) Machine and (3) Representation. A musical instrument in the sense that it allows a performer to explore and play with it, a machine in the sense that "it is under the control of complex computational and algorithmic layers" and interpretation as integrating the previous categories, allowing the composer to "define events, write scores and specify the computational and algorithmic layers while performers can apply gestural controls and adjust parameters". The authors exemplify composed instruments through the pieces realized at IRCAM by Philippe Manoury, *Jupiter* (1987) for MIDI flute and real-time processor, and *En Echo* (1993-1994) for voice and real-time processor; by Brice Pauset, *Perspectivae Syntagma I* (1997) for piano and processing; by Jonathan Harvey, *Advaya* (1994) for cello and electronic set-up; and, by Yan Maresz, *Metallics* (1994-1995) for trumpet and live electronics.

Although this concept is interesting and in the right line of thought, I believe

it fails when dissected. Battier's characterization of composed instruments as musical instrument, machine and representation becomes a contradiction. He fails to identify the composed instrument as a whole where even though the gestural interface and the sound production unit are decoupled, the composed instrument precisely couples them again through its 'representational' layer. The composed instrument ceases to be a group of elements as in the practical description of the digital musical instrument and the variation of its mappings doesn't make it multiple instruments or a schizophrenic instrument. The decoupling that makes a composed instrument possible disappears when composed, it acquires a life of its own; an identity.

As a different approach to that of the concept of the composed instrument we find George Lewis' composition *Voyager*. Lewis (2000) describes *Voyager* as a "nonhierarchical, interactive musical environment that privileges improvisation". Lewis analyses his computer music system in terms of Robert Rowe's taxonomy of 'player' and 'instrument' paradigms (Rowe, 1993). "In Rowe's terms, *Voyager* functions as an extreme example of a 'player' program, where the computer system does not function as an instrument to be controlled by a performer.... I conceive a performance of *Voyager* as multiple parallel streams of music generation, emanating from both the computers and the humans - a nonhierarchical, improvisational, subject-subject model of discourse, rather than a stimulus/ response setup." (Lewis, 2000)

Although *Voyager* is continually sensing a maximum of two players performing into the system, it possesses the ability to create generative behavior independent from the performer so that "decisions taken by the computer have consequences for the music that must be taken into account by the improviser. With no built-in hierarchy of human leader/computer follower - no 'veto' buttons, footpedals or physical cues - all communication between the system and the improviser takes place sonically." (Lewis, 2000)

Lewis goes further in assessing the program:

For me, what Jerry Garcia called the anti-authoritarian impulse in improvisation led me to pursue the project of de-instrumentalizing the

computer. If the computer is not treated as a musical instrument, but as an independent improviser, difference is partly grounded in the form of program responses that are not necessarily predictable on the basis of outside input. As we have noted earlier, Voyager's response to input has several modes, from complete communion to utter indifference. (Lewis, 2000)

This idea of de-instrumentalization and the programming of the computer as an independent partner is of interest here. Undoubtedly, this partner possesses the usual capacities of a human, the ability to perceive and to act, but also the ability to think and decide. This coincides with the *interactionist* cognition model, where an 'entity' or 'agent' is a "functional and structural whole. This separates it from the 'environment' which may or may not contain other 'entities'. The demarcation of the entity from the environment acts as a 'kind of interface,... the interface allows the entity to act upon the environment and the environment to influence the entity." (Kim and Seifert, 2006) In this sense both human and computing agents have an input interface with which they perceive each other (the ear and the microphone, analogously), and both of them possess as well a sound making ability. The question of agency arises, the same question that Artificial Intelligence has been trying to overcome. And the fact of the matter is still grey. On the one hand the computer needs to be programmed or composed, but the program intends the creation of an agent responsive to an improviser, but with a high degree of independence.

Lewis reminds us that Rowe's classification of 'player' and 'instrument' paradigmatic roles in the creation of interactive music should be viewed as a continuum along which a particular systems model of computer-human interaction can be located. Even within compositions we can find states of 'player' and 'instrument'. The key aspect is interactivity.

Whereas the initial approaches to musical controllers were thought of as extensions of 'players' or 'instruments' as in the case of the *Hyperinstruments* of Machover and Chung (1989) or *extended instruments* of Pressing (1990), Rowe's approach is different in that the whole system is viewed as an 'instrument' and/or 'player'. In this

sense, the stress is on stating the action-perception loop as opposed to a ‘enhancement’ or ‘extension’ of human gestural expressivity.

Kim and Seifert (2006) conceive of the sensed body as doubled into a physical and semiotic body. For him the semiotic body is “The virtual body, called the ‘data body’ by the German philosopher Sybille Krämer, embodies the physical body in a manipulable data form.” Kim goes further in his ideas:

The aspect of embodiment of the physical body in the form of a data body becomes the focus of attention within the scope of bodily based algorithmic sound generation since interaction takes place between the data body and sound. Yet this does not immediately imply a disappearance of the physical body. In contrast, the physical body remains a condition on which the virtual body will be able to act. This metamorphosis of the physical body into an actually existing symbol structure opens the possibility that the data body not only acts, but also interacts with other semiotic entities within a virtual environment (Krämer, 2001). Thus the physical body is doubled into a physical and a semiotic body through embodiment in an algorithmic space. The embodied body that is a physical and, at the same time, a semiotic body presents a condition for interaction in which relationships of internal and external, natural and artificial, as well as biological and technical processes become interwoven. Kim and Seifert (2006)

This duality, becomes important in the sense that the body isn’t extended, but somehow paralleled or represented as a symbol structure and in this sense is demarcated. Only through this demarcation, is it able to demarcate its environment and interact with it. What Kim stresses in this statement is the fact that the physical body per se, doesn’t interact with the computer; the data body does and the data body can only exist through sensing.

This brings us to another element that is present in both Waisvisz’s instrument/composition, Manoury’s composed instrument and Lewis’s Voyager system. That is the presence of a human performer; the presence of a physical body.

*‘Live’, seeing and the body*

Auslander (1999) states the following: “Prior to the advent of those technologies (e.g. sound recording and motion pictures). There was no such a thing as ‘live’

performance, for that category has meaning only in relation to an opposing possibility.” pointing us then to Baudrillard: “the very definition of the real is that of which it is possible to give an equivalent reproduction.” The term first appeared to distinguish radio broadcasts, because in radio, it was impossible to distinguish when performers were actually performing and when they were playing recordings, since in our living room, the fact that it was a recording was evident. In this sense, the term ‘live’ grew as a need from technology itself to distinguish between recorded performance and ‘live’ performance and this distinction grew exclusively from the broadcast technology. “As a consequence of the circumstances under which this vocabulary was instated, the distinction between the live and the recorded was reconceived as one of binary opposition rather than complementarity.” (Auslander, 2002)

The similarity of broadcasting and playback, in the sense that the sources aren’t visible and the sounds come out from speakers is of interest here, since they have been the rule with which we have labeled acousmatic music as recorded and live performance as ‘live’; labels that mean much more than their evident meaning. These cultural meanings are present through several value judgments like Attali’s: “Stated very simply, representation in the system of commerce is that which arises from a singular act; repetition is that which is mass-produced. Thus a concert is a representation, but also a meal a la carte in a restaurant; a phonograph record or a can of food is repetition.” (Attali, 1985) Benjamin on the other hand claims that in reproduction the ‘aura’ of a unique object is destroyed: “To pry an object from its shell, to destroy its aura, is the mark of a perception whose “sense of the universal equality of things” has increased to such a degree that it extracts it even from a unique object by means of reproduction.” (Benjamin, 1936)

From this it can follow that the ‘aura’ can’t be captured. Benjamin claims that this contemporary decay of the ‘aura’ rests on the desire of contemporary masses to bring things ‘closer’ and the desire of ‘permanence’; or put in other words, these masses are willing to sacrifice the ‘aura’ for proximity and permanence, pairing uniqueness with transitoriness and reproduction with permanence.

Using Auslander's terms we can view 'live' electronic music as mediatized performance: "as soon as electric amplification is used, one might say that an event is mediatized". (Auslander, 1999) In a way, mediatization in electronic music performance was present since the 'live electronics' pieces, be them tape and instrument compositions, Cage's *Imaginary Landscapes* or Stockhausen's *Mikrophonie*. "It is only since the advent of mechanical and electric technologies of recording and reproduction, however, that performance has been mediatized." (Auslander, 1999) This is particularly clear in the processing of live sounds with the computer with its characteristic lag caused by the block size of audio engines. Even if this block size were to be reduced to one sample, it could never exist at the same time.

So in this mixed media, of 'live' and mediatized, can we claim the birth of a new aura? Does this new performance setting conceal our desire for closeness? Is this a new singularity, a new kind of uniqueness?

For Auslander, there are two conventional explanations to the continued interest in live performance: that it "appeals broadly to the senses and that it creates community".

Much criticism has been made to studio electronic music for its acousmatic nature and specifically for the absence of a 'live performer' and therefore for a 'visual element'. While there is abundant literature on the history and hierarchy of the senses, few authors touch on the multimodal nature of perception in a deep and meaningful way, especially in the field of human-computer interaction. Cognitive science,

It is my intuition that when electronic sounds are controlled live by a human body, both the performer and the audience are provided with sensory stimulus across senses that allows them to adhere sounds, or features of those sounds, to gestures, or features of those gestures.

Although the concept of the embodied mind as a mind which perceives and acts upon the world is central to interactionism, multimodal perception requires us to explore the concept of synthesis. For that we'll need to draw concepts from cognitive science and phenomenology.



The great majority of experiments performed in cognitive science have been performed with highly artificial conditions often in a single sensory modality. This impoverished stimuli has little correspondence with the way in which we perceive in real-life conditions, that is through all our senses through time. Although these experiments can claim interesting facts about our perceptual system, they cannot fully address the way we perceive across senses through time:

...few real-life perceptions are instantaneous, or involve only one sense uninformed even by proprioceptive awareness of bodily orientation. In most human perception, unconscious mental processes synthesize a complex temporal manifold of successive signals, coming in through more than one sensory channel, into conceptualized representations of persisting (and often, changing) objects in space and time. (Stevenson, 2000)

Multimodal perception implies the existence of the body as it provides us with multiple modes of perception. But the need for a ‘mind’ arises when we perceive a single object like a ball through multiple senses. As we see a ball and touch it, we don’t perceive two balls, but one; that is, we synthesize information from several senses into the perception of one object. This process has been called synthesis and since it is multimodal, it has to happen at a superior level from those of the senses, but through a temporal and spatial coordination between them.

“Synthetic Unitites of Experience” by Stevenson (2000), reviews the philosophical literature on synthesis inspired mainly by Kant, Merleau-Ponty and Sellars. Stevenson’s basic premises are two:

An experience of a complex is not a complex of experiences.  
An impression of a complex is a complex of impressions.

The process by which we create conceptualized representations of the world has been a long lasting problem in philosophy, psychology and cognitive science. Stevenson points us to Kantian concept of *Erfahrung*:

Kant argued that *Erfahrung*, i.e. human conceptualized experience, is the result of preconceptual, sub-personal, mental processing. If we

are to think of perceived objects and make judgments about them, there must be psychological processes (which he called "synthesis", "working up", or "combination") by which the manifold and disparate elements of raw sensory input (i.e. the immediate effects of the world on the sense-organs) are somehow integrated, organized or processed into conscious, conceptual representations of objects and states of affairs. (Stevenson, 2000)

Although we won't delve into the depths of conscious and subconscious experience and many other ontological and phenomenological problems, for Kant, the world comes to us as a set of sense-impressions which need to be constructed as representations or conceptualizations of them. For him, these take place some 'time/place' before they reach conscious experience.

Since the advent of computers in art, there have many attempts to fulfill the dream of synesthetic art. The ability of computers to produce multiple synchronized media is what has inspired these attempts. In our case, what concerns us is the ability of the computer to receive sensory input and convert it into sound. This is perceived by the performer as haptic, visual and sonic feedback to his actions and by the audience as visual and auditory information. In both cases the process of synthesis is necessary.

On one side, the computer acts as a sort of human mind/body, in the sense that it synthesizes the input of several modes of perception such as visual and auditory (when using cameras and microphones) to produce one or many sonic events. This is done through continuous computation of incoming data.

This data on the other hand is the 'data body' as termed by Krämer, which is the doubling of the physical body of the performer that the performer itself, performs and that the audience sees. This data body, which interacts with the computer algorithms (would we call this a composition?) then is synthesized into sound. We on the other side, are presented with these phenomenon to synthesize them into one experience.

The question of perceptual fusion has long concerned computer music. While Mathews and Pierce (1980) demonstrated how sine waves were perceived as one sound when their frequencies were multiples of a fundamental frequency, Chowning (1999)

demonstrated that we perceive them as belonging together, as a single entity when they behave in the same way over time.

In the same way and following Stevenson's premise that "A perceptual experience of succession is not reducible to a succession of perceptual experiences" we can deduce that when the audience is presented with a sound that varies in the same way as a gesture over time, they are perceived as the same, or as belonging together. Or on the other extreme, like in Lewis' *Voyager*, the synchronized response of the computer with sound responses that vary differently than the gestures can be perceived as not belonging together. It is this contrast between gesture and body on one side and sounds in the other which allows us to assert the belongingness of sounds to those gestures, belonging together when they vary in the same way through time: "A cross-modal perceptual experience involves more than a simultaneous conjunction of single-modality perceptions caused by the same object. Incoming perceptual information is synthesized into a single scene - a representation of the world as perceived, from a spatio-temporal point of view." (Stevenson, 2000)

Michel Chion invented the term *synchresis* meaning the following: "(a word I have forged by combining synchronism and synthesis) is the spontaneous and irresistible weld produced between a particular auditory phenomenon and visual phenomenon produced at the same time. This join results independently of any rational logic." (Chion, 1994) Chion in this case comes from the theory of audiovisual media and specifically from cinema. As our study pertains the seeing of gestures and the hearing of sounds, many of cinema's theories are pertinent in our study of the synthesis of synchronized input from these two modes of perception.

An important concept developed by Chion is that of *de-acousmatization* of sounds. For him this process occurs when we are presented with sounds that we don't know the source of and whose source later appears in the screen. Similar concepts include on-screen, offscreen and *nondiagetic* sounds:

In a film an *acousmatic* situation can develop along two different scenarios: either a sound is visualized first, and subsequently *acousmatized*,

or it is acousmatic to start with, and is visualized only afterward. The first case associates a sound with a precise image from the outset. This image can then reappear with greater or lesser distinctness in the spectators mind each time the source is heard acousmatically. It will be an 'embodied' sound, identified with an image, demythologized, classified.(Chion, 1994)

This embodiment Chion makes reference to, is a direct reference to the fact that sounds belong to a movement. In this sense we can talk about embodied sounds. In many live electronics pieces, like Manoury's *Jupiter* or *Neptune*, the computer captures sounds from the live performance, and then presents them back transformed. This mode of presenting acousmatic sounds with direct referentiality to a previous 'embodied' occurrence, reduces the strangeness or remoteness of the sound. It is somehow a familiar gesture. In Cinema terms, it is an onscreen sound later rendered as offscreen. We can also relate this phenomenon to delay based effects.

Let's briefly review Chion's take on embodiment, for him embodied sounds are visualized sounds: "What can we call the opposite of acousmatic sound? Schaeffer proposed 'direct', but since this word lends itself to so much ambiguity, we shall coin the term 'visualized' sound i.e. accompanied by the sight of its source or cause." (Chion, 1994)

But in electronic music the source of the sound is never really viewed; the computer acts as a black-box where a invisible phenomenon occurs that we can't really asses. Still, once presented in the phenomenon of synchresis, for Chion, it is context and verisimilitude what gives them their final belongingness, and context and verisimilitude are different phenomena in music than in film. So to what extent does a viewed gesture with a sound that behaves in the same way de-acousmatizes the sound? Or is it that the source of the actual sound is not important and what is important is the seeing of its control?

This idea of context and verisimilitude is also a long standing problem in phenomenology. This is what Merleau-Ponty calls being in the world and the concept of the unity of experience:

When Gestalt theory informs us that a figure on a background is the simplest sense-given available to us, we reply that this is not a contingent characteristic of factual perception ... It is the very definition of the phenomenon of perception ... The perceptual 'something' is always in the middle of something else, it always forms part of a 'field'.

The tacit thesis of perception is that at every instant experience can be co-ordinated with that of the previous instant and that of the following, and my perspective with that of other consciousnesses - that all contradictions can be removed, that monadic and intersubjective experience is one unbroken text. (Merleau-Ponty, 1945)

This unity of experience creates a link between everything that happens in a space and through time. In this sense, all of our experiences are just one experience from our point of view. Still, in ordinary experience we are able to identify experiences within our experience of the world. In this sense, what Chion calls 'field of vision' we can call 'field of audition' for music, or even more precisely, 'field of experience' to that which happens in the concert hall and we unite as a specific place. In the same sense, our experience of time is delimited in performance. Performances begin and end and we are able to unite them as a single experience. Our memory then acts as a unifier, as a source for context.

Another important point to take into account is the way in which input from one sensory mode informs and complements information in the other. Many studies in speech perception stress the importance of seeing a face in the intelligibility of speech. The well known McGurk effect (McGurk and MacDonald, 1976) demonstrated in the 1976 paper called "Hearing lips and seeing voices", showed that when a subject was presented with an image of a face pronouncing the phoneme 'bah' and a sound of the phoneme 'dah', the person perceived 'bah'. That is, the visual element influenced the sound to the extent that the sound information was partially overwritten.

Similarly in Cinema, there are many cases where the contrary effect happens. The concept of added value coined by Chion explains this: "By added value I mean the expressive and informative value with which a sound enriches a given image so as to create the definite impression that this information or 'expression' naturally comes

from what is seen, and is already contained in the image itself.”(Chion, 1994) Chion’s comments on kung-fu movies are also informative, attributing the effect of sound on image to a difference of speed between the senses:

Basically the ear analyzes, processes, and synthesizes faster than the eye. Why, for example, the myriad rapid visual movements in kung fu or special effects movies create a confusing impression? The answer is that they are spotted by rapid auditory punctuation, in the form of whistles, shouts, bangs, and tinkling that mark certain moments and leave a strong audiovisual memory.

In this sense visual gestures inform sonic gestures and vice-versa. Elaboration on these issues could lead to a theory of gesture and sound. In a way, when we treat the visual aspect of musical performance as an image similar to cinema, as the image of music, we become aware of the extent to which our experience of the world through mediatized content is shaping our perceptual system. In reference Auslander’s concept of mediatized performance and Bolter’s concept of re-mediation, we could even claim that musical performance as a whole is re-mediating cinema, and acoustic music is re-mediating electronic music: “...new technologies remediate older ones, as film and television both remediated theatre” (Auslander, 1999) - “earlier technologies are struggling to maintain their legitimacy by remediating newer ones” (Bolter and Grusin, 1998)

Finally, what is it about the body that makes us want a live human performer on stage. Is there a hidden or perhaps explicit cultural connection between the concept of ‘live’ and life? How does our body condition our cultural preferences?

For Blacking, all social activity, like music, is biologically determined and requires social interaction. We learn the possibilities and limits of our bodies, like speaking or singing or even moving, by social interaction, determined in turn by a historical and ethical context. A crucial factor in this development of cultural forms is that of the possibility of shared *somatic states*, that is, the ability of a group of people to experience and behave in a similar way. For him this sharing is permitted by “the structures of the bodies that share them, and the rhythms of interaction that transform commonly experienced internal sensations into externally visible and transmissible forms.” (Blacking,

1977) Or in another words, our ability to perform and perceive with our bodies, is a shared condition that allows us to experience things in a similar way.

This primarily points to the fact that we all perceive with the same cognitive system, and potentially we can achieve a shared somatic state, that is, that our bodies can potentially have the same experience: “...every normal member of the species possesses not only a common repertoire of somatic states and a common potential for the altered states of consciousness... ”. But on the other hand, this potential, this shared body is also the reason why we want to see a human performer: “... what people can do and what they are capable of doing. If some humans can perform certain skills, it should be possible for any member of the species to do so, given the appropriate social and cultural environment, and similar opportunities and incentives...”(Blacking, 1977) by extension, our individual bodies possess particular limits, but the limits of ‘the body’ are not shaped by our own capabilities, but by the capabilities of other bodies. The bodies of others shape the limits of ‘the body’; of ‘our bodies’.

That is, even when composers exercise their conscious minds to produce a unified musical structure, the finishing touches of the work are often done by the body; and because it is a human body that is creating symbolic order with the minimum of cultural interference, the music resonates with a variety of people who may share little or nothing of the composers background.(Blacking, 1977)

So can we claim that our familiar bodies de-acousmatize electronic sound? Can we claim that electronic sound that is shaped by the body becomes humanized to the human eye? Does this ‘embodiment’ of sound compensate for the removal of the listener from the black-box that is creating the music?

I must stress that these questions can’t be answered asserting them as true or false problems. In a way, most composers and electronic artists will just claim that something works or it doesn’t. But the richer the sensorial experience, the more the body is addressed, the more tools we have to apprehend these ‘remote’ sounds with surrogate gestures. The body - seeing the body - makes them somehow familiar, it allows us to experience sounds through it.

# Chapter 5

## Conclusions

The body has historically been a crucial factor in the development of culture and social interaction. Music hasn't escaped this and the tradition of acoustic instruments present in any culture require the body to perform with them. Our ability of shared somatic states is determined by the structure of our bodies both as cognitive and performing entities.

The action-perception paradigm of the interactionist view in cognitive science combined with the fact that our perceptual system is endowed with a capacity to understand independently from the mind gives us the possibility of re-mediating remote and surrogate gestures in relation to the bodily gestures of acoustic performance. This can be done through the use of what have been called digital musical instruments, but more generally through the use of computers that can sense the world and synthesize those sense-impressions into sounds.

Controllers and interactive computer systems have provide us with an interface that allows us to re-contextualize electronic sounds in the context of live music performance; an interface into sounds or the ability to double our body into the algorithmic space of sounds where we interact with them. This practice provides the audience and the performer with haptic, visual and auditory stimuli that fit better the multimodal nature of our sared perceptual systems, allowing us to attach meaning to sounds otherwise



detached from our gesture based musical cognitive system. The question of seeing becomes of great importance in the sense that to see is in a way to listen differently and vice-versa.

In this way, the use of the body acts as a medium to incorporate or embody electronic sound as it hadn't been done before. Computers acquire a dual nature as part of the 'environment' with which we interact, but also as the 'doubling' of our embodied minds.

This re-mediation of electronic sounds through the use of the human body acts as a legitimator of both electronic music in many cases perceived as cerebral, foreign or artificial, but also helps traditional instrumental performance in establishing its currency and ability to renew itself as a practice of human beings in the contemporary arena, a space where medatized practices like the cinema are transofrming our being in the world.

On a personal level, my own creative work is based primarily on intuition. I make the music I want to hear and create experiences that exalt me, hoping that by achieving such somatic states, they can be perceived by others.

# References

- Adorno, T., 1976: *Introduction to the sociology of music*. New York: Seabury Press.
- Anderson, M., 2003: Embodied Cognition: A field guide. *Artificial Intelligence*, **149**(1), 91–130.
- Appleton, J., 1984: Live and In Concert: Composer/Performer Views of Real-time Performance Systems. *Computer Music Journal*, **8**(1), 48–51.
- Attali, J., 1985: *Noise: the political economy of music*. Minneapolis: University of Minnesota Press.
- Auslander, P., 1999: *Liveness: Performance in a Mediatized Culture*. Routledge.
- Auslander, P., 2002: Live From Cyberspace: or, I was sitting at my computer this guy appeared he thought I was a bot. *PAJ: A Journal of Performance and Art*, **24**(1), 16–21.
- Battier, M., 1999: Les polarités de la lutherie électronique. Instrument, machine, représentation. *Méthodes nouvelles, musiques nouvelles, musicologie et création*, 307–318.
- Bayle, F., 2007: Space, and more. *Organised Sound*, **12**(03), 241–249.
- Benjamin, W., 1936: The Work of Art in the Age of Mechanical Reproduction. *First published*.
- Blacking, J., 1977: Towards an anthropology of the body. *The Anthropology of the Body*, 1–28.
- Bolter, J., and Grusin, R., 1998: *Remediation: understanding new media*. MIT Press Cambridge, MA, USA.
- Bonnet, M., Guiard, Y., Requin, J., and Semjen, A., 1994: Mécanismes généraux de la motricité. *Traité de Psychologie Expérimentale*. Presses Universitaires de France.
- Borgo, D., 2005: *Sync Or Swarm: Improvising Music In A Complex Age*. Continuum International Publishing Group.

- Brech, M., 2002: New Technology - New Artistic Genres. *Music and Technology in the Twentieth Century*.
- Cadoz, C., and Wanderley, M., 2000: Gesture-Music. *Trends in Gestural Control of Music*, 71–93.
- Chadabe, J., 2002: The limitations of mapping as a structural descriptive in electronic instruments. *Proceedings of the 2002 conference on New interfaces for musical expression*, 1–5.
- Chion, M., 1994: *Audio-Vision: Sound on Screen*. Columbia University Press.
- Choi, I., 2000: Gestural Primitives and the Context for Computational Processing in an Interactive Performance System. *Trends in Gestural Control of Music*, 139–172.
- Chowning, J., 1999: Perceptual fusion and auditory perspective. *Music, cognition, and computerized sound: an introduction to psychoacoustics table of contents*, 261–275.
- Clarke, E., 1993: Generativity, mimesis and the human body in music performance. *Music and the Cognitive Sciences 1990: Proceedings of Cambridge Conference on Music and the Cognitive Sciences, 1990*.
- Cusick, S., 1994: Feminist Theory, Music Theory, and the Mind/Body Problem. *Perspectives of New Music*, **32**(1), 8–27.
- Delalande, F., 1988: La gestique de Gould. *In Glenn Gould Pluriel*, 85–111.
- Delalande, F., 2003: Sense and Intersensoriality. *Leonardo*, **36**(4), 313–316.
- Goldstein, M., 1998: Gestural coherence and musical interaction design. *Systems, Man, and Cybernetics, 1998. 1998 IEEE International Conference on*, **2**, 1076–1079.
- Goudeseune, C., 2003: Interpolated mappings for musical instruments. *Organised Sound*, **7**(02), 85–96.
- Hunt, A., and Wanderley, M., 2003: Mapping performer parameters to synthesis engines. *Organised Sound*, **7**(02), 97–108.
- Hunt, A., Wanderley, M., and Kirk, R., 2000: Towards a model for instrumental mapping in expert musical interaction. *Proc. of the 2000 International Computer Music Conference. San Francisco, CA: International Computer Music Association*, 209–211.
- Hunt, A., Wanderley, M., and Paradis, M., 2002: The importance of parameter mapping in electronic instrument design. *Proceedings of the 2002 conference on New interfaces for musical expression*, 1–6.

- Iazzetta, F., 2000: Meaning in Musical Gesture. *Trends in Gestural Control of Music*, 259–68.
- Kartomi, M., 2008: Classification of Instruments: The classification system of Hornbostel and Sachs. *Grove Music Online*, Accessed [May, 2008], <http://www.grovemusic.com>.
- Kim, J., and Seifert, U., 2006: Embodiment: The body in algorithmic sound generation. *Contemporary Music Review*, **25**(1), 139–149.
- Krefeld, V., 1990: The Hand in The Web: An Interview with Michel Waisvisz. *Computer Music Journal*, **14**(2), 28–33.
- Lee, M., Freed, A., and Wessel, D., 1991: Real-Time Neural Network Processing of Gestural and Acoustic Signals. *Proceedings of the 17th International Computer Music Conference, Montreal*.
- Lettvin, J., Maturana, H., McCulloch, W., and Pitts, W., 1959: What the Frog's Eye Tells the Frog's Brain. *Proceedings of the IRE*, **47**(11), 1940–1951.
- Levitin, D., McAdams, S., and Adams, R., 2003: Control parameters for musical instruments: a foundation for new mappings of gesture to sound. *Organised Sound*, **7**(02), 171–189.
- Lewis, G., 2000: Too Many Notes: Computers, Complexity and Culture in Voyager. *Leonardo Music Journal*, 33–39.
- Machover, T., and Chung, J., 1989: Hyperinstruments: Musically intelligent and interactive performance and creativity systems. *Proceedings of the 1989 International Computer Music Conference*, 186–190.
- Mathews, M., 1963: The Digital Computer as a Musical Instrument. *Science*, **142**(3592), 553–557.
- Mathews, M., and Moore, F., 1970: GROOVEa program to compose, store, and edit functions of time. *Communications of the ACM*, **13**(12), 715–721.
- Mathews, M., and Pierce, J., 1980: Harmony and nonharmonic partials. *The Journal of the Acoustical Society of America*, **68**, 1252.
- McGurk, H., and MacDonald, J., 1976: Hearing lips and seeing voices. *Nature*, **264**(5588), 746–48.
- Merleau-Ponty, M., 1945: *Phénoménologie de la perception*. Gallimard, [Paris].
- Miranda, E., and Wanderley, M., 2006: *New Digital Musical Instruments: Control And Interaction Beyond the Keyboard*. AR Editions.

- Nono, L., 1984: Verso Prometeo. Frammenti di diari. *Luigi Nono: Scritti e colloqui I*, 385–396.
- Ostertag, B., 2002: Human Bodies, Computer Music. *Leonardo Music Journal*, 11–14.
- Pressing, J., 1990: Cybernetic issues in interactive performance systems. *COMP. MUSIC J.*, **14**(1), 12–25.
- Pritchard, R., Heron, W., and Hebb, D., 1960: Visual perception approached by the method of stabilized images. *Can J Psychol*, **14**, 67–77.
- Rovan, J., Wanderley, M., Dubnov, S., and Depalle, P., 1997: Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. *Kansei, The Technology of Emotion. Proceedings of the AIMI International Workshop*, 68–73.
- Rowe, R., 1993: *Interactive music systems*. MIT Press Cambridge, Mass.
- Schick, S., 2006: *The percussionist's art : same bed, different dreams*. Rochester, NY : University of Rochester Press.
- Schnell, N., and Battier, M., 2002: Introducing composed instruments, technical and musicological implications. *Proceedings of the 2002 conference on New interfaces for musical expression*, 1–5.
- Smalley, D., 2001: Spectromorphology: explaining sound-shapes. *Organised Sound*, **2**(02), 107–126.
- Stevenson, L., 2000: Synthetic Unities of Experience. *Philosophy and Phenomenological Research*, **60**(2), 281–305.
- Stubley, E., 1998: Being in the Body, Being in the Sound: A Tale of Modulating Identities and Lost Potential. *Journal of Aesthetic Education*, **32**(4), 93–105.
- Supper, M., 2004: Música Electrónica y Música con Ordenador. *Alianza Musica*.
- Tanaka, A., 2000: Musical performance practice on sensor-based instruments. *Trends in Gestural Control of Music*, 389–405.
- Viviani, P., 1994: Les habiletés motrices. *Traité de Psychologie Expérimentale*, 777–857.