Title
A computer music instrumentarium

Author
Oliver La Rosa, Jaime Eduardo

Publication Date
2011

UNIVERSITY OF CALIFORNIA, SAN DIEGO

A Computer Music Instrumentarium

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Music

by

Jaime Eduardo Oliver La Rosa

Committee in charge:

Miller Puckette, Chair
Amy Alexander
Gerald Balzano
Jim Hollan
David Wessel

2011

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

Chair

University of California, San Diego

2011

DEDICATION

A Paloma y Cristóbal.

EPIGRAPH

* * *

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

Waisvisz: ... 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 (Waisvisz in Krefeld (1990)).

* * *

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank Miller Puckette, for giving me the opportunity of working with him over the past 5 years, dedicating time to feed my work and think with me about the ideas set forth this dissertation. His musical work has often served as a point of departure for my own and his work designing and supporting Pd makes him a (perhaps involuntary) co-author of my work.

I would also like to thank David Wessel, for his inspirational work and for his interest, guidance and infinite ideas; Jerry Balzano, for his dedication to educate a curious musician in cognitive science and for all the great conversation; Amy Alexander and Jim Hollan, for their comments and insight into this dissertation.

I would also like to thank Roger Reynolds, for his advice, teaching and the innumerable layers of experimentation and craft in our fertile collaboration over the past 3 years. Many of the ideas for this dissertation emerge from our work together designing a performance instrument for the Seasons Cycle, Dream Mirror and Marked Music.

I would also like to thank Philippe Manoury, Dick Moore, Shlomo Dubnov and Kevin Larke for their work, teaching, help and interest along the way; as well as Steven Schick and Katharina Rosenberger, the computer music people and the truly unique music community at UCSD for their interest and collaboration.

Finally, I would like to thank my parents Jimmy and Polla, and my brothers Diego and Luciana, for their support; and very specially my wife Paloma and son Cristóbal to whom I owe everything.

VITA

2003	Composition, National Conservatory of Music, Lima, Perú.
2006-08	M.A. Computer Music, UC San Diego.
2008-11	Ph.D. Computer Music, UC San Diego.
2011-13	Mellon Post-Doctoral Fellow in Composition, Columbia University.

PUBLICATIONS

Oliver, J., and Jenkins, M., 2008: The Silent Drum Controller: A New Percussive Gestural Interface. In *Proceedings of the International Computer Music Conference, Belfast.* International Computer Music Association.

Oliver, J., 2010: The MANO Controller: A Video-based Hand Tracking System. In *Proceedings of the International Computer Music Conference, New York.* International Computer Music Association.

ABSTRACT OF THE DISSERTATION

A Computer Music Instrumentarium

by

Jaime Eduardo Oliver La Rosa
Doctor of Philosophy in Music

University of California, San Diego, 2011

Miller Puckette, Chair

This dissertation reviews representative works of the history of electronic and computer music from the point of view of instrument design. It traces the way in which artists conceived of their systems with respect to traditional musical instruments and practices. While making evident the inherent differences between the mechanical media of the past and the new electronic and computer media at their hands, artists have forged a new kind of instrument.

This instrument is presented in contrast to the traditional concept of musical instrument as a passive tool; as an object that provides a stable timbre over which pitch and amplitude can be articulated. In contrast, the computer-based instrument has an active role in determining the shape and sound character of a composition.

The traditional conception of music, as closed works contained in scores, shifts to open environmental structures that can only be perceived and experienced through interaction. Theories of embodiment in cognitive science are surveyed to understand the nature of this interaction. Concepts from media theory are used to understand the

process whereby new instruments imitate older instruments, while at the same time revealing what is unique to them as new media.

The musician then finds himself to be a hybrid between composer, luthier and performer. However, he cannot start from scratch in every new work of music. In the not so brief history, we find archetypical practices that guide us; communities around computer programs that offer a body of knowledge and code that can be analyzed, appropriated and modified. Because of its relational nature, code contains part of the instrument (and the composition); its exchange advances music as a cultural practice.

An Instrumentarium suggests a set of available instruments from which the composer chooses. In the practice of making music with computers however, there are instead, sensors and interfaces, techniques for sound analysis, generation and manipulation, recording devices, hardware, software and speakers. An instrument is not contained in any one object, but consists of a series of elements that can be combined to form an open configuration. The computer music instrumentarium lies in all possible configurations.

Chapter 1

Introduction

In recent times, interest has resurged in making new musical instruments[1], and, with the availability of increasingly powerful new technologies like computers and a growing variety of sensors, designing instruments that use these technologies is an attractive idea.

However, building instruments based on computers, or what I will call computer-based musical instruments, confronts us with wider questions about our musical practice, like for example, what a musical instrument is in the first place? or, what do we expect a device to be able to do, in order to deserve inclusion in the category of musical instrument? or, going even further, is musical instrument the right term for the devices and resources we use in these new musical practices? Awareness of these wider issues informs our design and guides our exploration. This is the subject of this dissertation.

For Kartomi (2008), "musical instrument is a self-explanatory term for an observer in his own society." This dissertation is an intellectual engagement with the work and writings of composers and theorists of electronic and computer music, however; Kartomi's definition has guided its design. In other words, it is left to our intuition as cultural beings, to recognize an instrument and to intellectual analysis to understand

---

[1] Two examples of this new interest are the creation and growth of the *New Interfaces for Musical Expression (NIME)* conferences since 2001 and the *Guthman Musical Instrument Competition* at Georgia Tech Center for Music Technology since 2009.

how this new instrument is similar or different from the ones we know.

Technological innovation in cultural practices generates, and responds to, change. We can observe aspects of this change in the objects that this practice produces. In musical practices, these objects are usually scores and musical instruments. Much attention has been paid to musical scores as a means to understand the nature of the changes in music over time, but very little attention has been paid to musical instruments. Most studies of musical instruments are organological, that is, concerned with instrument classification, as collectors of instruments began to flourish in Europe and North America (Kartomi, 1990). However, understanding the evolution of instruments provides us with an insight into the aesthetic values and needs of a culture.

Paying attention to mechanical musical instruments, that is, instruments that do not use electricity, provides us with a benchmark. On one hand, since these were the only instruments we had, they configure what we, as a culture, understand musical instruments to be. On the other, they offer an opportunity to determine what features and qualities they share with the new electric, computer-based instruments, in order to find a broader, or perhaps deeper, way of defining and understanding musical instruments in general.

To talk about musical instrument design from the 20th century to date, is to talk about the use of electricity and computation technologies in music. One of the aims of this dissertation is to identify the ways in which composers and theorists have constructed the concept of musical instrument.

In order to construe electronic music as new, its inventors and practitioners in the seminal studios of the 1950's, presented it as a stage beyond the old instrumental music. In this context, a musical instrument was understood as a physical and mechanical system (performer included) that imposed multiple limitations. In this sense, an instrument was the opposite of electronic music which was construed as a practice that is *free* of the physical limitations of instruments and performers. But as soon as that premise is accepted or even posited, electronic music becomes the practice of designing the new instruments that let us achieve that *freedom*.

This is, I will argue, a recurring paradox in electronic and computer music: we search to free ourselves from the perceived constraint of instruments only to find ourselves making the instruments that permit this freedom. As a result, new ways of doing things are discovered.

In developing these new practices, that is, in attempting to overcome the limitations of instruments, we have transformed or extended them. These extensions are not about making violins of wider range or flutes that can play glissandi. These extensions or transformations are about what the role of the instrument is in the practice of music as a whole. By analyzing the conception and realization of representative works of electronic and computer music, I intend to explore the way composers design (and often perform) their own instruments as an inherent part of their compositional practice. I will argue that these instruments contain part of the composition and the musical thought, or theory, out of which this composition is realized. As music shifts from an ordering of pitched notes to sound objects, processes and systems, our instruments change too.

The traditional view of instruments as passive tools is therefore gradually overwritten by new instruments with bigger roles in determining sonic as well as structural aspects of the music made with them. In this context, the specialized roles of composer, performer or luthier, or the functional distinctions between score, instrument and hall, become blurred and give way to hybrids. Moreover, these social and medial relations are now open to be composed.

While the interface, gestural energy, timbral identity and sound production are all independent in the sense that they are not part of the same object as it was the case in wooden tubes or stretched strings; in these new instruments they are arbitrarily, or perhaps compositionally re-organized. Instruments in this sense are configurations; a disposition of elements and a set of relations that bound a system yet allowing the performer to explore and learn it.

In tape composition, we begin to hear the recorded sounds of voices and instruments, without seeing the bodies that produced them and in this sense they are disembodied. The sounds fixed on the tape, do not offer any opportunity for interaction and whatever performance existed it stopped at the final fixing of the sounds on the tape.

These new instruments and works of music are complex environmental structures, open to performers and audience for exploration. By environmental structures I mean that they offer the performer ways of altering its behavior and allow the sounds to be embodied; that is, the performer interacts with the instrument and through that interaction attunes his actions to the behaviors the instrument affords. We turn to theories of embodiment in cognitive science, as they provide a rich framework to examine the way we interact with the world.

However, there are significant differences between the physical, material nature of mechanical instruments and the plastic, malleable character of computer-based instruments. As I will argue, computer-based musical instruments produce a different kind of invariant features from mechanical ones. These invariances emerge from a loop of design and interaction as a compositional process.

Instrument design has become an inherent part of the work of composers of electronic music and a part of the compositions themselves. In this sense, these instruments are often inseparable from the works done with them. These instruments cannot be found in any one object; they are not the interface, the software, the sounds, and so on; they are complex configurations of elements.

However, in designing these new instruments, we tend to remediate or imitate the older mechanical instruments. Our conception of instruments as mechanical tools with stable timbres and fixed morphology, is so pervasive that we reach back to it as a model. Devices that fit the model are considered instruments; ones that do not are not. In imitation, we confront the new medium and by interacting with it, we begin to discover what is unique to computation and digital media.

It is not my intention to write a history of electronic and computer music. If it were, it would be incomplete and imprecise. I am not looking for the first person that did something, but for certain examples of thought, work and practice that better embody a practice. In contrast to most historical accounts of electronic music, this one attempts to understand computer music practices from a genealogical approach, one based on conceptual resonances instead of lists of events, dates and equipment manuals.

Although there have been and still are numerous names for the practice of media-based music[2], I will use *electronic music* to refer to the generic practice of making music with analog equipment and *computer music* to that which uses digital computers. Unless specified or determined by historical context, electronic music and computer music should be seen as parts of the same practice.

What I hope to demonstrate, and expect this dissertation to establish, is that "musical instrument" remains a valid concept towards achieving an understanding of contemporary computer music practices. While we try to re-invent musical instruments all the time, there are clear lines of thought and practice aimed in this direction. The practices of Theremin; Moog and Buchla, Carlos and Subotnick; Mumma and Martirano; Mathews and Moore, Tenney and Risset; Puckette, Manoury, Waisvisz and Lewis, all contribute to configure the new computer music instrumentarium.

---

[2] Such as ether-wave music, electric music, musique contrète, electronische musik, electronic music, tape music, sound art, sonic art, and so on.

Chapter 2

Prelude: Instruments as Media

A reason to begin a dissertation on computer-based musical instruments by analyzing mechanical instruments in terms of media theory, is to understand the way musical instruments have evolved over time.

New Media emerge within established social practices and therefore raise expectations as to what functions and roles we expect these media to fill. In this sense, new media are expected to advance these social practices by being better versions of older media. However, evaluating new media in terms of older media, obscures their unique qualities.

In the light of these ideas, I will attempt to explore the way mechanical musical instruments have been conceived within Western musical practices. By tracing a rough sketch of the evolution of these instruments, I will attempt to draw out what we expect musical instruments to be. This conception of musical instruments is the conceptual template into which we expect newer musical instruments to fit.

* * *

Models of Music Practice

Every music needs to account for a way of producing, shaping and articulating sound. We refer to the actions that produce, shape and articulate sound as *gestures*,

Figure 2.1: Composer-performer-listener model of music

and to the medium that allows us to transduce our gestures into sounds, as a *musical instrument*. We call the act of making music in this way *performance*.

Music always belongs to a larger cultural environment. Before the electric age, music could only be experienced in live performance. All musical practices have in common the fact that music is an activity; it is not an object but "something people do" (Small, 1998). And as Small would argue, it is an activity in which not only the performing musicians have an active role, but one that exists in interaction with a larger social context.

In many cultures, audiences participate by dancing, singing, or expressing themselves in some way or another. The music that is played is not necessarily authored by a composer, but drawn from a traditional repertoire or improvised as a recombination of a musical culture shared by performers and audience alike and embedded somehow in the instruments they use.

In the Western world, music is traditionally conceived as a uni-directional, linear, sequential communication process. The composer gives the music to the performer, and the performer plays it for an audience. Two media are used to enable this process: scores and musical instruments. Figure 2.1 shows what I will be referring as the composer-performer-listener model.

In his *Introduction to the Sociology of Music*, Adorno (1976) states that "[musical] 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 music is an object and not an activity; that the work is complete in itself and independent from its sound realization. This view is reinforced by a tradition of music analysis that focuses exclusively on scores; the score medium contains the totality of the work.

From this traditional standpoint, the score is the composer's encoding of the music and is, in fact, the music. The performer is a body that executes the score and can only aspire to an accurate realization; the audience can perceive and experience with certain degrees of accuracy, conditioned by their musical knowledge and by the accuracy of the performance.

It follows from this view that the performer is a mere transmitter, who uses his instrument to communicate or express the composer's music to an audience. In this model, the performer and his instrument are construed as *passive tools* and the audience as a passive receptor. Passive in this case, refers to the fact that there is only one interpretation and therefore, the performer, his instrument and the audience do not contribute to the meaning of the work. Furthermore, as the term accurate suggests, there is only one meaning.

In what follows, the term *agency* will be used as the way in which a person acts to produce a particular result. In the case of media, it will be used to reflect how design enables or restricts possible behaviors. The uni-directional sequential linearity of the idealized view reflected in the composer-performer-listener model reduces the level of agency of the various actors and media as we progress in the system. In this passive role, the instrument needs to have a predictable outcome and the ideal performer is a machine, capable of the most accurate realization of a score.

We also find in this model the concept of specialized labor. Composers make scores, performers reproduce them, audiences listen. To support the model we find the score publisher and instrument luthier, and in recent times the recording engineer and producer, instrument dealers, auditorium managers, agents, etc.

* * *

Media, Message and Content

Perhaps Marshall McLuhan’s most famous phrase is “the medium is the message”. For McLuhan, the message of a medium was not the content it was expected to communicate, but instead “the ‘message’ of any medium or technology is the change of scale or pace or pattern that it introduces into human affairs” (McLuhan, 1963).

McLuhan was interested in understanding the way technology or media altered social order. For him, people seemed to ignore the effect that media had on them. For example, the message of television was not the programming of content, but the effect it had on how we perceived and related to our environment. Following this logic, the message of print was not the text it contained, but the establishment of new logics where the “typographic principles of uniformity, continuity, and linearity had overlaid the complexities of ancient feudal and oral society.” (McLuhan, 1963)

In other words, the message of any particular technology (or medium) of music, is the change or changes it introduces into musical practice.

In his 1954 essay “The Question Concerning Technology”, Heidegger (2003) points us to the greek term *techné* which he defines as “the name not only for the activities and skills of the craftsman but also for the arts of the mind and the fine arts”. In this sense, we can speak of technologies of music not only when we refer to material objects like musical instruments and notation, but also to the theories we use to think about music.

Today, technology is conceived as exclusively, or perhaps even taken to be a synonym of, electric and digital technology. Much like Heidegger, McLuhan does not restrict technology to tools and machines, but includes technologies of thought. In this sense, language, the alphabet, numbers (or musical scales for that matter) are technologies that we use to think.

McLuhan conceived of technologies as “extensions of man”. For example, clothing extended skin, housing extended clothing, cities extended housing. Following a similar logic, he stated that “the content of any medium is always another medium”, which he exemplified saying that "the content of writing is speech, just as the written word is the content of print, and print is the content of the telegraph" (McLuhan, 1963).

McLuhan believes that content is not only irrelevant to the message of the medium, but that it actually "blinds us to the character of the medium". While most electric media, like radio, television or the phonograph, were used for some sort of communication, McLuhan found in electric light (and power) examples of media without apparent content. Only when electric light is used to "spell out some verbal add or name", we become aware of it as a medium. The appearance of content (in this case print) is now confused with the message of the medium, however, "the content or uses of such media are ineffectual in shaping the form of human association." In other words, what light spells is less important than the fact that we can see it at night or at longer distances than a book. This is precisely McLuhan's point: that the user of a medium is seduced by its content to such an extent that he fails to acknowledge how this medium shapes his ways of perceiving and associating. Therefore, the medium is the message.

> For the "content" of a medium is like the juicy piece of meat carried by the burglar to distract the watch-dog of the mind. The effect of the medium is made strong and intense just because it is given another medium as "content" (McLuhan, 1963).

But how can media have strong and intense effects if we are unable to consciously understand its message?

> Everybody experiences far more than he understands. Yet it is experience, rather than understanding that influences behavior, especially in collective matters of media and technology, where the individual is almost inevitably unaware of their effect upon him (McLuhan, 1963).

Bolter and Grusin coined the term *remediation* (Bolter and Grusin, 1998) to refer to the principle whereby any medium always has another medium as its content. As we will see later, Bolter sees remediation as both what is "unique to digital worlds and what denies the possibility of that uniqueness" ((Bolter and Grusin, 1998), 50).

* * *

The Voice

It is not the object of this dissertation to find the origins of music. However, it is quite probable that the first musical expressions used the voice and the sounds of our bodies exciting objects in the environment. It is impossible to determine when, where and how instruments were first used. Regardless of these facts, logic suggests the voice precedes the use of tubes or strings; and a single string or tube existed before multiple string or tube mechanisms. In other words, I would propose that the content of monophonic instruments is the voice and the content of polyphonic instruments is the monophonic instrument.

Sounds are vibrations with frequencies greater than approximately 20 to 30 hertz. In contrast to the arm-hand system, the muscles in the speech system are able to generate sound vibrations of audible frequencies without the manipulation of objects. Because of this fact, the voice possesses the ability to produce sounds of great gestural and sonic detail: the strength of the air pushed from the lungs through the vocal cords translates into amplitude, the length of the vocal cords controls frequency, the shape of the mouth cavity acts as a filter and the tongue, teeth and mouth cavity allow for the articulation of consonants. The voice is then a very plastic system able to continually transform timbre (vowels and consonants) as much as frequency and amplitude.

The richness of the voice lies in our unmediated, continuous, multidimensional and interdependent control of all of these muscles over time. Moreover, the voice is part of the identity of the singer or speaker. However, while in speech people have a recognizable voice quality to which we refer as ‘their voice’, in singing, and particularly in western vocal tradition, singers seem to aspire towards an ideal voice.

While the voice is free to produce a wide variety of sounds, cultures tend to converge on specific musical behaviors: rhythmic and pitch structures and preferred colors or timbres. Cultures develop styles, techniques and repertoires of these structures that are constantly recombined by musicians. The voice retains its freedom, but it becomes attuned to the musical culture to which it belongs.

The construal of an idealized singing voice different from the speaking voice is what we could call an instrument-voice. The instrument-voices one would use for opera, death metal, beat-boxing and jazz, it could be argued, are all different instruments.

Traditional definitions of musical instrument refer to "implements used to produce music, especially as distinguished from the human voice" (Kartomi, 1990). These definitions come from and reinforce the separation between vocal and instrumental music. I propose however to use the term musical instrument in a wider sense. That is, that the musical instrument is not an object, but a practice.

* * *

Monophonic Instruments: Remediating the Voice

The emergence of acoustic musical instruments can be seen as an attempt to formalize cultural structures into material objects by standardizing scales, tunings, ranges and constraining timbral possibilities. What the singing voice attains through the acquisition of technique and attuning to musical practices is, in the case of the monophonic instrument, partially materialized into a concrete material form. By doing so, musical instruments embody a music theory by design and this is what makes them unique as a medium with respect to the voice.

Western music theory is generally understood as the system of principles that govern the use of tempered tuned pitches. In what follows, I will use the term music theory in a much broader way to refer to a system of principles that govern what a music is made of and how it works. Therefore, when I say that a musical instrument embodies a music theory, I am in effect saying that all musical instruments have as their content musical technologies of thought like scales, notes, and so on.

Monophonic acoustic musical instruments have as their content the human voice. They remediate the monophonic nature of the voice and its ability to control fine gestural inflections. In wind instruments the tube is coupled with the speech system through an embouchure; in bowed strings, the instrument is coupled with our arms, hands and fingers and, with the help of a bow, continually shape the sound over time.

Continuous control gives these instruments their voice-like character. The coupling of the performer's body to the instrument allows for the continuous adjustment of his actions, modulated by the feedback received from the instrument. These fine gestural inflections, commonly associated with pitch bends, vibrato, tremolo, envelopes, and the like, reflect the voice as the content of these instruments.

Winds and bowed strings are mechanical systems in the sense that they work through the control of physical forces. Different string and tube lengths produce different pitches. The acoustical length of the instruments can be varied by pressing a string against the fingerboard or opening and closing tone holes or valves, but the actual size of the string and tube remains the same. This property enables the performer to produce very fast changes in pitch.

These pitch changes are perceived as discrete elements or notes. While the voice is able to perform and articulate discrete elements, the instrument is capable of faster and more complex articulations of these elements. Scales, trills, arpeggios and 'ornaments' of many kinds emerge as instrumental features.

Gesture can now be thought of in two layers: as fine gestural inflections within discrete elements (gesture-as-morphology) and as the phrasing or articulation of these elements (gesture-as-articulation).

As argued earlier, acoustic monophonic musical instruments have as their content the singing voice. In this sense, they seem to extend the capabilities of the voice through bigger dynamic and pitch ranges. Most notably however, they provide a mechanical means to articulate pitch, duration and amplitude over a timbre conceived and idealized as fixed. As a result, the concept of note is born as the elementary building block of music; a block with attributes of pitch, duration and amplitude on a given timbre.

It is this last idea that became the dominant concept of musical instrument in the Western world: a fixed timbre over which pitch, duration and amplitude are articulated.

What was unique about mechanical instruments was not their ability to extend the properties of the human voice. While these features certainly changed the music that was produced with them, this music can be seen as its content (much like the content of the electric light is what it illuminates). The message of these instruments was that they embodied a music theory or that they contain musical technologies of thought.

The instrument amplifies some of the faculties of the singing voice, while dampening or even silencing others. The flute, for example, reinforces the voice's ability to produce pitches, but drastically reduces its ability to produce different vowels. This is what gives instruments their character and their message. In other words, what is important about mechanical musical instruments is not what they amplify, but the fact that they do amplify and dampen.

* * *

Polyphonic Instruments and the Ensemble

If monophonic instruments remediate the voice, polyphonic instruments remediate the ensemble.

Because musical instruments embody a music theory, and because this theory usually emerges from culture, we demand from instruments multiple levels of standardization. It is expected of instruments that they can perform specific ranges, scales, have an expected timbre, and so on; this means that mechanisms, materials, lengths and shapes need to follow certain standards. Standardization allowed for ensemble playing and the forging of communities with shared musical practices.

An ensemble consists of a group of musicians performing together. The ideal of the ensemble is the construction of an organization that is conceived as a unified whole and not as a sum of parts. This is the case of the string quartet, the choir or the gamelan; they are *instruments made of instruments*. In the Western world, this ideal reaches its maximum realization in the symphonic orchestra.

The aim of these organic wholes is to extend properties of the monophonic instrument, such as to enlarge pitch and dynamic ranges, but also to provide new timbral resources. When large ensembles play together, the fine gestural qualities of the monophonic instrument are diluted in the whole; gesture is sacrificed for texture.

In the soprano section of a choir or in the violin section of an orchestra, a number of similar individual units are added endowing the resulting group with a larger dynamic range. At the same time, because these ensemble sections consist of adding a number of individual units, what makes each voice unique, the fine-gestural inflections that they are capable of performing, becomes blended into the whole. When the gestural detail of each unit is added, a texture is now perceived; a timbre emerges.

In the mechanical era, the search for loudness leads to new timbres; the search for new timbres leads to loudness.

* * *

The Organ: Mechanical Energy and the Interface

The hydraulis built in third century B.C. Alexandria, by the Greek engineer Ctesibius, is believed to be the first organ (Apel, 1948). It consisted of an air pressure system which had to be manually pumped, a set of pipes tuned to specific pitches and a key to allow and stop the flow of air to each pipe. The hydraulis was a relatively small instrument and had one pipe per key and therefore one perceived timbre.

The transition from the monophonic wind instrument (a pipe with holes for pitch control) to the polyphonic wind instrument (a set of pipes each with a pitch) configures a significant, multi-faceted break in musical instrument design.

The serial spatialization of the pipes, and consequently of pitch, was already present in panpipes, where pipes of increasing length succeeded each other in space. However, the problem of blowing in multiple pipes at the same time or in fast succession, and therefore of polyphony or fast articulation, rendered the instrument less versatile than its monophonic counterpart or the ensemble. To address this problem, the performer's blowing energy is replaced with mechanical energy.

In the hydraulis, as well as in later designs of pipe organs, the amplitude is determined by air pressure; pressing the key softer or stronger does not affect the amplitude of the sound. Spatialized pitch combined with an independent energy source, demanded the creation of an interface that would be able to link them, leading to the design of the console, consisting of keyboards (also called manuals) and later pedalboards, stops and swell pedals.

In the evolution from the hydraulis in the third century B.C. to the modern pipe organ, the driving force of design is the gradual increase in the number and size, and change in the material and shape of pipes. This growth responded to the mutually reinforcing ideals of loudness and timbral exploration and demanded interfaces with growingly complex mechanisms to address different combinations of pipes from a central console. Because performance gestures are not in control of the amplitude of the sound, amplitude increase is achieved through the addition of pipes: the more pipes, the louder the sound. The practice of choosing and combining pipes is called registration.

A full set of pipes is called a rank, and the biggest pipe in the rank determines the octave in which the rank will sound. A rank of 8' is considered the native or central pitch range, but there are octave (double or half the length) and unison ranks that span from 64' to 1/2'. Ranks of pipes that produce sounds at non-octave intervals are called mutations and are tuned to the harmonic series of the fundamental. This allows the organist to combine ranks, as if adding elements of the harmonic series, to obtain different timbres.

Organ pipes can be of the flue or reed types, can be made of metal or wood, and can be open or closed. Flue pipes use a vibrating air jet, as in the recorder or flute, to generate the sound. Depending on whether the tube is open or closed, as well as on its shape and material, flue pipes are classified as diapasons (or principals), flutes, and strings. Examples of the names of different flue pipes include Principal, Flute, Piccolo, Vox Céleste, Viola, Dulciana, amongst many others. Reed pipes use a vibrating reed to generate the sound and are coupled to a resonator. The shape of the resonators modify the spectrum of the sound and they usually receive names such as Oboe, Clarinet, Trompette, Tromba, Vox Humana, and so on.

As these names suggest, the different kinds of pipes have as their content a monophonic instrument. The different kinds of pipes can be used to make ranks of different sizes. The design of the pipes - the choice of sound generator, material, shape - responds to timbral emulation, and each timbral rank, works as if it were an element of the orchestral palette. Timbral construction becomes even more complex, not only adding elements of the harmonic series, but choosing timbral identities.

The console (main interface) features several stops, called *speaking stops*, that connect a keyboard (manual or pedalboard) to specific ranks of pipes. The keyboard controls the airflow into the ranks with an open stop, while in the ranks with a closed stop, the airflow into the pipes will be literally stopped. The ranks that can be played by any particular keyboard are called a division, and large organs tend to have several keyboards. There are also *coupling stops* that literally couple one keyboard to another, so that playing in one of them activates the other, allowing the performer to sound two divisions from only one keyboard.

Other stops include the *celeste*, which features a slightly out of tune rank that generates beating patterns (tremolo) when combined with in-tune ranks, and a *tremulant* stop, which varies periodically the air pressure, creating slight variations in pitch (vibrato).

Finally, one or more divisions are enclosed inside a large wooden space or box (swell division), which has several shutters that can be opened and closed from a pedal called the balanced swell pedal, formerly known as a ratchet swell lever. Closing the shutters in the box encloses the pipes and reduces their loudness - and certainly acts as a filter that changes its timbre - while opening them allows the sound to leave the box. Again, loudness and timbre are coupled.

* * *

The Message of the Organ

The content of the organ is the ensemble. Individual ranks of pipes, each with a stable timbre, take the names and timbres of monophonic instruments; ranks of voices, flutes, strings, reeds, trumpets, and so on, are all summoned to the command of the keyboard and are combined to form organic wholes, with tremolo and vibrato textures.

The organ-as-orchestra is the seductive content, but what is the message?

The use of mechanized air pressure led to the decoupling of energy, sound-production device, and interface. The first element of the message of the organ is that each of these elements are considered independently and they can be modified and re-combined as they often were from place to place, enabling schools of organ builders and styles of registration.

The organ constituted the first instrument where performance gestures, in other words, the pressing of keys, could lead to any sound, and perhaps most importantly, to no sound. The keyboard interface is now construed as capable of any timbre, of any loudness and over which pitches could be articulated up to the limits of the performer's ability and beyond. In the organ, the serially spatialized keyboard could be mechanically linked to pipes and to other keyboards; but with an independent energy source, it opened the door to future automatic machines, like barrel organs and player pianos.

The dominant concept (and ideal) of the Western musical instrument, a fixed timbre over which pitch, duration and amplitude are articulated, is challenged by the organ's diversity and versatility of timbre. Organs, in contrast to the standardizing tendency of musical instruments are constantly evolving; no two organs are the same. Not only their design and assortment of pipes and stops varies, but the instrument's sound is greatly determined by the acoustic of the space in which it resides. An organist, when confronted with a new instrument, needs to explore the instrument and work on registrations before playing. Standardization works now at higher levels: the unifying element of all organs is their interface and technique.

The keyboard is the material amplification of the technologies of tempered tuning, the diatonic scale and of the note as the smallest building block of music; it is the symbol of the musical mechanical age and the stable formalization of its theoretical technologies.

The keyboard interface was socially constructed as a mechanical wildcard, leading to its use in multiple kinds of mechanical instruments like the hurdy gurdy, clavichord, harpsichord and piano-forte. In this last one, the keyboard reaches a further step in its development; in the piano, a key not only determines the pitch of a sound, but also its amplitude. Later, in the electric age, it is used in organs and synthesizers of many sorts; other functions, such as aftertouch, are added. Finally, it emerges as the abstract MIDI keyboard interface: a USB device that makes no sound on its own[1].

* * *

How Media Affect other Media

Most studies of Media do not fully engage in the complexity of musical practices. Media studies have historically focused on visual and mass communication media. Even in his analysis of radio and the phonograph, McLuhan does not address their full potential and actual effect on musical creation and practice.

One should not infer that media succeed each other in linear evolution. It would be an oversimplification to think that the voice had developed its full potential (or that such potential exists), before monophonic instruments appeared, took it as its content and created change. When a new medium appears it not only takes previous media as its content, but it also induces those media to adapt in response.

There are many examples of this phenomenon. For example, Mark Katz (2002) explores the sudden increase of vibrato in violin playing as recording and the phonograph became more common; Philip Auslander (1999) explores the changes that live

---

[1] This abstraction also led to the keyboardist: a musician who plays keyboards. In other words of a performer with no particular instrument, but with a particular interface.

performance underwent with the appearance of mass media, first of theatre in response to television and the movies, and second, of the rock concert in response to the music video. Many argue that the emergence of photography pushed painting towards abstraction. Media do not succeed each other in linear evolution, but they co-evolve.

It would also be unreasonable to suppose that any medium only has one medium as its content. For example, Lev Manovich (2002) believes that media like the web page and CD-ROM are “largely made up from elements of other, already familiar cultural forms” like cinema, the printed word and general-purpose human-computer-interface.

In the same way, musical instruments have as their content earlier instruments, but more importantly they too have as their content musical technologies of thought such as scales, tunings and so on; they contain the multiple media or technologies of music. Scores, instruments, scales, notes, rhythmic and harmonic structures - in short, all that we use to make and think about music -affect and reinforce each other.

It could be argued that the explorations of environmental sounds and of timbre in electronic and computer music since the 1950s shifted the interests of acoustic music composition towards timbral exploration. This tendency is reflected in the work done by composers like Nono, Ligeti or Stockhausen, to name a few, who worked both with electronic and acoustic media.

The emergence of voice techniques different from the traditional lyrical, operatic Western voice is due in part to the emergence of the microphone and its ability to capture quiet intimate sounds, but also because the sounds that are considered musical material have changed. Although the same singer might able to sing a work by Monteverdi and Berio’s Sequenza, the voices used for each piece are different instruments.

A similar argument can be made for the emergence of extended techniques in acoustic mechanical instruments. Composer Helmut Lachenmann conceives of his music as ‘Musique Concrète Instrumentale’ in direct reference to the Musique Concrète tradition in 1950’s France. Lachenmann believes that composition means instrument building:

> That’s what I mean when I sometimes say, ‘composing means building

> an instrument'. Composing means discovering and revealing a new, invented imaginary instrument. In my case the problem is that such an imaginary instrument doesn't exist before I develop it by composing the piece ((Heathcote, 2002), 340).

* * *

The Western Concept of Musical Instrument

Analyzing mechanical musical instruments in the light of McLuhan's theories on media, allows us to see the way we conceive musical instruments and their function within broader musical practices. Understanding the way we construe mechanical musical instruments also provides us with a framework to understand the way electronic and computer-based musical instruments work.

Following McLuhan's outline, I will argue that electronic and computer-based instruments have as their content mechanical musical instruments. For this reason, we expect and project onto the former the properties of the latter.

As I will argue, the theremin and the synthesizer are idealized as universal instruments that can be easily played by anyone, that can produce pure sounds of any loudness and any timbre we desire; the studio and the computer are seen as precise machines that can accurately perform very complex music and therefore replace human performers and their instruments. In all cases, composers, theorists and instrument designers attempt to situate or use metaphors from or even alter the composer-performer-listener model.

Electronic and computer-based instruments are designed, explained and evaluated in terms of mechanical instruments rather than finding what is unique about them. Naturally, this creates unreasonable expectations (like getting tactile feedback from a theremin) and when these new media do not conform with the older mechanical media, they are considered as crippled designs or as a step towards an ideal western instrument yet to come.

Margaret Kartomi (1990) introduces the expression *concept of instruments*, which she defines as “the dominant or competing views in a society of the meaning and significance of instruments as cultural phenomena”. As it follows from the analysis of mechanical musical instruments as media, the Western concept of musical instrument is that of a fixed timbre over which pitch, duration and amplitude are articulated. This is not more than another angle to a premise that has supported western music theory for centuries, namely, that the elementary building block of music is the note, and that a note is an abstract set of attributes that consist of pitch, duration, amplitude and timbre.

It is hard to determine if this concept of instrument leads to the idea of note or vice-versa. As seen earlier, instruments are conceived as standardized passive tools with no agency over the music that is performed with them. By this logic, instruments are built and used to realize scores. However, I will argue in this dissertation that instruments have agency and therefore their design enables or restricts possible musical practices or behaviors.

As I have been arguing, mechanical instruments are used to define all instruments and therefore delineate the Western concept of musical instrument. However, deeper analysis of the evolution of mechanical instruments points to deeper facts about what these instruments do, or what I’ve been referring as their message.

The first point to be made is that all instruments embody a music theory or in other words, that they contain technologies of musical thought. Returning to an earlier idea, mechanical monophonic instruments amplify some of the faculties of the singing voice, while dampening or even silencing others. However, what is important about mechanical musical instruments is not what they amplify or silence, but the fact that they do so.

The second point to be made is that the energy source, the sound-production device, and the interface can be decoupled. Through this decoupling, these elements of a musical instrument can be considered independently and therefore the instrument is not contained in any one object, but in an assemblage. Instruments can be abstracted into energy, sound-generation, and interface and each of these elements can be modified and re-combined to form new instruments.

Chapter 3

THEREMIN: Taming the Wireless Howl

> An invisible plectrum, sounding the music of the 'spheres', will come true in New York next tuesday evening with the first American demonstration of the 'ethereal music' of professor Leo Theremin. (Lemuel F. Parton, NY times, 1928. in Glinsky (2000))

In 1927, Lev Sergeyevich Termen or Leon Theremin, began a tour of demonstrations and concerts through several German cities, Paris and London, finally arriving in New York on December 20. His new instrument, the *etherphone*, which would later be known as the *theremin*[1], intruded in the musical world causing a commotion in the press and attracting the attention of scientists, the cultural elite and the general public. The instrument provoked passionate reactions, receiving both praise and criticism. Nevertheless, there seemed to be a general sense of awe among the public, captured well in the words of Andreas Lunas in the Parisian *Le Courier Musicale* in 1927: "everyone shivered as if they had seen God inflate the dead instrument with life" (Glinsky, 2000)

The public reaction to the invention exemplified what the music of the future was expected to be like, and what everyone seemed to agree on was that it would be electric.

---

[1] I will use theremin, with lower case t to refer to the instrument, and Theremin, with upper case T to refer to the person.

Although it was not the first musical instrument to use electricity[2], the theremin captivated the attention of much of the Western World, because it operated in an unusual manner. While it used electricity, it was not operated with buttons, knobs and switches and although it was a musical instrument, it did not look like any conventional musical instrument, but instead like a piece of furniture with antennae.

Lev Termen created the theremin around the year 1920, which at the time he named the etherphone to "'distinguish it from products of the contact or keyboard method' the reference also, of course, was to the 'ether waves', an especially trendy notion in the delirious new era of broadcasting - one the layperson appropriated to explain and sometimes mythologize, the magic of radio" (Glinsky, 2000).

The etherphone measured capacitance in an electromagnetic field around two antennae, which was largely understood as a means to sense the distance of the hands to the antennae. One antenna controlled the amplitude of the sound and the other one the pitch. By changing the frequency of a variable oscillator that was mixed - or heterodyned - with a fixed oscillator, it created difference waves that lied in the audible range.

Electric circuits like the theremin had essentially a single configuration and therefore were not subject to changes, unless modified and re-soldered. Ranges could be adjusted with variable resistance and 'timbres' were chosen through the use of 'stops'. Timbre changes were achieved through the use of filters although there isn't sufficient information about the nature of these filters:

> These effects were achieved by using the inherent irregularities in the response curve of the loudspeaker to accentuate different overtones. Filters for this purpose were turned on and off with five switches located on the side of the instrument (Glinsky, 2000).

Several expressions of wonder about the instrument alluded to religious, magical and ethereal qualities. The enchantment that the world experienced from seeing Theremin make music without touching an object comprised both the message and the content of electric music.

---

[2] In his paper *Early Electronic Instruments*, Curtis Roads finds 5 instruments spanning 1899-1918

* * *

Thereminvox: the voice and the bowed string

> ... soon the man and his invention fused, and the papers began referring to the etherphone as "Termen's voice" - or the "Termenvox" (Glinsky, 2000).

In its first concert-demonstrations, the theremin was usually performed with piano accompaniment. Gradually original compositions for orchestra and theremin emerged, thanks to the birth of theremin virtuosi like Lucie Bigelow Rosen and Clara Rockmore. The use of the theremin in these contexts, immediately created associations and comparisons to existing acoustic instruments. The theremin was advertised and recognized as a musical instrument; for some the instrument of the future, for others an invention that still had to prove itself. The new instrument provoked questions about the place it would take in the established musical practice of the era.

Two of the most common associations were to the singing voice and to stringed instruments. There were several reasons for these associations. First, the theremin presented pronounced portamento and vibrato, second, it used an unobtrusive (if not invisible) interface with hand playing technique. As we will see later, there were other reasons for the association to the singing voice; it provided an unusually 'clean' sound, devoid of the noises of physical mechanisms, but also, the playing technique required the performer to continually adapt his actions contingent upon sound feedback.

The thereminist stood on stage without holding an instrument, like a singer would do, and yet he used his hands to perform like one would in an instrument. Furthermore, pitch was the result of moving through an (invisible) line, like with unfretted string instruments. As a thereminist, Rosen commissioned several pieces written especially for the theremin and described the instrument in the following manner:

> "When composers ask me what general instruction is needed to write for the theremin, I would say they should think of a song; a song for an

> archangel's voice, of five octaves, and incredible power and sweetness, that can dive to the rich low tones of a cello, and include the thin high harmonics of the violin; that can be heard in great spaces without effort, through and above a great orchestra, blending with all other instruments and voices." Lucie Bigelow Rosen (Glinsky, 2000)

Portamento and vibrato generated associations with the voice and with bowed strings, because both instruments had the ability to perform them. These techniques were associated with expressivity, but also allowed the performer to adjust pitch without being noticeably out of tune.

The theremin's characteristic portamento seemed initially inevitable because, to move from one pitch to another, the performer had to continually traverse the intermediate space. Accomplished thereminists like Rosen and, in particular, Rockmore, a former professional violinist, fought to avoid the general perception that portamento was idiomatic to the instrument.

Vibrato was also a perceived characteristic of the theremin. Much like with the voice and the violin, it was very hard to begin in, or arrive directly at, a specific pitch. Several music critics often remarked the instrument's difficulty for precise intonation. Again, Rosen and Rockmore fought against this perception by thorough practicing.

Both Rosen and Rockmore asked Theremin to make adjustments to the instrument, first, to obtain a faster left hand (amplitude) in order to make rapid attacks; and second, to increase the pitch range to have pitches closer in space and allow for 'fingerwork'. Rockmore developed a finger technique that significantly reduced the portamento effect to a minimum and after years of work on the instrument gave her first recital on October 27, 1938. The critics reacted:

> "Thanks, no doubt, both to the new model and to Miss Rockmore's mastery of its technique," the *New York Herald Tribune* wrote, "the recital marked a significant forward step in the development of the theremin as a concert instrument, especially in the swifter pace which it is now able to attain and the possibility of playing detached notes, as compared with the too much present portamento and the limitations of speed which were still in evidence at its last concert demonstration." ... "That the theremin can respond cleanly and nimbly in speedy passages when the performer

> has sufficient command of its resources," the *New York Times* declared, "became patent in Miss Rockmore's agile treatment of the second movement of the César Franck violin sonata, which was taken at as high a rate of speed as violinists ever adopt for its allegro sections." for the *Herald Tribune* reviewer, "the program showed that the theremin has accomplished a marked gain as an artistic medium." The critic for the *New York Post* concluded that the instrument "appears to posses no technical limits" (Glinsky, 2000).

The theremin did not attempt to change the position of the musical instrument in the composer-performer-listener model, but to insert itself in it.

* * *

Pure Sound

The forcibly continuous changes in pitch and amplitude[3] made it clear that the interface and the articulation of the sound of the theremin were inevitably related. However, the timbre it produced was for the most part independent of the gestures of the performer. Since the interface was an invisible electromagnetic field around a visible antenna, no mechanical sound was produced; the interface was silent and the sound strictly electric.

Public reaction to the theremin often referred to its sound as pure. A critic in London's *The Daily Telegraph* "found the experience ... quite uncanny. The purity of the tone he extracts is remarkable" (Glinsky, 2000). In New York, the singer Elizabeth Rethberg praised it in the following manner: "I believe it has a great future in music. It was surprising to me to find that it had no mechanical tone. It seemed to have a soul. I could sing with it in duet. I was amazed." (Glinsky, 2000)

Traditional vocal technique in the western world has tended towards an ideal of purity and the use of vibrato and portamento in soloist singers is a commonplace expressive technique. The theremin was commonly associated with a near-sinusoidal

---

[3] The theremin could not articulate attacks nor it could make discrete steps of pitch, everything in it was continuous.

sound, along with the portamento and vibrato mentioned earlier. The absence of the mechanical noises of blown air, bows and keys reinforced the association with the voice as a sound that was pure.

Already in 1922, Theremin had a fingerboard model of the instrument and by 1929, he had developed a fingerboard model commissioned by the conductor Leopold Stokowski and inaugurated in the same year although not as a soloist, but as part of the Philadelphia Orchestra. Stokowski thought it was "like a cello, but without strings" and the composer Joseph Schillinger described it as having an "idealized cello tone" (Glinsky, 2000).

The 'fingerboard' model, in contrast to the 'space-control' model, alluded directly to the cello, not only through its sound, but also through the shape and technique of its instrument. While the object looked like a cello and the performer acted as he was playing a cello, it had no strings. The absence of strings represented the transition from mechanical forms of sound production to electric ones.

However, the sound of the string and the bow were in fact the sound of the cello and not the "idealized cello tone" of the theremin. In this respect, Schillinger's allusion to the ideal appeared to have two consequences. First, that acoustic instruments like the cello aspire to the purity of tone of the voice and therefore, all mechanical sounds or noises produced while performing it are "extra-musical". Second, that electricity is what allows us to achieve the ideal.

* * *

Electric Power

Theremin had already declared that his instrument was loud enough "for 150,000 people to hear" (Glinsky, 2000) and had played it for an audience of twenty thousand people in Coney Island Stadium.

Schilinger composed the *First Airphonic Suite*, which constituted the first theremin concerto. It was premiered by the Cleveland orchestra in 1929, conducted by Nikolai

Sokoloff and with Theremin himself as the soloist. Music critics were baffled mainly by the loudness of the instrument rather than by its musical qualities.

New York Telegram's Oscar Thompson described that "at times [the theremin was] enveloping the whole orchestra as with a gigantic human voice" (Glinsky, 2000). The press talked of "colossal outwellings of sound" and of the disturbance of "seismographs in every observatory this side of Tokyo". The orchestra, the Western world's largest musical organization was surpassed by "a grave and slender figure in evening clothes" who "made the slightest motions in front of two bars." The theremin (and in some way, the electric realm) was being revealed as a space where the energy input and the energy output were disproportionate.

Sokoloff described the theremin as occasionally producing "the most unearthly, ear-splitting shriek" (Glinsky, 2000) which in the first rehearsal caused the first horn to faint and so did, in the premiere, a large lady in the front row. *The New York Times'* Olin Downes went as far as to say that the theremin having "horribly magnified sonorities" that preluded "terrifying instrumental days which are just before us", worrying about the wellness of "auditory nerves" (Glinsky, 2000).

Violinist Joseph Szigeti noted that "a singer is limited by his lung capacity, and a violinist by the length of his bow, but there is no limitation of on this instrument." Once again, the limitations of the mechanical and physical world seemed to be surpassed. While the theremin was limited to the amount of time the performer could hold his hands in front of it, it could well be programmed to play 'for ever'. In fact, if there is no body capacitance in the amplitude antenna, the theremin will go forever.

Finally, as noted earlier, the theremin did not have 'mechanical noises' because the sound production circuit was exclusively electric. The timbre of the instrument was not dependent on any mechanical means of sound production, but on the design of the circuit.

Electric power allowed for a disproportionate relationship between input and output energies. "The slightest motions" could indeed create great output, furthermore, if the performer went away, sound would be as loud as the amplification system would allow it. Energy and gesture were therefore dissociated.

* * *

Timbre

The theremin had continuous controls for pitch and amplitude, but buttons and switches for timbre. In some of his many designs, timbre worked through 'stops', similar to the stops of organs.

The RCA theremin, the first commercial model, did not have timbre stops, but only one timbre. According to Glinsky, "the characteristics of the individual tubes were designed to allow a pleasing combination of overtones in the sound". Whether choosing over a set of timbres as in the first and latest models, or limiting the instrument to one timbre as in the RCA commercial model, a timbre had to be designed.

This timbre would be "the timbre" of the theremin, or at least of the RCA theremin. The choice was to create a harmonic sound that alluded to known instruments like the voice or the violin, but even beyond that, it alluded to an ideal timbre. Theremin told a *Chicago Tribune* reporter in Paris that he could "produce equally well the sounds of a violin, an alto or a trumpet. The quality of the sound of my apparatus is even better than that of the instruments themselves" (Glinsky, 2000).

In his 1927 concert demonstration in London, the press reported not only demonstrations of violin, cello and human voices sounds, but also of "bellows, moans, child-like whimperings, a very realistic bleating of sheep, a motor-horn and a factory syren." An interesting description came from the *Birmingham Post*: "some examples of mere noise . . . highly suggestive of the range of tones obtainable from the 'taming' of the wireless 'howl."' The reaction of the audience was of laughter, but Glinsky points us in the right direction:

> The real significance of the instrument was not in the drippy recital of romantic miniatures they were about to hear. The genuine revolution in this technology was that the raw materials of sound were now exposed

> and could be molded in every dimension. If electricity could beam voices and music through the air and transmit moving images, it could also become an architect with the bricks and mortar of sound itself. ... It was this property, and this property alone, that fired up the critics of vision in the major cities. Even the most awestruck and deferential reviewers cautioned that the thereminvox was not an end in itself, but a first step toward some future horizon (Glinsky, 2000).

As a result of having an identifiable timbre, the theremin not only positioned itself as a regular instrument, that is, one where pitch and amplitude were articulated over a fixed timbre, but as one which solved the problem of producing a good quality sound. Theremin would commonly say things like “If a youth has the spirit of music in him, he can play with my instrument, in a fortnight, what a violinist can play only after two years’ training” (Glinsky, 2000).

* * *

The Interface

In 1927, Theremin had declared to a *New York Times* reporter in Berlin, that “ether wave music is created with a simplicity and a directness matched only by singing. There is no keyboard to obtrude itself, no catgut, no bow, no pedal, nothing but simple expressive gestures of the hands” (Glinsky, 2000). Theremin’s remarks were constantly echoed by the press and were central to RCA’s marketing strategies. Very often, the public was told that if they could sing or whistle, then they could easily learn to play the instrument.

The RCA advertising brochure went as far as to predict that the instrument was “destined to be the universal musical instrument; people will play it as easily, and naturally, as they now write or walk.” The German press had even declared that the theremin “shortened the path from the brain of a human being to matter” (Glinsky, 2000) to reflect on the directness with which theremin performance was perceived.

However, the instrument was extremely hard to play. Because the instrument worked not by sensing the hand, but capacitance in an electromagnetic field, any movement the performer made affected this capacitance. Clara Rockmore told Theremin “you can hardly breathe on that instrument without affecting something,” (Glinsky, 2000) and later stated: “you cannot register any of your internal emotion at all. You cannot shake your head, for instance, or sway back and forth on your feet. That would change your tone” (Glinsky, 2000).

The interface was literally invisible except for the referential position of the antennae. What actually controlled the sound was capacitance in an invisible electromagnetic field, a fact that rendered the instrument both spectacular and magical, as well as extremely hard to play.

Again, a transition from the mechanical to the electric becomes manifest. The transition to electric organs was not as traumatic as with the theremin, because although it provided a different kind of feedback, the interface remained the same and with it the music that keyboards were expected to produce.

But with the theremin all references to the traditional interfaces disappeared; there was no key, string, catgut, bow, pedal and therefore no tactile feedback. The performer was expected to develop his own technique and to adjust, by auditory feedback and practice, the way something was to be played. It was this fact which provoked the metaphors to singing and whistling: continuous adaptation through feedback.

The theremin space-control model based on the two antennae is certainly the most well known. But the interface had several variations. Goldberg, an assistant to theremin, had developed a version with pedals to control amplitude and a series of buttons to control articulation. Theremin himself also built the fingerboard model described above, a keyboard model, and his *terpsitone*, a floor surface which was to be played by dancing. All of these models used the heterodyning principle and so did the inventions of Mäger, Givelet, Martenot and several other instrument builders of the time. In fact, public disputes about who had invented what and legal debates about patents surfaced repeatedly.

The interface was separate from the sound production mechanism, or in other words, many interfaces could have the same sound.

* * *

A Market, A Standard

At the time, the press would talk about ether wave music, and only gradually adopted the term *electric music*. Theremin's lecture-concerts attracted artists, scientists and the cultural and economic elite; he performed in established concert-halls and private gatherings, while being able to attract thousands of people to Coney Island Stadium.

This quasi-hypnotic effect of the instrument, its ability to attract large crowds, the immediacy of a pleasing timbre and its associated ease of performance, added to the prospect of low fabrication costs, entertained the idea of commercial success. Seizing on the opportunity, in March 12, 1929, Theremin signed an agreement with RCA to develop a commercial version, and on September 23 the RCA Theremin was announced and sales began.

The RCA Theremin had a fixed timbre and pitch range and represented the first mass-produced, standard version of the instrument. Standardized models allowed for a recognizable instrument.

RCA expected to sell theremins for every home. They produced a first set of 500 instruments to test the market and although all of the units were sold, it generated a loss due to patent costs paid to Theremin and to merchandising expenditures. However, problems springing from technical malfunctions, legal suits about patent violations, and the beginning of the depression era, prevented them from pursuing the instrument's commercialization any further.

Glinsky was able to track an internal document from George H. Clark, manager of RCA's show division to "RCA executives summarizing his observations at radio fairs, based on discussions with the public" and provides us with an analysis of his own:

> Premised on my firm belief that we could make the Theremin a home musical instrument of great acceptability to the public - in addition to the apparent present plan of having it a high-grade orchestral device - I suggest the following:
>
> 1. That such a home device be made so that the player may operate it while he or she is sitting.
> 2. That a control natural to the average person be used, i.e., that the hand be moved to the right for a note, or, in the piano analogy, that the hand be moved to the right for a hight note, and vice versa.
> 3. That some scale be provided for indicating specific notes.
> 4. That some form of control be added whereby finger playing may be substituted for hand-waving. ...
> 7. That a form be developed wherein volume control is taken away from manual operation, both hands thus being left free. A second oscillator to be furnished, with hand control, so that two-part music can be obtained.
>
> Clark also recommended an automatic means for creating vibrato, “thus reducing the effort of playing and leaving the mind free to concentrate on note production.”
>
> As Clark systematically dismantled the qualities that made the theremin unique, he essentially wound up with an electronic organ and discovered in the process that the American public leaned toward conventional electronic keyboard instruments. These kinds if devices were growing in popularity anyway, and they began to point the way for commercial success in the field of home electronic musical instruments. Clark, perhaps without fully realizing it, had begun to sound the death knell for the RCA Theremin (Glinsky, 2000).

While RCA stopped producing the theremin, the 500 units that had been sold up to 1930 had created a market, perhaps not as massive as the one they had envisioned. People all over the US played the theremin in all sorts of social events. Several posts in the *New York Times’s Classified Ad* or *Public Notices* section featured announcements of people looking to buy and sell theremins, as well as thereminists offering their services as entertainment, generating a small scale economy for the instrument. Its later use in movies, radio, TV shows and certain successful LPs managed to keep the theremin in the imaginary of the public and therefore kept interest in the instrument.

The absence of a mass commercial manufacturer opened the door to individual people to make their own versions. The instrument boiled down to several similar schematics, which the builder could appropriate and change. The schematics were occasionally published in electronics magazines, including a 1949 article used by Robert Moog to build his first theremin, and a 1954 article by Robert Moog published in *Radio and Television News* (Glinsky, 2000) (Pinch and Trocco, 2004) with his own version of the instrument. At this time, Moog became a small-scale theremin builder, hand-building each unit upon order and offering the first commercial model in 15 years. Moog began to improve the theremin offering different models, some of which included an “overtone selector” and a “synthetic formant”.

A mass produced theremin would not have generated the success that electric organs did. However, a different “product” would prove more able to establish to the market: the schematic. From that moment on, a person with electronics skills could build his own version of an instrument instead of buying it.

* * *

Universal Instrument

In the mid 1920’s the Russian press wrote in reaction to a theremin demonstration: “The problem of producing the ideal instrument is solved.” Electric instruments were often perceived as the route to reaching a universal instrument. For RCA, it was the instrument every one could play: “A child .. an elderly lady ... a skilled musician ... a blind man ... all can learn to play this incredible instrument with exactly the same facility!” (Glinsky, 2000). There also seemed to be a perception that electric instruments would be able to reproduce the timbre of any instrument and therefore all acoustic instruments could be replaced.

The instrument led to comments from the Russian press on the lines of “a musical tractor coming to replace the wooden plough”, or “Termen’s invention has done almost what the automobile has done for transportation” and members of the audience saying “We might as well throw our violins aside” (Glinsky, 2000).

As a consequence of these ideas both Theremin and Stokowski began thinking of an electric orchestra whose instruments would, in Theremin's view, "have the sound character of the various orchestral instruments." Stokowski believed this orchestra would "begin a new era in music, just as modern materials and methods of construction have produced a new era in architecture, of which the sky scraper is one phase." In his view the electric instruments would be added "to the present orchestra only for the music of the future" (Glinsky, 2000).

Theremin went on to create a fully electric orchestra called the *Theremin Electronic Symphony Orchestra* which had its first performance on April 1, 1932. On March 27, the *New York Times* reported that "because a variety of timbres can be obtained at the will of the performer, only sixteen instruments are needed for full orchestra." Stokowski was forced to postpone his plans because of objections from the union.

Theremin had even a more radical vision for the electronic orchestra: "The time is coming when an entire orchestra will play without instruments. Then, before each musician there will stand only a music stand with music and on it two antennae, and through waves of the air an entire orchestral work will be played - string, winds, drums - all" (Glinsky, 2000). Again, the idea of a universal interface that could perform any timbre was recalled.

* * *

Social Construction

The reactions of the press and the general public often included the words uncanny, mysterious, unearthly and otherwordly. The once praised purity of its sound, devoid of the noises that linked sounds to the physical, mechanical world, suggested that they were in fact the sounds from some other world. Even the sounds of the electric world of radio and the phonograph had a documentary character; that of storage and transmission of the sounds of this world, but electronically generated sounds, like that of the theremin, had the quality of the non-physical or "super-natural", which in the popular imaginary meant outer space or inner space as psychological states.

Composer Miklos Rozsa used the theremin for the soundtrack of the 1945 movie Spellbound by Alfred Hitchcock which was received with great success. Thereafter, Dr. Samuel Hoffman became the thereminist of choice for science-fiction and thriller movies, reaching the status of "Man of the Week" in 1950. In Glinsky's words:

> The sound relegated the theremin to a specialized music tool, like an infrequently played percussion instrument, and not the all-purpose expressive medium Theremin hoped it would become. The instrument began to be typed in the popular imagination as a harbinger of the strange and weird - associated with a world outside the boundaries of everyday emotional experience (Glinsky, 2000).

* * *

The Micro-Tonal Promise and Alternative Paths

> Most inventors of electrical instruments have attempted to imitate eighteenth- and nineteenth-century instruments, just as early automobile designers copied the carriage. ... When Theremin provided an instrument with genuinely new possibilities, Thereministes did their utmost to make the instrument sound like some old instrument, giving it a sickeningly sweet vibrato, and performing upon it, with difficulty, masterpieces from the past. Although the instrument is capable of a wide variety of sound qualities, obtained by the turning of a dial, Thereministes act as censors, giving the public those sounds they think the public will like. We are shielded from new sound experiences (Cage, 1961).

Most instrument builders and composers in the Western world saw in electric instruments a way to create micro-tonal music. Theremin told Walter Kaempffert from the New York Times that his "apparatus frees the composer from the despotism of the twelve-note tempered piano scale, to which even violinists must adapt themselves. The composer can now construct a scale of the intervals desired. He can have intervals of thirteenths, if he wants them. In fact any gradation detectable by the human ear can be produced" (Kaempffert, October 5, 1927).

In a dominating system where pitch was the most significant aspect of a sound, micro-tonal music promised an unexplored terrain for pitch relations. Yet it was not this feature that attracted experimental composers of the time.

Edgard Varése used two fingerboard theremins built by Theremin himself in his piece *Ecuatorial* premiered in 1934. The theremins are used throughout the piece using the full register, glissandos are commonly used and vibrato is hardly heard. In the absence of fingerboard theremins in the 1950's, Varése re-scored the piece for two ondes martenot.

In 1935 Percy Grainger adapted his *Free Music No. 1*, originally for string quartet so it would be played by theremins and wrote *Free Music No. 2* for six theremins. His works essentially explored the glissando feature of the theremin. But he had other ideas in mind:

> Too long has music been subject to the limitations of the human hand," he wrote, "and subject to the interfering interpretations of a middle-man: the performer. A composer wants to speak to his public direct. Machines (if properly constructed and properly written for) are capable of niceties of emotional expression impossible to a human performer. That is why I write my Free Music for theremins - the most perfect tonal instruments I know. (Grainger in Glinsky (2000))

He envisioned and even planned with Theremin the construction of automated models of the instrument. Grainger joined with these statements a group of composers interested in automatic performance, another ideal projected onto technology since the mechanical age. Henry Cowell had been working with Theremin on an instrument that was able to automatically play multiple complex polyrhythms at the same time, using different frequencies from the harmonic series, called the *Rhythmicon*. Automation took a new role in the electric age. The theremin not only created visions of music without instruments, but of instruments without performers.

* * *

Conclusions: The "message" of the theremin

The *Birmingham Post's* expression "the 'taming' of the wireless 'howl"' captured the ethos of the history of the theremin. Taming in this sense meant that the resources made available by electronic (mainly radio) technology would be made to comply with the idea of music of the beginning of the XXth century and the *Western concept of musical instrument*; in short, articulating pitch, amplitude and duration over a fixed timbre. The theremin's occasional malfunctions in the form of "loud, unearthly shrieks", that is, the untamed, sporadically announced the possibilities that lay underneath the controlled environment that Theremin had constructed.

The separation of energy from gesture, of the sound production device from the actions of the performer, and the introduction of an invisible interface, opened new spaces of action. Part of the theremin's "message", as we will see later, was that the gestures that were converted into two voltages that controlled pitch and amplitude, were but one of the many functions a gesture could have; many "pure tones" were yet to be discovered. What the theremin had missed by dedicating time and effort in imitating the instruments of the past now opened unexpected paths.

But these new paths were only possible because of the imitation of the instruments of the past. This is what made the theremin such a strong medium. The popularity of the instrument and of the inventor himself were the catalysts of future changes. In other words, what attracted the great audiences of the 1930's was the lure of attaining the ideals of Western music: the pure tone, the loudness that the orchestra and organ did not achieve, the instrument everyone could play, the possibility of realizing all timbres in one device, a mass market; in short, the appearance of a new medium that would match and surpass what music media had allowed until then. What they got was a schematic.

The schematic, a graphical representation of an electric circuit, converted the instrument into a code for exchange of information and a source of technological innovation. Theremin himself had appropriated schematics and techniques from radio technology, including De Forest's Audion, Armstrong's use of it as an amplifier, and of loudspeakers. De Forest's claims of patent violations by the RCA theremin was one of the reasons RCA discontinued production of the instrument.

What the schematic opened up was the possibility of not only reconstructing the theremin, but of modifying it and/or recombining it with other electric devices; in short, it opened up the doors (0) for the exact replication of the instrument, (1) for converting physical forces such as human gestures into voltages, (2) for any voltage to control any other voltage, any device to control any other device, (3) for making these processes audible and (4) to convert these processes into a schematic too.

Chapter 4

TAPE: The Sounds of Instruments Never Built

The two founding currents of electronic music in the 20th century were *Elektronische Musik* in Germany and *Musique Concrète* in France[1]. The former was associated with a formalist tradition concerned with principles of sound organization of synthesized sounds, while the latter was associated with a spectralist tradition concerned with timbral manipulation of recorded sounds.

In both currents, the unifying technique was the use of magnetic tape as a storage medium to fix sound. Fixing sounds on tape had multiple functions.

One of its functions was to serve as a buffer between the processes used to build a sound. For example, if one wanted to build a sound consisting of 30 sinusoids with one sine wave generator, each sinusoid would have to be recorded independently and then mixed. If only a four-channel tape recorder was available, intermediate mixes would need to be made in order to achieve the final mix. If the composer wanted to filter the 30 sinusoids, a tape containing the mix would be used to play it through the filter, while recording the resulting sound on another tape. We can think of this process as buffering, or temporal storage, in between processes that are perceived as happening at the same

---

[1] I will use the terms *Elektronische Musik* and *Musique Concrète* to refer to these practices instead of the term electronic music which I will reserve for broader meanings.

time, but that are actually done in separate stages in the studio.

Another function of magnetic tape was to serve as a sound container. Any sound element would have its corresponding tape segment or space. These segments would later be cut and spliced in a desired order to obtain a sequence of sounds. For example, when a melody had to be built, each note would be spliced one after the other to achieve the desired sequence of sounds. We can think of this process as buffering, or temporal storage, of the elements that form a sequence or succession over time.

Other functions included tape-based processing techniques like loops, playback-based pitch changes, and delays, amongst others. Tape recorders also allowed the composer to collect sounds as material for processing.

However, the crucial aspect of these studio-based electronic music practices, is that they rendered pieces of music that were fixed on magnetic tape as a final product, hence the (mainly American) term, *tape music*. We will refer to all practices whose final product is a vinyl, tape, or digital sound file, as “fixed” music.

The Radio Studio offered a different environment, where the facilities for sound recording and playback were already available. Sound editing and mixing was already possible and eventually this equipment was re-purposed by musicians and engineers to generate and modify sounds for musical purposes. The electronic music instruments built in the first part of the XXth century had been built in the image of mechanical instruments. This was one of the reasons why composers felt they could not obtain the degree of control that they needed, both over structure and timbre. This is also the reason why the studio became an ideal environment for this type of music.

Every feature of an electric instrument existed as a circuit. At the time, if a composer wanted a sound made of 30 sinusoids, mixed and filtered, then each of the sine generators had to be built including a mixer, a filter and a means to control all of the equipment. To build an instrument like this one, of arbitrary complexity for each composition, was not only an expensive undertaking, but also would have made for large machines with extremely specific uses.

Composers found in the studio a means to construct ‘as sound’ the instruments that were impossible or inconvenient to build as physical devices that would produce those sounds. I will call the sounds of these instruments that were never built *instruments-as-sound* or *pseudo-instruments* as Pierre Schaeffer came to call them. In achieving these *instruments-as-sound*, composers expanded the sense of what a musical instrument could be expected to be.

* * *

Elektronische Musik

Much of the theoretical writing associated to the elektronische musik movement was formulated in the journal called *Die Reihe*, or “The Row”, in direct allusion to the serial row of twelve-tone compositional technique. Edited by composers Herbert Eimert and Karlheinz Stockhausen it was published in German and in English. The first volume of the English edition features “What is Electronic Music” by Herbert Eimert as its first article. Eimert begins the article by saying that electronic music was composed “without recourse to any instrument or microphone” and clearly stated: “This electronic music is not ‘another’ music, but is serial music” claiming that the “revolutionary thought of Anton Webern” was to be extended into the “total organisation of the electronic sphere” (Eimert, 1957). In this “manifesto” of elektronische musik, its main principles and conceptions are set forth.

The first element concerns the *note*, the *sinusoid* and *timbre*. For Eimert, the note is “the most fundamental music-forming element” *(ibid)* and from it is derived all musical organization. He further specifies that “the composer determines each note by its pitch, duration and intensity”. In a similar way, Eimert states that the “sinus tone is a fundamental element of musical perception” (Eimert, 1957). Elektronische musik is claimed to have started with the construction of a spectra out of “pure tones” by Stockhausen[2], (Eimert, 1972) thereby invoking the purity of tone once ascribed to the

---

[2] Toop (1979) associates the idea of pure tones to theological views. This is not new in western music, for example, vocal technique in western music aims at a clear, pure sound free of noises and corresponds to the ideal of pure tone of the theremin as explained in the previous chapter.

theremin. Although formal structure is a fundamental characteristic of elektronische musik, it was clearly concerned with spectral qualities of sound as well.

This concern for timbre is further reinforced by Stockhausen who essentially claimed that there was a fundamental unity in all sound. In this view, melody and harmony are consequences of a higher range of frequencies, meter and rhythm of lower frequency ranges, and form and composition of even lower ranges. That is, an event repeating at more than approximately 30 times per second produced pitches, if less it produced rhythm and if it repeated even at lower frequencies, it was not perceived as rhythm any more and corresponded to form. Deviations in the periodicity of pulses led to timbre and inharmonicity at the scale of melody and harmony and to syncopation in the case of rhythm and meter. Because of this “fundamental unity”, Stockhausen believed that all aspects of a musical composition could be controllable from few if not one global parameter (Stockhausen and Barkin, 1962).

A second element concerns *control*. Well aware of the instruments of Trautwein and Mager[3], magnetic tape is considered indispensable: “the decisive means of maintaining and operating sound only became available [...] with the discovery of the means of recording sound on tape. Electrically generated sound could only be utilised[4] as a genuine compositional element when this technique had been invented” (Eimert, 1957). For Eimert, the problem of performing electronic music was a problem of control: “the number of individual sound elements is so great that any attempt to find means of instrumental realisation is doomed to failure” (Eimert, 1957).

Again, given the complexity of these sounds, building an instrument able to produce them would have been costly if not impossible. Supposing however that one could have built such a device, the problem of controlling it constituted an even bigger hurdle. The crafting of electronic music pieces not only created an ideal, metaphoric

---

[3] Trautwein and Mager, were both electric music designers in Germany contemprary to Theremin. Trautwein's *Trautonium* became a popular instrument, for which Hindemith, amongst other composers, wrote music. Eimert (1957) mentions them in his texts, but sees their work as insufficient.

[4] I am keeping the original spelling in all citations. We should expect to find discrepancies with american english when the author or translator is british.

musical instrument, but also an ideal performer capable of accurately executing whatever complexity was needed. “Genuine composition” therefore consisted of the liberty of composing without the constraints of physical devices and the limitations of human performers.

A third element concerns *specialized knowledge*. Operating an electronic music studio required specialized knowledge beyond music theory: “The composer is required to have a certain amount of acoustical knowledge” (Eimert, 1957). The expression heralds the birth of a new kind of hybrid music practitioner concerned with cognition, acoustics, electronics and eventually programming.

A fourth element is concerned with *space*. “The stereophonic distribution of sound transmitters is a further element of the form of electronic music. The various loudspeaker systems around the hall are the ‘concerting instruments’ - a conception similar to the distribution of orchestral and choral forces in church or concert hall” (Eimert, 1957). With electronic music, space becomes a formal element to be addressed by the composer.

Eimert implicitly mentions the duality that would characterize most electronic music practices. As the composer worked directly with sound the process of composing becomes confused with the act of performance. The studio is both a “compositional means” and an “instrument”:

> The normal studio technique of broadcasting is transformed into a compositional means. Tape recorder and loud-speaker are no longer “passive” transmitters; they become active factors in the preparation of the tape. This is the essential secret of electro-acoustical technique. One might say that today we have perfected a “keyboard” of this elaborate and differentiated sphere of radio transmission; now we lack only the virtuosi to master it (Eimert, 1972).

For Eimert, the studio (the elaborate and differentiated sphere of radio transmission) has a “keyboard” (tape recorder and loudspeaker) with which it can be played. Eimert construes the elektronische musik studio as a musical instrument; as a sound device and an interface. Paradoxically, in this metaphor, this virtuoso performer is the composer. Curiously, the keyboard is chosen as the metaphor of an interface.

Finally, a question of material arises. Elektronische musik was concerned with the construction of complex spectra out of pure elements called sinus tones and with the musical organization of fundamental blocks called notes. Some of the most important theoretical writings of this movement reveal a self-referential conception of music; of a system where musical meaning is constructed exclusively by the organization of musical sounds: sinusoids and notes.

* * *

Musique Concrète

Theoretical writings on *musique concrète* were left mainly to its main figure, Pierre Schaeffer. Schaeffer's main theoretical ideas were collected in his *Traité des Objets Musicaux* or *Treatise of Musical Objects* (Schaeffer, 1966), published 18 years after his first experiments with the transformation of recorded sounds. Although the treatise is one of the first and few that attempt to theorize the practice of musique concrète, it can be, as Landy puts it, "a bit difficult to follow" ((Landy, 2007)). Michel Chion published an exegesis of the treatise under the title *Guide des objets sonores: Pierre Schaeffer et la recherche musicale* (Chion, 1983), (Chion, 1995). After Schaeffer's original publication, other authors such as Smalley, Wishart, Delalande and Bayle, provided adaptations, extensions and personal interpretations to the theory of sound organization set forth by Schaeffer.

The *Traité* is a key work to understand the foundations of musique concrète, a term which Schaeffer would propose to change to "experimental music". Instead of a manifesto, the Traité became a 700 page treatise that attempted to develop a theory in a practice where traditional music theory was not sufficient or adequate; as a new practice, it demanded a new vocabulary. New words and re-definitions are scattered throughout the Traité.

Musique concrète was born out of the recording medium. Schaeffer gave the following explanation of the term 'concrète' in the Traité:

> When in 1948 I suggested the term 'musique concrète', I meant, with this adjective, to signal an inversion[5] in the practice of music. Instead of notating musical ideas in the symbols of traditional music theory, and entrusting their realization to known musical instruments, I wanted to gather concrete sound material, wherever it came from, and extract from it the sonorous musical values which it potentially contained (Schaeffer, 1966).
> ... he wanted to emphasize that this new music came from concrete sound material, sound heard for the purpose of trying to abstract musical values from it. And this is the opposite of classical music, which starts from an abstract concept and notation and leads to a concrete performance (Chion, 1995).

In these definitions we find the central premises of the theories contained in the Traité. These premises are contained in the phrases "concrete sound material", "extract from it the sonorous musical values", "heard for the purpose of trying to abstract musical values", " entrusting their realization to known musical instruments" and "the symbols of traditional music theory".

* * *

Concrete Sound Material: Reduced Listening and the Sound Object.

Concrete sound material meant fixed or recorded sound. What recording afforded for Schaeffer was the ability to isolate "the sound from the 'audiovisual complex' to which it initially belonged" (Chion, 1995). The term *acousmatic* was adopted to refer to the phenomenon of listening to sounds of which we cannot see the source, a practice that in Scheffer's view led to a method he called *écoute reduit* or reduced listening. Interpreting Schaeffer, Chion defined reduced listening as "the listening attitude which consists in listening to the sound for its own sake, as a sound object, by removing its real

---

[5] In the original translation by Dack and North, they use the term "u-turn", but in my opinion it misleads the reader to thinking of a "going back" instead of the inversion of the compositional process.

or supposed source and the meaning it may convey" (Chion, 1995). Reduced listening is permitted by the fixing on a medium, by recording, which allows for a sound to be repeated:

> By repeated listening to the same recorded sound fragment, the emphasis is placed on variations of listening ((Schaeffer, 1966), 94) ... if curiosity about causes remains in acousmatic listening (and it can even be aroused by the situation), the repetition of the recorded signal can perhaps 'exhaust' this curiosity and little by little impose 'the sound object as a perception worthy of being listened to for itself' ((Schaeffer, 1966), 94), revealing all its richness to us. ((Chion, 1995), 11)

Reduced listening, as enabled by repetition (and therefore by recording), was a listening strategy that provided a way of stripping the sounds from their sources and meaning. A second practical approach to achieve reduced listening was:

> "...the recognition of a timbre was not linked as much as was thought at the time to the presence of a distinctive harmonic spectrum, but also to other characteristics in the form of the sound (particularly the attack). As an "experiment in interruption", isolating a sound from its context, manipulating it, and thus creating a new sound phenomenon which could no longer be traced directly to its cause, the experiment of the cut bell together with the closed groove encouraged people to practice "reduced listening" and draw out from it the notion of the sound object ((Chion, 1995), 13).

"Experiments in interruption" and "repetition" are then presented as techniques that allowed the composer to practice reduced listening. It is inevitable to interpret reduced listening as a kind of "purification" process; of repetition as a means to suppress or "decondition" our instinctive search for sound sources, causes and meanings. From reduced listening we extract a new concept, that of "sound in itself". Schaeffer's borrowing of the term *époché*, or bracketing, from Husserlian phenomenlogy, formed the conceptual basis of reduced listening: "a putting in parentheses of 'natural' perceptual intentions" ((Chion, 1995), 29)

> ... époché represents a deconditioning of habitual listening patterns, a return to the "original experience" of perception, enabling us to grasp the

> sound object at its own level as a medium, an underlay of the perceptions which use it as the vehicle of a meaning to be understood or a cause to be identified. The "putting in parentheses", which is what reduced listening is, and is thus an époché, leads us then: ... to "put to one side" the consideration of what the sound refers to, in order to consider the sound event in itself; to distinguish this perceived sound event from the physical signal to which the acoustician attributes it, and which itself is not sound ((Chion, 1995), 28).

As suggested several times in the previous quotes, reduced listening had its "intentional object" and its "correlate" in sound objects:

> Reduced listening and the sound object ... define each other mutually and respectively as perceptual activity and object of perception ((Chion, 1995), 30).
> ... [the] sound object is the meeting point of an acoustic action and a listening intention ((Schaeffer, 1966), 271).

In Schaeffer's writing (and in Chion's), the sound object is constantly defined in negative terms, that is, by stating what it is not, rather than what it is. In these definitions the sound object is not "the physical signal", "a recorded fragment", "the same as the more or less accurate written symbol which 'notates' it", "its real or imagined causality", and so on. One of the few things that is positively defined is that the sound object is a product "of our hearing and relative to it", stating already in this concept a primacy of hearing as opposed to abstract formal plans. In this spirit, the sound object is posited as a fundamental unit:

> The name sound object refers to every sound phenomenon and event perceived as a whole, a coherent entity, and heard by means of reduced listening, which targets it for itself, independently of its origin or its meaning. ... It is a sound unit perceived in its material, its particular texture, its own qualities and perceptual dimensions ... it is a perception of a totality which remains identical through different hearings; an organised unit which can be compared to a "gestalt" in the psychology of form ((Chion, 1995), 31).

In musique concrète then, the sound object signified to a certain extent a breakthrough from the note paradigm still present in the German tradition which saw itself as an evolutionary continuation of serial thought. This point is made explicit by Chion when asserting that “a harp arpeggio on the score is a series of notes; but, to the listener, it is a single sound object.” ((Chion, 1995), 32)

* * *

A Composition “Method”

In his ambitious Traité, Schaeffer attempted to set forth a method of composing. Through reduced listening, the composer was to identify and derive “musical values”[6] in sound objects. The next step in the process was “a question of choosing the sound objects which are judged suitable to become musical objects” ((Chion, 1995), 70). The “values” of these musical objects became the basis for what Schaeffer called “calibrations”, which allowed “an evaluation of the degrees and their intervals as absolute values and which form scales” which enabled the composer to create “abstract relationships” ((Chion, 1995), 45). In other words, identifying values allows the composer to form scales or calibrations based upon these values.

Schaeffer did not only conceive of sound material as calibrated. He also envisioned a discrete/ continuous axis, but he, and particularly Chion, considered the continuous values as less truly musical:

> ... a MUSICAL STRUCTURE is not necessarily a discontinuous structure of interconnected but discrete objects, but that it can also be a CONTINUOUS structure, made of a macro-object whose internal variations form a whole discourse.
> ... music made of combinations of discrete objects, is probably the more truly musical. The other, made from continuous structures, can better be described as “plastic”. ((Chion, 1995), 59)
> ... only calibrations of criteria are capable of creating abstract relationships and not dynamic or impressionistic (“plastic”) relationships ((Chion, 1995), 45)

---

[6] Values are defined as: “the relevant features, which emerge from several sound objects placed in a structure and form the elements of the abstract musical discourse properly speaking; the other aspects of the object which are not relevant in the musical structure but which form its concrete substance, its matter, are grouped together under the name of characteristic” ((Chion, 1995), 73). For example, in traditional western music theory, pitch is a values and everything else a characteristic.

Let us reanalyze Schaeffer's idea of "inversion of the musical process"[7], in the light of his method. Schaeffer's method proposes what Landy (2007) termed *bottom-up composition* and Emmerson (1986), termed *abstracted syntax*; terms that refer to a process where the structure of a piece is derived from its concrete sound material. Their opposites, *abstract syntax* or *top-down composition*, refer to a process where the material must adapt to pre-conceived abstract formal structures.

However, Schaeffer's method also proposed to use sounds without sources or meaning from which to extract parameters and form "calibrations" or scales, as a means to an aesthetics of sound-in-itself. In this sense, what Schaeffer was proposing was also a self-referential system. A "more truly musical" system made of "combinations of discrete objects" that form "abstract relationships", instead of the "plastic, dynamic or impressionistic relationships" made from "continuous structures". Faced with a new medium, both Schaeffer and Eimert framed their musical practices in terms of older musical media. However, in doing so and in trying to theorize about it, new concepts emerged.

A fundamental change occurred with the concept of *sound object* as contrasted with the note. The note was an abstract entity that could be parameterized. Pitch, amplitude, duration, and even timbre, were objects that could be serialized. In other words, abstract formulations would determine the parameters of a note or group of notes, and therefore, one could arithmetically operate on them. In contrast, the sound object was a signal, and as such, simple arithmetic operations did not apply. A simple operation like pitch transposition would be attempted through variations in playback speed only to find the sound had changed. Instead of being transposed, they were transformed.

Although reduced listening and bracketing (and therefore sound-in-itself) appear to be somewhat artificial, Schaeffer proposed the "primacy of the ear" as an approach

---

[7] When in 1948 I suggested the term 'musique concrète', I meant, with this adjective, to signal an inversion in the practice of music. Instead of notating musical ideas in the symbols of traditional music theory, and entrusting their realization to known musical instruments, I wanted to gather concrete sound material, wherever it came from, and extract from it the sonorous musical values which it potentially contained ((Schaeffer, 1966), 23).

to composition. Through listening, the composer had to determine what was relevant or musical in a sound, therefore containing the proposition that any "sonorous material" could potentially become musical material, opening the door to new approaches.

* * *

Known Musical Instruments and the pseudo-instrument

In contrast to Eimert, Schaeffer was deeply concerned with the problematic of the instrument and of live execution. In 1948, Schaeffer defined musique concrète in opposition to the practice of entrusting the realization of notated music to "known musical instruments". Eimert too defined elektronische musik as "without recourse to any instrument" (Eimert, 1957). What they could achieve in the studio was not possible with physical instruments and human performers. However, by 1966, when the treatise was published, Schaeffer perceived a crisis.

On one hand he saw the "most well-known contemporary experimental composers" leaving the studios to return to work with orchestras, but on the other, "a world-wide multiplication of studios" ((Schaeffer, 1966), 24). Schaeffer believed that electroacoustic music [8] had suffered of a series of excesses and defects: composers had entrusted their music to intellectual and scientific procedures and models that were divorced from human perceptual experience.

> It is in electroacoustic's own excess of possibilities that it escapes equilibrium, in which the virtuosity of a musicality with no interest or a sonority that does not go beyond the description of its origin predominates. The contemporary paradox then, resides in the instrumental failure of a great number of powerful media. Without live execution from a credible instrument, there is no way out of the crisis ((Schaeffer, 1966), 44)[9].

---

[8] A term he used to group both elektronische music and musique concrète as the boundaries between them dissipated.

[9] My translation.

In 1966, proposing live execution of a media-based practice supposed an impossible task: "for the moment we do not do more than perturb the space" ((Schaeffer, 1966), 45). The tape studio allowed the composer to work on a sound for a long time, yet even the simplest act like placing one sound after another implied a time-consuming task. For Schaeffer, live execution seemed to imply a way of achieving equilibrium; a way to counter-balance the seemingly infinite possibilities of electroacoustic music by establishing limits in what an instrument can do and a performer can play.

Schaeffer found that musical instruments provided the foundation for musical practice. Concrete experience allowed abstract theorization. In the absence of musical instruments to provide these foundations, in electroacoustic music, abstract theorizations had taken a flight of its own.

Schaeffer conceived a musical instrument as "every device from which a varied collection of sound objects - or a variety of sound objects - can be obtained, whilst keeping in the mind the permanence of a cause" ((Chion, 1995), 52). What allows a person to recognize "that various sounds come from the same instrument" ((Schaeffer, 1966)) is timbre.

In his initial questioning about the nature of instruments, Schaeffer reached the paradox that although each key of a piano has an independent timbre, the piano as a whole has a generic timbre. Mechanical instruments then show a "permanence of characteristic" (timbre) and a "variation of value" (pitch, duration), which coincides with the traditional view of musical instruments. Schaeffer was searching for a definition that would include mechanical instruments while leaving open the possibility of including new devices.

On the other hand, Schaeffer recognized that electroacoustic music pieces seemed to create the illusion of a common source, a phenomenon he called pseudo-instruments:

> In these musics [concrete and electronic], however, a "sequence of well-formed suitably registered1 objects" can give rise to relationships of permanence such that these sound objects seem to come from the same instrument, which is imaginary, and which can be called a "pseudo-instrument" ((Chion, 1995), 55). Without the real instrument, present

> in the execution, the pieces offer "invariants" to the listener that work as pseudo-instruments as (apparent) permanence of a common character, while other elements vary ((Schaeffer, 1966), 44).

At the founding of the Groupe de Recherches Musicales (GRM), re-founded[10] by Schaeffer in 1958, Battier (2007) identifies in a series of works a new approach:

> A new way of electroacoustic composition began characterised by an effort to create a unity of the material by inserting it into a formal, coherent discourse. Schaeffer created two études which had as their point of departure the desire to make a work exploring a limited collection of objects, assuring a form integration in the most musical way possible (Battier, 2007).

Schaeffer was looking for ways of providing a sense of instrumentality to musique concrète, whether by "exploring a limited collection of objects", or through his compositional method. Based in reduced listening, Schaeffer's compositional method demands the denial of "real or imaginary" sources; however, it is precisely this method which leads the composer to the pseudo-instrument, a permanent "imaginary source", bringing into question the concept of sound-in-itself.

* * *

Studio Instruments

The studios at Radio Television Francaise (RTF) in Paris and West Deutscher Rundfunk (WDR) in Cologne mainly consisted of equipment for radio broadcasting. The first act of instrument design for the studio was one of repurposing:

> When the gramophone changed its status from being an apparatus for reproduction to an instrument of production, an artist has, by thought or deed, reinvented the apparatus. ... This machine, in the Club d'Essai's studio, was above all represented by the turntable, but it was also the potentiometer of the mixing desk. In Schaefer's hands, the turntable in

---

[10] Founded as Groupe de Recherche Musique Concrète (GRMC) in 1951, it was refounded to be more inclusive. The term Musique Concrète was changed to Experimental Music

> effect became the generator for unheard behaviours of sound. They were thoroughly explored and classified, and became the vehicle for the research which led Schaeffer towards the conception of musique concrète: closed groove (made into a loop), speed variation (transposition of pitch and tempo), reversal, removal. Using the mixing desk's potentiometers, one's action leads to a new dynamic outline, which, in some cases were able to mask the identification of the source: the sound sheds its envelope and becomes disembodied (Battier, 2007).

The Club d'Essai was the studio where Schaeffer was working in 1948 and was mainly equipped with turntables and mixing equipment. In 1951, he formed the Groupe de Recherches Musique Concrète (GRMC) and switched to magnetic tape as storage medium. Looking for new and better ways of creating music in the studio, Schaeffer associated himself with engineer Jacques Poulin, with whom he developed a series of new equipment for the studio. The main objective of these instruments was to adapt the tape recorder to achieve better ways of transforming recorded sounds.

According to Francois Bayle in (Desantos et al., 1997), these were conceived as "sound-transformation-instruments" , conceived for sound manipulation in the studio and not for live performance.

The first instrument was concerned with playback speed and was called the Phonogène (a tape-based transposer and time-stretcher). The phonogène had two versions. There was a chromatic phonogène or phonogène à clavier, which featured a keyboard and was able to transpose sound by changing the playback speed at discrete diatonic steps. There was also a slide phonogène or phonogène à coulisse, which allowed for the continuous transposition of a sound (Battier, 2007).

A later version called the universal phonogéne "allowed one to work independently on speed variation, resulting in an acceleration or a slowing down, as well as the transposition of pitch. Its manipulation was, however, very delicate."This machine was an adaptation of "the principle of the machine developed in Germany by Axel Springer, the Springer Tempophon." used by Eimert in the Cologne studio (Battier, 2007).

Another kind of sound-transformation-instrument was the Morphophone, which "would read a loop of magnetic tape with the aid of ten heads which one could place along the loop and be regulated at will: one might thereby produce repetition according to a chosen rhythm" (Battier, 2007).

As multichannel tape recorders and players became available composers began exploring the use of space. The WDR included a table with which a speaker was rotated surrounded by four microphones, creating a quadraphonic distribution. Space was performed in the studio and fixed in to the final recording.

While these sound-transformation-instruments were conceived for use inside the studio, they enabled a new kind of performance, which was recorded, mixed and presented fixed in a concert. Some of the first electroacoustic works were presented through radio transmission and others in concert halls, where the composer had to leave the tranquility of the research studio and face an audience. Over the years, a practice of "sound diffusion" has emerged which is concerned with the distribution of sounds through multiple speakers in space and which has its origins in a device designed by Poullin called pupitre d'espace in 1951.

The GRM gradually began using computers and shifting from making tape-based devices to programming software. In words of Franois Bayle:

> Each decade we invent a new and more refined Phonogène and Morphophone, which allows us to concentrate on the manipulation of sound morphology. At present, this means the GRM Tools software for the Macintosh (Desantos et al., 1997).

In some sense, the sound-transformation-instruments above were software prototypes. The computer, whose main characteristic is its flexibility, has absorbed all of the techniques of the studio. Personal computers are becoming virtual personal studios displacing the traditional studio as conceived in mid XXth century (Schedel, 2007).

* * *

The "message" of TAPE: the Pseudo-instrument

Schaeffer found in musical instruments a model of musical practice that emerged from experience: "concrete precedes abstract" ((Schaeffer, 1966)). But in the absence of the physical objects, he felt the need to ground his practice in a compositional method, based in the experience of listening. This method looked for ways of extracting, from a recorded signal or sound object, the criteria for its processing, finding values which could change and characteristics that would remain stable, either in calibrations or scales or in continuous transformations. In this way, the perception of a common cause was possible and the pseudo-instrument emerged.

The pseudo-instrument depended too on the choice and production of sound material. The use of recorded sounds contained an intentionality of performance and recording. On one hand, Schaeffer talked of “the person who records sound as a performer” and of the microphone as an instrument. On the other, a choice of sound objects had to be made. Schaeffer and the GRM composers began “exploring a limited collection of objects” as a way “to create a unity of the material”.

Finally, Chion calls our attention to the fact that electronic manipulations have a timbre, which Smalley (1994) calls processing timbre, whereby sound processing techniques become “recognizable by a particular ‘colour’ such manipulations give to these disparate objects” ((Chion, 1995), 47). The sound-transformation-instruments developed in these studios provided a unifying aspect to the sounds they produced.

The pseudo-instrument has a dual nature: it is both the result of compositional decisions about material and its transformation, and its technological realization. It is important as it reveals the fact that in the electroacoustic domain, the composer is always forced to create an instrument as part of his compositional work.

If there is a pseudo-instrument however, there is a pseudo-performer. This pseudo-performer was constructed out of the vestiges of gesture contained in the recordings of the material, in the nature of its transformation and on the way these were articulated in composition. His performance could not be more accurate as the piece itself was the score.

However, storage media such as vinyl disc and magnetic tape, were socially construed as documentary. A record of a band or an orchestra presented us with access to generally live performances that occurred in a different time and place. While in the studio, tape was used as buffer and as a production medium; the final tape was a fixed piece of music and carried with it the illusion of continuous sound. Presenting a work recorded on tape in the concert hall, instead of the live performance of this work, resulted in what we could perhaps call a *pseudo-performance*: the presentation of the document of a performance instead of the performance.

As Schaeffer said,“for the moment we do not do more than perturb the space” ((Schaeffer, 1966), 45). And indeed the space was perturbed. Composers spent long periods of time in the studio producing compositions and brought to the hall a document. In this way, Electroacoustic music proposed a drastic transformation of the traditional model of music production: the composer interacted with the studio in private, where the audience was removed from the process.

Chapter 5

CIRCUITS: Hard and Soft Wires

* * *

The Effects of Electronic Music as a Tape Studio Art

As the electronic music studio established itself as the facility of choice for the research and production of electronic music, large institutions like academic departments, radio companies and large corporations financed their construction and operation. Composers began to demand studio time; however, these facilities offered multiple limitations. Because studios were large and expensive, only large institutions were able to afford them. Furthermore, composition in the studio was a long, labor-intensive process, with constant re-processing of material and laborious tape cuts and splices.

Computers had even more limitations. As with studios, they were large and expensive and therefore were only found in large institutions. Computers required not only knowledge about sound, but also about the specialized skill of computer programming. As processing power was relatively low, there was no immediate feedback and corrections had to be made after hours or days when the results of a computation would be obtained. Computers were flexible, multi-purpose machines and musicians had to share the resource with other researchers. In this sense, computers were initially less appealing than tape studios.

In the 1950's, almost all electronic music was being produced in well-known, established tape studios and composers had to find a way of accessing, building or acquiring electronic equipment in order to participate of this new music practice. As many composers were not able to afford much of the standard studio equipment they were forced to make their own alternative studio designs. Engineers too, saw a possibility of creating equipment that catered to the needs of tape studios and musicians. In the process of re-designing the studio, several trends can be identified.

In order to be accessible to independent composers and smaller institutions, new or alternative studios had to be *cheaper* and *smaller*. At the same time, studios offered opportunities for improvement, both in terms of better quality of equipment and particularly in terms of achieving compositional processes that were *less laborious* than re-processing, cutting and splicing tape. A less laborious process implied some level of automation which afforded composers to hear results *immediately*, instead of going through multiple steps of processing and tape assembly, creating a *smaller feedback loop*.

These two trends (cheaper/smaller and automated/immediate) lead to a listening experience of electronic sound as live in the studio. As studio creation of electronic sounds became more common, a need for live performance emerged, which required *portable* equipment. The creation of electronic music confronted a traditional music practice based on known musical instruments and therefore required the creation of an instrumental identity for each composer and even for each composition, which in turn demanded studios that could be *customized* depending on aesthetic philosophies.

In short, the approach to the new studio would require: cheaper, smaller and portable equipment, that would provide immediate results (smaller feedback loops) and less laborious production, but that still allowed for customization to each composer's personal aesthetics and style.

As a result of these trends, several new ways of producing electronic music emerged, particularly in North America: (1) several independent or small new studios, (2) the modular, voltage-controlled synthesizer and (3) custom built machines.

* * *

New Studios: Ann Arbor - San Francisco

In 1958, Gordon Mumma and Robert Ashley established the *Cooperative Studio for Electronic Music (CSEM)*, in Ann Arbor, Michigan as their alternative to institutional studios. Mumma believed there were “too few studios to accommodate all the composers interested in making electronic music” and the composer had the alternative of “building his own studio and assuming the engineering and financial responsibilities himself” (Mumma, 1964).

For Mumma, building a studio was comparable to buying an instrument. One of his design aims was to have the cost correspond to the “investment that a composer would make in a conventional musical instrument”, putting as much “consideration and planning” into building his studio as a “professional musician would direct toward choosing his conventional performing instrument”. Furthermore, he claimed that the cost of duplicating his studio “could be held below $1,500.00” which compared “favorably with the cost of a grand piano” (Mumma, 1964).

In his article called *An Electronic Music Studio for the Independent Composer*, Mumma (1964) lays out the principles he applied in designing the CSEM in 1958. The article offers the reader an insight into the way such a studio could be built, detailing what “high-fidelity industry products” were available and preferred as well as the kinds of modifications that were possible in order to optimize them. However, the most important aspects of the article relate to the design choices made in terms of choice of equipment and its configuration in space.

The first Margaret (Kartomi, 1990), for choosing equipment depended “upon the manipulation procedures the composer is likely to apply to his music”. Looking at this aspect from another angle, the composer had to realize how his design choices would determine the music he would make. The second criterion was that a composer should choose “a maximum of ‘neutral’ equipment ... designed to operate as separate components ... placeable anywhere in the equipment configuration ... Thus, the composer has considerable freedom to pursue whatever compositional technique or aesthetic he desires". The conscious choice of modular design offered combinatorial configurations and therefore, customization.

The configuration of the equipment in space had the aim of designing the optimal interface for the composition and realization process. It reflected both the customized sound processes as well as the optimal way to manipulate those processes. Optimal in this sense, meant the reduction of tape cutting and splicing in favor of live execution in the studio:

> ... a configuration can be arranged which will put the entire studio within arm's-reach of the comfortably seated composer [providing him with] greatest access to the processes of electronic manipulation. As a result, Ashley has developed facile techniques of magnetic tape composition which rely almost exclusively on switching and mixing procedures and make little use of splicing or cutting of the tape itself (Mumma, 1964).

However, the most important aspect in the design of the CSEM resided in Mumma's ability to modify purchased equipment and to design and build sound modification equipment. Mumma's knowledge of electronics allowed him to realize "electronic and mechanical modifications" as well as "to construct special devices". Many of these devices were built or modified, because they were "not commercially available", because they were too expensive or to extend or achieve particular ranges.

In this sense, the studio was not a monolithic concept and certainly not a standard product. Designing a studio depended on the resources available and the aesthetic purposes of the composer. The CSEM's design reflected a move towards live performance, even if it was achieved primarily as a means to simplify the production of magnetic tapes.

A different approach to confront the scarce access to electronic music studios was through collaboration and resource sharing. The *San Francisco Tape Music Center (SFTMC)* was created in 1961 as a form of collaboration between several composers including Ramon Sender, Morton Subotnick and Pauline Oliveiros. The SFTMC has been understood as a place "to pool their equipment into one studio location" (Pinch and Trocco, 2004). In Subotnick's words, this "collaboration was not in creative works, but in the studio itself - a center for the creation of new works. The mutual thread between us was a distinct break from the post-Webern serial tradition as we saw it at that time" (Roads and Subotnick, 1988).

In Ramon Sender's *An Overview of the Tape Music Center's Goals, Autumn 1964* he defines the SFTMC as "a nonprofit corporation developed and maintained by a group of composers and creatively oriented around engineers" (Bernstein, 2008). While the composers in the SFTMC did not have the engineering skills that Mumma had, they collaborated with engineers such as Bill Maginnis and Donald Buchla. Buchla played both a creative and engineering role.

In another 1964 report, Sender defined the Center as a cultural agency as much as an independent and experimental studio. Because of its collaborative nature and ethos of musical experimentation, the studio attracted composers such as John Cage, David Tudor and Karlheinz Stockhausen to give lectures and concerts. The SFTMC grew and began to appeal to new audiences, later receiving a Rockefeller Foundation grant which would fund, among other things, the collaborative research of Buchla, Subotnick and Sender.

The collaboration between Buchla, Sender and Subotnick began in 1962 and was conceived as the development of a "composer's blackbox." (Roads and Subotnick, 1988). It aimed to "move away from cutting and splicing to get something that was more like an analog computer" (Pinch and Trocco, 2004). The work towards this *Blackbox* led Buchla to the voltage controlled synthesizer, to which we now turn our analysis.

* * *

The Synthesizer

While their were several predecessors[1], the invention of analog, modular, voltage-controlled synthesizers is usually credited to both Robert Moog and Donald Buchla,

---

[1] Perhaps the most important would be found in Bode (1961) as he specifically mentions modular designs.

because they offered the first commercially available modular systems. Its simultaneous, yet independent, invention in both the east and west coasts of the United States in the years 1964 and 1965 speaks to the way in which the multiple conditions for its development had consolidated.

The advent of the transistor allowed engineers to design significantly smaller equipment with stabler behavior. As electronic music continued to become a central part of the musical establishment through the creation of institutional and independent studios around the world, a growing market demand emerged, which provided a fertile ground for collaboration and design based on practical and aesthetic needs; in other words, cheaper, smaller, portable, immediate and customizable equipment.

In very few years, the synthesizer became widely used inside studios and gradually as an instrument in live performance. The main reasons for this warm reception were that it was, at least partially, able to satisfy the practical and aesthetic needs outlined above. The synthesizer proposed to contain the studio inside one machine, into an object one could call an instrument. In his first catalogue, Moog states that "the synthesizer should perform all of the basic generating and modifying operations of the classical studio, and provide additional resources for the state of the art" (Pinch and Trocco, 2004). In this sense, the content of the synthesizer was the classical analog studio. In order to do this, tape as buffering media, began to be replaced by different kinds of automated control mechanisms.

Like the studios it attempted to contain, analog synthesizers were not a standard object; quite on the contrary, there were multiple modular and integrated models with significant variations in control strategies and interfaces. A central concept spanned all synthesizers: voltage control.

Voltage Control

While studios already had ways of interconnecting different devices, these connections were only available from the audio output of a device to the audio input of another one. The control of these devices depended on knobs, switches and the like.

This kind of manual control reduced the number of variables that could be controlled at any one moment by the composer in the studio. Therefore, a sound had to be recorded on tape to be re-processed on a subsequent step and when individual sounds were ready, they had to be pasted together to form a sequence in time. Often, a sequence had to be processed further, re-initiating the process until the final piece was assembled on a master tape.

Instead of just connecting the audio signal output of a device to the audio input of another one, the logic of voltage control allowed the composer to use voltages to control parameters as well. Therefore, the devices that were originally controlled by knobs and switches could then be controlled by the voltage output of another device.

voltage control was aimed at improving the classical studio. Moog would speak of voltage control as simplifying "both the generation of complex, dynamically varying sounds and the arrangement of these sounds into a composition", or stating that "it is generally easier to change a voltage rapidly and precisely than it is to reset panel controls with equal speed". For Moog, these improvements proved "the relative inefficiency of tape editing in the composition of music" (Moog, 1967). Buchla was more succinct about his aims in developing voltage control: "I guess I was basically looking for a replacement for tape" (Bernstein, 2008).

Because the synthesizer attempted to replace tape, it re-introduced the problem of control. Moog believed that composers would benefit from "*play[ing]* spontaneously, more in the manner of a conventional musical instrument than of a code-controlled apparatus" (Moog, 1965b). Voltage control signals could come from a gestural interface, like a theremin or keyboard, or through automation devices like envelope generators and sequencers. As Buchla explains, "when you invent voltage control, you invent ways of generating voltages for control" (Bernstein, 2008).

Designing Control: Gesture, Automation and Culture

Because the human body establishes both possibilities and limitations for manual control, to be able to get rid of tape, and therefore, to perform live, certain control signals would have to be generated live by a human operator or performer and others would have to be automated. For synthesizer designers like Moog and Buchla, choices needed to be made regarding what should be controlled through gestures and what through voltages or at least, what modules should be made available for those purposes. In an overly simplistic view, Moog decided to use a keyboard and an envelope generator while Buchla designed touch-sensitive pads and sequencers.

The choice of using keyboards or not had several implications and created heated debates. As explored earlier, the keyboard can be seen as the material amplification of the musical technologies of thought of tempered tuning, the diatonic scale and of the note as the smallest building block of music. Since the keyboard interface was socially constructed as a mechanical wildcard utilized in a large variety of mechanical instruments like the pipe organ and the piano-forte as well as in electric instruments like the ondes martenot and a plethora of electric organs, it constituted the glue that connected them all.

In Moog's first experiments in voltage control with composer Herb Deutsch, a keyboard was used to generate the control voltage for pitch. However, if pitch, were to be controlled by a keyboard, the rest of control variables would need to be, at least partially, automated. As a consequence, Moog endowed each key with the secondary function of producing a voltage when the key was depressed and to stop producing it when it was released. According to Pinch (2002) Moog's later interaction with Ussachevsky "led to the standard way to describe the main functions of an envelope generator" or ADSR, that was triggered upon pressing a key and sustained until its release[2].

Moog's designs tended towards a recurrent model where the user would choose a particular kind of generator that endowed the sound with an initial harmonic content which was transformed with ADSR-controlled filter and amplifier. The nature of the generator and the filter and amplitude envelopes determined the timbre of the instrument over which the keyboard-controlled pitch varied. In other words, Moog synthesizers ap-

---

[2] The envelope generator was later dubbed ADSR in the ARP synthesizers as an acronym for attack, decay, sustain, decay

proached the traditional *Western concept of musical instrument*. This meant that the Moog synthesizer could speak the language of notes as an abstract set of attributes consisting of pitch, duration, amplitude and timbre, and more importantly, it could imitate mechanical musical instruments.

Ussachevsky was a member of the Columbia-Princeton Electronic Music Center which also included composers Luening, Babbit and Sessions. The Center was associated with the Post-Webern serial tradition, more akin in this sense to the elektronische musik tradition of Cologne. Ussachevsky opposed the use of the keyboard and advised Moog to avoid it in his synthesizers. Both of them however, co-designed the envelope generator as it allowed their equipment to speak the language of notes. Pinch and Trocco (2004) note that Ussachevsky had an engineering education that allowed him to contribute to the design of the RCA Mark II Sound Synthesizer owned by the Columbia-Princeton Electronic Music Center . The RCA was programmed with paper tape realizing the note ideal beyond the keyboard. The paper tape was used to "program" the control voltages of the synthesizer, eliminating live input completely, and allowing the music to be "recorded in real time in its entirety ... almost completely eliminat[ing] the necessity of splicing" (Moog, 1967). In this sense, the ideal studio production of electronic music aspired to the automation of score realizations.

Buchla rejected the total automation of scores, but he too refused the use of the keyboard, which he thought was an "unnatural" way of controlling a voltage controlled system (Bernstein, 2008) :

> A keyboard is dictatorial. When you've got a black and white keyboard there it's hard to play anything but keyboard music. And when there's not a black and white keyboard you get into the knobs and the wires and the interconnections and the timbres, and you get involved in many other aspects of the music, and it's a far more experimental way. It's appealing to fewer people but it's more exciting (Pinch, 2002).

Buchla initially developed two main kinds of gestural interfaces originally called "Touch Controlled Voltage Sources" and "Sequential Voltage Sources".

“Touch Controlled Voltage Sources” were “arrays of touch-sensitive metal pads housed in wooden boxes that he [later] called the *Kinesthetic Input Ports*” (Pinch, 2002). Buchla and Associates offered several kinds of these ports. Model 112 is described in Buchla’s 1966 catalogue as:

> Model 112 Touch Controlled Voltage Source
>
> Touch activated keys produce one of twelve preselected voltages at each of two outputs. A third output voltage is proportional to finger pressure, and a fourth output is a pulse generated whenever a key is activated. Generally used to initiate notes and control their pitches (Buchla, 1966).

Model 112 could output a programmable voltage, but also a voltage proportional to finger pressure as well as a pulse. The pulse worked as a trigger which could initiate envelope generators or other automated behaviors, but finger pressure voltage allowed the performer to create a continuously changing envelope, generating events with different shapes and durations. In contrast to keyboard controllers, Kinesthetic Input Ports simultaneously provided both discrete and continuous control.

“Sequential Voltage Sources” were later known as sequencers. The Buchla catalog describes them in the following way:

> Model 123 Sequential Voltage Source
>
> Produces a sequence of two to eight programmed voltages at each of three outputs. Switching is accomplished by applying a pulse, usually from a Model 140 pulse generator. Indicator lamps show which of the 24 potentiometers are in control. Eight pulse outputs are energized as corresponding segments are switched. Unit may be used to simultaneously program pitch, amplitude, and duration of single or repetitive sequences of notes (Buchla, 1966).

In Buchla’s terms, the sequencer could be seen as part of “looking for a replacement for tape”, however, it partially replaced both magnetic and paper tape. As we will see below, instead of realizing a detailed pre-determined composition, it provided an opportunity to generate and control processes, which could be controlled through the manipulation of global variables.

Customization: Modularity and Integration.

The design and concept of modular synthesizers were constantly evolving. In his 1964 paper, Moog only presented the voltage controlled oscillator and amplifier modules. He later added a filter and gradually added other modules. For Buchla, design was a gradual process too: “it didn’t occur all at once. It was a modular system. One module developed into the next” (Bernstein, 2008). For this reason it is hard to establish when the voltage control synthesizer was invented although most writers suggest 1964 or 1965. But more importantly, the design of any part of the system suggested the design of other parts. Customers were able to order whatever modules they considered interesting and within those choices, multiple combinations could be made.

Buchla’s modules could receive external sound inputs and the performer was able to control mixers, which not un-commonly mixed tape players live; modules like microphone or instrument pre-amps permitted to process live audio inputs as well. In this way, recorded and live instrument sounds could be included in the system. The Buchla also provided modules to operate on control voltages with arithmetic functions like adding, inverting them and re-routing them through matrices. In other words, Buchla systems offered a wide variety of modules that rendered an extremely flexible system, which allowed customization to aesthetic needs.

The SFTMC began to commercialize the modules produced by Buchla and Associates. Ussachevsky was amongst the visitors and clients, buying three Buchla systems. Ussachevsky considered Buchla’s system “logically arranged to be more accessible to composer’s thinking ... than Moog’s synthesizer”, which he felt “somehow it did not have enough flexibility” (Pinch and Trocco, 2004). In a pamphlet from the *Artists’ Research Collective (ARC)* promoting Don Buchla, the following statement from Bob Moog is used:

> I have to admire what Don Buchla has done. He hasn’t allowed himself to limit the complexity of his instruments to meet the demands of the so called ‘market place’ ... As a result, I think of all the systems that are available, his can be organized in arbitrarily complex ways most easily (ARC, date not available).

While Moog's first models were modular as well, his designs gradually moved towards integrated models such as the 1970 Minimoog, designed while the R.A. Moog Company was confronting deep financial trouble. Integrated models aimed to satisfy a market demand for cheaper and portable equipment, but were also simpler and easier to use, at the expense of being less customizable. In the Minimoog there was no possibility of re-patching or adding modules; these were hard-wired to one of their possible interconnections. The Minimoog focused on the idea of keyboard-controlled pitches, with a waveform selector and ADSRs to control filters and amplitude. As integrated designs like the Minimoog found great commercial success, they became the archetypal synthesizer. Synthesizers remediated mechanical instruments.

The monophonic keyboard was improved by including "pitch bend" and "modulation" wheels for expression, alluding perhaps to the bends and vibratos of electric guitars and in some way to the singing voice. At this point, the left hand was concerned with expression and the right hand with pitch control, leaving timbre as a fixed element to be controlled by automation.

Switched-On Bach and the Resurgence of the Ideal

In 1967 Moog had revived the ideal of the universal instrument:

> The "ultimate" performance instrument, on which the performer can produce any sound or combination of sounds in real time, is yet to be developed. The main problem is the design of manual-control devices which permit the performer to continuously specify all the parameters of the sounds. This problem is one of "human engineering" and will undoubtedly take years of experimentation, on the concert stage as well as in the laboratory, to solve completely (Moog, 1967).

Social recognition of the synthesizer as a legitimate musical instrument began with the extremely successful commercial release of the album *Switched-on Bach* by Wendy Carlos in 1968. The album cover presented a keyboard synthesizer with a baroque character in front of it; the cover announced: "Wendy Carlos Performing on the Moog Synthesizer". Switched-On Bach presented the synthesizer as a keyboard instrument that was able to imitate the orchestra, and thus it remediated the organ. The elaborate tape overlaying and mixing created the illusion of live performance on a polyphonic instrument. In an interview in the *New York Times*, Moog called the album "the most stunning breakthrough in electronic music today" (Schonberg, February 16, 1969).

In his 1969 article in the New York Times, the critic Harold Schonberg believed the Moog instrument to be of "unparalleled virtuosity and flexibility" and that "the betting is that, with refinements in the new instruments, even the human voice will be capable of being synthesized. This could literally mean a new era in music" (Schonberg, February 16, 1969).

The ideal instrument seemed to return: the pure tone, the loudness, the instrument everyone could play, the possibility of realizing all timbres in one device, a mass market; in short, the appearance of a new medium that would match and surpass what music media had allowed until then. Like the theremin, the synthesizer was legitimized, or at least recognized, as a musical instrument through the performance of traditional classical repertoire. Furthermore, Pinch reported that Moog "entertained hopes that his synthesizer would find its home in classical music" (Pinch and Trocco, 2004). Even a live Moog quartet emerged to perform classical repertoire and popular music. Gradually a wide array of rock musicians started to use Moogs live and in the studio.

The 1970 Minimoog was the Moog model that achieved the highest sales thus far; synthesizers and other instrument-imitating organs would later become one of the largest electronic instrument markets.

As the ideal of the universal instrument re-emerged, unions of performing musicians began to see the instrument as a threat; as another case of machines taking over people's jobs. The American Federation of Musicians even reached an agreement with advertisers agencies and producers in New York City in 1969, that temporarily banned the use of the Moog synthesizer in commercial work (Pinch and Trocco, 2004). In Moog's words:

> Basically the union didn't understand what the synthesizer was. They thought it was something like a super Mellotron. All the sounds that musicians could make somehow existed in the Moog - all you had to do

> was push a button that said 'Jascha Heifetz' and out would come the most fantastic violin player!" (Pinch and Trocco, 2004)

Switched-On Bach and the emergence of the synthesizer market, reinforced the belief that an instrument is legitimized as a result of its production of revenue. Paraphrasing a joke by Bob Moog, "real music makes real money" (Pinch and Trocco, 2004). The R.A. Moog Company was a business in search of profits and Moog catered to his clients, adding features in response to market demands.

* * *

Composing Circuits

Buchla, Subotnick and Silver Apples of the Moon

Buchla on the other hand was not only an engineer. He belonged to various art groups and performance groups like the Artists' Research Collective and the Electric Weasel Ensemble. He commonly made installations, performances and lectures and was commonly described as a musical instrument designer, a multi-media composer and designer of high level music composition languages. As a Guggenheim Fellow, he conducted research in interactive performance-oriented computer music languages. He later invented several interfaces for live performance such as the Thunder (1990), Wind (1994), Lightning II (1996) and Marimba Lumina (2000) (Buchla, 2005).

Buchla's instruments, synthesizers and interfaces embodied musical technologies of thought different from the keyboard. Unfortunately, Buchla's writing is scarce and his interview responses are extremely concise.

Silver Apples of the Moon, by Morton Subotnick, was realized in a Buchla modular system and released in 1967, enjoying moderate success in the classical and underground audiences. While it sounded elaborately produced in the studio, it actually was closer to live performance than Carlos' Bach album. Subotnick chained sequencers to create processes with “directionality”. He was able to modify the global variables of the processes, while the details of the sounds and the sound structures were automated.

> ... the sequencer was built into the whole concept. My keyboard was not a black-and-white keyboard but was a programmable keyboard right from the beginning. The sequencer was conceived as a programming device. I used three sequencers: two 16-stage and one 8-stage sequencer. They were interconnected in long elaborate patches that were not loose and unstructured, but rather forward-moving programmed processes (Roads and Subotnick, 1988).

Buchla’s first system already included a *Dual Random Voltage Source* (Model 165), which he would later call *The Source of Uncertainty*. Random voltage sources, combined with the multiple arithmetic modules, provided the system with the ability to generate random and probabilistically determined voltages. As these generators were controlled by and fed back into automated modules like sequencers, the sequences could be continually modified and the synthesizer revealed its power as generator of processes; as a partially autonomous system.

Moog had overlooked this property of sequencers:

> “The sequencer can be pre-programmed for only as many events as the number of sequencer stages, after which it repeats the sequence. Thus, the sequencer can eliminate much, but not all, splicing. The next step in programming capability is the continuous programming of control voltages via paper tape (Moog, 1967).

In Moog’s statements resurfaced the ideal of a performer-instrument coupling as predictable, passive and accurate reproduction of the composer’s score, in other words as a machine. In the Buchla system, as used by Subotnick, the instrument was instead indeterminate, active and endowed with compositional agency. The traditional model of music was transforming again: the composer was forced to also take the role of the performer as well and the score gave way to a system configuration that worked as an algorithm in an analog computer.

Mumma and Cybersonics

In his paper *Creative Aspects of Live Performance Electronic Music Technology*, Gordon Mumma (1967) declared that his work had evolved "from its primacy as a magnetic tape medium to an almost complete preoccupation with live-performance electronic means."

Mumma regarded his composition *Medium Size Monograph 1963* as his first "cybersonic" composition and also, as the "most elementary". The piece consisted of a piano and a circuit. The sound input of the circuit was obtained through pickups in the soundboard of the piano. The circuit would "compress the attack transient, spread that energy over into a later portion of the piano envelope, and expand the final part of the envelope" (Mumma, 1967). While simple, the circuit incorporated a logic beyond, for example, modulation. Mumma however, saw in this piece a major innovative feature: he could "carry a small box with which, in a relatively simple way, ... could (effectively) make a 'new' instrument out of a conventional one" (Mumma and Smoliar, 1971).

The term 'effectively' seems to be used in this sentence to stress that the new instrument exists as a matter of fact and not in a metaphorical sense. With this box, "the performer was able to make a different timbre out of the piano simply by changing its amplitude" (Mumma and Smoliar, 1971) producing "a sound that obviously comes from the piano, but is at best a distant acoustical cousin." (Mumma, 1967) The instrument is neither the piano, nor the circuit, but the system that emerges form the piano-circuit configuration:

> I am concerned with "system-concepts": configurations which include sound sources, electronic modification circuitry, control or logic circuitry, playback apparatus (power amplifiers, loudspeakers, and the auditorium), and even social conditions beyond the confines of technology. I suggest that the most important creative aspect of live-performance electronic music is not this or that circuit innovation, but rather the total configuration itself (Mumma, 1967).

The idea of "total configuration" is central to Mumma's work, but can be applied to interactive systems in general. The creative work of the composer does not exclusively consist of writing a score to be executed, but also in 'composing' a system of relations; a situation; a process.

The concept of total configuration reached a deeper realization in his work *Hornpipe* from 1967. The work is described as “a composition for solo French Horn”, but it uses a waldhorn, valvehorn (both with horn and double-reed mouthpieces) and a “cybersonic console” (Dewar et al., 2009). The piece consists of a horn player with a microphone, “a series of vertical pipes, containing their own microphones and resonant at different frequencies” and a “sound modifying circuitry” to which these microphone inputs are applied “at different points in the configuration” (Mumma, 1967). It seems from later accounts of the piece that the vertical pipes were abandoned for just microphones, using the resonances of the space instead of the resonances of the pipes as these pipes are not mentioned after the 1967 paper.

The idea of total configuration now extends its reach to include not only a performer-circuitry coupling, but now also include the space. The performer chooses form sequences of material found in the score, which he has the freedom to re-order. The aim of both the performer and the circuitry is to find the resonant frequencies of the space and gradually narrow down the pitch material to “those things that are relevant to the space” (Mumma and Smoliar, 1971).

It is interesting to compare the narrative created by Mumma and its technological realization:

> The performer wears a small box of circuitry attached to an amplifier and loudspeaker. The piece starts as a solo, during which the box is listening. It has within it several resonant circuits which move about on their own until they find those sounds made by the horn which are most resonant within the space. As soon as these resonant circuits have locked onto these different spatial resonances they begin to store information about the number of times that those resonances are hit. After a certain amount of time, that storage opens a gate. In effect, one or more of the several amplifiers in the box is actuated, and the space is able to respond to the loudspeakers. Further, the performer is in a situation where he can add more sounds from the French horn in such a way as to turn off the amplifiers. So the result is, effectively, a duo between the performer (and his hardware) and the space itself. There are time circuits on the amplifiers that will shut them off on their own, so the performer can turn

> them off in advance according to what he plays or he can extend them. He can even turn them back on again (Mumma and Smoliar, 1971).

On the other hand, we find his description of the technological realization:

> The hardware is realized mostly with differential amplifiers ... There are eight such differential amplifiers, and they have resonance circuits in their feedback loops which are voltage-controlled as to what their resonances are. This is accomplished through the resistance between emitter and collector or a transistor, which varies according to what's on the base. The voltage stored on the base is what tunes each of those amplifiers. Those amplifiers are then followed by a gated amplifier, essentially a voltage-controlled amplifier which turns on with a Schmidt trigger. After certain voltages are released, it turns on ... The memory is capacitors' charge on the gates. There are two storage parts: first, an electrical storage to tune the resonant circuit; then, following the resonant circuit, is a voltage-controlled amplifier which has a charge added on to a second capacitor storage until it reaches the point that the gate opens. As it opens, it drains out the second capacitor after a certain period of time until it closes the gate. It can also be turned off, as was mentioned, by playing pitches that detune the previous resonant amplifier which will make that amplifier drop out by changing its resonant point (Mumma and Smoliar, 1971).

In a way, the circuit is also a configuration of devices, serving a purpose within that total configuration. However, the narrative of the piece ascribes human values to the technology such as “the box is listening”. At other times, he gives the circuitry attributes of intelligence such as “find those sounds” and later in the paper refers to the circuit as making “decisions”. The circuit is believed to “store information” in the form of voltages. For Mumma, the circuitry is a “kind of elementary computer-like function ... It's a special-purpose computer - of no good to mankind except to make a bit of sound” (Mumma and Smoliar, 1971).

Mumma began using the term *cybersonics* to refer to a “facet of electronic music ... very closely related to the field of artificial intelligence”. Mumma directly referenced the term *cybernetics*, coined by Norbert Wiener (1948) to refer to the communication between man and machine, machine and man, and machine and machine. A central concept in cybernetic theories is that of feedback, which, as we will see later, is also central in the theories of embodiment and ecological perception.

Cybernetics was associated with the way in which complex systems could be self-controlled or controlled by humans. These control scenarios were conceptualized as a series of feedback loops, where there is a level of control performed by a human and others that are automated by a machine. Feedback loops form a critical part of Subotnick and Buchla's work as well as of Mumma's. Feedback is viewed on one hand as both audio and voltage feedback loops, used to generate new material or to make comparisons and "decisions". On the other hand, feedback is generated when these decisions "affect what the performer does outside of the device itself" (Mumma and Smoliar, 1971) which, through his actions and sounds, re-enter the system and affect it, forming another loop.

For Mumma then, building the circuitry for his pieces was not like making a tape. The circuitry was constantly presented in a duality. It is presented as a system that is performed by playing a mechanical instrument into a microphone; the system is dependent upon the input of the performer. However, the system is also presented as an "other" who performs as a duo with the performer; an "other" that is partially unpredictable, capable of making decisions and endowed with a "personality". In this sense, the circuitry is an instrument, but to some extent, it is also a performer and a composition.

Mumma considered that his "designing and building of circuits is really 'composing'". The design of the equipment was part of the process of composing: "I am really like the composer who builds his own instruments, though most of my 'instruments' are inseparable from the compositions themselves ... My "end-product" is more than a package of electronic hardware, it is a performance of music (Mumma, 1967). In contrast to integrated electronic equipment like the Minimoog, designed to be independent from the music made with it, Mumma's circuits fused with the piece.

> ... why didn't I go and get myself [a commercially produced synthesizer]? I was good friends with Bob Moog, I could have got myself

> one...I was very impressed with things they did...but I'd been doing all that myself! I'd built stuff that did that already. What Moog had done that was impressive was that he put the combinations in one thing – he built for studio situations – that didn't work for me, I was building for the road, so my stuff was either simpler or smaller, or unique to each situation (2005 interview with Dewar, in Dewar et al. (2009)).

For Mumma, the work of some of his contemporaries like his fellow members of the Sonic Arts Union, Ashley, Behrman and Lucier and others like "Neuhaus, Martirano, ... Rosenboom, Tcherepnine, Lunetta" (Appleton and Perera, 1975) was establishing a practice of instrument building:

> Very few composers consider the creative design of electronic circuits as a requirement of their craft, though it is already clear that some of the most important innovations in electronic music have been contributed by electronically educated composer (Appleton and Perera, 1975).

Salvatore Martirano and the Sal Mar Construction

Salvatore Martirano began working on the *Sal Mar Construction (SMC)* in the summer of 1969, when he developed "a strong interest in devising customized, electronic circuits with which to emulate many of his compositional procedures, particularly various combinatorial manipulations, logic operations, and ideas for sound generation" (Rosenboom, 1996). After learning principles of digital logic, "Sal made several subsequent constructions with various technical collaborators, finally leading to the Sal-Mar Construction, his ... customized, hardware realization of compositional procedures" (Rosenboom, 1996).

The SMC was therefore initially conceived as a way to emulate compositional procedures. It was "an interdisciplinary project involving Martirano, computer science graduate student Sergio Franco, and ILLIAC III designers Rich Borovec and James Divilbiss", and it was "designed, financed and built in 1969 - 1972" at the University of Illinois (Martirano, 1989). Being part of an academic institution, Martirano was able to secure funds to get faculty from other departments, as well as graduate students, involved. Some of the funding was obtained through the university, some of it from large foundations with which he met great obstacles explaining to them how his work was composition. The SMC "weighs 1500 lbs crated and measures 8'x5'x3"' (Martirano, 1989). It included a system of 24 speakers that were hung in the space. While it was not portable in the sense of fitting in a suitcase, Martirano would crate it and then drive it or ship it.

In 1971, during the construction of the SMC, Martirano wrote *Progress Report #1: Electronic Music Instrument Which Combines the Composing Process with Performance in Real Time* (Martirano, 1971). In this text he outlines many of the design choices that were being made as the instrument was constructed. It also reflected Martirano's conception of his system as an *Electronic Music Instrument*.

The SMC presented a new approach to the control and automation of voltage-controlled equipment: using digital circuits and memory. The SMC was a hybrid system where "logical circuits (small and medium scale integration) drive analog modules, such as voltage-controlled oscillators, amplifiers and filters." (Martirano, 1989).

The SMC can be thought of as consisting of three layers. The first is an "Input System" and a "Central Control System", which controlled a second layer consisting of four separate "voices" or programs. Each of thee voices consists of a digital control unit that controlled analog equipment which was then mixed and spatialized. The third layer consisted of 24 speakers or the "Output System".

The voices or programs were truly hybrid and innovative in their design. They were able to produce timbres and envelopes different from other synthesizers. One of the most interesting innovations of the SMC was contributed by Sergio Franco (1974) who designed a Digitally-Controlled Oscillator (DCO) to replace the Voltage-Controlled Oscillators (VCO) of most synthesizers, and which were too part of Martirano's original design:

> In commercially available music systems sound is synthesized by oscillators with five waveshapes, which are then added together to produce composite waveshapes. A common complaint of many composers re-

> garding the possibilities of timbal variety available with such a system is partially due to the fact that waveshapes are produced in this manner.
>
> Our method is different. In conjunction with a counter and a 2-64 bit Random Access Memories (RAM's), the DCO can assume a variety of arbitrary waveshapes in real time. ... The period of a sound is divided into 16 equal time slots, and within each slot one may specify one of 16 possible voltage levels that the waveshape can assume. ... decoded by means of a Digital-to-Analog converter (DAC). ... in a sequence which repeats itself every 16 input pulses (Martirano, 1971).

These digitally synthesized waveforms were then filtered with a Voltage Controlled Filter (VCF) to smooth "the sharp edges of the staircase". In this way, Martirano, used a digital sequencer to convert a sequence of voltages into a waveform, claiming that it was "possible to obtain a rich class of waveshapes which cannot be achieved by other methods of waveshape generation."

Actually, two waveshapes were always generated and sent to a sort of crossfading mixer he called a see-saw. The see-saw circuit consisted of two Voltage-Controlled Amplifiers (VCAs) teach corresponding to a waveshape. When one of them was silent the other was audible. Martirano used comparators to sense the control voltage of the VCAs so that when one of them was silent, it triggered a change in the waveshape of the oscillator and the process repeated so that they would be "blended at the output. By putting an envelope on the control, you can make a large variety of waveshapes and get time-varying spectra" (Vail, 2003).

The nature of the RAM changes is not entirely clear in 1971, but his experiments were geared towards an algorithmic process. A constraint was that the power of the mixed signal would not drop more than 10% of the previous state to achieve a signal with constant power: "If the comparator is in the enable state, the processor performs a sequence of DATA swaps, and then continues to the next memory in the sequence. The result, hopefully, is that the loudness of a sound remains unaffected by timbal changes." (Martirano, 1971) Martirano's approach to timbral construction revived the ideal of an instrument that could produce any timbre "... because of the fact that 2 RAM's are driven by one oscillator, any two waveshapes can be mixed to produce more than 8,000,000 combinations. With variable amplitude and phase relationships the availability of the entire spectrum of waveshapes is approached" (Martirano, 1971).

These time-varying spectra of constant power were the "timbre" of his voices. Each of these voices were then subjected to another VCA to impose an envelope. Martirano criticized most commercial systems and took as an example Moog's envelope generator based on the ADSR model. His main criticism was that this generator did not allow to generate the small variations present in most mechanical sounds: "it is impossible to program the dimples and pimples in a shape that custom has conditioned our ears to expect in a musical sound." He therefore used RAM to program rise and fall times, but also these small variations in envelope:

> In our system 12 bits of control allow one to produce an enormous number of envelopes ranging from .02 seconds to 20 seconds. ... A pseudo-random set can be used to program the least significant bits in order to produce dimples and pimples to one's heart's content (Martirano, 1971).

Martirano also used feedback loops to vary the timbral variations and the kinds of envelopes that controlled the voices. Similar processes controlled the spatialization system.

These voices were controlled by a Central Control System in charge of generating durations and pitches. Martirano describes it as a "system ... for duration or clock rate which allows independent control of six clocks." The rates of these clocks are said to vary between "60 seconds per cycle to 8000 cycles per second." These clocks are then logically combined in several ways:

> The 6 clocks can be "ANDed" and "EXCLUSIVE Ored" in a variety of ways. For example, 6 groups of 4 clock rates can be ANDed in the following manner: 1,2,3,4; 2,3,4,5; 3,4,5,6; 4,5,6,1; 5,6,1,2; and 6,1,2,3. A common denominator frequency which is the product of 4 clock rates lends coherence to the durations produced" (Martirano, 1971).

One can infer that when the different clocks at their multiple rates satisfied a logical test, certain actions would be taken. The slow rates suggest concerns with form, while faster rates suggest concerns with rhythm and duration. Martirano had designed a matrix, which was stored in memory, to generate multiple 12-tone pitch rows through several combinations and transformations from which the pitches were drawn. More importantly, this clock system made the SMC a constant process.

All of these modules, the “Central Control System”, the “Voices”, their “Waveshapes” and their “Envelopes” were controlled by the performer through the “Input System”. The input system had no continuous input at all, neither in the form of knobs or faders nor in the form of an audio signal. Instead, “an array of 291, 2-state switches can be controlled by logic circuits or played manually. The keyboard of the instrument features TOUCH CONTROL (no moving parts) which set or reset a flip-flop when the TOUCH CONTROL circuit is closed to a current source” (Martirano, 1971).

It seems that the functions of these switches was extremely diverse. For example, Martirano stated that “one mode of interaction possible between the performer and automatic circuitry ... [happens when] ... TOUCH CONTROLS function as the upper rank of a shift register. The performer has access to a parallel input on each bit of the register and can program a ‘1’ or a ‘0’.” In other cases, combinations of switches would have different meanings. For example, when the position of two switches was [0 0], the system chose 12 tempered tones per octave, when it was [0 1], 16, when it was [1 0], 20 and when it was [1 1] a variable number. Another example refers to amplitude control. Certain switches would force a voice to REST, while others would HOLD it to “inhibit the evolution of the program” or to implement “amplitude modulation”. As a final example, “set combinations [of pitches] can be biased to stress certain pitches by repeating them. And as well by withholding certain other pitches for a time, provide intrinsic contrast which can be used at will” (Martirano, 1971).

Many more examples of control by switches are provided and commonly each switch would have a specific role in the system. A crucial aspect of the SMC was that it possesses “circuits which allow the performer to decide whether a particular function is under manual or automatic control” (Martirano, 1971) and even in the evolution of these automatic controls “ You can also contour the field in which pseudo-random sequences operate.” (Vail, 2003)

Martirano described the SMC as “based on the idea of ‘zoomable’ control - being able to apply the same controls at any level, from the micro-structure of individual timbres to the macro-structure of an entire musical composition” (Walker et al., 1992). The concept of being able to shift control from timbre to structure is central to Martirano, and is one of the most salient features of electronic and computer music instruments. “A pianist remains at one level only, that is, on the keys” (Martirano, 1989).

The level of complexity of the SMC was daunting. As in most of the examples of complex systems mentioned so far - Subotnick’s Silver Apples, Mumma’s Hornpipe and, also in the Sal Mar Construction - the composer is the performer and is deeply involved in the design of the system of music generation which allows him to explore it in deeper ways. The instrument becomes a complex system that has to be designed, learnt, explored in order to be performed.

For Martirano, the SMC was “played and programmed” and playing it involved “the selection of melodies and sequences generated by logic circuits.” (Vail, 2003) Performance was not about continuously shaping sounds with his mouth or his hands, but the control of a series of interconnected processes:

> When I touch switches, I’m actually making a circuit. Each switch can be both logically and manually driven. That allows machine functions to interfere, or for me to interfere with the machine functions (Vail, 2003).
>
> It was too complex to analyze. But it was possible to predict what sound would result, and this caused me to lightly touch or slam a switch as if this had an effect. Control was an illusion. But I was in the loop. I was trading swaps with the logic. I enabled paths. Or better, I steered. It was like driving a bus (Chadabe, 1997).

For Martirano, an instrument was performed, but at the same time a piece of music was composed. Patching linked input to audio equipment, as well as the latter to the spatialization system. For Martirano, patching seemed to be a compositional act:

> ... an analogy to traditional composition exists, considering that particular patching patterns are developed over a period of time. In performance a composer can improvise, better said, compose in real time, within a large set of musical possibilities (Martirano, 1971).

For him, the "instrument is ... congruent with the development of compositional systems." (Martirano, 1971) Indeed, the instrument executed the logical formalization of compositional systems for durations, pitches, timbres, envelopes as well as structural aspects of a work, and therefore, both building it and playing it were complementary aspects of composing. The first aspect consisted in determining how the instrument contains compositional thought. The second concerns how this thought is realized in real time as a concrete form.

In a 1977 letter to Howard Klein from the Rockefeller Foundation, Martirano struggled to get this idea across:

> I know that the Rockefeller Foundation's policy in the past has been to not support projects that involve hardware costs. This may also be the case with the NEA. However, I think that in this project composition, performance and hardware design are related conceptually rather than toward flexible instrument for composition (Martirano, November 2, 1977).

* * *

Conclusions: The "Message" of the Circuit

In essence, the theremin was a circuit system with one configuration, or in other words, a hard-wired voltage-controlled synthesizer; and in this sense, not much different from a Minimoog. Moreover, Buchla's Model 117, called "Dual Proximity Sensor", was offered with "theremin-style antennas", (Buchla, 1966) enabling the user to patch a theremin by combining the right modules.

The voltage-controlled synthesizer was initially conceived as a portable analog studio. It attempted to reduce the laborious process of tape cutting and splicing by creating sound live. Fully automated reproduction systems used paper tape as an idealized, accurate and precise, score medium that rendered the ideal performer-instrument coupling a machine.

Modular design allowed composers in the studio to configure the synthesizer according to their needs. Dealing with voltage-controlled equipment, combined with a desire to reduce the use of tape as buffer, demanded the use of automation. While in studios like Columbia-Princeton the initial solution was based on full automation, other systems like Moog's combined gestural input (keyboard) with automated behavior (ADSR).

In the case of Buchla, Mumma and Martirano, it was automation, along with "sources of uncertainty" like random generators, human choices and live signals, that allowed for the creation of feedback loops and the emergence of complex processes. These processes evolved to the point of rendering the final tape an unnecessary step; furthermore, the act of performance was gradually becoming central to the works.

In the generation of a process, we find the concept of "total configuration" set forth by Mumma. In this kind of musical situation, the work cannot be contained on a score, nor the instrument in any one object. The circuit becomes a configuration. Interactive circuits embedded compositional logic and composition required performance.

The traditional music model, with clearly demarcated roles of composer, performer, audience and delineated and specialized media is radically re-formulated. In the context of the "total configuration" the model itself is bound to be composed.

Exchange

Although the theremin did not manage to remain as a performance instrument in the traditional music circuits, and there were no commercial models available for purchase, circuit schematics were sporadically published in electronics magazines. Moog learnt from a schematic to build a theremin and improved it proposing modifications for new models. Moog's sales of theremins established the beginning of his commercial endeavor. Their reputation attracted composer Herbert Deutsch with whom the synthesizer was developed. Moog published various theremin schematics, as well as schematics for his voltage-controlled oscillator, amplifier (Moog, 1965b), filter (Moog, 1965a), frequency shifter, (Bode and Moog, 1972) amongst many others. Central to his company was the publication of the magazine Electronic Music Review. Publications like *Electronotes Newsletter*, first published in 1972, served the purpose of exchanging information including audio processing techniques and publishing schematics and tutorials for building sound equipment.

Both Moog and Buchla belonged to an era of “do-it-yourself stereo kits, ham radio sets, model rockets, chemistry projects, and so on, epitomized in the DIY philosophy of the by-then venerable Popular Mechanics magazine (which was first published in 1902)” (Dewar et al., 2009). A philosophy that constituted a social attitude towards technology is present too in Mumma and the members of the Sonic Arts Union. Gordon Mumma, David Behrman and David Tudor exchanged letters about electronics and the equipment they were building. They commonly bought kits and sabotaged them or re-purposed surplus equipment and junk. Mumma’s tutorial on building a studio offered the independent musician a way of entering the world of electronic music. Mumma would also publish diagrams about the way his circuits were realized. Martirano even stated “we have schematics for two commercially available units and a study of these can serve as a point of departure” (Martirano, 1971).

These sorts of exchanges and interventions on technology pre-configure the way we exchange information today about software and hardware. If, as I am arguing, instruments contain technologies of thought or music theories; if they formalize compositional techniques, or personal styles:

What is the effect of this exchange in terms of instrument building and music creation?

Chapter 6

COMPUTERS: To Solder or Not to Solder

* * *

A Set of Instructions

After the initial success of the NWDR and RTF[1] studios in Germany and France respectively, important studios for electronic music were opened in several countries including Italy, Japan and the United States. Parallel to this growth, research was being conducted at Bell Labs by Max Mathews who was later joined by F. Richard Moore on the use of digital computers in music.

Mathews was aware of the techniques used in the European studio models and saw in the computer an opportunity to improve them:

> Two great technical difficulties facing composers of electronic music are the generation of waveforms of arbitrary complexity and the assembling of multiple sounds with precision and without addition of noise. A source of sound which is free of constraints on waveforms and some problems of assembly is the digital computer (Mathews and Guttman, 1959).

---

[1] North West Deutscher Rundfunk and Radio Television Française as referenced earlier

Imprecision and added noise were noticeable problems of cutting and splicing tape, but also a reflection of intensive, time consuming, manual labour. Many of the patches and processes in the studio, such as adding sinusoids, were designed to achieve a particular waveform, which the computer could produce by compiling commands.

"A computer is very precise" ... "A computer is very flexible" ... "A computer is very easy to use". These were the beginnings of some of the paragraphs in (Mathews et al., 1962). The ideals of an instrument capable of all sounds and a performer capable of performing without any limits re-emerged:

> We believe that with a program such as that described in the paper any or almost any distinguishable sound can be generated. This does not mean that any waveform can be generated but that any waveform can be approximated beyond the ability of the ear to make distinctions. Present limitations lie in lack of understanding of what waveform will produce a given subjective effect rather than in lack of ability to create a specified waveform. There is no limitation whatever on speed of execution (Mathews et al., 1962).

This time however, the ideals were limited by our knowledge. The computer can only do what you can tell it to do. This realization provided the fuel for an interest in psychoacoustics and synthesis. A composer needed to know what to tell the computer, and therefore composers like James Tenney and Jean Claude Risset (Tenney, 1969) (Tenney, 1965) (Risset, 1969) would spend time researching psychoacoustics, stochastic processes and synthesis as part of their positions as resident composers (Risset, 1987).

Mathews specified one more aspect he saw as valuable in the use of computers to generate music:

> It is quite possible to relieve the composer of some of the tedium of composition by allowing him to store not just notes themselves but instructions as well. Then, with a few instructions, the composer may modify a much larger set of notes and, in mid-composition, alter or replace stored waveforms and parameters. He may also allow the computer itself to do its own composing. Perhaps ideally, the composer will arrange the structure of the composition, and the computer will construct the details. This

> is analogous to the once common practice of setting aside portions of a composition for improvisation by the performer (Mathews and Guttman, 1959).

The idea of storing instructions instead of notes or waveforms established a native role for the computer in music, a role that was being sought after with diverse devices for automation. Instead of the automation of envelopes or of scores, Mathews proposed a set of instructions to modify those envelopes or change the scores and other forms of feedback. More importantly, he proposes that the activity of a composer can be to design processes or systems rather than constructing the details of a piece, that instead of using a note or a sound object as the minimum element of music, one could also use instructions.[2] In fact, the program that Mathews was describing in this 1959 paper, and which he was already programming, was also a set of instructions.

In his 1963 essay entitled *The Computer as a Musical Instrument*, Mathews devised two strategies to understand the role of the computer as a means to make electronic music with computers. The first one consisted in making programs that generated number sequences that were converted into sounds (computer sound synthesis as described so far) and the other one consisted in making programs with which the computer controlled analog sound devices through voltages (a hybrid system). About the former he said the following:

> The electronic equipment (computer and output equipment) has been constructed once and for all. There are no soldering irons, tape-splicings, or even knob-twistings involved, as there are with other electronic equipment for producing music. No manual dexterity is required. Instead, one writes down and gives the computer a sequence of numbers (Mathews, 1963).

* * *

MUSIC-N: A First Generation of Computer Music Programming Languages

---

[2] A precedent of course existed in the work of Lejaren Hiller in 1955, but we are not intending historical thoroughness

In contrast to studio practice around the world, Mathews proposed a radical change: the "virtual" studio. By virtual I mean that the studio does not consist of physical equipment in space, but exists in code. There was no need for reconfiguring patches, rotating knobs, mixing and splicing tapes, as all of these changes could be specified in code and computed. MUSIC I, the first in Mathew's MUSIC-N programs[3], was written in 1957 and incorporated several concepts rooted in traditional musical practice:

> In the first part the composer specifies, in computer language, the characteristics of a set of musical instruments. ... He then prepares a score consisting of a list of notes to be played on the instrument-units he has created. The samples of sound wave are generated by putting the score, in a form the machine can read, into the computer, together with the instrument-units, and turning on the computer. ... The interconnected blocks of program which make up the instrument-unit are called unit generators. Once the composer has supplied specifications for the orchestra, he must prepare a score giving the parameters of the notes he wishes played (Mathews, 1963).

Mathews concept of *unit generators* became a standard element in music programming environments to come. Unit generators provided a parallel to the methods of elektronische musik, where sounds were built up of "pure" elements and controlled through parameters; and while it preceded them, it also provided a conceptual parallel with voltage-controlled analog synthesizers. Unit generators were conceived as the building blocks of instrument-units. A sound in elektronische musik, or a patch in a voltage-controlled synthesizer, could also be seen as an instrument-unit. In contrast to these practices however, there was no limit to the number of voices that could be used. Anything that could be specified in code was possible. However, there was no interface for live input and no immediate feedback was possible: computing sound required great amounts of computing power and therefore it was far from being real time, taking hours, days or weeks to hear the resulting sound of a program.

By incorporating the concepts of instrument, orchestra and score, Mathews embedded in his programs the concept of the note as the unit of musical organization and

---

[3] The term MUSIC-N will be used to refer to the five versions of these programs, ranging from MUSIC I through V. MUSIC V was already a compiling program, and therefore it was not computer specific.

of the instrument as spectra. These analogies also became a paradigm which many music programming environments followed. By presenting the computer in terms of traditional practices, the ideals of the musical instrument re-emerged.

However, much like in the analog studios, the composer had to create his own instruments in order to write a score for them. In this case, an instrument was a formalization of a sound in terms of unit-generators:

> How the instruments are to be designed is thus one of the first questions the composer must answer when he begins to program a composition using the computer, and an understanding of this aspect of the process is necessary for an effective utilization of the medium (Tenney, 1963).

Tenney here is pointing us to the fact that instruments had to be designed with the composition in mind; that the design of any part of the system had to have in mind the whole system. This system was a computer program which consisted of a score that was made of notes that referenced instrument-units that in turn referenced unit-generators.

These MUSIC programs had the characteristic of being a "complete description of the sound structure" (Risset, 1987) and in this sense they were the perfect realization of the score-performer-instrument ideal. More importantly however, these "complete descriptions" could be "studied, replicated, and varied" (Risset, 1987). Risset had developed an *Introductory Catalogue of Computer Synthesized Sounds* (Risset, 1969) in which several sounds and sound structures were described theoretically and with its realization in code.

Unknowingly perhaps at the time, Risset was opening the door for a kind of exchange that would become a determinant factor in the practice of programming software and composing computer music. In contrast to the artisanal ways in which studio music was created and the difficulty of modifying soldered circuits, computer code offered a way of revisiting one's own work and of understanding the practice of another person:

> My examples, of course, were meant as instances, as points of departure for developing timbres or sonic processes, and by no means as models. I believe such documents can be very useful. I often find it difficult to get started making a sound. Within a certain class of timbres it is

> easier to tune an instrument to one's specific desires. ... I had been impressed with the efficiency of communication when John Chowning left his data at Bell Laboratories. The input data for programs like Music V give a thorough record of the physical structure of the sounds and of their combination - a genuine score for the control of the sound structure (Risset, 1985).

Risset came to advocate the flow of information as a way of advancing a shared practice:

> This issue of passing information and know-how is crucial. A MUSIC V score (this also applies for scores of other synthesis programs such as CSOUND or CLM) gives complete information about the sounds and their elaboration. Analyzing scores is an essential part of teaching - or learning - composition. MUSIC V, CSOUND or CLM scores can be analyzed by composers interested in recipes for sound synthesis and in the compositional elaboration of sonic structures. Computer synthesis of sound would remain very primitive and unmusical if explorers had not shared the know-how they developed (Risset, 2003).

In this way, compositions or instruments from one person could be used as points of departure for the compositions or instruments of another.

* * *

"Live" Before real time

In an attempt to achieve real time control of sound, Mathews and Moore decided to embark on the second strategy proposed in Mathews (1963): to make programs whereby the computer would produce voltages to control analog sound devices. Systems that combined digital and analog equipment were called *hybrid* systems. Composers like Martirano arrived at digital circuits as a solution to the problem of controlling analog equipment. In contrast, Mathews and Moore arrived at voltage-controlled equipment as a solution for real time generation of sound. In other words, the voltage-controlled equipment was an attempt to compensate for the long computation time, and obtain immediate feedback.

However, the design of a hybrid system implied a compromise with the configurability that Mathews' MUSIC-N programs afforded. For Mathews, the disadvantage of using analog equipment instead of computation was that "the only sounds that can be generated are those produced by the particular electronic apparatus employed, and hence the generality of the sampling process is not attainable." (Mathews, 1963) Furthermore it implied that the analog equipment had to be patched for each particular application.

This hybrid system was called *GROOVE* which was an acronym for *Generating Realtime Operations On Voltage-controlled Equipment*. It was considered by the authors as a "man-machine system" (Mathews and Moore, 1970). Much of the language used to describe and conceptualize GROOVE came from the theories of Cybernetics proposed by Wiener (1948) and was probably the first attempt at conceptualizing music control as a feedback loop:

> ... the control of the body itself, speaking, and playing music can all be characterized by a suitable set of time functions
> ... People inevitably use their sensory inputs to control their motor activities in a feedback operation. In creating time functions on the computer, GROOVE provides opportunity for immediate feedback from observations of the effects of time functions to computer inputs which compose the functions. Thus he is able to modify the functions instantaneously as a result of his observations of their effects (Mathews and Moore, 1970).

The idea of time functions had clear differences with the note paradigm. It proposed continuous and multidimensional control through the use of an interface consisting of multiple knobs and a 3-dimensional "linear wand" or joystick as interfaces. Mathews and Moore decided not to include a keyboard in their input devices because of the "danger of imposing on the computer the limitations of the organ". The choice of continuous control with real time feedback reflected their desire to "make possible the nuances of real time performance in computer music" (Mathews and Moore, 1970). While the "precise" computer was able to create an ideal score-instrument-performer, it lacked the nuance of real time performance.

Because of the size of the computer, the analog equipment and the control environment (which consisted of a few rooms), it was not conceived as a portable performance system for the concert-hall. However, the user was able to hear the effects of his gestures in real time and capture, store and edit these gestures:

> ... with the addition of a human link, we can currently do more complex tasks ... the authors, who wrote the program, enjoy not only being in the loop but retaining command (Mathews and Moore, 1970).

GROOVE was defined as a program for composing, storing and editing functions of time in real time (Mathews and Moore, 1970) (Moore, 1969). The user was able to record functions of time through live performance at the input devices or design functions of time through mathematical descriptions. The system was able to combine the functions of time mathematically: “It is easy to imagine relatively complex control signals which are really only the sums, differences, products, etc., of several simple functions of time.” (Mathews and Moore, 1970). After this process “The resultant functions of time may be saved in bulk memory, so they need only be calculated once, and thereafter are available for either ‘replay’, or further combination” (Moore, 1969). The choice of how to combine these functions of time and to assign the output of the computer to the input of the analog sound equipment was left to the user, and in recent literature has come to be known as a *mapping layer*.

The mapping layer was achieved through a patchboard where the output voltages of the computer were linked to the input voltages of analog audio modules with wires. To be able to store and retrieve patches for each user, piece or for a particular configuration within a piece, the GROOVE system implemented a removable patch bay. Although it might seem a minor detail, the removable patch bay permitted users to change the configuration by using their own patch boards which could be literally plugged into the main patch bay. In this way, the computer extended its configurability into the analog domain. Although it still required a manual action to load the previous configuration, the removable patch bay acted as a kind of “computer memory” of the interconnections of one signal to the other. Figure 6.1 shows the room containing the analog equipment and the removable patch bay.

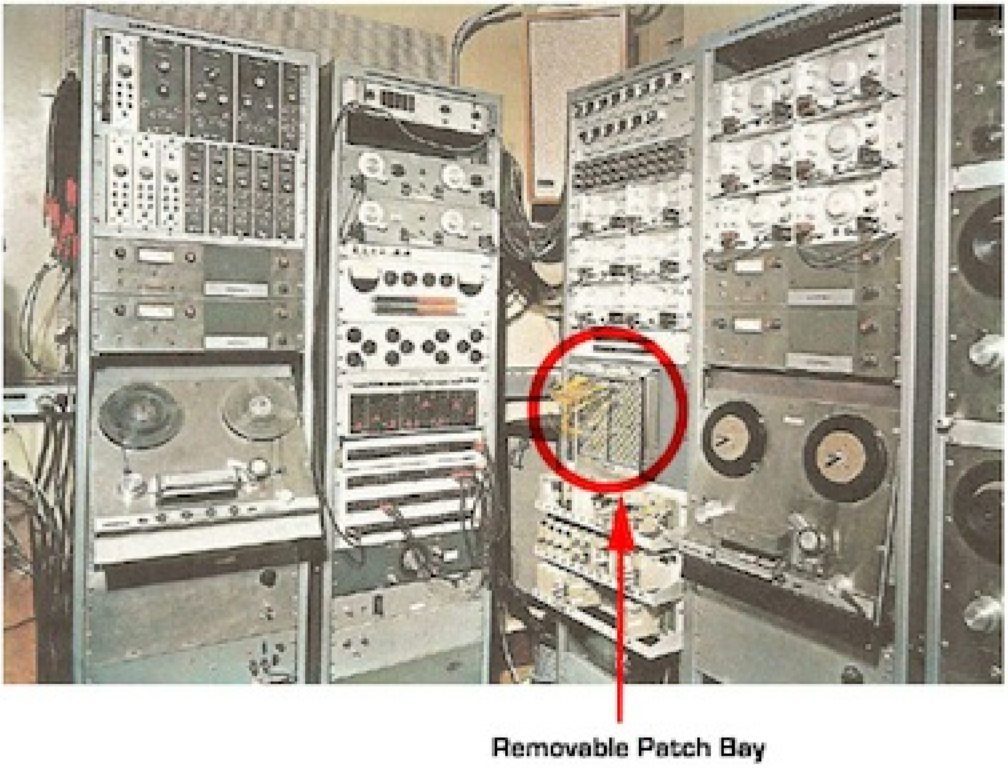

Figure 6.1: Analog Equipment for the GROOVE System and Removable Patch Bay (Mathews and Moore, 1970).

This was a big conceptual break with studio practices as sound did not have to be fixed on magnetic tape as a buffer in the process of its construction. Instead, the sound was stored in computer memory as functions associated with inputs to the voltage-controlled equipment which would generate the sound. In studio practices, sounds were recorded as audio signals on tape and therefore they could only be transformed through tape modification techniques. In GROOVE, since the sounds were stored as functions of time in digital memory, that is, as instructions and data, sounds were modified by editing the functions.

The overall diagram in the GROOVE paper in 1969, shown in Figure 6.2, can serve as a basis to conceptualize all digital musical instruments. All diagrams of elec-

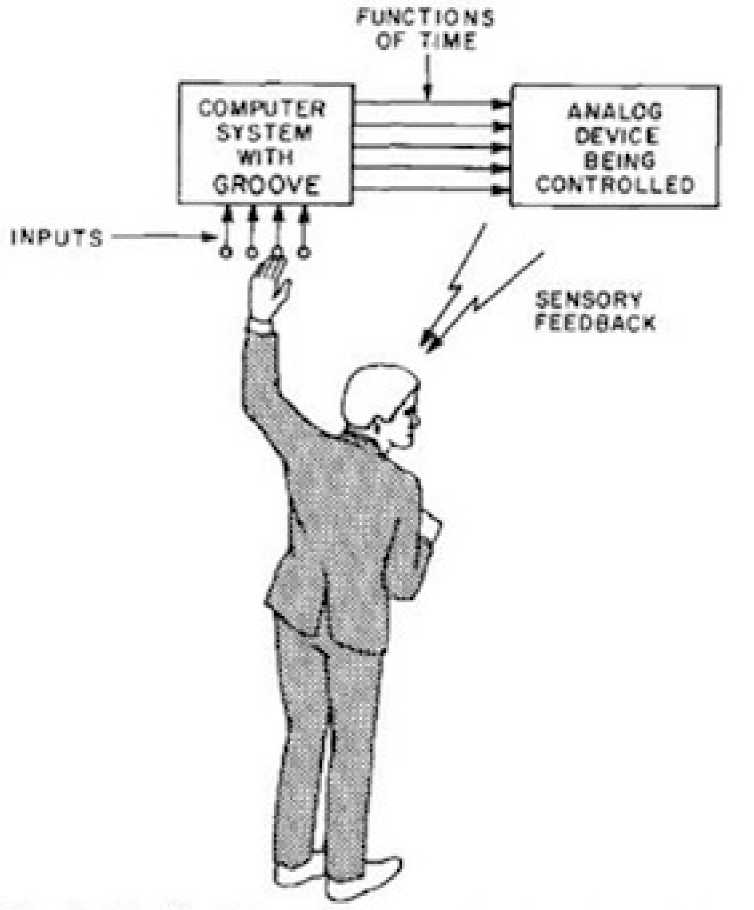

FIG. 1. Feedback loop for composing functions of time

Figure 6.2: Global diagram for the GROOVE system (Mathews and Moore, 1970)

tronic musical instruments are essentially variations on it: a human being inside a feedback loop which was central to the cybernetic theories and also to the theories of embodiment in cognitive science.

Although GROOVE established the model for the design of digital musical instruments, Mathews and Moore did not conceive of it as a musical instrument, but as a "conductor program":

> Further thought convinced us that the desired relation between the performer and the computer is not that between the player and his instrument, but rather that between the conductor and the orchestra. The conductor does not personally play every note in the score; instead he influ-

> ences (hopefully controls) the way in which the instrumentalists play the notes. The computer performer should not attempt to define the entire sound in real time. Instead, the computer should have a score and the performer should influence the way in which the score is played. His modes of influence can be much more varied than that of a conventional conductor who primarily controls tempo, loudness, and style. He can, for example, insert an additional voice of his own, or part of a voice such as the pitch line while the computer supplies the rhythm. He should also be able to modify or edit the score. The computer should not only remember the score, but also all the conductor's functions, so when he achieves a desired performance, it can subsequently be replayed by the computer from memory (Mathews and Moore, 1970).

Mathews and Moore incorporated the metaphors of musical function developed in the MUSIC-N languages to the GROOVE system, seeing the latter as an analogy for the control of those metaphors. In that sense, the metaphor of the conductor was made to fit into the logic of computer performance. However, there were certain ideas that did not fit into the traditional idea of conductor: that the conductor could have "a voice of its own" or that he could be able to "modify or edit the score". This "conductor" began to shape the computer-performer, able to control higher level parameters or a voice of his own. The GROOVE system also configured a "musical instrument" that was not only composed of multiple timbres, remediating the orchestra, but also an instrument that would have "memory" and an embedded "score", blurring the boundaries between composition, instrument design and performance.

The key to understand this passage lies in the phrase: "The computer performer should not attempt to define the entire sound in real time". This resonates with Eimert's statement about the impossibility of controlling "such a great amount of sound elements" and with Schaeffer's basic law of instruments and law of all musics: values vary while characteristics remain the same. Contemporary with the development of analog synthesizers and other hybrid systems like Martirano's, certain elements were controlled by the performer and others were left to automation.

For Mathews and Moore, the computer performer controlled certain aspects while it left others to the "score"; live performance meant a compromise between "live" (inputs) and “fixed” (automated) signals. As I will argue, this compromise lies on one of the continua that characterizes computer music performance in general. In mechanical instruments, although the performer has a great degree of control over multiple dimensions, the performer does not define the entire sound in real time either, much of it is left to the physical properties of the instrument, properties that Wishart calls intrinsic morphology (Wishart and Emmerson, 1996).

* * *

MUSIC-N and GROOVE

GROOVE created a situation. The computer received gestures and produced voltages that produced sound. The user would input gestures and receive the sound output as feedback. In this sense, GROOVE was a musical instrument and the user was a performer. Yet this same user commonly had the ability to design the behavior or processing of the control signals, the routing of these signals to analog equipment and the way this configuration would evolve over time. In this sense the user was a composer.

Mathews and Moore (1970) decided to make a man-machine system and to do so, they needed real time computation. They could certainly compute control signals with the available processing power, but computing audio waveforms was still not attainable.

To Solder[4] meant that the “generality of the sampling process” would have to be sacrificed, in order to gain live-control, and with it, “make possible the nuances of real time performance in computer music.” The use of analog equipment restricted the possible sounds and number of voices to those available in the assortment of analog devices.

Not to Solder on the other hand, gave the composer the ability to specify in computer code a “complete description of the sound structure”. In this way, “the generality

---

[4] In allusion to Mathews’ (1963) opposition of the computer to: “soldering irons, tape-splicings, or even knob-twistings involved, as there are with other electronic equipment for producing music.”

of the sampling process" was attained, but the "nuances of real time performance in computer music", were lost.

Computing power has been a constraint since the beginning of the use of computers. Code must be optimized, processes prioritized, and there is a limit to the amount of processes that can be attained live. Restricted by computing power, Mathews and Moore were forced to make a decision.

What Mathews saw as two options over which a compromise should be reached, became the ideal configuration: retaining the ability to completely describe sound processes, while leaving the system open to receive and process complex control signals. In other words, the ideal was a GROOVE program controlling a MUSIC-N program (if it were possible to execute it live); a program for control and a program for sound synthesis, both expressed in shareable code.

Chapter 7

LIVE: Real Time Interactive Computer Systems

The 1970s witnessed the creation of two important research centers. In Amsterdam, 1971, the *Studio for Elektro-Instrumentale Muziek, STEIM*, was founded with the aim of creating new interfaces for musical interaction and control of live electronics. The *Institute de Recherche et Coordination Acoustique/Musique, IRCAM* started functioning in the year 1977 in Paris, and had as one of its central research areas the use of computers for real time transformation and synthesis of sounds.

The 1980's saw a technological revolution with the introduction of the silicon microchip. Computers became more powerful, smaller and cheaper, leading to the personal computer. It also saw the birth of the *Musical Instrument Digital Interface (MIDI)* protocol in 1982. This protocol standardized messages for communication and synchronization allowing the interconnection of various kinds of equipment including computers, interfaces and synthesizers. MIDI was designed to encode control messages, accommodating particularly well the traditional keyboard interface and enforcing a trigger logic into the music made with it. The MIDI protocol, as an industry standard, enabled the growth of a market. To the artist looking for personal ways of making music, it offered a fast way of communicating devices, but it also established multiple limitations.

The limitations of the 1982 MIDI protocol, which is still widely used today, were already well documented by Moore (1988) in his paper *The dysfunctions of MIDI*. In this paper, Moore evaluated its inadequacy for complex musical projects, and warned that the actions of live musical performers were being “MIDIfied”:

> A fundamental motivation for achieving a real time performance capability in computer music ... is to recapture this level of “visceral control” of musical sound ... The more a musical instrument allows such affects to be reflected in the sound spontaneously at the will of the performer, the more musically powerful that instrument will be. Consider the inflections of a human voice. Consider the intimate nuances of a violin. Consider the plaintive saxophone. Now let us consider a MIDI-based synthesizer. Moore (1988)

In the four years spanning 1984 through 1988, one software and three live computer music works emerged from IRCAM and STEIM, and established themselves as models of live computer performance. These models can be conceptually traced to certain predecessors and many of today’s practices can in turn be traced to these models. These three performances, or performance systems, were George Lewis’s *Voyager* system, 1986-8, Michel Waisvisz’s *The Hands*, 1984-6, and the *Sonus ex Machina* series by Philippe Manoury and Miller Puckette, which began with *Jupiter* (1985-7), for flute and computer, and which featured the new computer music software Max.

Interactive computer music systems are defined by Robert Rowe as “those whose behavior changes in response to musical input” (Rowe, 1993). He classified these systems according to three dimensions. The first dimension runs along an axis of score-driven and performance-driven systems. The second dimension distinguishes response methods as transformative, generative, or sequenced. The third dimension distinguishes between the instrument and player paradigms. His conceptions of instrument and player paradigms are described as follows:

> Instrument paradigm systems are concerned with constructing an extended musical instrument: performance gestures from a human player are analyzed by the computer and guide an elaborated output exceeding

> normal instrumental response. Imagining such a system being played by a single performer, the musical result would be thought of as a solo.
>
> Systems following a player paradigm try to construct an artificial player, a musical presence with a personality and behavior of its own, though it may vary in the degree to which it follows the lead of a human partner. A player paradigm system played by a single human would produce an output more like a duet ((Rowe, 1993), 8).

Both (Lewis, 2000) and (Rowe, 1993) conceive of these paradigms as ends in a continuum. As such, we will use the three pieces mentioned above to explore this continuum.

* * *

The Hands

Michel Waisvisz was an instrument designer and performer working at STEIM. His best known instrument and piece was *The Hands* (Waisvisz, 1985). The Hands used analogue-to-MIDI converters to connect its sensors with digital synthesizers. The interface allowed the performer to navigate presets with previously designed sounds (therefore changing timbres). Most of the control features of The Hands were discrete (buttons and switches) except for a sonar sensor that measured the distance between the two hands.

However, the instrument also provided a "scratch" function (on/off) which forced every switch that was in the "down" position to send new "key-on" events at every change perceived by the sonar detector, which measured the distance between his two hands. With this feature Waisvisz converted all his discrete buttons in continuous control features:

> This is one of the more interesting possibilities of control with the hands. Continuous controllable timbral changes become possible ... makes them expressive under the "scratch" control ... By moving one's arms while having keys pressed one "bows" the sounds. ... Actually most MIDI

> systems stall when the "scratch" function is used excessively; "MIDI buffer overflow" messages were easily obtained on the DX7 (Waisvisz, 1985).

This passage reveals the subversive use of the MIDI system (and therefore of the Yamaha DX7 synthesizer) devised by Waisvisz, or at least, it reveals the inadequacy of the MIDI protocol under the conditions to which it was exposed with The Hands.

Wasivisz' work could be genealogically traced back to Theremin's, Buchla's, and Mathews and Moore's approach to interfaces. His 1985 paper begins by stating that the modern keyboard provided "rather poor translation of the rich information generated by hand/arm gestures and finger movements/pressures" (Waisvisz, 1985). His use of arm-hand gestures in the air link his interface to the theremin. However, unlike the theremin, he is not controlling one device that is hardwired or sonically fixed, but instead, one that is reconfigurable: the computer.

* * *

Sonus ex Machina

The *Sonus ex Machina* series by Philippe Manoury comprised the pieces *Jupiter* (1987), *Pluton* (1988), *Partition du Ciel et de l'Enfer* (1989), *Neptune* (1991), and *En Echo* (1993-4) all for (usually solo) instruments and live electronics. Developed at IRCAM, these pieces worked by using score followers[1] to track the player's position in the score, enabling the performer to trigger events and transformations of the live signal as specified in the score.

> The idea was to give the flutist the liberty of changing tempo, as against playing with a tape where the player has to follow the tempo of the tape.

---

[1] Score followers are computer programs that use the live signal to determine where in a score a performer is, by matching patterns in performance to patterns in the score. The score followers used in the Sonus ex Machina series were pitch based.

> ... the spectral evolutions of the electronic sounds were synchronized to the flute in pitch and timbral quality. My idea was to establish a connection between a soloist and a virtual orchestra but the orchestra would keep the characteristics of the soloist. Consequently, perhaps seventy percent of the sounds of the piece were flute sounds. ... we used frequency shifting, harmonizers, and other kinds of transformations. Manoury in Chadabe (1997)

Manoury and Puckette used the sounds of the flute as gestural inputs for discrete and continuous control of sounds; in this sense a strategy similar to Waisvisz'. The input was used on two levels. On a first level, discrete events (notes) were extracted from the continuous live signal to match the performance to the score. Whenever specified in the score, the performer was able to trigger events by playing a pitch at a particular moment in the score. On the second level, the input was subjected to live transformations retaining in the resulting sound the sonic and gestural content of the live sound, that is, continuous control.

The first barrier to seeing this work as an instrument is posed by the fact that the performer has his own instrument, in this case, the flute. This barrier is perpetuated even further by the label "flute and live electronics".

In the spirit of Mathews and Moore's conductor example, the flutist has his own line, which doesn't preclude him/her from also playing the computer, as in fact, he/she does. That is, playing the flute into a microphone is playing the computer. The flute player has the task of playing his line and playing the computer at the same time. The computer not only uses the live sound signal of the flute to trigger events, but it is actually transforming the live sound of the flute. Playing the flute into the microphone then is performing the computer. The interface for this instrument is the interface of the flute and its relation to the microphone.

A second barrier to seeing the computer music system as an instrument is posed by the fact that the computer changes functions constantly as specified in the score. Schnell and Battier (2002) saw, in *Jupiter*, an example of what they call a *composed instrument*:

> The term of the *composed instrument* underlines 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 (Schnell and Battier, 2002).

The idea of a *composed instrument* was already present in Mathews and Moore's conception of a score as an intermediate software layer in the GROOVE system. Even before that, studio compositions could be seen as composed instruments; that is, as instruments that change and transform during a piece according to a score.

* * *

Voyager

George Lewis, composer and trombonist, wrote the software "Voyager" between 1986 and 1988 at STEIM and later at IRCAM (Lewis, 2000) . This software 'listened' through a microphone to Lewis' trombone improvisation and reacted as an improvisation partner.

> First, the Voyager program is conceived as a set of 64 asynchronously operating single-voice MIDI-controlled "players," all generating music in real time. 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 ... Each new [computer] ensemble chooses not only a distinct group sonority, but a unique response to input, deciding which [human] improvisors - one, both or none - will influence its output behavior ... options include imitating, directly opposing or ignoring the information coming from the improvisors. With no built-in hierarchy of human leader/computer follower - no "veto" buttons, footpedals or physical cues - all communication between the system and the improvisor takes place sonically ... Voyager does not need to have real time human input to generate music (Lewis, 2000).

Although all communication between the system and the improvisor takes place sonically, the computer codifies, interprets and generates music at a symbolic (or at least numerical) level. The computer analyzes the live incoming sound signal through pitch followers, devices that “try to parse the sounds of acoustic instruments into MIDI data streams” (Lewis, 2000). The continuous sound signal is discretized into MIDI information, which means that the musical input is represented as pitch-velocity pairs at specific moments in time. Although Lewis does not use the continuous live input of the trombone signal, he works with gesture at a higher level. In this case, gesture does not refer to the direct mapping of the continuous flows in the performer’s actions and sounds, but to the articulation of these sounds in intervalic, rhythmic and dynamic terms.

This information is statistically analyzed over time and used to generate global parameters and data sets. These parameters and data sets are used to generate output MIDI information. The system processes input MIDI streams at the low level (MIDI raw data) and also by a “mid-level smoothing routine that uses this raw data to construct averages of pitch, velocity, probability of note activity and spacing between notes” (Lewis, 2000). The computer uses this information to decide how “each ensemble will respond to elements of the input, such as tempo (speed), probability of playing a note, the spacing between notes, melodic interval width, choice of primary pitch material (including a pitch set based on the last several notes received) octave range, microtonal transposition and volume” (Lewis, 2000).

It is interesting to note again how the author conceived of the system he programmed. On one hand, he talks about his system as a “computer partner” and many of his performances are understood as a duo; this construal of an “other” is reflected by the fact that there is one global listening strategy, one microphone and one global routine that specifies how many “ensembles” will play. On the other hand, he conceives Voyager to be an interactive “virtual improvising orchestra” made of “ensembles” and taking the model of the “Javanese gamelan ensemble”. So what instrument does this “other” play? Is this other an ensemble?

* * *

Inter-Agency

Can an instrument Listen?

Waisvisz' *Hands* have an incorporated microphone to capture sounds and transform them, or to capture and transform his live voice. Manoury's *Jupiter* listens to the flute to follow a score and to transform the sounds of the flute. Lewis' *Voyager* system listens to a live improviser to "parametrize" his recent activity. All of these "listening" strategies open the system to its surrounding environment, much like Mumma had accomplished in the 1960's. Beyond the fact that they use a microphone, the computer is listening for something, whether it is pitches or attacks, and this enables it to react to the environment.

As in the case of Jupiter, the performer or performers of Voyager have their own instrument that they use as input into the computer system. It is clear that control is not direct if we can speak of control at all, as in the Hands. A question arises, can we perform an instrument that exhibits independent behavior?

Wishart distinguishes *intrinsic* from *imposed* morphologies (Wishart and Emmerson, 1996). Intrinsic morphologies are those which the instrument exhibits when receiving an input; it is the part of the sound that corresponds to the instrument and is usually associated with the idea of resonance. Imposed morphologies correspond to systems where there is continuous input as in the flute, where the air stream of the player imposes a morphology on the continued sound.

"Most musical instruments have a stable intrinsic morphology. When energy is input in a steady stream or as an impulse, they produce a sound-object of the attack-resonance type" (Wishart and Emmerson, 1996). But not all instruments have a predictable, stable intrinsic morphology. Certain multiphonics in a trombone (and many wind instruments), windchimes, gongs (particularly when bowed) all exhibit unpredictable and sometimes semi-random behavior. In the electric domain, audio feedback (for example in an electric guitar) is an example of an unstable system.

In computer systems, Waisvisz (1985) proposed to use algorithms in The Hands that decided to "go wild" and ignore his control actions and Manoury and Puckette introduced Markov chains to create generative sound structures based on the input of the performer. It is therefore not entirely uncommon to find or design responses that exhibit independent or unpredictable behavior in response to human input.

The question then needs reformulation. Can an instrument exhibit independent behavior? Can it have agency over the music done with it?[2]. These questions confront us with a line of recurring and ongoing questioning on whether machines have agency, conscience, and so on, which is not within the scope of this dissertation. However, we can call into the discussion the concepts of distributed agency and inter-agency:

> Actions emerge out of complicated constellations that are made of a hybrid mix of agencies like people, machines, and programs and that are embedded in coherent frames of action. The analysis of these hybrid constellations is better done with a gradual concept of distributed agency than with the dual concept of human action and machine operation. The push on the button, the foot on the brake, and the click with the mouse trigger the activities between several agencies that more or less guide the machine ... [they] confront the user with unexpected offerings and assistance because the profiling programs have made the user into an object. The user of this type of advanced technologies is neither the master of the machine nor the slave of the technological system, neither the sovereign of his action nor the victim of media's manipulation, case by case. The wider concept of inter-agency replaces the narrow one of instrumental use and of the perversion of means and goals (Rammert, 2008).

Voyager confronts the user with unexpected offerings; the performer listens to the system and the system to the performer. In Rammert's view, instrumental use is narrow and is replaced by that of inter-agency, but this statement presupposes that instruments are strictly causal. Instead of extending the concept of instrument away from its traditional passive conception, as he does with machines, he retains that meaning in order to make his point. However, his complaints on traditional views that see machines as purely operational are the same as the traditional views that see instruments as strictly causal and passive.

---

[2] In this dissertation, the term agency will be taken to mean "a thing or person that acts to produce a particular result" (McKean, 2005). In this sense, when I say that an instrument has agency, I refer to the many ways in which an instrument determines the music made with it. A simple example would include the way the piano interface makes certain scales and intervals possible and others impossible.

In the continuous axis that runs from instrument to player paradigms proposed by Rowe to classify computer music systems, we can place Waisvisz's Hands closer to the instrument paradigm, Manoury's Sonus ex Machina somewhere in the middle and Lewis's Voyager closer to the player paradigm. A computer system can even shift paradigms during a performance or exhibit both paradigms at the same time. Both Rowe and Lewis, see these paradigms as ends in a continuum.

As we have seen throughout this dissertation, the design of musical instruments embeds concepts of music theory or technologies of musical thought. Scales, tunings, articulative possibilities, stable or variable timbre, all of these factors determine what music can be made with an instrument. For Rowe, Interactive Systems can be considered "applied music theory":

> Music theory, in its best form, is the scholarly attempt to describe the process of composing, or listening to music. Computer systems able to implement this work in real time allow the musician to assess the validity of the intellectual enterprise by hearing it function in live musical contexts ... The construction of formal processes is judged by hearing it function ... When a theory has been brought to the point of interactivity, it can be applied to the production and analysis of music in its native environment - that is, performed and experienced as live music has always been (Rowe, 1993).

One of the characteristics of the interactive systems described above is that they are highly personal or piece specific. To the extent that it was programmed by a person, and that Voyager is a computer program, it will have a signature behavior; it will be an applied music theory and its agency and unpredictability will emerge from it. Although they offer freedom and afford interaction to the performer, the programs for the Sonus ex Machina series can only play those pieces and they are composed in that sense. Waisvisz' controller, although probably considered the most expressive example of a computer music instrument was essentially only played by him and was not commercialized or replicated.

* * *

Max: A Second Generation of Music Programming Languages

The three systems presented previously shaped the field of interactive computer music. To achieve the interactions of Jupiter, Miller Puckette (1991) developed the prototype of a piece of software which he named *Max*, in honor of Max Mathews. To be able to debug and to provide Manoury with a more intuitive way of using the software, he created a graphical user interface called *Patcher*. The new piece of software did not use metaphors from traditional music, although as we will see later, Puckette understood Max as a musical instrument. Instead it borrowed the concept of Unit Generators from Mathews' MUSIC-N languages and used the terminology and concepts of analog synthesizers: modularity, patches and patch cords. Patcher reflected this metaphor in having boxes (unit generators or modules) that could be connected by lines (patchcords) which signified the flow of sound and control signals.

This new piece of software was fast adopted by musicians. For some musicians, like Peter Desain and Henkjan Honing, Max could not be considered a music system as it did not include "musical" features. This was part of a set of criticisms they made to the software and to which several composers and musicians, including Puckette, responded. R. Rowe and B. Garton edited the responses to these criticisms and published them under the title *The Mins of Max* (Rowe et al., 1993b). Below are excerpts of the criticism by Desain and Honing and excerpts of the responses by Dannenberg, Lippe, Settel and Puckette:

> Max is claimed to be a music system or musical language, but ... it does not provide primitives (e.g. notes, chords, ornaments) or control structures (e.g. repeat or slow down) that may be familiar or useful to musicians ... Even in the most unconventional music, organizational and structural aspects are essential. Max lacks them.
>
> Peter Desain and Henkjan Honing
>
> I find the claim that "Max is not a language for music" a feature. I'd rather search for and implement my own elements of musical structure than pick them from a menu. The danger of building in musical concepts is that they can preempt other ways of thinking and structuring.

> Roger Dannenberg
>
> Our belief is that any embedded “musical knowledge”or musical primitives may tend to bind us to pre-conceived notions about what someone else thinks music is or should be. One of the strong points of Max, from our point of view, is that it is general enough (and stupid enough) to allow us to create our own musical world, without any dependence on another’s definitions or preconceptions.
>
> Cort Lippe and Zack Settel
>
> One thing Max was not intended to do was operate on a musical level. In my own experience using “musically knowledgeable” tools, the tool always has an opinion about music that is different from mine. I would object if my piano started “helping” me interpret the music I was trying to play. Why would I want the computer to do differently? ... I think of Max as a musical instrument, not a musical assistant. ... does Max know about music? No. Does the shell? No. Should they? No.
>
> Miller Puckette
> (Rowe, 1993)

Indeed, Max did not embed “musical primitives” or “musical control structures” in its design. Its power resided precisely in an openness which contrasted the notions of “note”, “score” and “instrument” still present in the MUSIC-N languages. As the first in a second generation of music software that prioritized real time, it took what was powerful from the MUSIC-N predecessors and left out higher level “musical” features. Puckette thought of Max as a musical instrument. Although it used metaphors from analogue synthesizers, Max was a way of directly controlling the computer and therefore its main feature was its flexibility. Instead of using someone else’s software to control the computer, it allowed the user to make his own software. Reflecting Mathew’s unit generators, Max provided building blocks to construct flexible configurations. Although it offered several primitives or *internals*, these primitives did not reflect a theory of music; on the contrary they allowed each user to make their own “theory” by combining them with each other.

However, in the same as instruments have agency over the performer, the composer, and the composition, software too conditions our way of thinking. Certain pieces of software enable certain designs and though processes while harnessing others; certain designs are as ‘natural’ to a particular software, just as certain techniques are ‘idiomatic’ to a mechanical instrument.

Computers allowed the programmer/composer to add a sine generator or a filter to a program with code instead of soldering. In this sense, Max embodied the flexibility of the computer; it could re-mediate previous media like the theremin, the *musique concrète* studio, *elektronische musik* studio, a modular synthesizer, Mumma’s cybersonic consoles, Martirano’s Sal Mar Construction, and so on, if programmed to do so.

* * *

The “Message” of real time

Max, later *Max/MSP*, belongs to a group of real time programming languages such as Supercollider (McCartney, 1996) and Pure Data (Puckette, 1996).[3] The two latter programs are open source platforms and all of them have enabled the creation of a community of users.

As in the MUSIC-N Languages, the users were able to, in Risset (1987)’s words, study, replicate, and vary them. A *patch*, which is the name of a user-created program that is programmed and run in Max, borrows its name and paradigm from the action of “patching”, or connecting modular analog synthesizers with wires or *patchcords*. In this case however, like Mathews’ unit generators, the modules and the cords that connect them are “virtual”.[4] As virtual equipment, they are expressed in code and are therefore exchangeable.

That this software and these three pieces occur in the 1980’s is partially a consequence of the increase in processing power. These works reconcile the two procedures Mathews had devised, the real time-ness of GROOVE and the ability to fully describe sound through instructions of the MUSIC-N languages. The reason why both

---

[3] Also known as Pd, it is an open source software similar to Max, also designed by Puckette.

[4] Again, by virtual I mean to say that the modules or cords do not consist of physical equipment in space, but exist in code.

approaches, GROOVE and MUSIC-N, existed separately in the first place was computational constraints, but now, with increased power, sound synthesis could be attained live. Naturally, as in Waisvisz' "buffer overflow", this power was not nearly enough. While power increases, so do the needs for computation.

One of these new needs was that of live sound analysis or *listening*. With sensors and microphones for input, the computer was given several tasks. The first task is that of obtaining continuous functions of time that can be used to control sound generation processes. The second task, to obtain discrete pieces of information, like pitches or onsets, that can be operated on to follow scores or choose processes.

As in GROOVE, the computer was part of a system; of a larger environment. In contrast to Risset's characterization of the MUSIC-N languages as a "complete description of the sound structure" (Risset, 1987), the systems for these three pieces were contingent on input from the performer. Some sounds or parts of the sounds followed an internal "score" or form of automation, but its final realization remained open, awaiting input from the performer.

What these works did achieve was a complete description of the system. As Mathews and Guttman (1959) pointed out, one could "store not just notes themselves but instructions as well". The program for each of these pieces describes the rules of interaction using computer code.

Chapter 8

FIXED: Contemporary Fixed Practices and the Fixed-Live Continuum

While live computer practices became possible in the 1980's, many composers from that time on have insisted on working on fixed media or *tape music*. There are mainly two kinds of reasons for this.

The first subscribes and extends the Schaefferian ideas of reduced listening, acousmatic practices and sound-in-itself, whereby fixed music is ontologically different than live music practices (Chion, 1991), (Smalley, 1996), (Smalley, 2007).

The second is based on the believe that live practices still are not powerful enough or that the interfaces are inadequate to provide the composer the degree of refinement that the studio offers (Wishart and Emmerson, 1996), (Smalley, 2001).

New positions have however emerged where live and fixed practices are seen as sharing a similar aesthetic impulse and are analyzed in the same light (Landy, 2007).

As we have seen, many of the practices of live computer music have emerged from attempts at improving, and therefore remediating, the tape studios of the 1950's. Now we turn our eyes to see the way tape music adopted strategies from live practices opening the sound worlds to new spaces and meanings.

* * *

Post Schaefferian Aesthetics.

The *Traité des Objets Musicaux* (Schaeffer, 1966) was the first treatise to talk about composition with sounds, and the foundational document of *musique concrète*. While certain terms like sound object received a lot of attention and certainly became common parlance amongst composers of electronic music, the ideas for a compositional method did not seem to have a great impact. This lack of impact is reflected by the fact that there has not been much theoretical writing about compositional methods, using the terms the *Traité* proposed, or demonstrating the application of the compositional methods found in it. This can be due to many reasons, of which the first is that there is no english translation of the 1966 *Traité* and only recently we can find a translation of Chion's 1983 *Guide to Sound Objects* (Chion, 1995) which is still unpublished. A second reason is that it uses a terminology that is extremely personal and founded upon the french meanings of these terms, making its translation into other languages a harder problem.

It is a general problem that electronic and computer music practices have created new sound phenomena for which there is no generally accepted terminology. A later generation of, mainly British, composers and theorists such as Simon Emmerson, Leigh Landy, Denis Smalley and Trevor Wishart, have taken up the challenge of theorizing about electronic music. As in the case of Schaeffer, the need for an adequate vocabulary to talk about their music, has given rise to a great number and variety of terms. Since these terms form a kind of "private language", their writings are not immediately accessible and do not always share a common terminology.

Another characteristic of these writers and composers is that their main output of electronic music has been as "fixed" or "tape" music composed in the studio. This studio however usually comprises a real time computer and some composers, like Wishart, have written their own software. Furthermore, Smalley's and Emmerson's critique of "live" practices usually conceives of them as pieces for mechanical instrument and tape, or mechanical instrument and live electronics. For example, in a book dedicated exclusively to the practice of live performance of electronic music called *Living Electronic Music*, Simon Emmerson (2007) touches on electronic musical instruments, and even live electronics, only tangentially.

On the other hand, only after discussing in great length the aesthetics of several "fixed" or studio practices, Landy (2007)points out that:

> What holds all of the live areas together is the fact that they are slowly but surely offering the same opportunities for "chemistry" as Trevor Wishart likes to call it that traditionally have occurred as part of studio-based work. It is indeed the highly skilled improviser or new music technology performer who can mould with the same finesse as the studio artist. This being the case, another form of convergence is taking place.
>
> More important, the implication of Wishart's chemistry being applied in these contexts raises the question to what extent this subsection really needs to be kept separate from the first six[1]. The performance dimension is important, but the act of performance itself is not necessarily the distinguishing factor. The sound of studio and performance works is converging as our processors become more sophisticated and performance opportunities continue to grow. This is not necessarily a bad thing ((Landy, 2007), 160).

What Landy is essentially saying is that "live" performance is not different from "fixed" music in terms of aesthetics or language. This is probably one of the first times an electronic music theorist has ventured to say this; hence the disclaimer: "This is not necessarily a bad thing" (Landy, 2007).

* * *

Dynamic and Spectro- Morphologies

There is a term in the previous quote by Landy which is of great importance, it is the word "chemistry" (Wishart, 1993) (Wishart, 1994). This word was proposed

---

[1] The first six are all Fixed Media.

by Wishart to designate music that relies on molding instead of architecture. The term builds on a previous concept from 1985 which he called dynamic morphologies:

> An object will be said to have a dynamic morphology if all, or most, of its properties are in a state of change - I use the word properties rather than parameters here, because I feel at this stage that it is important to view sound-objects as totalities, or gestalts, with various properties, rather than as collections of parameters. The concept of a musical event as a concatenation of parameters arises directly from lattice-based musical thought4 and is singularly inappropriate to the musical structures we are about to discuss. In general, sound-objects with dynamic morphology can only be comprehended in the totality and the qualities of the processes of change will predominate in our perception over the nature of individual properties ((Wishart and Emmerson, 1996), 93-94).

While reaffirming the sound object, Wishart marks a significant break with Schaefferian thought. Instead of the Schaefferian permanence-variation, Wishart asks for continuous evolution of all properties; instead of using lattices (scales or "calibrations") to build "architectures", Wishart asks for processes in a state of continuous change. In contrast to Schaeffer's ideal of abstracting features of sound objects to build scales and pseudo-instruments that use them, we reach a case where there seems to be few instrumental metaphors. For Wishart, the archetypal instrument is utterance, the human voice: "voice, the immediate source of intellectual-physiological gesture... The keyboard on the other hand represents the ultimate rationalisation of a lattice-based view of music" ((Wishart and Emmerson, 1996), 29). This idea of dynamic morphology is also taken by Smalley in his term Spectromorphology (Smalley, 1986), (Smalley, 2001):

> I have developed the concepts and terminology of spectromorphology as tools for describing and analysing listening experience. The two parts of the term refer to the interaction between sound spectra (spectro-) and the ways they change and are shaped through time (-morphology). ... spectromorphology is concerned with motion and growth processes... (Smalley, 2001).

Although a bit more constrained than Wishart's definition (at least in the sense that it proposes a duality of spectra and shape over time), the idea of motion and growth processes also suggests continuous change. In both accounts, pitch is not a central aspect of their accounts. Both however, regard pitch as an extremely important element in music, which therefore, has to be dealt with carefully. Pitch is culturally engrained into our perception of music and our ideas about what music is to such an extent that it needs, on one hand, careful choice so that it does not dominate the discourses or narratives, and on the other, a link must be found with the familiar world of pitch:

> ... the principal problem facing music which focuses upon sound-objects of dymamic morphology - and here I am talking particularly of objects whose pitch is not stable - is how one achieves some kind of coherent connection with more traditional pitch architecture ((Wishart and Emmerson, 1996), 107).

Windsor (2001) sees pitch not only as able to create musical meaning by reference to other pitches (self-referentiality). He believes that wider conceptions of sound, such as sound objects, spectromorphologies and dynamic morphologies, allow us to explore pitch in an ecological view. In this view, pitch is seen as having greater potential for meaning when juxtaposed with, for example, environmental sounds.

* * *

Gesture

Smalley finds the origin of spectromorphology in gesture, which he defines as "an *energy - motion trajectory* which excites the sounding body, creating spectromorphological life" (Smalley, 2001). For Wishart, gesture is "the articulation of the continuum" ((Wishart and Emmerson, 1996), 17) and, as we have seen, the continuum is the essence of dynamic morphologies. Furthermore, for Wishart "musical gesture is evidenced in the internal morphology of sound-objects and also in the overall shaping of groups, phrases, etc" ((Wishart and Emmerson, 1996), 17). In other words, gesture is associated with human action and is not only reflected in the shape of individual sound objects, but also in their articulations over time.

For both authors, gesture is the source of life in sound and the listening experience is conditioned by our knowledge of sound sources and sound gestures. Both authors recognize the recognition of sources as a critical aspect of listening to acousmatic music. Instead of "bracketting" them out like Schaeffer did, Smalley proposes awareness to build on those expectations, while Wishart even proposes to use source recognition and meaning as part of the compositional strategy. This leads Smalley to talk about a spectromorphological referral process:

> We should not think of the gesture process only in the one direction of cause-source-spectromorphology, but also in reverse - spectromorphology-source-cause. When we hear spectromorphologies we detect the humanity behind them by deducing gestural activity, referring back through gesture to proprioceptive and psychological experience in general (Smalley, 2001).

These views resonate more with the ecological and embodied theories of perception that will be developed later and will be used to extend the concept of affordance. But what happens when the source of a sound is not recognizable? Smalley coined the term gestural surrogacy to refer to the process of increasing the remoteness of a sound. He envisions four orders of surrogacy:

> Original, *primal gesture*, on which sounding gesture is based, occurs outside music in all proprioceptive perception and its allied psychology. *First-order surrogacy* projects the primal level into sound, and is concerned with sonic object use in work and play prior to any 'instrumentalisation' or incorporation into a musical activity or structure.
>
> *Second-order surrogacy* is traditional instrumental gesture, a stage removed from the first order, where recognisable performance skill has been used to develop an extensive registral articulatory play.
>
> *Third-order surrogacy* is 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.
>
> *Remote surrogacy* is concerned with gestural vestiges. Source and cause become unknown and unknowable as any human action behind the sound disappears (Smalley, 2001).

Smalley and Wishart deal with this problem by appealing to imagined sources. Smalley also introduces the term source bonding to mean "the *natural* tendency to relate sounds to supposed sources and causes, and to relate sounds to each other because they appear to have shared or associated origins". For him music has intrinsic and extrinsic links. While his *Spectromorphology* approach is concerned with intrinsic features of music, and in this sense, he is more akin to the ideas of sound-in-itself, he recognizes that "Music is a cultural construct, and an extrinsic foundation in culture is necessary so that the intrinsic can have meaning. The intrinsic and extrinsic are interactive" (Smalley, 2001). For Wishart and Smalley, this extrinsic link might call upon non-musical or archetypal structures like bird flocks or stones falling; a kind of mapping of non-sounding phenomena into musical processes: "Quite often listeners are reminded of motion and growth processes outside music and the terms selected are intended to evoke these kinds of connections" (Smalley, 2001).

* * *

Space and Landscape

Wishart takes the concept further. He relies on the meanings and associations of sounds to build landscapes, which he defines as "virtual acoustical space" ((Wishart and Emmerson, 1996), 134)

> We may in effect break down our perception of landscape into three components which are not, however entirely independent from one another. They are:
> 1) the nature of the perceived acoustic space;
> 2) the disposition of sound-objects within the space;
> 3) the recognition of individual sound-objects
> ((Wishart and Emmerson, 1996), 140)

Wishart uses the recognizability of sound objects to create meaning by placing real recognizable objects in unusual or imagined spaces, or imagined or unusual objects in recognizable spaces; a process he calls the composition of landscape. In Wishart's landscapes, the concepts of aural-image and sound-image are of great import; they relate to our reliance on the use of context and of the "visual medium to assist in the recognition of objects and events". Wishart proposes the use of sound transformations to move from one image or landscape to the other. Finally, Wishart calls the concepts of metaphor and myth as organizational principles in a composition, using Lévi-Strauss's analysis of myths, which "moves from a description of the surface structure to an elucidation of the deep structure of the myths concerned" as a parallel to the way music is structured. For Wishart, meaning and source recognition is of the utmost importance: "We must use sound transformations and formal structures with both sonic impact and metaphorical import. We must be both sonically and metaphorically articulate" ((Wishart and Emmerson, 1996), 166).

This brings us to the idea of perceived space or virtual space. One of the most important aspects of electronic music is its ability to construct virtual spaces, that is, of placing the sound events of a piece inside a space. The sense of space was inevitably present in any recording, and, as Wishart notes, afforded a compositional use. This leads to a duality, the internal space of the piece, where the sound events take place, and the external space, which is the place where listening experience takes place. This is a fact that music has dealt with since its origins and lead to the concert hall as a sort of "bracketted" space. This duality is recognised by most authors and is usually called in several ways with multiple intentions as for example "composed vs listening spaces" by Smalley (Smalley, 1996), (Smalley, 2001), (Smalley, 2007), "local vs field" by Emmerson (2007), "internal vs external space" (Chion, 1991), amongst many others.

* * *

Texture

Traditional definitions of texture, oppose it to that of gesture in the same manner that shape is opposed to background. Space is construed in a similar manner. To this traditional view, belongs the following definition by Schaeffer:

> There are times when one sound is heard; there are times when many things are heard. Gesture is the name we can give to the unique event, the

> solo, the specific, the noticeable; texture is the generalized aggregate, the mottled effect, the imprecise anarchy of conflicting actions ((Schaeffer, 1966), 159).

Smalley however brings a new light in the role texture as a structural element in music:

> Gestural music, then, is governed by a sense of forward motion, of linearity, of narrativity. ... If gestures 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 on a human scale and events on a more worldly, environmental scale. At the same time there is a change of listening focus - the slower the directed, gestural impetus, the more the ear seeks to concentrate on inner details (in so far as they exist). A music which is primarily textural, then, concentrates on internal activity at the expense of forward impetus.

This idea of internal activity, of the ability to attend to micro-structural elements of music or "inner details" is a salient characteristic of electronic music. It bridges the connection between sound objects and the spaces they inhabit. While the idea of space connects us to the idea of place, texture situates us inside the sound. The use of texture as space is called texture-setting by Smalley, but his view on texture is not only as that of a background, but one which can be a driving force for musical discourse.

* * *

Aspirations in the Fixed-Live Continuum

It is one of the aims of this dissertation to establish that the practices of 'fixed' and 'live' electronic music share a same language and aesthetics. The practices developed in the studio in their early age as exemplified by Eimert and Schaeffer and in a later stage as characterized by Wishart and Smalley have construed a new kind of pseudo- musical instrument and performance practice. An instrument that is carefully constructed and of which we only have its recording.

Aspirations

Fixed electronic music is the set of practices that create music by fixing it on a medium and is usually associated with the work performed in a studio. It is composed and realized outside the time of performance and in this sense it is an out-of-time practice. This out-of-time-ness allows for a careful choice of materials and transformations. Since real time responses are not needed, processing power is in a sense infinite.

Electronic music became the idealization of the compositional process as the composer could bypass the performer and "transmit directly" to the listener, trade inexact interpretations for sound itself, escape the bounds of what is physically possible, and so on. However, although it represents the ideal compositional practice, fixed music aspires to the condition of "live": it looks to construct a sense of gesture and the resonance of physical spaces; it is played in concert halls; it seeks a sense of instrumentality, cohesiveness and unity.

Fixed music was established as a concert practice by playing back the final composition in specialized concert halls with multi-channel spatialization systems confronting the audience and the music itself with the duality of an internal and external space. While composers enjoy great freedom in the creation and organization of sound, audiences loose the experience of performance: they are removed from the sound generation process and left with its fixation. It is precisely the careful construction of gestures, spaces, pseudo-instruments and so on what creates in the audience a sense of being left only with the recording of something they did not get to experience live.

Live computer music, on the other hand, is the set of practices that use human input (gestures in the widest sense) to control computer sound organization, synthesis and processing. It is also composed and/or programmed (at least partially) outside of the time of performance; there is careful planning and rehearsal, but the piece is actually realized in the act of performance.

In live performance, sounds are produced in the presence of the audience, who sees the gestures and actions that create them. Sounds are in some way explained through their gestural sources. The computer performer becomes flexible and achieves the ability to play with other performers.

Although it achieves the condition of liveness, performance is constrained by computer processing power and the body limits of the performer. In this sense it aspires to an unconstrained ideal performance that can only be achieved through fixed techniques: to the "infinite" processing power of non-real time composition, to the multiple control layers achieved through automation, etc.

While in the strictly acoustic world all music could only exist through performance, computer music exists in the continuum that lies between fixed and live; between ideal sounds and structures made with unconstrained resources on one side, and the engagement of live performance as a cultural practice on the other.

Performance

Computer music instruments are devices which use sensors to measure human gestures and apply them to the control of sound synthesis and processing. Mechanical instruments are shaped according to the acoustic properties of the sound they make (tubes, strings, resonant boxes) and to accommodate our bodies (our mouths, hands, etc.); that is, they reflect the sounds they produce and our anatomy.

Computer music instruments do not need to accommodate a sound producing mechanism since sound is produced digitally, however, they do need to accommodate our bodies, that is, to allow us to input signals into the system. The interfaces of computer instruments therefore need to reflect our anatomy. These interfaces are called controllers and are separate of the sound producing mechanism, which is linked arbitrarily (or might we say compositionally) by a control or mapping layer.

In fixed electronic music, the performance action consists of pressing the button that starts the piece. Can we call this button an instrument? this action a performance?

Much of music's richness as an art lies in the relationship between human gesture and sound result; in the information a sound carries about the gesture that created it. The most well known paradigm for computer music instruments is the keyboard.

This interface however brings with it its cultural baggage: a tempered pitch scale, tonal repertoire and most crucially, a trigger representation of the piano keyboard. Essentially, the keyboard paradigm offers the model of discrete control: a series of buttons.

What is the other option?

Continuous control.

The world, its sounds and gestures are continuous in time and space. Even the hammering of a piano string is the result of a continuous movement. As we will see in chapter 10, theories of embodiment in cognitive science propose that perception is intimately linked to action. When performing, a violinist holds a note in a delicate equilibrium of movement and hearing; his gestures and his hearing are linked in one interrelated system. There are no buttons. In the act of performance, performers execute gestures in this constant balancing of expectation and sound result. Audiences share these expectations and they engage them in the actions of performance.

Computer music instruments need to achieve continuous control (Wessel, 2006). It not only allows the performer to achieve fine control over the sound he produces, but also engages the audience in the process of sound and music generation. But as we saw earlier, a performer cannot define every aspect of a sound and he certainly does not need to in order to achieve convincing musical performances. There is always a layer that is left to software - a score, an algorithm - certain parameters that are fixed. Computer music instruments also lie in the continuum between fixed and live.

The axis between the action of pressing a button and the chemistry of sound moulding suggested by (Wishart and Emmerson, 1996), is both practical and conceptual. In so far as the button reveals that the same sound is not heard in the same way twice, fixation is an illusion and to the extent that performance attempts to reach an ideal, its variability is an illusion as well.

Can Live music be Acousmatic?

The acousmatic condition is achieved when we hear something of which we cannot hear the source. As we have seen however, sounds usually have at least two sources: (1) a gesture, energy-motion trajectory or imposed morphology and (2) an object, resonant body or intrinsic morphology. Processing a recorded sound of which we cannot see the source can be thought of as a second source.

With electronic music instruments, a performance gesture may control a synthetic sound, the transformation of a recorded sound, or an algorithmic layer that controls either or combinations of them. This gesture becomes a source. It does not necessarily reveal how the sound is being computed, but it may reveal how that computation is being controlled. These gestures are equivalent to the gestures that are constructed in the studio, but they now become visible.

Is it still acousmatic? Well, to a certain degree. Although the source of the transformation and its nature is visually present in performance, the source of the sound that is being transformed or synthesized is not shown.

But why is this relevant? Once the source is recognized, reduced listening is not possible, as the sound cannot be "bracketted". Sound cannot be heard for itself, but the exploration of that processing reveals something about the sound. Landy (2007) points us to an insight by McCartney:

> [W]hen we listen to a processed real-world sound, and recognize it as such, we regard the composer as "doing" something to familiar material. Processing becomes an activity that guides, and changes, our previous understanding of the source; it offers an interpretation. . . . [I]n offering a new interpretation of something that, nevertheless, remains "known" from reality, real-world music invites us to deploy, and develop, "ordinary" listening skills; it encourages us to feel that we are involved, and participating, in the creation of a story about real life (McCartney, 1999).

This makes the composer-performer accountable.

Chapter 9

Total Configuration: Building an Environment

* * *

Percussion and the Setup

In the chapter *Instruments as Media*, I attempted to analyze the voice, monophonic, and polyphonic instruments as examples of instrument paradigms, but percussion was systematically left out. While percussion has been a central practice in several cultures, the Western world has not given it a central role in music until the XXth century.

Percussion practice extends and challenges some of our ideas of what mechanical, or pre-electronic, instruments are. Many percussion 'instruments' are inharmonic and un-pitched and in this sense, they contrast the traditional musical material drawn from tubes and strings. Moreover, these instruments are often drawn from other cultures or are found objects of many sorts. Handmade or unavailable as market products, these instruments are often not standardized like the flute or the piano, but unique objects appropriated by percussionists.

The problem of defining what is *the* percussion instrument, stems out of the fact that there is not a single object that can be called as such; there is not even a single kind of mallet. Aside from semi-standard configurations, like for example the drumset, the percussion instrument remains undefined:

> 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).

For Schick (2006), it is impossible to define percussion through the myriad physical objects available for the practice. It is also impossible to define the instrument through the reduced contemporary percussion repertoire and proposes the following: “The most succinct definition of percussion comes from the German, Schlagzeug; Schlag means ‘hit’ and Zeug means ‘stuff’.” The activity of percussion is then better defined by the gestural activity of percussion performance; by a gesture-object compound. However the problem of what the percussion instrument is, remains unsolved and perhaps problematized even further as just about every material object qualifies as “stuff”.

A common practice in percussion is the creation of a *setup*, which is a spatial arrangement of objects in space. The main criterion for these arrangements is that objects that need to be played in sequence are always within arms’ reach of the percussionist. In large setups, the performer is able to walk around and therefore the possible sounds are constrained by the place of the performer in the setup and the objects within his reach.

I propose that *the* instrument in percussion is variable and it varies with every setup. The instrument-as-setup is defined not by any of its objects but by the possible relations between them, making certain combinations easier or harder; possible or impossible.

* * *

The Electronic Setup as Total Configuration

Very much in the spirit of the percussion setup, we can speak of an electronic setup. The setup in this case refers to what Mumma (1967) called the *total configuration*, and from whom I quote again for convenience:

> I am concerned with "system-concepts": configurations which include sound sources, electronic modification circuitry, control or logic circuitry, playback apparatus (power amplifiers, loudspeakers, and the auditorium), and even social conditions beyond the confines of technology. I suggest that the most important creative aspect of live-performance electronic music is not this or that circuit innovation, but rather the total configuration itself (Mumma, 1967).

Let us begin to develop this idea, by considering the use of the *ring modulator* in the piece *Mantra* by composer Karlheinz Stockhausen, as an example.

Ring modulation is a technique used to modify the spectra of an incoming sound by multiplying it by a waveform, usually a sine wave. Ring modulators were quite popular amongst composers in the 1960's as circuit boxes that allowed to obtain inharmonic spectra out of common instruments. Ring modulators were emblematically used by Stockhausen for the live transformation of an orchestra in *Mixtur* (1964), choral voices in *Mikrophonie II* (1965) and two pianos in *Mantra* (1970).

Can we call the ring modulator an instrument?

Much like a woodblock or a cowbell, the ring modulator (as well as other devices like filters and oscillators) has a specific function. While in percussion objects the function is that of sound production, in the ring modulator it is signal processing. Nevertheless, in both cases the most interesting uses of these elements emerge from the ways they can be used in relation with other elements.

A ring modulator can be implemented as an analog circuit or as a software program that runs in a computer. In both cases, it is an element or unit that is positioned with respect to other elements or units. In the case of *Mantra* for example, a microphone is connected to the ring modulator which in turn is connected to a speaker; the piano is played into the microphone and the pianist controls the frequency of the modulating oscillator by rotating a knob. While simple, this configuration reflects a logic that operates at multiple levels of the system.

A level above the ring modulator (what is connected to and from the ring modulator) we find the piano, microphone and speaker. Going down a level (what is connected to make the ring modulator) we find the circuit or software program, and in them, several elements or units with the same property of connectivity or coupling[1].

The circuit or program that performs the ring modulation is placed between its input (sensors) and output (speakers). The knob, input and output jacks, frame the circuit. Inputs and outputs place the circuits and programs[2] in a relational position. What makes the circuit powerful is the flexibility of interconnection: any voltage can control any other voltage. What makes the computer even more powerful is that none of these circuits have to be soldered, but exist virtually.

In the example of *Mantra* the ring modulator has two kinds of input.

The first is an *audio signal* that undergoes the modulation and consists of the sounds of the piano as captured by the microphone. The second are *control signals* obtained from the knobs that regulate the amplitudes of the input and output signals and the frequency of the modulating signal.

At this point we can ask, what is *the* instrument in *Mantra*?

The sound of *Mantra* cannot be explained by the piano, the ring modulator or any of their components alone. This sound is the convergence of the sound of the piano and its ring modulated version. This sound can only be explained by the configuration of all of these elements, the position of each knob, of the microphone and the speakers. It is the total configuration of the system what can ultimately be called *the* instrument.

* * *

Input and Control Signals

Let us reconsider the global diagram for the GROOVE system presented in Fig-

---

[1] These units are equivalent to the unit generators in Mathews (1963) or to objects in Max or Pd. Each of the constituent elements of the ring modulator have inputs and outputs.

[2] As well as the unit generators that are used to build them.

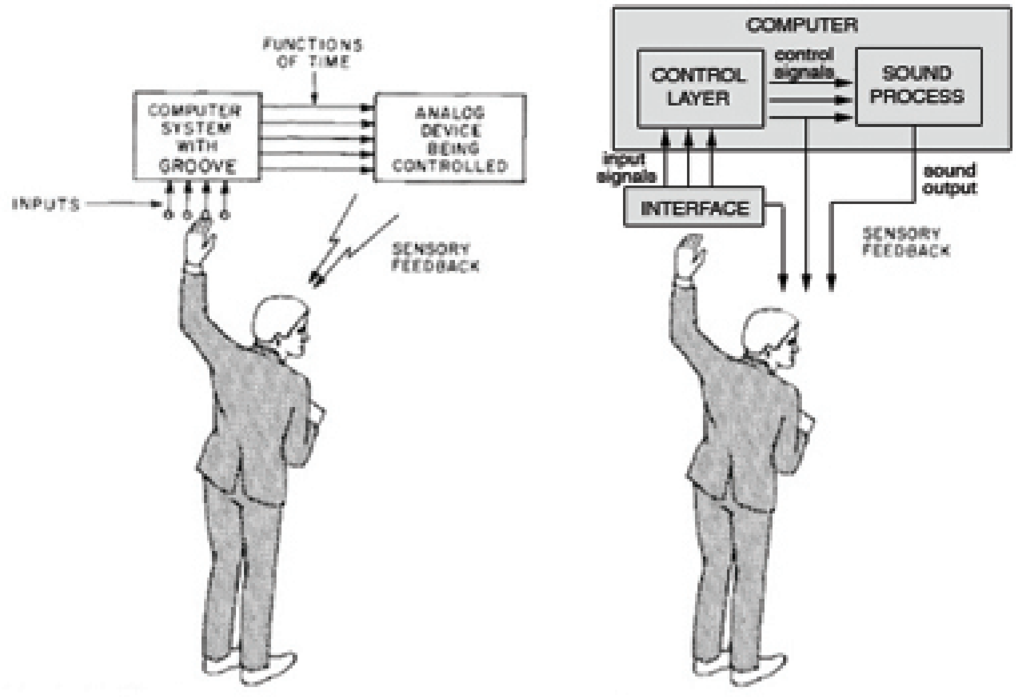

Figure 9.1: Updating the GROOVE Model

ure 6.2, this time, element by element to find how the diverse systems presented in this dissertation contribute to extend it. As shown in Figure 9.1, instead of using analog devices for sound production, modern interactive music systems or computer-based musical instruments generally accomplish both the control layer and sound processing within one computer.

In this chapter we will focus primarily on the design of the control layer. The role of this layer is to convert *input signals* that are produced by the manipulation of an interface into *control signals* that govern sound processes. In the following chapter, we will focus on how the *interface - control layer - sound processing* configuration is explored and learnt by a performer, as understood by theories of embodiment in cognitive science.

Input / Interface

We assume that the main output of a circuit or program in a musical system is a voltage or voltages that generate sound by controlling the movement of a speaker[3] or the movement of actuators or robots that will excite vibrating objects. Speakers are the most common solution as they can produce any waveform provided the right voltages. A musical system then is designed as a means of outputting sound to a performer and an audience.

Considerable attention has been given to designing interfaces. An interface is a way of coupling the performer[4] with the system through the use of sensors. Sensors are devices that produce voltages that are analog to a feature in the physical world, like air pressure, light or capacitance. A microphone is also a sensor, but because the ultimate aim of a musical system is to produce sound, the signals of a microphone ar already sound and are considered separately. The main function of an interface then, is to transduce features of the physical world, such as gestures, into *input signals*.

Input signals into a program can be of several kinds. One one hand, input can be a sound or *audio signal* captured with a microphone, or it might be a gesture (or another non-sound source) or *control signal* captured with a different kind of sensor. On the other, control signals can be either *continuous* or *discrete*. Figure 9.2 reflects this simple input classification system.

A system might interface with the world in multiple ways. For example, the flute used in *Jupiter* (1987) by Philippe Manoury and Miller Puckette, originally included a MIDI system. Because of this, it was able to input sound signals (from which continuous control signals were derived), but also discrete information such as the keys pressed. In contrast, a traditional keyboard mainly produces on/off triggers, that is, discrete and non-sound input signals; the capacitance sensors of a theremin produce voltages, that is, continuous non-sound input signals. Finally, a system can include in its "total configuration" several interfaces and receive multiple kinds of input.

---

[3] A voltage-controlled cone after all.

[4] Or other elements of the physical world.

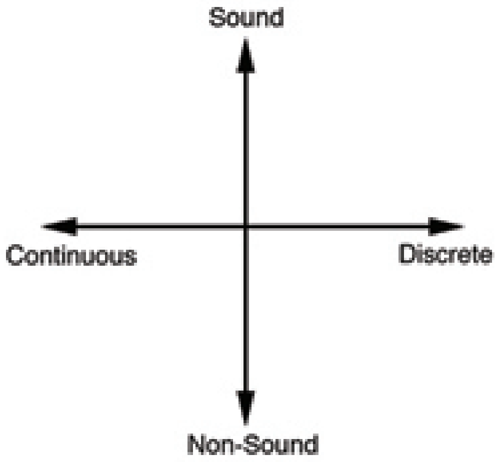

Figure 9.2: Sound / Non-Sound and Continuous / Discrete Input Signals

* * *

The Control Layer

Sound Process

As specified above, the role of the control layer is to transform *input signals* produced by the manipulation of an interface into *control signals* that govern sound processes.

Figure 9.3 shows a sound process model. Thick black arrows represent the flow of audio signals and hollow white arrows the flow of control signals. A sound process is controlled by *sound-control parameters*. The task of the control layer is to produce adequate control signals that control these parameters.

Sound processes can be of different kinds. A live sound source might be modified, pre-recorded sounds might be transformed live or the sound can be created through a variety of generators in live performance.

Analysis

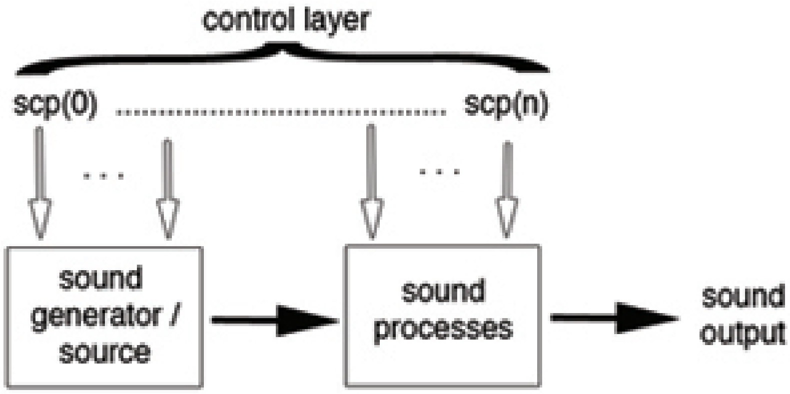

Figure 9.3: Sound Processes and *sound-control parameters (scp)*

One of the functions of the control layer is to analyze incoming input signals to derive control signals. As seen in the preceding chapters, analysis processes can be used to derive control signals from sound or non-sound input signals. When analysis is performed in sound or audio signals, it is called a machine or computer listening.

The primary role of the analysis process within the control layer is to extract information from the input signals that is used to produce and organize control signals.

An audio signal can contain any sound. Composers design listening strategies to recognize individual sounds and obtain features from them. Listening strategies can then provide the system with discrete information like pitches or attacks or with continuous control signals like those obtained from envelope followers.

Non-sound input signals can also be analyzed to produce discrete and continous control signals. As with audio, a signal or group of signals can be analyzed to extract features. For example, in my own work (Oliver and Jenkins, 2008) and (Oliver, 2010), a video signal is analyzed to obtain both discrete and continuous information.

As interfaces grow in complexity, *multidimensional* and *interdependent* control signals are becoming available. This means that there is richer information to be analyzed, but it also means that input is already rich before reaching the control layer.

Intrinsic Morphology and the Preset Paradigm

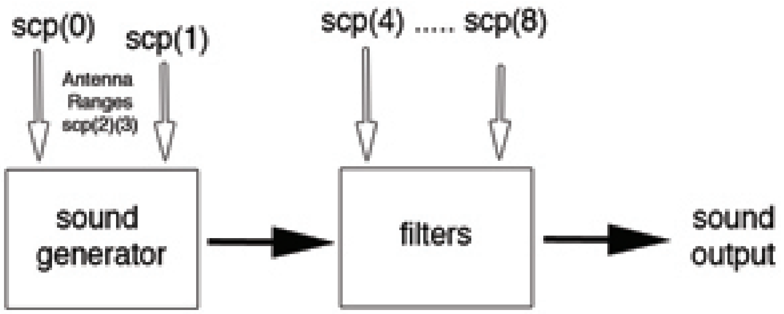

Figure 9.4: A Control Model of the Theremin

The simplest kind of control layer does not have an analysis process. The reason the theremin is such an archetypical electric musical instrument is because it has the simplest configuration that can resemble a mechanical instrument. Figure 9.4 shows a control model of the theremin.

The interface of the theremin consists of two electro-magnetic fields that output two signals that are analogue to the capacitance in these fields. Two potentiometers are used adjust the range[5] of the antennae and five switches used to turn on and off five filters that modified the 'timbre' of the instrument[6].

This system then had 9 input signals, but it also had 9 sound-control parameters. The five switches are discrete/non-sound input signals and control scp 4 through 8 which turn each filter on and off. The two potentiometers are continuous/non-sound input signals and control scp 2 and 3, which control the range of the antennae. The antennae produce continuous/non-sound input signals and control scp 0 and 1, which control the frequency and amplitude of the sound generator.

As we can see, for every sound-control parameter there is an input signal. Some sound-control parameters, like the ones that adjust the range of the voltage outputted by the antennae to the range of the sound control parameters are constant, and therefore

---

[5] This range refers to the size of the electromagnetic field.

[6] Ranges could be adjusted with "variable resistance" and "timbres" were chosen through the use of "stops". "These effects were achieved by using the inherent irregularities in the response curve of the loudspeaker to accentuate different overtones. Filters for this purpose were turned on and off with five switches located on the side of the instrument" (Glinsky, 2000).

hidden and fixed in the instrument.

The control layer is therefore extremely simple. It consists exclusively of range adjustment and on/off switching, but these parameters are set in advance and only the control of frequency and amplitude is performed live. When controlling frequency and amplitude live, one with each hand, the performer is then literally out of hands. Part of the power of the theremin, was due to the fact that the performer was felt as being in *direct control* over the sound.

To be able to attain *direct control* over the sound, certain variables must be set in advance. This is akin to what Trevor Wishart calls *intrinsic morphology*, that is, the features of the sound that correspond to the instrument[7]. For example, the shape and material of a tube, the position of the wholes, its tuning, are all set in advance and determine "the features of the sound that corresponds to the instrument".

The *Minimoog* and probably all of the commercial keyboard synthesizers follow this *preset paradigm*. In other words, a preset is loaded and then the performer plays. This model remediates the pipe organ although in the pipe organ console, there are stops in between keyboards to achieve fast changes as a means to gradually change timbre and amplitude. The *Minimoog* offered much more configurability than instruments like the theremin, with as many as 14 switches and 27 selectors and potentiometers that determined the sound that the instrument would produce as the keys were pressed and depressed. However, instead of controlling the contour of each note with a continuous gesture, the *Minimoog* used two ADSR envelope generators to automate the shapes of a filter and "loudness contour".

This *preset paradigm* however, is not exclusive to commercial instruments. For example, in *Mantra* the frequencies of the ring modulators are set in advance and the performer has the occasional task of changing them in between, and exceptionally in parallel to, their piano playing. The input signals for a pianist are first, two sound-control

---

[7] "Most musical instruments have a stable intrinsic morphology. When energy is input in a steady stream or as an impulse, they produce a sound-object of the attack-resonance type" (Wishart and Emmerson, 1996).

parameters - that determine the frequency of the ring modulator and the amplitude of the processed sound - and second, the actual piano playing. Because the system is geared towards producing sound, audio signals are commonly conceived as sound sources and control signals as fixed or constant values.

Extending the Preset Paradigm

Let us now consider a system with a larger amount of variables. For example, a system that produces sound with 30 oscillators. In this case, we have at least 30 frequencies, 30 amplitudes, and 30 phases: that is, 90 parameters to control. If we add a global amplitude for the sum of all or of groups of oscillators, and any other additions, we come to realize that a simple system can become cumbersome to control as the number of parameters increases.

Re-visiting Eimert's opinion about control in the early elektronische musik studios, where only one sine generator was available and tape had to be used as a buffer, we are forced to agree with him about the fact that "the number of individual sound elements is so great that any attempt to find means of instrumental realisation is doomed to failure" (Eimert, 1957). Even when these oscillators are virtual, producing individual input signals to control each of the parameters independently becomes very difficult.

To be able to address a large number of parameters, a bigger control layer is necessary. As with GROOVE, "complex control signals" can be generated from "the sums, differences, products, etc., of several simple functions of time" (Mathews and Moore, 1970). Instead of the theremin's immediate mapping, a live input signal could be combined with other live or stored signals, but also with various other functions, thus opening the possibility of *interdependent* control signals.

For example, if we were to use the theremin antennae to control this system, we would have two input signals (I_a and I_b) to control 30 frequencies (f_n) and 30 amplitudes[8] (a_n). We could for example have control signals that are multiples of an input signal so that for oscillators O_0 through O_{29}, $f_n = I_a * n$ and $a_n = I_b * (1 - n/29)$. This

---

[8] Let us assume a constant phase of zero for all oscillators.

would generate increasing frequencies and decreasing amplitudes as the index *n* of the oscillators increases. In this case, each input is still in separate control of frequencies and amplitudes.

However, we could also think of a model where $f_n = I_a * (n * (1 - n/29)) + I_b * (n * (n/29))$ and $a_n = |I_b - I_a|$. In this case, because both control signals are combinations of both input signals, these control signals are interdependent and the results are less predictable and better learnt through experimentation than mental calculation.

This kind of trivial mathematical operations were amongst the many important (and certainly non-trivial) features of the GROOVE system. With a layer like this, a small number of input signals can produce a large number of control signals; control is still direct, but complexity emerges:

> Complexity derives from a multivariable system of sub-systems which are affecting and changing each other through their various relationships (Chadabe, 1977).

Zooming and Automation

The *preset paradigm* as presented above, is an attempt to remediate the traditional musical instrument in several ways by alluding to *direct control*.

By pre-setting values to processes, instruments like the theremin, the *Minimoog* and the configuration of *Mantra*, attempted to provide an ‘intrinsic morphology’. As Mathews and Moore (1970) remind us, “the computer performer should not attempt to define the entire sound in real time”. Because of this fact, some features of an instrument are automated (or intrinsic) and others are left for the performer to control (or imposed). In this way, when a pianist presses a key, one can metaphorically say that the key ‘automatically’ displaces a hammer that sets a string in vibration. In a mechanical instrument like the piano, the string is set in vibration by a physical force, but with computers these behaviors must instead be automated. The *preset paradigm* is *automation-as-intrinsic-morphology*.

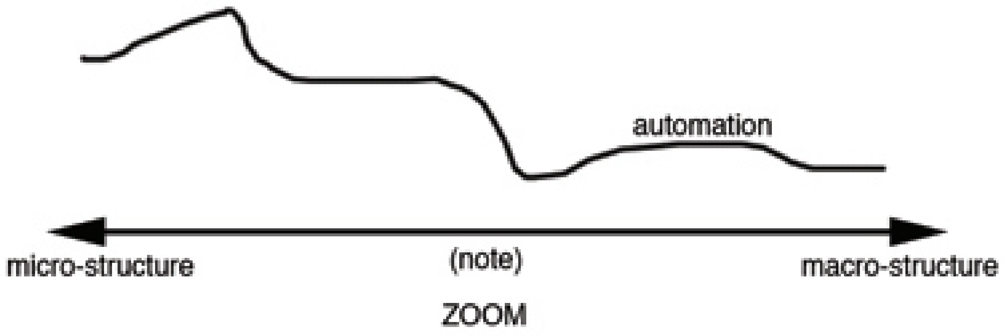

Figure 9.5: Structural Control and Automation

The theremin and the *Minimoog* present us with the illusion of a stable timbre over which pitch and amplitude are articulated. The theremin offers the performer the ability to control the shape and evolution of each note, but the *Minimoog*, because of the use of the keyboard, sacrifices continuous control for predictable, repeatable, automated notes.

From the analysis of mechanical musical instruments as media, the Western concept of musical instrument, is that of a fixed timbre over which pitch, amplitude and (therefore) duration are articulated. Its conceptual correlate, the note, is conceived as an abstract set of attributes that consist of pitch, duration and amplitude over a pre-defined timbre.

In a piano for example, the performer chooses a key (that results in a pitch) and controls how hard or soft it is played, thereby determining the amplitude. The pianist can hold the note for a duration shorter than the time the string can vibrate on its own or wait for it to end.

In figure 9.5, we see that in electronic and computer based musical instruments, the note must have some sort of intrinsic morphology or automation. However, we also see that control does not need to be at the note level exclusively.

Martirano (1989) reminds us that "a pianist remains at one level only, that is, on the keys." In this context, composers' choices are about which key or keys should

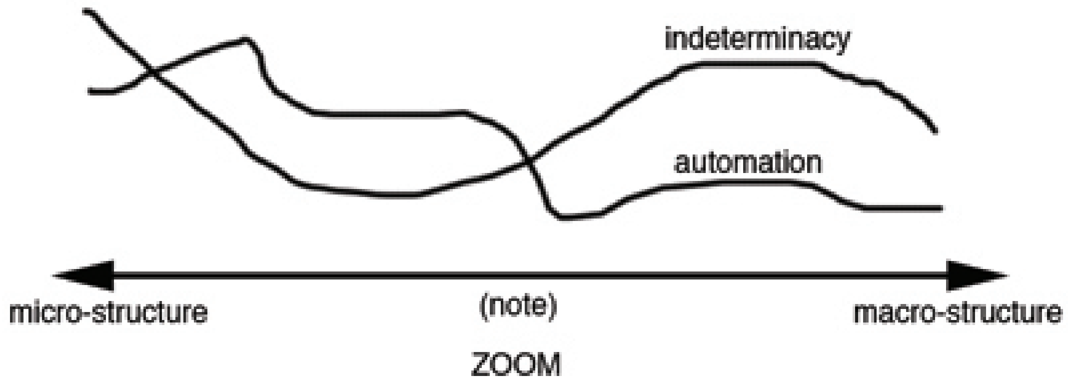

Figure 9.6: Structural Control, Automation and Indeterminacy

be played, when and how loud. Martirano described the SMC as “based on the idea of ‘zoomable’ control - being able to apply the same controls at any level, from the micro-structure of individual timbres to the macro-structure of an entire musical composition” (Walker et al., 1992).

Electronic and computer music instruments offer the opportunity of controlling aspects of sound other than pitch and amplitude. However, in order to control certain aspects of a sound, certain other aspects need to be fixed or automated. In other words, in order to control aspects of the macro-structure (phrase), the structure (note) and micro-structure (inside the note) need to be automated; to control some elements of the micro-structure, others must or might be automated.

Designing a musical instrument as a part of the compositional process then, implies designing what is the control space of the performer and what is the nature of the instrument; what is fixed and what is open for variation.

Automation, Indeterminacy and Process

Not all notes are the same.

Materials gradually degrade over time, no action is perfectly repeatable, ambient conditions change, and so on. No action or sound - even notes as conceived in mechanical instruments - is perfectly repeatable over time. Figure 9.6 attempts to reflect this fact as it opens a new dimension in our control layer.

In the theremin, we are presented with a state, we are always in direct control as with the voice, a flute or violin, and as such, the small variations in each gesture over time are present in the sound. The *Minimoog* offers us a different model. Pressing and depressing keys sets envelope automations in motion. In this sense, control in the *Minimoog* is not as direct as in the theremin, because the performer is mainly in charge of triggering the start and end of notes, closer to the piano and percussion. Gesture in the case of the *Minimoog* is about the way notes are articulated instead of the construction of notes themselves.

When Mathews and Moore (1970) remind us that “the computer performer should not attempt to define the entire sound in real time”, they have in mind the conductor-orchestra metaphor: “The conductor does not personally play every note in the score; instead he influences (hopefully controls) the way in which the instrumentalists play the notes”. In this case, control is achieved at the macro-structure level, that is, on how organized notes are executed.

But already in 1959, as Mathews conceived of the MUSIC-N languages, he devised ways of altering the organizations of notes themselves:

> ... with a few instructions, the composer may modify a much larger set of notes and, in mid-composition, alter or replace stored waveforms and parameters ... Perhaps ideally, the composer will arrange the structure of the composition, and the computer will construct the details (Mathews and Guttman, 1959).

In Mathews’ view, the ‘instruction’ reveals a larger potential role for the computer: the ability to modify not only the organization of notes, but their ‘timbre’. In contrast to the emphpreset paradigm, the system can modify some of its own parameters. Presets and automation are therefore set at different levels in the “zoom” or structural axis of Figure 9.6.

A system of this sort moves further down in the indeterminacy axis. If the computer is able to “alter or replace stored waveforms and parameters” and “construct the details", then the composer-performer does not know what the details of the output will be, but knows the space in which they can happen. The work of Buchla, Mumma or Martirano use feedback systems, and random generators and statistical processes, to modify themselves. Structural aspects of music, or in Martirano's words, aspects of the macro-structure are determined by these processes. The role of the composer - now to some extent also instrument designer and performer - is to 'compose' a process that 'composes'.

As Chadabe (1977) notes, "there is another dimension to complexity. It derives from a system's ability to reorganize. By 'reorganization', I mean a changing of the connections between sub-systems" . As the role of indeterminacy and automation increases, in the sense that results are not completely expected by the performer, the system begins to be perceived as an "other" that has independent behavior. The perceived coherence of this other's actions lies in the design of the system.

A system can begin to loose touch with a performer's actions or with a perceived coherence. In the extreme, the sense of performance or composition is lost and instead, we perceive complete indeterminacy.

* * *

Environment Design

When we say that an instrument is "a fixed timbre over which pitch, duration and amplitude are articulated", timbre is posited as the instrumental identity and the other features as the abstract values of music. In other words, the composer chooses from the available timbral palette, which timbre will a particular note sequence have. Instrument is then equated to timbre.

In electronic and computer based musical practice, it is common to equate timbre to waveform or spectra. However, since (Grey, 1975) and (Wessel, 1979), timbre is expanded to a larger, multivariate dimension of sound. In an expanded view of timbre, it is not only spectral content, but the "dynamic morphology" of a sound (to use Wishart's term), that is, the way the multiple dimensions of a sound vary over time that gives an instrument its identity.

This is precisely the role of the control layer (and therefore also of the gestures of the performer), to give the system a character that is independent of the sound sources used. In Chadabe's words:

> The music produced by these systems is identified more by the nature of its process than by its details. Since the details of process automation music will vary from performance to performance, one's attention shifts from the details for the process itself and to the system which produces the process. Designing the system thus comes into focus as the first act of composition (Chadabe, 1977).

Similarly, Lewis extends instrumental identity beyond the concept of timbre as spectra in reference to Voyager:

> ... although the chain can be said to possess a "timbre", timbre is not the only sonic cue available to a listener. Rather, aspects of how the keys are played, how the sound varies in its performance over time, seem to have an enormous impact upon one's perception. The same can be said of computer music instruments, where one computer program might perform with a given "timbre" (in the commercial MIDI sense of that term) in quite a different way from a person playing a keyboard, or even another program that operated differently (Lewis, 1999).

It is not that the actual sound-sources or sound-processes are unimportant, it is rather, that a sound-process can be controlled, and a sound-source can be transformed, in many different ways, and that these ways contribute to the instrumental identity. The composer may choose what interface and input signals to use, what aspects of sound will be controlled live, and what other aspects will be automated and with what degree of predictability. These design choices endow the system with invariant behavior. Schaeffer's concept of pseudo-instrument re-emerges with new meanings:

> Without the real instrument, present in the execution, the pieces offer "invariants" to the listener that work as pseudo-instruments as (apparent) permanence of a common character, while other elements vary ((Schaeffer, 1966), 44).

The composer / performer / luthier composes a system that is open to performance, and it is through performance that the instrument reveals its invariants. These kinds of instruments are proposed not as “objectively structured things and meaningful in themselves, things that invite analysis and can be perceived and experienced with different degrees of accuracy” (Adorno, 1976), but instead as environmental structures that invite to, and can only be perceived and experienced through, interaction.

Chapter 10

Embodiment, Media and the Audience

> ‘Listening to’ the composer/performer’s decision-making process as merged with his/her physical movements will become a crucial part of the audience’s musical experience.
> This is how one can *dig* the music.
> ‘The fingers are the brain; the brain is the body.’
> (Waisvisz, 1999).

As humans we seem to admire expert ability. In western culture arts, this has translated into words like mastery or virtuoso. To master something, to be a virtuoso at something, is taken to mean that while for some a task is practically impossible, it is very easy for someone who masters the task or tool used to achieve that task.

The term embodiment has, in recent times, been increasingly used to describe the nature of someone’s mastery of something. When someone masters something, we say that he embodies it. When something is embodied, we believe it to be beyond a process of thought, a direct and complete engagement of the person.

Theories of embodiment in cognitive science give us a framework to understand the process of embodying something. What we embody however is not a physical object in the world, but instead a media object. Nevertheless, mechanical instruments provide us with a reference point or benchmark to understand how the process of embodying computer-based instruments differs and overlaps.

* * *

Inside - Outside No. 1: Mind - Body or Action in Perception

A mantra in the introduction of papers and books that deal with embodiment is that of the mind-body problem. We attribute to Descartes a long standing belief in western thought that we can only reach the truth through reason (the mind / brain), for the body and the senses cannot be trusted and constitute an obstacle in our search for the truth. In this sense, the mind / brain complex is considered internal and the body (and the rest of the world) is considered external to the subject. These positions have been termed as internalist or neurocentric.

In Varela's terms, disembodied cognition is "the representation of a pre-given world by a pre-given mind" (Varela et al., 1992). In this traditional view of cognition, the senses are channels of sensation that transmit raw data to the brain/mind about a pre-given world and that it is the role of the mind/brain to build an internal representation of this world and process it as information, to achieve knowledge. This is the traditional view of cognition or thought. We will talk about the term pre-given in the next section. In this view, "perception and action are seen as separate from each other and as peripheral" (Hurley, 2001) and it is the mind which interfaces between them. At the core of the mind, we find cognition, forming a perception-cognition-action model which Hurley calls the *sandwich model*: "The mind is a kind of sandwich, and cognition is the filling" (Hurley, 2001).

An embodied approach on the other hand states that it is only through our body that we know the world; that our minds are embodied, and therefore our bodies cannot be considered external to us. Our minds are not embodied in the sense that they are "trapped" inside a body, but in the sense that our ability to act and receive feedback from the environment (to interact) is central to the way we acquire knowledge about the world.

Since the perceiver is active, his experience is shaped by his sensorimotor structure, which is, in Varela's terms, the manner in which the perceiver is embodied: "This structure ... determines how the perceiver can act and be modulated by environmental events” ((Varela et al., 1992) 173).

Nöe makes a similar point by saying that genuine perceptual experience depends not only on the character and quality of stimulation, but on our exercise of *sensorimotor knowledge* ((Nöe, 2005) 10), a term he defines as a “practical grasp of the way sensory stimulation varies as the perceiver moves” ((Nöe, 2005) 12). Nöe establishes a potentiality for cognition based on our sensorimotor skill, that is, the fact that potentially we can access the world by acting on or interacting with it.

Our bodies or sensorimotor structures not only “contain” and shape our minds, but also enable us to navigate and interact with the environment. We, as individuals, are mind-body indivisible entities; furthermore, the brain is just another organ.

* * *

Inside - Outside No. 2: Environment and Individual

One of Gibson’s crucial contributions was to shift attention to the environment as the place where perception is situated. “The environment consists of opportunities for perception, of available information, of potential stimuli. Not all opportunities are grasped, not all information is registered, not all stimuli excite receptors” ((Gibson, 1966) 23). For Gibson, an organism recognizes in the environment opportunities for action or *affordances*, which are dependent on the sensorimotor structure of the organism.

Furthermore, for Gibson, organisms and environmental niches fit together and coevolve, forming a tight perceptual *attunement*. This co-evolution to which Gibson refers is similar to what Varela calls *history of structural coupling*. This is the first account of the role of history in this paper and occurs in an evolutionary time-scale.

Varela expanded this idea by saying: “the environment is not pre-given and separately determined from the organism, but a reflection of the biology of the species” (Varela et al., 1992). Our sensorimotor structure is a reflection of the environment and the environment a reflection of our sensorimotor structure.

In light of these mutual reflections it is harder to establish the boundaries of inside and outside, of where we as individuals end and the environment begins.

* * *

Environment: Musical Instruments

Let us first return to the idea of affordances we previously defined. A monkey might perceive in a tree the affordance of being climbable, but a cow will not. In a similar manner, a luthier might recognize in the tail of a horse the affordance of friction or in certain woods that of resonance; a flutist might recognize in a bottle the ability to blow to produce sound, etc.

Gibson extended the environment so as to include technology in what he termed the "manmade environment" ((Gibson, 1966) 17) as well as a social environment: "If the response of one animal to another is considered a social response, then the stimulus from the other animal is a "social" stimulus..." ((Gibson, 1966) 23) calling a succession of stimuli between two animals a behavioral loop. These terms and concepts suggest the idea of culture and are interestingly applied to musical performance and music or technology in general[1].

Instrument builders or luthiers and instrument players or performers could be considered a behavioral loop. We can see the evolution of traditional western acoustic instruments like the double bass or the trumpet as a series of stimuli between luthiers and performers (and composers and listeners...) co-evolving in time as a cultural niche.

There is a widespread and somewhat romantic belief that the ultimate embodiment in musical performance happens when the instrument becomes part of the body of the performer. Many people have understood embodiment in musical performance as (or stated an agenda of artistic/research production in terms of) extending the body of the performer. This has also been reflected on the idea of cyborgs, or humans extended

---

[1] It is worth noting that this is precisely one of the main criticisms to Gibson, as he did not expand on the concepts of cultural and social environments.

by technology. This belief is not gratuitous. When we see a gifted performer play his instrument it seems as if there was no boundary between him and his instrument, as if the instrument had become part of his body. Instruments however, are objects from which we can detach, a fact that stands in contrast with any of our body parts.

Musical instruments could instead be considered as part of the cultural environment. They are built to fit a performer's sensorimotor structure or body. It is this structure that determines the size and weight of a double bass or trumpet, the thickness of the neck or the position of keys and valves, the range of pitches, etc.: it needs to fit the height and hand size of the performer's body and limbs, produce sounds in our audible range, etc. In short, musical instruments are "a reflection of the biology" of their performers (and composers and listeners...). Musical Instruments are tools, technology; they are cultural objects. Again, the line between inside and outside is diffused: instruments are reflections of our bodies and through technique and training, our bodies seem to reflect our instruments, but more on this later.

We not only recognize affordances in instruments, but over time and practice we attune to them, we become tightly coupled with them, giving us the ability of performing better. The term attunement means that animals become *directly* sensitive to certain features of the world or affordances - the term directly refers to the idea that there is no representation mediating between the performer and the instrument. In Varela's terms we could again speak of a history of coupling and in Gibson's of a co-evolution to refer to this process of attunement. Rowlands brings forth the concepts of history and proper function.

> The proper function of some mechanism, trait, state, or process is what it is supposed to do, what it has been designed to do, what it ought to do... The proper function of an item is determined not by the present characteristics or dispositions of that item, but by its history ((Rowlands, 2006) 127-8).

The importance of history, now in the scale of a lifetime and/or of cultural practices, is highlighted in all of these accounts. The process of attunement to a musical instrument is the result of a history of interaction. Be it the result of formal training and/or trial and error (a history in both cases), the proper function of our sensorimotor system for a particular instrument is what we call technique. It refers to the set of abilities we use when encountered with the mechanisms and physical shape of a particular instrument.

Embodying an instrument is to attune to it.

* * *

Representation

For Rowlands the problem of representation is the following:

> How is it possible[:] for one item to represent another? ... to possess another as its content? ... to mean another? ... to take another as its intentional object? ... to be about another? (Rowlands, 2006)

The concept of representation is usually associated with internal or mental representations. Traditional views of cognition state that that the brain builds an internal, detailed representation of the world, information that we then process. However, in embodied theories of cognition, this conception of representation is questioned since the environment is directly accessible though our actions, by moving our head, our hands: Why should we build a detailed representation of the environment in our minds when the environment is immediately there? Most cognitive scientists will agree that representation has a role in the cognitive process. In Nöe's words:

> It makes good evolutionary and engineering sense to off-load the representations ... The claim is not that there are no representations ... The claim rather is that the role of representations in perceptual theory needs to be reconsidered ((Nöe, 2005) 22).

Both Noe and Rowlands recognize explicitly that there is a need for internal representations. These become particularly relevant when the object of cognition is distal (spatially or temporally) or counterfactual, that is, when the environment is not immediately available or when dealing with hypothetical cases. However, when the environment is available, the need for representations is reduced, because we can access it directly.

* * *

Representing*: Representation in Action*

Rowlands however goes further in making a distinction between representation to the category of the word and to the category of the deed.

Representation to the category of the word is the way Rowlands characterizes our traditional conception of representation, where we think of "representations as items that sit in the mind/brain of a subject" ((Rowlands, 2006) 10). In this conception, representations are internal configurations of a subject about some content, which is "typically extrinsic to it". Like words, their meanings are established in a context and therefore mean something only in that context; they require interpretation. These representations are activated when we need to represent that content and therefore have a genuine duration and are passive.

Rowlands also proposes the conception of representation to the category of the deed, where instead of passive items sitting in the mind of a subject waiting to be activated, we think of "representations as something that a subject does or achieves", of representations as actions.

But not all actions are deeds; actions can be classified according to their intentionality. Intentional actions correspond to the *strict conception of action*, actions "caused ... by a prior intention, belief-desire couple, or trying on the part of the agent" ((Rowlands, 2006) 95). On the other extreme of the spectrum we find *sub-intentional acts*, which have "the property of not being intended" and of which we aren't aware ((Rowlands, 2006) 98). Finally, there are *pre-intentional acts* or *deeds*, which are situated somewhere in the middle of the spectrum between actions in the strict sense and sub-intentional acts. Deeds are defined by Rowlands in the following way:

> ... the act can be broken down into a succession of subacts... Each of these subacts consists in a succession of online, feedback-modulated adjustments, where incoming sensory information is correlated with the required motor response. And these subacts, although clearly being things we do or perform, do not fit into the strict conception of action outlined earlier.
>
> In this sort of cases, what we have is:
>
> (i) a general antecedent intention (to catch the ball, to play the piece, or some variant thereof), and
>
> (ii) an array of on-line, feedback-modulated adjustments that take place below the level of intention but, collectively, promote the satisfaction of the antecedent intention.
>
> Let us call the events identified in (ii) as deeds ((Rowlands, 2006) 103).

Rowlands' claim is that these pre-intentional acts or *deeds* are *representing*. There is a careful choice of words to avoid confusion. *Deeds* are used instead of the generic 'actions', and *representing* is used instead of 'representation' to reflect the fact that actions represent over time. In the context of this paper, whenever I say that actions are representational, I mean that *deeds* are *representing*.

An example continually used by Rowlands is that of a person attempting to catch a ball thrown at him. In this example, the content of the representation is the trajectory of the ball. The general antecedent intention is to catch the ball. The deeds are the many feedback-modulated adjustments in body-arm-hand-finger positioning that are done pre-intentionally in order to catch the ball. Our actions are shaped by the environment and therefore represent it.

Rowlands identifies a set of conditions for an action to count as representational:

- *Informational condition* - This condition refers to the fact that our actions carry information about the environment. In the previous example, the time-varying repositioning of our body, arm, hand and finger positions contain information about the trajectory of the ball.

- *Teleological condition* - This condition refers to the history of our gestures mentioned earlier, which we develop by training, imitation, trial and error, etc. and how this history endows gestures with a proper function. This proper function means that we perform certain gestures in the presence of particular environmental structures. That is, all the balls we have caught or seen someone catch, have given us a repertoire of gestures that correspond to particular ball trajectories and therefore, in virtue of that history, represent or are about them.

- *Decouplability condition* - Our actions are decouplable from the content they are representing when changes in our actions do not change that content. That is, the movement of our body-arm-hand-finger system does not change the trajectory of the ball.

- *Misrepresentation condition* - if our actions can represent their content, they must be able to misrepresent it as well. That is, when our actions misrepresent the ball trajectory, we fail to catch the ball.

- *Combinatorial condition* - actions are constituted by lexical elements or behavioral units, which are recombined in several ways depending on the variations a similar task presents.

The attempt made in embodiment theories to “off-load” representations is probably correct in terms of representations to the category of the word. This dismissal of the use of representations in perceptual and cognitive processes is usually presented in opposition to the fact that “we enact our perceptual experience; we act it out” (Nöe, 2005). In Varela’s words:

> ... instead of representing the world they enact a world ... ((Varela et al., 1992) 139) ... the world we cognize is not pre-given, but enacted throughout a history of structural coupling ((Varela et al., 1992) 198).

What Rowlands claims is that deeds are representational, that our actions are *about* the world, they contain information about it, that is, they represent it. We do not need to make detailed (internal) representations of the (external) world, because the world is directly available to us, we can access it through our actions, and our actions are representational; our *deeds are representing*. Interacting with the world is a representational act; the content of this representation is the world itself. So it is not necessary to establish representation and action as mutually exclusive, but instead, we need to make a categorical distinction between at least these two conceptions of representation: to the category of the word and to that of the deed. To enact a world, to act it out, is a representational act; our actions are representing - they just do not represent in the classically intended manner in which “words” represent.

* * *

Inside - Outside No. 3: The Content of our Representations...

There is also the problem of establishing if representations are internal. We have already established that our bodies are not external to ourselves, moreover, we cannot even consider a person as subdivided in mind and body: it is the person who thinks, not his brain; it is the person who acts, not his body. So where are the boundaries? Where does the subject end and the world begin? When we say that we are representing our environment, we are somehow stating that it is exterior to us. Since Gibson, we see ourselves as manipulating external information-bearing structures (the environment) to obtain information. We also manipulate internal information-bearing structures to obtain information, be it through thought/reason or by virtue of our sensorimotor knowledge. To obtain information about the environment, we can touch, move, think, and so forth.

Rowlands points to the fact that mental content has worldly constituents and is therefore unbound; it does not stop inside the mind/brain, but instead it extends out into the world. In this sense, cognition cannot be viewed as a strictly internal process for although it might happen in the mind, its content extend it onto the world, a world which he can still explore. Our actions, being representational also have worldly constituents and are too, unbound; they don’t stop inside our body.

This view of cognition by Rowlands, extends Hurley's ecumenical view of cognition[2] in the sense that mental and internal processes are brought back to the process of cognizing, but always in a dialectical relationship with the environment. In Balzano's words, "an alternative ecological view is to treat knowing as a process (rather than treating knowledge as a thing) and consider it a continuously ongoing transaction of participation and coordination between an organism and its environment" (Balzano and McCabe, 1986).

In the act of cognition, the 'dividing' line between internal and external is diffused:

> Cognition occurs when inner and outer are caught up in a complex and dynamic dance, a web of interdependencies and interrelations, and, consequently, there does not exist the separation of inside from outside necessary for what is on the inside to qualify as representational ((Rowlands, 2006) 44).

* * *

Feedback and Feed-Forward Models

> Operations around the feedback loop do not occur sequentially; rather, control requires all variables to be changing continuously and simultaneously (Hurley, 2001).

As it has to be clear by now, feedback plays a central role in all of these authors' accounts of embodied cognition. This concept has become central to the idea of the action-perception loop, standing in opposition to the input-process-output model of information processing characteristic of traditional cognitive science.

---

[2] Hurley's ecumenical view of cognition is her way to balance the weight that Gibson put on the environment. While she believes that ecological approaches to perception were right in shifting attention to the environment, she considers their dismissal of brain and mental processes as an overreaction to the weight that traditional views of cognition have put on internal processes. "The right response to Gibson is ecumenical: both movement through real environments by whole organisms and brain activity play essential roles in extracting information from the environment and enabling a creature to have a perceptual perspective" (Hurley, 2001).

The use of the term “feedback”, however, seems to have lost part of its meaning in much of the literature concerning computer instruments and in the abundant models of control theory. In many cases, feedback has become just an arrow going from output to input, keeping the traditional input-process-output model almost intact. This has probably happened because the use of the word feedback in electronics means that some output (or part of it) is literally fed back to the beginning of the process which led to it.

Gibson’s term for feedback was action-produced stimulation, although the cultural and scientific spirit of the era perpetuated the term feedback:

> Action-produced stimulation is obtained, not imposed - that is, obtained by the individual not imposed on him. It is intrinsic to the flow of activity, not extrinsic to it; dependent on it, not independent of it. Instead of entering the nervous system through receptors it re-enters. The input is not merely afferent, in the terminology of the neurologist, but re-afferent - that is, contingent upon efferent output. The favorite modern term for this action-produced input is one borrowed from electronic circuitry, namely, feedback ((Gibson, 1966) 31).

Gibson’s original term *action-produced stimulation* seems to mean more than just feedback. It implies that feedback is somehow obtained instead of encountered, which is something that the term feedback by itself does not mean. This seems particularly relevant when we think of, for example, a clarinet player, obtaining vibrations from his actions that will not only feed back into his ears, but will also make his lips, jaws, head and the whole of his body vibrate in resonance. The clarinetist doesn’t casually find these vibrations, he obtains them.

We find in Rowlands the idea of a person engaging in *on-line, feedback-modulated adjustments with respect to relevant structures in its environment*. Keeping in mind Rowlands’ definition of deeds, we can analyze the example of a clarinetist: the *general antecedent intention* he is attempting to fulfill is to play a particular phrase and the deeds are the *feedback-modulated* actions he executes towards achieving that intention. But these actions are modulated with reference to what?

Borrowing from control theory, Wessel suggests: “A *forward model* characterizes the relationship between the commands sent by the controller and the resulting sound output" (Wessel, 2006). Jordan and Rumelhart define a *forward model*[3] as: "...an internal model that produces a predicted sensation ... predicts the consequences of a given action in the context of a given state ... [it] can be learned by comparing predicted sensations to actual sensations" (Jordan and Rumelhart, 1992).

The concept of a *forward model* seems to be more in the spirit of an *action-produced stimulation* and particularly adequate to the example of the clarinetist. His general antecedent intention produces a set of actions of which he already predicts an outcome. This expected outcome is what the clarinetist predicts the sound will be in terms of timbral quality, amplitude and pitch (if it is intended) as well as the way it will resonate in his body, giving way to a succession of *feedback-modulated* actions as a result of the comparison between the perceived (feedback) and predicted (forward) outcomes.

It is tempting to understand the forward model in terms of a detailed internal representation. We culturally are internalists are after all and Jordan and Rumelhart's use of the word internal reinforce this idea. However, the idea of the forward model is precisely what Nöe calls *sensorimotor knowledge*, a "practical grasp of the way sensory stimulation varies as the perceiver moves" ((Nöe, 2005) 12); a practical grasp of how our instruments stimulate us in response to gestures: *action-produced stimulation*.

This relates to the concept of *efference copy* developed in motor theory. This concept is defined as:

> ... efferent or motor output signals are projected back to other processing areas of the brain, including to perceptual processing areas. Such processing areas may then receive both efference copy and reafference from the same movements. Efference copy has a smaller orbit or loop than does reafference. Together, efference copy and reafference make a record of movement available to the nervous system (Hurley, 2001).

---

[3] "The learner is assumed to be able to observe states, actions, and sensations and therefore model the mapping between actions and sensations. A forward model is an internal model that produces a predicted sensation y[n] based on the state x[n-1] and the action u[n-1]. That is, a forward model predicts the consequences of a given action in the context of a given state vector. As shown in figure 3, the forward model can be learned by comparing predicted sensations to actual sensations and using the resulting prediction error to adjust the parameters of the model" (Jordan and Rumelhart, 1992).

The concept of efference copy resonates strongly with that of a forward model and it is described in Hurley's ecumenical view of cognition. In this view, instead of one single feedback loop between action and perception which goes through the environment, she demands a complex dynamic network of feedback that allows for internal feedback loops as well:

> The validity of Gibson's insights about the important role of external feedback in perception is not compromised by the role of internal feedback. The external ecological feedback loop and the internal efference copy feedback loop can both have an essential role. They may play complementary and interacting roles within an overall theory (Hurley, 2001).

Our actions seek to obtain stimulation; our actions are forward models in constant renewal within a complex dynamic network of feedback loops.

* * *

So how does this all fit...

We have reviewed the key ideas of current theories of embodiment in cognitive science to conceptually situate gestures in musical performance. My claim is that the gestures we execute in musical performance are representational.

What are Gestures Representing?

Gestures, as actions, represent over time. On a first level, gestures represent the interface in which they occur and on a second level a sonic structure we want to perform.

As we have seen before, we acquire technique through a history of interaction. When we see a guitar player performing, his posture, his gestures are about the guitar, he is tightly attuned to it. Traditional piano technique learning includes practicing all possible scales, chords and arpeggios; the tradition of etudes directs the learner to explore multiple combinations of those basic exercises (teleological condition). When a piece needs to be performed, actions are drawn from this repertoire of gestures, from technique, and combined to achieve it (combinatorial condition). Gestures are grounded in the interface through which they are performed.

On a second level, gestures represent an intended sonic structure; the sounds we want to produce. What is the size of an intended sonic structure? Intended sonic structures could be notes, phrases, sections, movements, pieces, etc.: articulated sonic elements. They are therefore hierarchically organized so that we can extract certain gestural units or *gestural primitives* as defined by Choi (2000). We could conceptualize them as beginning with an intentional and continuing pre-intentional acts. Intended sonic structures can come from a composition or from improvisation. The performer has a general intention (to perform a sonic structure), which is then modulated by a series of pre-intentional acts or *deeds*.

My claim is that we are representing a sonic structure that we intend to execute on an instrument. Traditionally, musical ideas are conceived as products of the mind. At a first glance, this seems to fall into an internalist position: there is an idea inside our head that we want to exteriorize through performance. These conceptions can be traced to the Composer-Performer-Listener model.

Are intended sonic structures internal or external? Has the sound we intend to perform existed before as an environmental signal or is it an abstract idea? In many cases, what we intend to play might have been an environmental signal in the past: we have played it ourselves or seen someone else play it. But even when we have never seen anyone play it or played it ourselves, through the combinatorial property of behavior, we can identify lexical elements from past environmental signals that we can combine to reconstruct it. This entails of course that there is no intended sonic structure, no musical idea, that has not existed before: they can all be combinatorially re-constructed. We acquire sonic structures through our experience of the environment of the physical, animate and cultural environment; sounds do not come from the mind of the composer, because there are no immaterial, abstract sounds. Now more than ever, composers meet with performers or experiment themselves to see what is possible with the instrument and build an inventory of sounds. All sounds we imagine have a reference to the environment; they are combinations of our experience of environmental signals ((Rowlands, 2006)).

Intended sonic structures then have worldly constituents, they are reconstructions of a *body playing an instrument in the world.* As opposed to abstract musical ideas, a performer's intended sonic structures are always concrete: they imply their body, their instrument and a space. For a violinist, an intended sonic structure has a particular timbral and articulatory quality: it is a violin played by a human in a space; an intended sonic structure is translated into sensorimotor knowledge, into a forward model that 'predicts' the resonance in his ears, hands, arms, chin, etc.

So let us apply Rowlands' conditions for actions to count as representational, to musical gestures:

- *Informational condition* - gestures contain information about the sonic structure we intend to perform or are performing.

- *Teleological condition* - In traditional instruments, gestures are passed throughout a history of interaction (instruction, imitation, stimulus enhancement and trial and error) as part of a technique or tradition. This gives gestures a proper function, so that in the presence of a particular sonic structure that is to be performed, these gestures are recalled. When a pianist sees a particular melody in a particular key, fingerings come "automatically", as if he had played the piece before. In the case of new musical instruments these gestures are acquired primarily through trial and error or through what Wessel (2006) has termed *babbling*. Perhaps at this point we could talk about a technique for theremin, ondes martenot or electric guitar and its standard pedals.

- *Decouplability condition* - Can we claim that our gestures are decouplable from their content? If the content is, as I am arguing, an intended sonic structure, the actual performance does not change the original intended sonic structure.

- *Misrepresentation condition* - If we misrepresent we don't achieve the intended structure, but something else; this is what we usually consider an error, which could range from a wrong note to a little too much vibrato or being slightly out of tune.

- *Combinatorial condition* - A gesture is constituted by several lexical elements or behavioral units. These are recombined in different ways depending on the variations a similar task presents. Again, in trying to play a phrase in the piano, the phrase can be divided into several lexical elements, each of them with a proper function. So that melodic scales, trills, etc. are each gestural primitives, which are combined to form a larger unit.

We can therefore claim that performance gestures are representational.

Why is the fact that gestures are representational important?

Performed electronic sounds seem to be different (not necessarily better) than non-performed electronic sounds; they seem to be perceived differently by the audience. Some computer instruments have the ability to capture (I will not use *represent* yet) the performer's gestures in a rich way and remap them onto sound.

There is a long-standing concern in electronic music about the sources of a sound. While some claim that "a problem" with tape music is that we cannot recognize the source of a sound, which distracts us from hearing the actual music, others closer to the Schaefferian practice of reduced listening, will claim that sources distract us from hearing sound for itself. However, when we are presented with performed electronic music, we are allowed to relate a gesture to a sound and to other gestures; to assign a source to a sound *teleologically* instead of acoustically: we are not informed about what object caused the sound, but what gesture made it or transformed it. Few people, even percussionists and acousticians, fully understand the physics behind the sound of a tam tam, but they learn through experience, through interaction, and that seems to be enough.

The information a gesture contains about a sound has important anticipatory quality. A gesture informs us about the sound we are trying to obtain. Depending on the interface in which they are performed, gestures contain the dynamic, timbral, pitch and articulatory structures of a sound through time. This is the reason why tracking gestures is interesting in the first place. Gestures are the ‘sources’ of sounds; accessing gestures is in this sense accessing sounds.

Finally, what interests us here is to understand why certain performances and instruments are good; our provisional hypothesis is that a performer attunes to an instrument to such degree as to embody it. It is in this process of attuning that gestures become more perspicuous representations of intended sonic structures, we embody them and we embody the instrument. This leads us to think that to embody is to represent more accurately the sounds we intend to make and the interface in which we make them.

* * *

Tracking, Mediation, Computation

Tracking

> A stimulus may specify its source, but it is clearly not the same thing as its source ((Gibson, 1966) 28).

From the previous section, it follows that tracking gestures provides us with information about both the interface and the intended sonic structure. This fact is therefore key to any interface designer. The interfaces of traditional instruments need to accommodate both our sensorimotor structures and the acoustic properties and physical dimensions of the sound producing mechanism. In computer instruments, this sound-producing mechanism is left to increasingly smaller computers and in consequence, our interfaces just need to accommodate our hands; our bodies.

The increasing attention paid to physical computing in recent years has provided us with better, cheaper, and more diverse ways of sensing gestures. Still, there is no way of capturing everything that is involved in a gesture, much less in real time. Gestures are complex phenomena that occur in the whole body of the performer, not just in any one part; every muscle action, every thought, potentially belongs to the gesture, and therefore also to the sound produced. As mentioned earlier, a person cannot be subdivided into sub-personal systems. It is the person who plays the piano, not his fingers, or his hands, or his arms, or his brain, and so on. It could be argued however that certain parts or systems are more relevant than others, when it comes to, for example, a hand gesture. This might be true, but our body is completely interconnected, it is a tightly interrelated system. Posture reflects on gesture, gesture reflects on sounds.

It is this fact, sometimes overlooked, that determines how effective or rich is the tracking strategy and the resulting gestural space. Below are some implications.

Gestures are analog signals that are captured through sensors, digitized and thereby become the result of a computational process. What the computer 'sees' or receives is not our bodies, but the data that the sensors and analysis algorithms provide it with. This transformation of our body into data has been called the 'data body' by philosopher Sybille Krämer[4] (Krämer, 2008). What a sensor sees is not our gestures, but their representation in data. This data contains information about the gesture, but also information about the sensor. For example, a camera and any data obtained from analyzing its image gives us information about the environment it is capturing, but also of the lens, focus and the characteristics of the color/light chip it uses. In this sense, the gesture represents the interface and the intended sonic structure it attempts to control and so the data too represents these two contents.

What are the gestures a guitar or a piano 'sees'? When a performer plays a piano, the only points of contact are the fingers on the keys and the feet on the pedals. This does not mean that the rest of the body does not play a role in the performance. On the contrary, the gestures of the whole body, the arm, the hand and fingers, are reflected in the gesture of the finger with which the performer is actually pressing the key; this

---

[4] We will explore this concept a bit further in the next section. Suffice it to say for now that the body is translated into a data structure which evolves over time.

gesture is also placed sequentially in time, it comes from a previous body configuration and goes on to the next one. The whole of the body is in some sense compressed into the fingers and it is therefore reflected in sound. The same parallel can be made with breathing in singers and wind players.

The whole body, in all of its dimensionality is somehow partially contained in each of its parts, reflected in the one gesture we track; the gesture contains information about the whole of the body while the one aspect of a gesture contains information about the whole of the gesture. We can understand this partial reflection of the whole into its parts as a kind of *dimensionality reduction*.

The choice of a tracking strategy needs to take dimensionality reduction into account to determine the quantity, kind and relationship between sensors. In other words, depending on the aesthetic goals of a particular work, it might be futile or even counterproductive to track the posture of a pianist instead of his actions on the keys. Furthermore, in analyzing the data, it might be necessary to combine data from sensors to obtain a parameter.[5]

Gestures exist in space and time and therefore are continuous. A guitar player does not perform a discrete pluck, but a finger trajectory; even a staccato note in a clarinet is the result of a convergence of continuous movements. While discrete (on/off) controls such as buttons and keys with instant velocity, can be a good representation of concrete states, they are very poor representations of a gesture[6]. The trigger paradigm that reigned over computer instruments so many years is now challenged by embodied or enactive approaches to tracking. As Wessel puts it, “triggering and subsequently shaping events” (Wessel, 2006).

Continuous control is important in that it allows for deeds or pre-intentional acts to occur. Shaping in time is what is representational about gestures and is what allows us to attune to an instrument.

---

[5] Since the camera was mentioned it is important to keep in mind that the camera itself is a matrix of photosensitive sensors.

[6] “The modern keyboard provides a rather poor translation of the rich information generated by hand/arm gestures and finger movements/pressures” (Waisvisz, 1985).

Certain interfaces, like multi-slider controllers, allow us to trigger events with a button and subsequently shape them with the sliders. A more embodied approach could be to obtain discrete controls directly by extracting features from the analysis of continuous data. That is, instead of having separate actions to trigger and to shape, we can extract discrete information (triggers) from a continuous stream. Furthermore, these triggers can be of different kinds and have different functions.

Computational Mediation

The word mediation comes from medium or middle. As seen in the beginning of this dissertation, medium is understood by McLuhan (1963) as technology. Three inter-related terms are relevant to us: mediation, immediacy and media.

The concept of medium is commonly used in communication theory as the channel through which communication takes place. For Hollan and Stornetta (1992), “media are simply what mediates communication”. We will not delve into a discussion about communication and music, however, musical performance consists in a way of a performer communicating with an audience through sounds; a model that resonates with the traditional composer-performer-listener model of music presented in Figure 2.1.[7]

The term mediation alone has been used in fields like Marxist theory, material anthropology, statistics, computer science, to name only a few. It is generally conceived of as “reconciling mutually interdependent, opposed terms as what one could loosely call ‘an interpretation’ or ‘an understanding of”’: in a way, a ‘translation’. It is possible to conceive several layers of mediation in music. Mediation in music implies that there is something in between the performer and the listener. Hollan and Stornetta (1992) conceive of

By definition, sound exists in and travels through air in a space, through a physical environment. Variations in air quality have an effect on sound quality; open spaces

---

[7] Although this passive model of communication theory with sender-message-receiver will help us now to think of mediation processes, we should not think of both performances and communication as univocal sending of messages, but as complex social interactions where audiences relate to their environment in more active ways than just receiving sounds.

and rooms have an effect on sound too. We will refer to the effects of space on the reception of sound as *sound reception mediation*. The effects of this mediation alter the signal so that it now contains information about the space that mediated it.

But we can also talk about a second layer of mediation; about what mediates between our actions and the sound they produce. We will refer to this as *sound production mediation*.

Sound is a periodic variation in air pressure faster than around 20 or 30Hz.; the speed of our gestures is generally well below that threshold and therefore do not produce sound. When we bow a string in a cello to sustain a note, we mechanically set a physical process in motion; we excite a vibrating mechanism. The gesture we apply is a slow one, but the friction of the bow in combination with the tension of the string results in very fast vibrations; fast enough to be audible. In a sense, we are mechanically transferring the energy from our slow motion into faster quasi-periodic motions. The instrument then mediates between our gestures and sound, it translates them into sound. In this case sound production mediation is *mechanical*.

There is however one ‘instrument’ in which this is not entirely true: the voice. In the case of the voice, nothing appears to mediate between action and sound. It is the most direct channel through which we can produce musical sound. We seem to have to say that there is no *sound production mediation*, because our muscles produce sound directly. There is a certain amount of *sound reception mediation*, since our voice is emitted into space (and space, through air, extends inside our bodies) and we can hear it as feedback, but all of our body resonates unmediated with our utterance.

As we stated before, in computer instruments the gestural interface is separate from the sound-producing mechanism; we do not directly handle the sound vibrating mechanism. Instead of mechanically transferring the energy of our gestures, these are tracked with a sensor, and converted into data through the use of an algorithm; then in the digital domain, they are analyzed and mapped onto sound algorithms and outputted through loudspeakers back to the ‘analog world’. This mapping or control layer, is a key process as it determines what tracked gesture data (input signal) is transformed and assigned to what variables of sound control (control signal). Tracking here implies an analog to digital conversion (ADC), while outputting to loudspeakers requires the opposite, a digital to analog conversion (DAC). These conversions bound the computational processes of analysis, control layer, and sound-process computation. In this case, sound production mediation is *computational* instead of mechanical.

The use of the term mediation is twofold here. The energy transfer from gesture to sound is not mechanical anymore and therefore it is not immediate, but mediate. In a functional sense, mediacy as opposed to immediacy means that the transfer is not direct, that there is something in the middle, a process that "translates" one into the other: computation. In a temporal sense, it implies that latency is introduced.

However, what is crucial in computational mediation is the transformation of gestures from energy to information, originally devised by Norbert Wiener (1948). As introduced earlier, for Krämer the body is converted into the *data body* through its digitization with sensors. This *data body* is the body's transformation into data structures, which is what allows us to manipulate them and to interact with other data, as is the case with audio manipulation and synthesis. The *data body* becomes a temporal stream of information (Krämer, 2008).

* * *

Media and Remediation

Media

Notation has been the medium which enabled western music to survive time and traditional music disciplines (counterpoint, harmony, analysis) to exist. Notation belongs to what the German philosopher, Friedrich Kittler calls "textual media":

> Textual media record some things while leaving others unwritten. What is noted is analogous to the repeatable chain of signification in speech; what eludes transcription is equivalent to a singular sound - just as musical notation records intervals but not the tone of the instruments (Krämer, 2006).

Textual media notate only that which falls inside the *'grid of the symbolic'* (Kittler in Krämer (2006)), everything else falls through the grid and is not represented in it. For Kittler, *'textual media transform the linguistic-symbolic into an operable code'*. This is what allows musicians to transpose, reverse, invert and perform other combinatorial and permutational operations to notated music. It is also what has allowed composers to transmit directions and content to performers.

Kittler also distinguishes technological media, which, "by contrast, attempt to select, store, and produce the physical realities themselves" (Krämer, 2006). While textual media records symbolic representations of the real, technological media records the material world of the real or the real itself. For Kittler, technological media (such as phonography, photography and film) are a revolution equivalent to the invention of the alphabet and literacy, which "mark the beginning of a development that ends with digitization and the computer" (Krämer, 2006).

> Textual media transform the linguistic-symbolic into an operable code; technological media, by contrast, transform the contingency-based, material, real itself into a code that can be manipulated ... This type of manipulation creates the possibility of reversing temporally-sequenced events (Krämer, 2006).

The main characteristic of media is that it *spatializes* time (Krämer, 2006). For example, in a musical score, time is plotted on the page of the score and in a sound recording, the samples that make a sound are placed one after the other throughout the length of the magnetic tape. This allows us to access temporally discontiguous events, breaking or reversing the natural temporal sequence:

> Storing is not merely a means of preserving but is also intrinsically connected to spatial order. Wherever something is stored, a temporal process must be materialized as a spatial structure. Creating spatiality becomes the primary operation by which the two remaining functions of data processing - transporting and processing - become possible at all (Krämer, 2006).

For Kittler then, what defines media is its ability to manipulate time. To perform media is to acquire the ability of using time as a another variable for control:

> ... the explanation of the technological as a modality of time management is precisely the 'main point'. The most basic experience in human existence - and this is relevant because man is, after all, a physical being - is the irreversibility of the flow of time. Technology provides a means of channeling this irreversibility. In media technology, time itself becomes one of several variables that can be manipulated.
> ... Data processing becomes the process by which temporal order becomes moveable and reversible in the very experience of space (Krämer, 2006).

As we have seen in the previous section, for Krämer (2008), digitizing the body by tracking it with sensors, converts it into a stream of data, which is what allows the performer to interact with media; the digitization of the body is a method of translating it into a format in which it can interact with the data and signals of textual and technological media respectively. The data acquires a continuous nature through the process of sampling continuous voltgesat high frequency constant intervals, but this high frequency provides the illusion of immediacy.

What allows for our embodiment of media is that our interactions feel temporally immediate:

> While it is true that the computer writes and reads, it does so in a way that is invisible to the writing and reading human. The operative logic of technological media is comprised precisely in structuring streams of data in such a way as to pass under the radar of the 'time of human perception'. The so-called 'real time reactions' can only emerge as a consequence of skipping over human perception. Real time analysis does not exist. Every step in computer processing takes time, albeit a span of time that is less than the smallest unit of time that can still be captured by the human senses (Krämer, 2006).

The operation of the computer at units smaller than that which we can perceive is what creates a sense of immediacy. The fact that the environment can be recorded and stored allows us to interact with it. Media - as the storage, manipulation and production of the "contingency-based, material, real" - extends the natural environment. The high speed of operation of the computer assures us that we experience this extended media environment as immediate.

Media is also “an information bearing structure we can manipulate to obtain information[8] about the environment”. Furthermore, its spatialization allows us to manipulate the media environment through the added variable of time. The manipulation of (real time) computer music instruments is the exploration of an environment with which we can develop a history of interaction, and therefore, to which we can attune. Interactive computer music systems are information-bearing structures that afford manipulation. Perhaps it is time to re-quote Gibson:

> “The environment consists of opportunities for perception, of available information, of potential stimuli. Not all opportunities are grasped, not all information is registered, not all stimuli excite receptors” ((Gibson, 1966) 23).

Remediation

As in cognitive science, most of the media studies literature is oriented to the visual realm. McLuhan (1963) pointed us to the fact that “the content of a medium is always another medium. The content of writing is speech, just as the written word is the content of print, and print is the content of the telegraph”.Bolter defined this fact pointed by McLuhan as *remediation* (Bolter and Grusin, 1998).

As with the visual arts, music needs a medium to exist. Both of them need a spatial medium, either formalized as a gallery or concert hall; in more experimental spaces, light and air respectively. Just as traditional visual arts have existed in materials such as painting and sculpture, music has had the voice and mechanical instruments. With the advent of electric technology, media expanded to include photography and phonography, animation and sound synthesis, that is, what Kittler calls technological media. With modern computers, visual and audible media coexist in the same machine, leading to techniques genuine to the digital domain in both fields, and to hybrid fields that comprise both of them.[9]

---

[8] Rowlands’ definition of cognition.

[9] A fact that is allowed by the use of representations as numerical data.

For Bolter, remediation works through a double logic of hypermediation and transparent immediacy and aspires to the real. He defines these three terms as follows:

> "Hypermedia and transparent media are opposite manifestations of the same desire: the desire to get past the limits of representation and to achieve the real. They are not striving for the real in any metaphysical sense. Instead, the real is defined in terms of the viewer's experience; it is that which would evoke an immediate (and therefore authentic) emotional response. Transparent digital applications seek to get the real by bravely denying the fact of mediation; digital hypermedia seek the real by multiplying mediation so as to create a feeling of fullness, a satiety of experience, which can be taken as reality. Both of these moves are strategies of remediation" ((Bolter and Grusin, 1998), 53).

A large amount of analytical literature explores visual remediation; to what extent does photography remediate painting? film remediate theatre? and so on. It is interesting to ask ourselves how musical media work in the context of remediation. Could we think of pitched instruments as remediating the voice? at least when they first appeared? Could we think of computer musical instruments as remediating mechanical ones?

As we saw in the previous chapter, the control layers of several electronic and computer-based instruments and interactive systems are designed with the model of a stable intrinsic morphology, and in this sense, these electronic and computer systems remediate older mechanical instruments.

The Logic of Transparent Immediacy

> The curves of the physical gestures performed by my fingers were directly translated into analog control voltages which immediately moulded the sound synthesis into equivalent curves. The application of my efforts' tension pattern was introduced as a dominant parameter to shape the music.
> The audience of the day could only interpret this system in terms of playing mechanical sound objects such as traditional instruments and extended instrumental techniques.
> So when human curves were applied to electronics, it made those people believe that what they were listening to wasn't electronics at all.

> They used the information of those movements as a way to (mis)judge the timbral information. (Waisvisz, 1999)

It would be valid to say that when we embody an instrument, the instrument 'disappears' to give way to an experience. When a performer is attuned to his instrument, the instrument fades into the background, and the listeners (including the performer) are left with the music. This fading into the background is precisely what the logic of transparent immediacy seeks to achieve. In computer music instruments we need to assess two barriers to transparency: (1) the computer and (2) the interface.

Latency is a barrier to immediacy. It is inherent to computation and to sound itself. Sound is air vibrations, and vibration by itself implies time. Sound travels in air at a speed of about 340 meters or 1100 feet per second. A person in a concert hall, 10 meters or 32 feet away from the instrument, experiences a latency of 34 milliseconds. Computational latency, for many processes, is only recently low enough as to be experientially immediate.

A small number of studies and papers[10] have attempted to determine our intermodal latency toleration boundaries, between, for example, touch and sound, or between vision and sound. Other studies also determine that our latency toleration varies depending on the mapping and that audience and performers develop anticipatory mechanisms to compensate for latency (Mäki-Patola and Hämäläinen, 2004). Some of these studies (Levitin et al., 2000) give us precise numbers for laboratory environments but not for audiences in front of performers. Finally, it is also important to account for *jitter*, which is defined by Wessel and Wright (2002) as the variation in latency, but no experiments have been performed in this matter. Latency and jitter are consequences of the resolution of our sensors, data transmission and processing power/speed. Keeping them below perceivable thresholds helps us forget that we are dealing with a computer and enable us to focus on the interface and the music we want to make with it.

Latency and jitter are not the only pre-requisites for transparent immediacy. As

---

[10] Particularly interesting are the works of Levitin et al. (2000), Mäki-Patola and Hämäläinen (2004), Mäki-Patola (2005), Wessel and Wright (2002).

it can be derived from previous sections, an interface should reflect our sensorimotor structure to allow us to feel that its manipulation is natural or to feel that it is altogether absent. This means that the perception of mediation is not only conditioned to real time responses, but to a transparent experience of those responses.

Connecting this idea, Dourish (2004) points us to Heidegger's concepts of ready-to-hand and present-at-hand:

> As an example consider the mouse connected to my computer. Much of the time, I act through the mouse; the mouse is an extension of my hand as I select objects, operate menus, and so forth. The mouse is, in Heidegger's terms, ready-to-hand. Sometimes, however, such as when I reach the edge of the mousepad and cannot move the mouse further, my orientation toward the mouse changes. Now I become conscious of the mouse mediating my action, precisely because of the fact that it has been interrupted. The mouse becomes the object of my attention as I pick it up and move it back to the center of the mousepad. When I act on the mouse in this way, being mindful of it as an object of my activity, the mouse is present-at-hand. (Dourish (2004) 109, italics mine)

The logic of transparent immediacy is related to this idea of having a medium (interface and computer) *ready-to-hand*. For Dourish (and for Heidegger), what determines this ready-to-handness is attention. For someone like Bolter, it is awareness:

> ... the user will move through the space interacting with the objects "naturally," as she does in the physical world ... In this sense, a transparent interface would be one that erases itself, so that the user is no longer aware of confronting a medium, but instead stands in an immediate relationship to the contents of that medium ((Bolter and Grusin, 1998) 23).

Dourish (2004) strongly supports the idea of interacting with computers as we do in the physical world. This approach has shaped some conceptualizations of computer music performance. For example, in Wessel and Wright (2002), we interact with musical instruments through the metaphors of *drag and drop*, *scrubbing* and *dipping*, actions we perform when we act in our everyday lives.

The logic of transparent immediacy seems to be at play at some of the most successful computer music performances such as Waisviz's. In contrast to the norm, which has a person in front of a computer staring at a visual display, transparent performances attempt to hide the computer (and its programmer) altogether or set it outside the focus of the audience's attention. Not using a visual display is a way of establishing a tight feedback loop where the performer relies primarily on audition to monitor his gestures, just as it is expected in most musical environments.

The logic of transparent immediacy is not that the performer, or the audience, believes that the interface literally 'disappears', but that in experience, the interface is not the center of our attention, it is the music made with it.

In the subject of interface design in cognitive science, the idea of transparency is closer to the concept of *direct manipulation* explored earlier, where the user, and the audience, feel in control of the sound itself instead of obstructed by the medium. (Hutchins et al., 1985)

The Logic of Hypermediacy

New media presents us with the ability to perform multiple tasks and processes with only one machine: the computer. The computer is not a single-task machine by definition, it is able to operate on anything that can be represented as data. Computers as a medium, or computational mediation, are or can be multiple media or multi-media. In other words, the computer-based instrument can mutate at the will of the performer or automatically by design. This variable nature can be seen as posing a challenge to transparency, because it makes the computer - as a medium - present to the performer and the audience.

For Bolter, it is the search for immediacy itself what leads to hypermediacy:

> Although each medium promises to reform its predecessors by offering a more immediate or authentic experience, the promise of reform inevitably leads us to become aware of the medium as a medium. Thus immediacy leads to hypermediacy ((Bolter and Grusin, 1998) 19).

We are constantly creating new instruments, with new interfaces and sound mappings and in each of these new instruments, we reflect our desire of embodying it, of making it more transparent and immediate. But it is precisely in this process of renewal, in the experience of the new that we are confronted with the instrument as a medium; of the medium as a medium.

Hypermedia is the experience of the medium as a medium. As opposed to transparent immediacy it does not attempt to erase the medium, but to emphasize it. Dourish points us to William Mitchell (1994), who states that hypermediacy "privileges fragmentation, indeterminacy, and heterogeneity and ... emphasizes process or performance rather than the finished art object". Because the medium itself is not transparent, it gets on the way of our experience. The following paragraph from Bolter illustrates it particularly well:

> Where immediacy suggests a unified visual space, contemporary hypermediacy offers a heterogeneous space, in which representation is conceived of not as a window onto the world, but rather as "windowed" itself - with windows that open on to other representations or other media. The logic of hypermediacy multiplies the signs of mediation and in this way tries to reproduce the rich sensorium of human experience ((Bolter and Grusin, 1998) 34) ... What characterizes modern art is an insistence that the viewer keep coming back to the surface or, in extreme cases, an attempt to hold the viewer at the surface indefinitely. In the logic of hypermediacy, the artist (or multimedia programmer or web designer) strives to make the viewer acknowledge the medium as a medium and to delight in that acknowledgement ((Bolter and Grusin, 1998), 41).

The logic of hypermediacy is also reflected in many computer music instruments and performance practice.

While Waisvisz (1999) finds it positive that people would "believe that what they were listening to wasn't electronics at all", reflects our desire to experience the new transparently. Although Waisvisz's instrument featured metal keys and wires hanging from his hands, evidencing the fact of mediation, he managed to engage the audience in a transparent experience.

Lewis' *Voyager* on the other hand takes the medium to the foreground persistently. The medium is not expected to be transparent. Quite on the contrary, it presents itself as an "other" that behaves independently. Both *Voyager* and Lewis are constantly listening to what the other does, to how the dialogue between them takes place. The audience then is persistently confronted with identifying how the medium imitates, opposes or ignores the improviser; the audience is forced to constantly enquire about an ever changing medium of multiple ‘timbres’, dynamics, scales and so on. But just as Lewis learns the many instantiations of his code through experience, the audience can at times, forget about the medium. The duo in *Voyager* reaches transparency when it fulfills the idea of the ensemble, the ideal when both of them behave as a whole where the units complement each other.

As in the example of Waisvisz above, an instrument presented live in performance is also a visual experience that provides the audience with cues on what is it that they are experiencing. For example, Ericson (March 7, 1971) described in *The New York Times* how Martirano’s then *MarVil Construction*’s[11] “upright panel decked out with hundreds of wires, resembled a white-on-white Jackson Pollock painting ... a slanting panel held a symmetrical array of blinking lights ... the sounds came out of 21 speakers hanging at various levels around the room.” The visual complexity of the device was in the spirit of patch cords in analog synthesizers. For example, Buchla had modules to control visual processes through voltages and much of the electronic music produced around the San Francisco Tape Music Center (SFTMC) carried with it complex visual images (Bernstein, 2008). Subotnick even flew Buchla to New York a few times to develop visual imagery for concerts (Roads and Subotnick, 1988).

Many new instruments and interfaces emphasize visual displays, both in the interface and in surrounding projections that remind us constantly that we are dealing with a computer, with media. In the practices of live-coding and laptop performance, performance gestures are about the computer and not about the sounds. The computer screen is projected so as to see what the program and the programmer are doing and interfaces for direct manipulation are rejected; in short, to remind us that we are listening to a computer; to be aware of the medium as a medium and delight in it.

---

[11] An early prototype of the SalMar Construction developed in collaboration with engineer James Divilbiss.

Mathews et al. (1969) proposed instruments that make playing easier, instruments that facilitate performance. Wessel (1991) has proposed instruments that learn; that adapt over their interaction to their user.

Many people claim that computer instruments need to have stable mappings:

> *The importance of mapping consistency over time.*
> It has been claimed (Wessel, 1991) that instruments that present the ability to adapt to the user's playing style can eventually become more expressive. The potential problem with adaptive instruments is the reduction of demands on the amount of effort the user needs to spend in order to master the instrument - and it has been claimed that effort and expression are somewhat related (Ryan, 1992). In other words, to what extent are dynamic mappings of interest to instrumental designers and to skilled performers of digital musical instruments? If in the context of interactive composition this question could perhaps be easier to answer due to the inherent temporal evolution of the musical material, in the instrumental case it may perhaps become counterproductive if it demands less effort from a performer. It could also be considered that a system which reconfigures itself is a case of 'constantly moving the goalposts'. In other words the human mind-body system is extremely good at adapting to fixed physical challenges, and learning their subtlety. How will this be affected if the physical challenge is constantly modifying itself? (Hunt and Wanderley, 2002).

This claim is supported on the fact that it is the only way to attune to them and to define the instruments with particular identifiable 'personalities'. This claim can be read at many levels. On one level, it expresses our desire to reproduce the qualities of - to repurpose[12] or remediate - acoustic instruments, both in terms of the stability of their acoustical properties - or intrinsic morphology - and of their cultural signification and status as instruments. On another, it reflects our desire for transparent immediacy, to give the performer and the audience a stable sound-space, and with it, the chance to learn the mapping until it fades out of attention and become ready-at-hand.

The desire for instruments with stable mappings as a desire for transparency is stated in a different way in Fels et al. (2002):

---

[12] "Repurposing" means to take a "property" from one medium and reuse it in another ((Bolter and Grusin, 1998) 45).

> Metaphor enables device designers, players and audience to refer to elements that are 'common knowledge' or cultural bases which we call literature. By grounding a mapping in the literature, it is made transparent to all parties. Metaphor restricts and defines the mapping of a new device. Through metaphor, transparency increases, making the device more expressive. (Fels et al., 2002).

Fels' claim is that if a new electronic instrument behaves as an old mechanical one does, then it will be perceived as transparent and consequently more expressive. But as we have seen, according to Bolter it is precisely the desire for transparent immediacy that leads us to become aware of the medium as a medium, for as stable as we intend the mapping to be it will still be relatively new. Computer instruments establish the ability of constantly renewing themselves and therefore of continually presenting themselves as new and therefore revealing their condition of media as media.

Electronic instruments presented themselves, historically, with relatively stable mappings. However, since the pipe organ we are able to change mappings during performance, although always within the confines of its built-in options, a characteristic that modern electric organs retained. The theremin is associated with a particular synthetic sound, the heterodyning effect, that many people are able to recognize as its timbre. The configurability of analog synthesizers was limited by being unable to change patch configurations while playing live and were therefore perceived as relatively fixed, although through the recording industry they managed to present themselves in multiple timbral personalities. But computers can re-patch and change samples at will; they can become an entirely different medium through a single command.

As stated earlier, one of the characteristics of computers is to be multiple media. In music, the computer is the digital version of the Paris and Cologne studios of the 1950's as well as every digital audio technique invented; it is all imaginable configurations of audio generation and transformation. With modern day interfaces, computer instruments are all imaginable mappings of control signals to all imaginable configurations of audio generation and transformation. In short, an extremely similar gesture can sound in strikingly different ways; a computer instrument as a medium is multiple instruments.

The continual re-definition of mappings in computer instruments is a reflection of hypermediacy and of the nature of the computer itself. The reason we make computer interfaces is to attain control over a medium that is really multiple mediums.

Embodying a Double Logic

In his book *Digital Mosaics: The aesthetics of cyberspace*, Steven Holtzman (1998), claims that repurposing analog media is a transitional step in establishing digital media as a medium with its own unique qualities:

> In the end, no matter how interesting, enjoyable, comfortable, or well accepted they are, the approaches [of repurposing] borrow from existing paradigms. They weren't conceived with digital media in mind, and as a result they don't exploit the special qualities that are unique to digital worlds. Yet, it's those unique qualities that will ultimately define entirely new languages of expression. And it's those languages that will tap the potential of digital media as new vehicles of expression. Repurposing is a transitional step that allows us to get a secure footing on unfamiliar terrain. But it isn't where we'll find the entirely new worlds of expression. Like a road sign, repurposing is a marker indicating that profound change is around the bend (Holtzman, 1998).

Bolter, on the other hand, finds that the nature of digital media is to exist in the double logic of remediation:

> However, like their predecessors, digital media can never reach this state of transcendence, but will instead function in a constant dialectic with earlier media, precisely as each earlier medium functioned when it was introduced. Once again, what is new about digital media lies in their particular strategies for remediating television, film, photography, and painting. Repurposing as remediation is both what is "unique to digital worlds" and what denies the possibility of that uniqueness ((Bolter and Grusin, 1998), 50).

Repurposing as remediation implies the double logic of remediation where on one hand we aspire to transparency and inevitably hypermediate.

Can we embody an always-changing instrument? Have we not gotten used to the nature of digital media as users and observers in this hypermediated environment? By now, a large number of people have used a computer consistently throughout their lives, managing multiple windows and jumping between multiple tasks. Computer interfaces now permeate our everyday lives in what Dourish calls *ubiquitous computing*, they are our telephones, media players, televisions; our computer is our mail, our bank, our entertainment; each device uses multiple media, and we jump from device to device. Dourish points us to the following statement by media theorist Erkki Huhtamo:

> Technology is gradually becoming a second nature, a territory both external and internalized, and an object of desire. There is no need to make it transparent any longer, simply because it is not felt to be in contradiction to the ‘authenticity’ of the experience (Huhtamo, 1995).

As Hollan and Stornetta (1992) point out: “to the extent that the goal is imitation, one will not be led to exploit the distinctive strengths of the new medium”. And even when the goal is not imitation, the concept that an instrument is an instrument only if it can articulate pitches over time, reflects the expectations of a culture towards new instruments. For the computer musician however, the comparison with older mechanical media is inevitable as the musical culture relies heavily on tradition. For this reason, it is easier perhaps to adapt to newer video games or mobile phones, because unlike music, what we expect to hear and see is not governed by centuries of tradition.

For example, a computer-based instrument might be able to control spatial distribution through a multi-channel system. However, when people are confronted with such instruments, a common complain is that they cannot locate the instrument based on spatial cues and demand that the computer-based instrument emits sounds from one source. Is this a remedial nostalgia of the mechanical instrument and its acoustic source? Paradoxically, there are not many modern cinema theaters that do not have a multichannel system.

It is my claim that we can attune to a constantly-changing device. Gibson suggests that we find affordances in the environment and we attune to them. These affordances present themselves to us as what is invariant. And it is precisely its varying nature, its mutability, what is invariant about computers.

As I suggested in the previous chapter, the processes that generate the sounds, or the control layer, can perceived as an invariant. We attune to its multiple manifestations; we develop an ability of navigating multiple windows or mapping spaces. We explore this multilayered environment.

An instrument may change mappings from piece to piece or within pieces, it may shift from a lesser amount of agency to overt independent behavior. But more crucially, a computer instrument and performer must have a way of addressing this issue: to control not only the sounds it makes, but also the actual instrument; in other words, to control the sound of the instrument, but also the ways in which the sound is made.

This is not to say that we can embody something that has no invariant features. The different systems that we have explored in this dissertation are characterized precisely because they have a way of reflecting the performer's actions in their sound output. Again, it is not that sound-sources or sound-processes are unimportant, but instead that a system has a way of responding that is even more stable than those sounds.

For example, Waisvisz would commonly record sounds produced by the audience and transform them live. What was invariance from performance to performance was the transformation, not the sound.

* * *

The audience

What the Audience Sees...

> These five perceptual systems overlap one another; they are not mutually exclusive. They often focus on the same information - that is, the same information can be picked up by a combination of perceptual systems working together as well as by one perceptual system working alone ((Gibson, 1966) 4).

In a musical performance, the audience and the performer perceive through several perceptual systems. Particularly relevant in the case of music performance are the visual, auditory and haptic systems. These have also been referred to as modes of perception and hence the term multimodal perception.

Sound is, as we stated earlier, a time-varying physical force; and music is the structuring of sound. Sound is therefore not only perceived by the ears, but potentially by the whole body, through our haptic system. Our sense of touch is not used in the active way in which we explore the texture or shape of something, instead, the audience is touched in a manner similar to the way wind touches us; structured, quasi-periodic pressure variations in the air. Most performers feel the sound not only with their ears, but also directly on their bodies, the most extreme case being that of singers, whose bodies are the sound producing mechanism and their chests and heads, the sound's resonance 'boxes'. Anyone who has ever sat right in front of a powerful low-frequency speaker and has been exposed to high-amplitude low frequency sounds, has felt the way sounds literally move us. And it is precisely this experience that informs us of how sound acts on us, it subjects us to its vibration.

The visual system however does not usually 'see' or 'feel' these vibrations. It sees the slower gestures we perform to make them and the ones we make as a response to them. How does input from one sensory mode inform and complement information in the other?

The effects of seeing images and hearing sounds at the same time have been the subject of a few studies. The most well-known example is that of the *McGurk effect* (McGurk and MacDonald, 1976). When a subject was presented with an image of a face pronouncing the phoneme 'bah' and a sound of the phoneme 'dah' being pronounced, the person perceived 'bah'. That is, the visual element influenced the sound to the extent that the sound information was partially "overwritten". In a similar study, Schutz and Lipscomb (2007) found that in percussion performance, long gestures produced after striking a bar in a marimba, make the viewer believe the sound is actually longer.

The concept of added value coined by Chion explains cases where auditory information "overwrites" or transforms visual information:

> 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... 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 (Chion, 1994).

Sounds and gestures "add value" to each other. We see the image in front of us (the body of the performer and his instrument) as transformed by the sounds it makes and we hear the sounds as transformed by the gestures that make them. This is not entirely surprising, since gestures are representational and sound is the result of gestures; gestures contain information about sound and viceversa; we see sounds and hear gestures.

Seeing gestures also reveals something about the ongoing dynamics of performance. Facial expressions and bodily posture reveal that the performer is also a listener:

> "In this way, visual aspects of performance signal that performers are not merely producers of sound but are themselves listeners, highlighting the musical activity as a shared experience between performers and listeners" (Thompson et al., 2005).

The dynamics of performance to which we are referring, are a performer in a tight complex dynamic network of feedback loops, where his actions are modulated by the sounds they produce. Facial expressions and gestures in general may reveal how the feedback he is receiving matches or not the intended sonic structure; they reveal something of what he intends the sound to be. The performer can thus be seen as relating to his instrument in the same way as all people relate to any actions through physical exploration of their environment.

For the anthropologist John Blacking, all social activity, like music, is biologically determined and requires interaction. We learn the possibilities and limits of our bodies, like speaking or singing or even moving, through social interaction, determined in turn by a historical and cultural context. A crucial factor in the 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 Blacking, sharing these states 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); our ability to act and perceive is a shared condition that allows us to experience things in a similar way.

> "...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 ... 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).

Our individual bodies possess particular limits, but the limits of 'our 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'.

Nöe makes a similar argument:

> To perceive you must be in possession of sensorimotor skill ((Nöe, 2005) 11) ... If perception is in part constituted by our possession and exercise of bodily skills ... then it may also depend on our possession of the sort of bodies that can encompass those skills, for only a creature with such a body could have those skills. To perceive like us, it follows, you must have a body like ours ((Nöe, 2005) 25).

Indeed, as Gibson pointed out, the animate environment affords more than the physical environment. As a part of the (social) environment, we attune to other peoples bodies, to particular performance techniques, to pieces of music, etc. All of these are part of the social and cultural environment and constitute what Gibson called the behavioral loop. We are modulated by other people's actions. The loop of course is less tight, but over time cultural forms are informed by social stimuli.

Extending the Concept of Affordance

As mentioned earlier, gestures allow the viewer to assign a source to a sound *teleologically* instead of acoustically; the gesture becomes the source of the sound as we learn about how they relate; as we learn about the way a gesture manipulates media. As stated in the quote by Huhtamo above, technology is becoming a second nature; our interaction with technology in our everyday lives gradually relieves us from attempting to understand the mechanics behind them. Just as few people fully understand the physics behind the tamtam, few understand the physics behind the computer. Furthermore, understanding the physics behind the computer would not necessarily make things clearer. This is an example of treating knowledge as an object instead of knowing as a process (Balzano and McCabe, 1986).

At this point it is relevant to extend the concept of affordance. We previously defined it as the recognition of an organism of opportunities for action as exemplified in the tree-climbability pair. The problematic of acousmatic listening lies in the fact that when a person listens to a recorded or electronically produced or transformed sound, he/she might wonder what possible object caused that sound. In an extended concept of affordance we could instead ask the question what actions and objects does this sound afford as a cause? What termination does this sound afford? What other aesthetically meaningful sounds does this sound afford next? What meanings does this sound afford?

When we hear and see, a sound and a gesture interacting over time, we are informed of what that instrument affords the performer in gesture-sound couples. Our long history of interaction in the world, coupled with the pieces of information that each new gesture reveals are in constant interaction. In extreme cases, big and forceful gestures might be found to be in perceptual dissonance if they produce quite delicate sounds, but in evident consonance if they make a loud noise. But there is a limitless area in the middle, for sounds to afford gestures and meanings. Our experience of these gesture-sound couples over time, helps us understand what gesture-sound couples the instrument affords. And this continuing build up of expectation is at the center of aesthetic and perceptual experience.

In this context, we can re-introduce the concept of transparency. For Jäger and Kim (2008) “the transparency of the medium is not a ‘property’ of the medium, but an aggregate state in which the mediatised semantics as silent knowledge is not communicatively ‘disturbed”’. In their view, performances go through states of transparency and disturbance. Disturbance is the “state in the process of a communication which has the effect that a medium (operatively) loses its transparency and is perceived in its materiality”. These disturbances work as “explications” that focus and concentrate on media “in the interest of creating (explicative) semantic effects which, when they occur, push the medium out of the focus of attention once again.” (Jäger and Kim (2008), 53) In short, a work or performance, reaches over time a state of transparency where the audience stops focusing on the medium because the medium is understood, but as the work changes, this state of transparency is disturbed until the next state of transparency is reached.

For Jäger and Kim then, performance with media occurs in a fragile state of equilibrium that is potentially disturbed at any time, bringing the medium to the foreground and through and act of explication pushing it out of attention once again.

Accountability

Music performance is a social practice that has evolved through time to become a ritual with different configurations across communities and cultures. The question of audience in computer music as with any other music performance practice is then crucial; they validate it as a practice and guide its evolution.

Alfred Schutz explored the problem of intersubjectivity: if our experience is fundamentally our own, how can we have shared experiences and meanings of the world? Schutz’s premise to solve the problem of intersubjectivity is the assumption of rationality; that “we work under the assumption that others are rational as we are, and that others’ experience is like our own” (Dourish (2004) 112).

As, for example in the case of language, communities and subsets of these communities share “a set of understandings of how to act, and how to understand action” (Dourish (2004) 112). We can think of music and computer music communities in a similar way. This brings us to the concept of accountability, developed by Harold Garfinkel. Accountability is considered a feature of action, whereby members of some setting are able to observe and report, that is, to understand actions in the context in which they arise. In this definition, being a member is sharing the set of understandings mentioned above: context situates the action temporally, spatially and symbolically.

Paul Dourish analyzes a passage by Garfinkel (1967) about the notion of accountability.

> Accountability lies in the reciprocality of action and understanding ... the methods of understanding and making sense of action and the methods for engaging in it are the same methods. ...being a competent member of some setting is being able to engage in action in ways that are recognizable to other members. The organization of action serves to demonstrate what that action is ... the accountable aspect of activity is never a “commentary” on the activity, standing separately from it; rather, it is an intrinsic and inseparable feature of how the activity is woven into the fabric of action and interaction ... the organization of action, as it arises in situ, provides others with the means to understand what it is and how to respond in a mutually constructed sequence of action. It turns our attention away from simply the perceived result or outcome of an action, to include how that result is achieved. We pay attention not just to the destination, but also to the route taken to get there. (Dourish (2004), 78-80)

This last idea of destination and route opposition is remarkably similar to Wessel’s desire to obtain continuous control of computer music sounds as opposed to the trigger paradigm, because in the latter there is no ‘route taken’ to achieve the sound except for the trigger action. The theory of accountability also resonates with the relationship of action and perception, with the representational quality of gestures, etc.

What is relevant about the concept of accountability is the relationship between the observing member (the listener) and the acting member (the performer); that performance actions are organized in such a way as to produce the desired sonic structure, and at the same time to account for its production. We use the same mechanisms to understand and produce a sound. Our actions seek to achieve a sound and in doing so represent it.

* * *

Conclusions

This chapter has attempted to situate computer music performance in terms of the theories of embodiment in cognitive science.

In this context, embodying a musical instrument is to explore it, identify its affordances and attune to it as we do with the environment. This attuning process is achieved through time, through exploration and training, forming a history of interaction. This history provides gestures with a proper function and representational status; gestures become and contain information about the sound structures they intend to produce. Representation in this case is something performer does or achieves and it needs to be distinguished from representations in the traditional sense, which are conceived as internal or mental configurations. Our actions seek to obtain stimulation and work as forward models within a tight complex dynamic network of feedback loops; they are an exploration of - and are modulated by - the environment.

Gestures, as actions, represent over time. On one level, gestures represent the interface in which they occur and on another a sonic structure we want to perform. In performed electronic music, we are allowed to relate a gesture to a sound; to assign a source to a sound teleologically instead of acoustically, we somehow stop caring what object caused the sound, but what gesture. This is the reason why tracking gestures is interesting in the first place: gestures are the ‘sources’ of sounds; accessing gestures is a way of accessing sounds. Embodied performance of computer music, is to represent more accurately through our actions the sounds we intend to make and the interface in which we make them.

The visual system does not usually ‘see’ or ‘feel’ sounds. It sees the slower gestures we perform to make them and the ones we make as a response to them. Computational mediation seeks to ’see’ and track these slow actions and translate them into sound. Gestures are complex phenomena that occur in the whole body of the performer, not just in any one part. What a sensor sees is not our gestures, but their representation in data or our ‘data body’. This data contains information on the gesture, but also information on the sensor. This data is presented as a dimensionality reduction of our actual body; the whole body, in all of its dimensionality is somehow partially contained in each of its parts and reflected in the one gesture we track. The digitization of our body through sensors converts our body into a “data body” which translates it into a format with which we can manipulate media and interact with real time computer systems.

Low latency and jitter are crucial to obtain real time, and therefore a pre-condition for transparent, mediations. Interfaces should reflect our sensorimotor structure, at least well enough as to allow us to feel that its manipulation is ‘natural’ or to feel that they are not present altogether. This is what we refer to as the logic of transparent immediacy. This doesn’t mean that the performer, or the audience, believes that the interface literally ‘disappears’, but that in experience, the interface is not the center of our attention, it is the music made with it.

The computer is not a single task machine by definition, but one that is able to perform any algorithmic operation on anything that can be captured and represented as data, in other words, media. Hypermedia is the experience of the medium as a medium. The logic of Hypermediacy functions as a reminder that we are listening to a computer; it makes us aware of the computer medium as a medium and lets us delight on it. Computers can re-patch and change samples at will; they can become an entirely different medium through one command; they have the ability of continually presenting themselves as new and therefore evidencing their condition of media as media. An extremely similar gesture can sound in strikingly different ways; a computer instrument as a medium is multiple instruments, multiple media. Furthermore, the medium can increase its complexity to become an ‘other’ altogether. The medium takes a voice of its own and interacts with the performer.

The reason we make computer interfaces is to attain control over a medium that is multiple media. Hypermediacy, to which we arrive in our desire for transparency, brings about an aesthetic of ‘windowed’ fragmentation that reflects this multiplicity of media. We attune to its multiple manifestations; we develop an ability of navigating multiple windows or mapping spaces. A computer instrument and performer must have a way of addressing this issue: to not only control the sounds it makes, but also the actual instrument; to control the sounds of the instrument and to ’switch’ between all the instruments a computer can be.

Transparency is not a property of the medium, but a state in performance. Performance with media oscillates between states of transparency and disturbance. The states of disturbance lead to acts of explication which make explicit what an instrument affords the performer in terms of gesture-sound couplings.

All social activity, like music, is biologically determined and requires social interaction. To perceive and to perform music like we do as a social group, we need to have the same kinds of bodies and therefore to share the ability to act and perceive in a similar way; that is, to achieve shared somatic states and explain intersubjectivity. Embodied computer music performance shifts attention away from the result to the process that produces it, to the accountable and representational nature of the actions of our bodies.

The meanings of the words perception, representation, action, cognition and so on, are gradually changing. In the spirit of Wittgenstein, meaning in language emerges from use, but it also arises in context. As in the case of the traditional conceptualization of musical instrument, the “linguistic baggage” that these terms carry with them, are information bearing structures which form a linguistic environment which affords inspection through constant cultural interaction, an interaction that materializes in our every day actions.

Chapter 11

Music Models : A Computer Music Instrumentarium

* * *

Music Models

As we go back to the first element of this dissertation, the current concept of musical instrument emerges from the way we have modeled our musical practice. Figure 11.1 re-produces the traditional composer-performer-listener model originally presented in Figure 2.1, but this time it introduces the figure of the *luthier*.

It is not the role of this dissertation to determine what were the roles and motivations of luthiers in the different aesthetic and social contexts throughout history, although it remains a promising subject of enquiry. However, we take for granted the fact that the luthier has partly shaped the sounds we hear today. Each instrument affords certain musical behaviors and styles while preventing or impeding others. The role of the luthier, while silent, has been to determine what sounds is music made of.

Theremin sought composers to write for his instrument, he had numerous pupils and ventured into commercial production. Like most instrument builders in the first half of the XXth century, saw his instrument as an improvement on existing instruments; as a

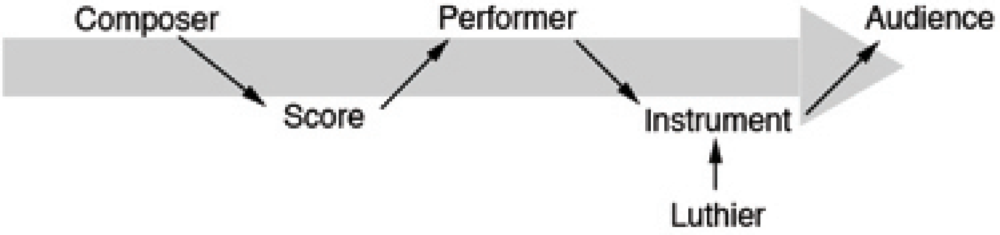

Figure 11.1: The Luthier in the Composer-Performer-Listener Model of Music

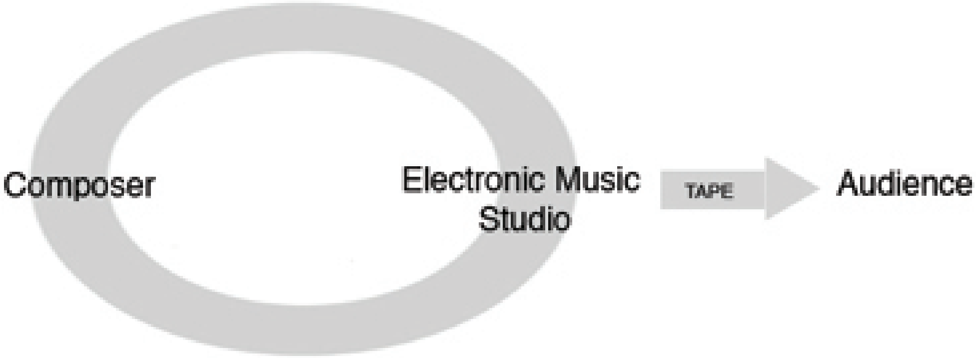

Figure 11.2: A Model of Music for the Electronic Music Studio

way of advancing musical practice within the model. In his view, his instrument would "perform the same functions as ... traditional instruments", however by being "able to produce sounds by responding readily to the free movement of the hands in space, it would have an advantage over traditional instruments" (Theremin and Petrishev, 1996).

As seen in Figure 11.2, the emergence of the electronic music studio did have a substantial effect on the model of music presented above. Because the composer worked directly with the actual sounds of the composition, there seemed to be no need for performers, instruments or notation; these were are all seen as limitations. Although no instruments were used to perform this music and the audience was presented with a finished tape, I have argued that instruments are still perceived. I have therefore stated that tape music presents us with the sounds of instruments that have not been built.

Schaeffer (1966) called these instruments pseudo-instruments and conceived of them as “invariants” that work as “permanence of a common character”. But if there are pseudo-instruments, there must be pseudo-performers, pseudo-luthiers, pseudo-scores, pseudo-spaces, etc. In the same way that in physics energy is never destroyed, but transformed, the pieces that emerged from the electronic music studio suggested not the disappearance of the instrument and the performer, but instead, their distribution.

In the absence of the mechanical instrument, the composer was now also a luthier, in charge of designing the sounds and possible articulations of the pseudo-instrument. These instruments however, constituted an integral part of the composition and in a sense were inseparable from it. The criteria and parameters used for the construction of the instrument and the composition were often the same.

The desire for more accessible, less laborious and customizable studios fueled the development of live practices. This time, instead of the closed work fixed on tape, the work was conceived as an open system. Nonetheless, the role of the composer as luthier remained.

Mumma (1967) considered that his “designing and building of circuits is really ‘composing’” and Martirano (1971) believed that “an analogy to traditional composition exists, considering that particular patching patterns are developed over a period of time”. The design of the instrument itself is seen as a compositional act. Because the instruments were designed as part of the composition process, they embedded compositional ideas and were “inseparable from the compositions themselves” (Mumma, 1967). In this sense, the composer is a composer-luthier, but the instrument is also an instrument-composition.

Performance is also confused with composing. Because the system is open to performance, the composition is completed in real time. Mumma (1967) believes that the “end-product” of his work is “more than a package of electronic hardware, it is a performance of music”; Martirano (1971) states that “in performance a composer can improvise, better said, compose in real time, within a large set of musical possibilities”. As the system has larger agency over the final shape of the work, the performer is seen less in the role of "accurate"[1] reproduction of a score, but instead as "steering"[2] the system, bringing a sort of open score into a concrete form.

In this sense, the composer is also a composer-performer or the performer, a performer-composer. The works of these authors shift the belief that a composition is self-contained in a score by art of the composer to an open system where the composition emerges from the interaction between a composer, luthier, instrument and performer.

In the work of Waisvisz and Lewis, the roles of composer and performer are confused even further into a composer-performer or improvisor, and the system is believed to have a large amount of agency over the work:

> The performer can be the composer, conductor and various of the soloists at virtually the same time. The instrument consists of a system that can compose and perform electronic music independently on stage or can be 'steered' by a composer/performer (Waisvisz, 1999).
>
> These works involve extensive interaction between improvising musicians and computer music-creating programs at the performance ("real time") level. In both theory and practice, this means that both human musicians and computer programs play central organizing and structuring roles in any performance of these works (Lewis, 1999).

The blurring of the bounds of the specialized skills of composer, performer or luthier and of the roles of score-as-composition and instrument-as-passive-tool configure a new kind of practice in computer music. Commonly, the creation of new works defy the traditional model of music production, by re-drawing it.

In consequence, composing interactive computer music systems is not an independent activity of performing them or designing them. As seen in Figure 11.3, the compositional process can be seen as a loop where an environment is constructed and modified based on the experience of interacting with it. In other words, the instrument and the composition emerge from the consistent experiencing of the system in advance of the actual performance.

While we attune to the instrument, we also tune the instrument to ourselves.

---

[1] (Adorno, 1976).

[2] "I steered. It was like driving a bus" (Martirano in Chadabe (1997)).

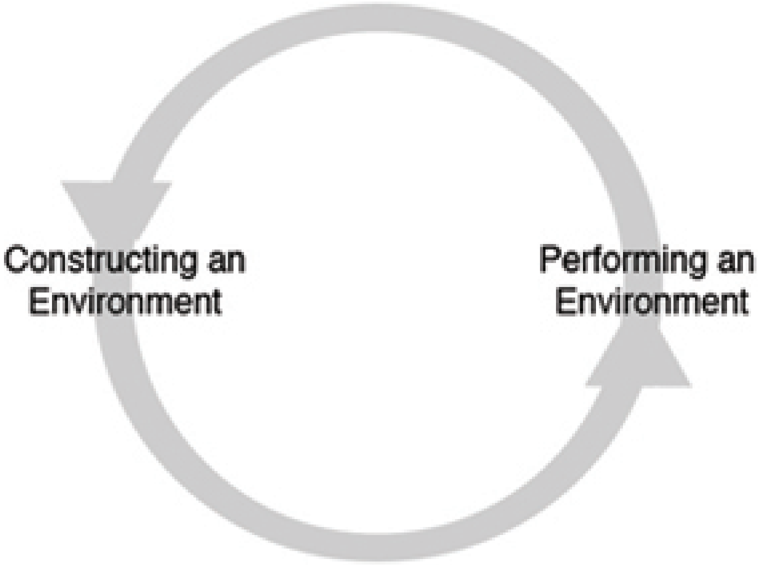

Figure 11.3: The Compositional Loop in Interactive Computer Music Systems

* * *

A Computer Music Instrumentarium

Exchange

If, as I am arguing, an instrument contains compositional thought; if it contains an applied music theory or technologies to think about music or about a piece; if the instrument is inseparable from the composition itself, then, what are the consequences of exchanging code?

The schematic opened the door to a new kind of exchange. The theremin appropriated the audion, but the theremin was also appropriated, and a modified version published, by Robert Moog.

Software allowed sound generating and processing modules to be realized in code. What was now exchangeable were instructions. Risset (1985) used the data that John Chowning left at Bell Labs, but he also produced a catalogue of sounds which provided code, mathematical and sound examples of certain sounds (Risset, 1969).

For Risset (2003), “analyzing scores is an essential part of teaching - or learning - composition” and computer program “scores” can provide composers with “recipes” and insight about the “compositional elaboration of sonic structures”. Moreover, Risset (2003) believes that “computer synthesis of sound would remain very primitive and unmusical if explorers had not shared the know-how they developed.”

Code exchange amounts to the advancement of a collective knowledge, but as with composition, it also becomes a way of advancing a cultural practice. If code contains compositional thought, to appropriate, transform and re-use code amounts to cultural quotation; to an author referencing another one.

Since Music V, computer music software allows for the exchange of programs or *patches*. Commonly, the users of these softwares exchange information over the internet through several mechanisms like mailing lists and sites of several kinds. These tools allow for the exchange of code developed by the users either as applications that use the software platform, smaller programs or the patches for a particular piece. For example, Puckette (2001) has published the complete code for Manoury’s *Jupiter*, but also realizations of Stockhausen’s *Mantra* and other pieces.

In an open source community it is very common to take the code from someone else, study, modify and appropriate it for their own purposes, commonly returning this piece of code back to the community. Because an instrument is composed, one’s use of that instrument is inevitably linked to the original composition.

For example, if lewis played his trombone into Mumma’s cybersonic console for hornpipe or into Manoury’s Jupiter patch, or the other way round, if Mumma played into the voyager software, how much will be Lewis, and how much will be Mumma’s or Manoury’s? Furthermore, how much would it be Mathew’s or Puckette’s?

It is impossible to determine how much of a particular result would each composer be responsible for[3]. We can however, understand that when code is exchanged, built upon, composed with, used as inspiration, recombined, and so on, there are conse-

---

[3] In this same line of thinking, we cannot determine either, to what extent has the programmer of the language in which the music is specified, framed the work altogether.

quences for authorship and the advance of a shared practice.

We are each other's luthiers.

The Instrumentarium

The word instrumentarium refers to the set of instruments needed for a particular practice. In the traditional music model where the instrument is an object, the instrumentarium consists of a set of objects that allow for a particular music to be achieved. In traditional mechanical music the composer chooses -from an available set of instruments - which ones are to be used to achieve a particular result.

In computer music, the instrumentarium, like the concept of instrument itself is distributed. An instrument emerges from the configuration of a system that consists of an interface, a control layer, sound processes and sound sources. Both the elements of the system and the system itself are placed in a relational position; in other words, instruments are not contained in any one object, but instead distributed. The system is then an open configuration that is completed in performance.

New instruments are commonly the result of the development and re-combination of interfaces and programs.

The instrument in computer music cannot be found or contained in any one object, but instead, in the total configuration that emerges from the practices of exchange, creation and appropriation of computer programs and physical interfaces.

The instrumentarium then, lies not in a collection of objects, but in the combinatorial freedom to build open configurations.

Much as in percussion all the available objects to a percussionist are potential instruments, in computer music, all possible configurations are the potential instruments.

* * *

My Own Practice .:. Final Remarks

As I reflect on my own practice as a computer musician, I realize that a central concern in my work in general, is to elucidate the hybridity of the role of computer musicians in general. The hybridity I refer to, consists not only of being a composer, luthier and performer, but also, of the inevitable concern for reaching an understanding about the nature of a new practice; to understand how computer sounds are cognized, how the physics of sound work, how we make computers listen and understand, and so on.

This selective historical account of electronic and computer music's most emblematic practices, has had the intention of tracing how the concept of instrument has evolved over time. It is my claim that these instruments do not simply contradict our traditional conception of musical instrument, but perhaps reveal a truer nature of musical instruments, and in consequence, of musical practice as a whole.

The examples used serve as paradigms of interactive music systems; of new kinds of computer-based musical instruments that listen, respond and exhibit agency. These works, are also examples of the way in which composers design their own instruments to embody their own theories of music, even their theories of musical practice. Acoustic instruments always embody theories of music. Computers allow us to move a step further, and articulate them over time. These theories can be local to a piece or to a patch, but they require the composer and performer to make their own theories through their interaction with the system.

As I have argued throughout this text, composers of electronic and computer music have since the very beginning, been concerned with the creation of instruments. This is not only a consequence of creative needs, but of the nature of the computer medium itself.

We are faced then, with the opportunity, and even perhaps the need, to reflect upon the nature of our common practice. This reflection need not be a theoretical dissertation. As we have seen, each of the works explored in this dissertation constructs not only an instrument or a composition. Each of these works is a statement about musical practice; each of them redraw what the composition is, what the instrument is: they are propositional.

In this context, the boundaries among composer, performer, instrument, audience and work become themselves subject to exploration instead of a stable ground over which one works.

It is my hope that my own creative work is an exploration of this kind, an exploration that takes both the form of scholarly enquiry and musical practice.

References

Adorno, T., 1976: *Introduction to the Sociology of Music*. New York: Seabury Press.

Apel, W., 1948: Early History of the Organ. *Speculum: A Journal of Mediaeval Studies*, 191–216. ISSN 0038-7134.

Appleton, J., and Perera, R., 1975: *The Development and Practice of Electronic Music*. Prentice Hall. ISBN 0132076055.

ARC, date not available: Don Buchla. *Vasulka Archive: Artists' Research Collective*.

Auslander, P., 1999: *Liveness: Performance in a Mediatized Culture*. Routledge.

Balzano, G., and McCabe, V., 1986: An Ecological Perspective on Concepts and Cognition. In *Event cognition: An ecological perspective*, 93–112. Lawrence Erlbaum.

Battier, M., 2007: What the GRM Brought to Music: from Musique Concrète to Acousmatic Music. *Organised Sound*, **12**(03), 189–202. ISSN 1469-8153.

Bernstein, D., 2008: *The San Francisco Tape Music Center: 1960s Counterculture and the Avant-garde*. University of California Press. ISBN 0520248929.

Blacking, J., 1977: Towards an Anthropology of the Body. *The Anthropology of the Body*, 1–28.

Bode, H., 1961: New tool for the exploration of unknown sound resources for composers. *The Journal of the Acoustical Society of America*, **33**, 861.

Bode, H., and Moog, R., 1972: A High-Accuracy Frequency Shifter for Professional Audio Applications. *Journal of the Audio Engineering Society*, **20**(6), 453–458.

Bolter, J., and Grusin, R., 1998: *Remediation: Understanding New Media*. The MIT Press.

Buchla, D., 1966: The Modular Electronic Music System. *Buchla and Associates*.

Buchla, D., 2005: A History of Buchla's Musical Instruments. In *Proceedings of the 2005 conference on New interfaces for musical expression*, 1. National University of Singapore Singapore, Singapore.

Cage, J., 1961: *Silence: Lectures and Writings*. Wesleyan University Press. ISBN 0819560286.

Chadabe, J., 1977: Some reflections on the nature of the landscape within which computer music systems are designed. *Computer Music Journal*, **1**(3), 5–11.

Chadabe, J., 1997: *Electric Sound: The Past and Promise of Electronic Music*. Prentice Hall.

Chion, M., 1983: *Guide des Objets Sonores: Pierre Schaeffer et la Recherche Musicale*. Paris: INA/Buchet/Chastel.

Chion, M., 1991: *L'Art des Sons Fixés ou La Musique Concrètement*. Editions Metamkine.

Chion, M., 1994: *Audio-Vision: Sound on Screen*. Columbia University Press.

Chion, M., 1995: *Guide to Sound Objects: Pierre Schaeffer and Musical Research. Unpublished translation by John Dack and Christine North 1995*. Unpublished.

Choi, I., 2000: Gestural Primitives and the Context for Computational Processing in an Interactive Performance System. *Trends in Gestural Control of Music*, 139–172.

Desantos, S., Roads, C., and Bayle, F., 1997: Acousmatic Morphology: An Interview with François Bayle. *Computer Music Journal*, **21**(3), 11–19.

Dewar, R., et al., 2009: *Handmade Sounds: The Sonic Arts Union and American Technoculture*. Ph.D. thesis, Wesleyan University.

Dourish, P., 2004: *Where the Action Is: the Foundations of Embodied Interaction*. The MIT Press.

Eimert, H., 1957: What is Electronic Music? *Die Reihe*, **1**, 1–10.

Eimert, H., 1972: How Electronic Music Began. *The Musical Times*, **113**(1550), 347–349.

Emmerson, S., 1986: The Relation of Language to Materials. *The Language of Electroacoustic Music*, 17–39.

Emmerson, S., 2007: *Living Electronic Music*. Ashgate Publishing Company.

Ericson, R., March 7, 1971: Martirano Explores Avant-Garde Music With a 'What Is It'. *The New York Times*.

Fels, S., Gadd, A., and Mulder, A., 2002: Mapping transparency through metaphor: towards more expressive musical instruments. *Organised Sound*, **7**(02), 109–126.

Franco, S., 1974: *Hardware Design of a Real-Time Musical System.* Ph.D. thesis, University of Illinois at Urbana-Champaign Champaign, IL, USA.

Garfinkel, H., 1967: *Studies in Ethnomethodology*. Prentice-Hall.

Gibson, J., 1966: *The Senses Considered as Perceptual Systems*. Houghton Mifflin.

Glinsky, A., 2000: *Theremin: Ether Music and Espionage*. University of Illinois Press. ISBN 0252025822.

Grey, J., 1975: *An Exploration of Musical Timbre*. Department of Music, Stanford University.

Heathcote, A., 2002: Sound Structures, Transformations, and Broken Magic: An Interview with Helmut Lachenmann. *Contemporary Music: Theoretical and Philosophical Perspectives. Max Paddison, Irène Deliège Eds.*, 331.

Heidegger, M., 2003: The Question Concerning Technology. In *Philosophy of technology: the technological condition: an anthology*. Wiley-Blackwell.

Hollan, J., and Stornetta, S., 1992: Beyond being there. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 119–125. ACM.

Holtzman, S., 1998: *Digital Mosaics: The Aesthetics of Cyberspace*. Touchstone.

Huhtamo, E., 1995: Encapsulated bodies in motion: Simulators and the quest for total immersion. *Critical Issues in Electronic Media*, 159–186.

Hunt, A., and Wanderley, M., 2002: Mapping performer parameters to synthesis engines. *Organised Sound*, **7**(02), 97–108.

Hurley, S., 2001: Perception and Action: Alternative Views. *Synthese*, **129**(1), 3–40.

Hutchins, E., Hollan, J., and Norman, D., 1985: Direct manipulation interfaces. *Human–Computer Interaction*, **1**(4), 311–338.

Jäger, L., and Kim, J., 2008: Transparency and Opacity. In *Seifert, U. and Kim, J.H. and Moore, A. Eds.: Paradoxes of Interactivity: Perspectives for Media Theory, Human-Computer Interaction, and Artistic Investigations*, 44. Transcript Verlag, Roswitha Gost, Sigrid Nokel u. Dr. Karin Werner.

Jordan, M., and Rumelhart, D., 1992: Forward Models: Supervised Learning with a Distal Teacher. *Cognitive Science: A Multidisciplinary Journal*, **16**(3), 307–354.

Kaempffert, W., October 5, 1927: Ether Wave Music Amazes Savants. *The New York Times*.

Kartomi, M., 1990: *On Concepts and Classifications of Musical Instruments*. University of Chicago Press. ISBN 0226425495.

Kartomi, M., 2008: Classification of Instruments: The Classification System of Hornbostel and Sachs. *Grove Music Online*, Accessed [May, 2008], http://www.grovemusic.com.

Katz, M., 2002: Aesthetics Out of Exigency: Violin Vibrato and the Phonograph. In *Music and Technology in the Twentieth Century*, volume 20, 174–185. Johns Hopkins University Press.

Krämer, S., 2006: The Cultural Techniques of Time Axis Manipulation: On Friedrich Kittler's Conception of Media. *Theory, Culture & Society*, **23**(7-8), 93.

Krämer, S., 2008: Does the Body Disappear? A Comment on Computer Generated Spaces. In *Seifert, U. and Kim, J.H. and Moore, A. Eds.: Paradoxes of Interactivity: Perspectives for Media Theory, Human-Computer Interaction, and Artistic Investigations*, 26. Transcript Verlag, Roswitha Gost, Sigrid Nokel u. Dr. Karin Werner.

Krefeld, V., 1990: The Hand in The Web: An Interview with Michel Waisvisz. *Computer Music Journal*, **14**(2), 28–33.

Landy, L., 2007: *Understanding the Art of Sound Organization*. The MIT Press.

Levitin, D., Maclean, K., Mathews, M., Chu, L., and Jensen, E., 2000: The perception of cross-modal simultaneity (or"the greenwich observatory problem"revisited). In *Proceedings of the American Institute of Physics Conference Series*, volume 517, 323–329.

Lewis, G., 1999: Interacting with latter-day musical automata. *Contemporary Music Review*, **18**(3), 99–112.

Lewis, G., 2000: Too Many Notes: Computers, Complexity and Culture in Voyager. *Leonardo Music Journal*, 33–39.

Mäki-Patola, T., 2005: Musical Effects of Latency. *Suomen musiikintutkijoiden symposiumin satoa*, 82.

Mäki-Patola, T., and Hämäläinen, P., 2004: Latency Tolerance for Gesture Controlled Continuous Sound Instrument Without Tactile Feedback. In *Proceedings International Computer Music Conference*, 1–5. International Computer Music Association.

Manovich, L., 2002: *The Language of New Media*. The MIT press. ISBN 0262632551.

Martirano, S., 1971: *Progress Report #1: Electronic Music Instrument Which Combines the Composing Process with Performance in Real Time*. Department of Music. Urbana: University of Illinois. Sousa Archives.

Martirano, S., 1989: *Program Note, White Street, July 23*. Unpublished.

Martirano, S., November 2, 1977: *Letter to Howard Klein, Rockefeller Foundation*. Sousa Archives.

Mathews, M., 1963: The Digital Computer as a Musical Instrument. *Science*, **142**(3592), 553–557.

Mathews, M., and Guttman, N., 1959: Generation of Music by a Digital Computer. In *Proceedings 3rd International Congress on Acoustics (Stuttgart, 1958)*. L. Cremer, Editor. Elsevier Publishing Company, Amsterdam.

Mathews, M., Miller, J., Moore, F., Pierce, J., and Risset, J., 1969: *The Technology of Computer Music*. The MIT Press.

Mathews, M., and Moore, F., 1970: GROOVE - A Program to Compose, Store, and Edit Functions of Time. *Communications of the ACM*, **13**(12), 715–721.

Mathews, M., Pierce, J., and Guttman, N., 1962: Musikalische Klänge von Digitalrechnern (Musical Sounds from Digital Computers). *Gravesaner Blätter*, **23/24**, 109–118/119–128.

McCartney, A., 1999: *Sounding Places: Situated Conversations through the Soundscape Compositions of Hildegard Westerkamp*. Ph.D. thesis, York University Toronto, Ontario.

McCartney, J., 1996: Supercollider, a new real time synthesis language. In *Proceedings of the International Computer Music Conference*, 257–258. International Computer Music Association.

McGurk, H., and MacDonald, J., 1976: Hearing Lips and Seeing Voices. *Nature*, **264**(5588), 746–48.

McKean, E., 2005: *The New Oxford American Dictionary*. Oxford. ISBN 0195170776.

McLuhan, M., 1963: *Understanding Media: The Extensions Of Man*. New York, USA. McGraw-Hill.

Mitchell, W., 1994: *Picture Theory*. University of Chicago Press.

Moog, R., 1965a: A Voltage-Controlled Low-Pass High-Pass Filter for Audio Signal Processing. In *17th Annual Meeting, Audio Engineering Society*. Audio Engineering Society.

Moog, R., 1965b: Voltage-Controlled Electronic Music Modules. *Journal of the Audio Engineering Society*, **13**(3), 200–206.

Moog, R., 1967: Electronic Music: Its Composition and Performance. *Electronics World*, **xxvii**(2), 42–6.

Moore, F., 1969: Technical Memorandum: GROOVE - A Program for Composing and Editing Functions of Time in Real-Time. *Bell Telephone Laboratories Internal Memorandum*, **Case 25952**.

Moore, F., 1988: The dysfunctions of midi. *Computer music journal*, **12**(1), 19–28.

Mumma, G., 1964: An Electronic Music Studio for the Independent Composer. *Journal of the Audio Engineering Society*, **12**.

Mumma, G., 1967: Creative Aspects of Live-Performance Electronic Music Technology. In *33rd Convention of the Audio Engineering Society*. Audio Engineering Society.

Mumma, G., and Smoliar, S., 1971: The Computer as a Performing Instrument. *Artificial Intelligence Memo No. 312, MIT*.

Nöe, A., 2005: *Action in Perception*. The MIT Press.

Oliver, J., 2010: The MANO Controller: A Video-based Hand Tracking System. In *Proceedings of the International Computer Music Conference*. International Computer Music Association.

Oliver, J., and Jenkins, M., 2008: The Silent Drum Controller: A New Percussive Gestural Interface. In *Proceedings of the International Computer Music Conference*. International Computer Music Association.

Pinch, T., 2002: Emulating Sound: What Synthesizers Can and Can't do: Explorations in the Social Construction of Sound. *Wissen und soziale Konstruktion*, 109.

Pinch, T., and Trocco, F., 2004: *Analog Days: The Invention and Impact of the Moog Synthesizer*. Harvard University Press. ISBN 0674016173.

Puckette, M., 1991: Combining event and signal processing in the max graphical programming environment. *Computer Music Journal*, **15**(3), 68–77.

Puckette, M., 1996: Pure data: Another integrated computer music environment. *Proceedings of the Second Intercollege Computer Music Concerts*, 37–41.

Puckette, M., 2001: New public-domain realizations of standard pieces for instruments and live electronics. In *Proceedings of the International Computer Music Conference*, 377–380. International Computer Music Association.

Rammert, W., 2008: Where the Action Is: Distributed Agency Between Humans, Machines, and Programs. In *Seifert, U. and Kim, J.H. and Moore, A. Eds.: Paradoxes of Interactivity: Perspectives for Media Theory, Human-Computer Interaction, and Artistic Investigations*, 62–91. Transcript Verlag, Roswitha Gost, Sigrid Nokel u. Dr. Karin Werner.

Risset, J.-C., 1969: *An Introductory Catalogue of Computer Synthesized Sounds*. Bell Telephone Laboratories.

Risset, J.-C., 1985: Computer Music Experiments 1964 - ... *Computer Music Journal*, 11–18. ISSN 0148-9267.

Risset, J.-C., 1987: About James Tenney, Composer, Performer, and Theorist. *Perspectives of New Music*, **25**(1/2), 549–561. ISSN 0031-6016.

Risset, J.-C., 2003: Computer Music: Why? *France-University of Texas Institute*.

Roads, C., and Subotnick, M., 1988: Interview with Morton Subotnick. *Computer Music Journal*, **12**(1), 9–18. ISSN 0148-9267.

Rosenboom, D., 1996: BC-AD and Two Lines: Two Ways of Making Music While Exploring Instability–In Tribute to Salvatore Martirano. *Perspectives of New Music*, **34**(1), 208–226. ISSN 0031-6016.

Rowe, R., 1993: *Interactive Music Systems*. MIT Press Cambridge, Mass.

Rowe, R., Garton, B., Desain, P., Honing, H., Dannenberg, R., Jacobs, D., Pope, S., Puckette, M., Lippe, C., Settel, Z., et al., 1993b: Editor's Notes: Putting Max in Perspective. *Computer Music Journal*, **17**(2), 3–11.

Rowlands, M., 2006: *Body Language: Representation in Action*. The MIT Press.

Ryan, J., 1992: Effort and expression. In *Proceedings of the International Computer Music Conference*, 414–414. International Computer Music Association.

Schaeffer, P., 1966: *Traité Des Objets Musicaux*. Paris: Editions du Seuil.

Schedel, M., 2007: Electronic Music and the Studio. In *Nick Collins and Julio dEscriván, Editors: The Cambridge Companion to Electronic Music*. MIT Press.

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.

Schonberg, H., February 16, 1969: A Merry Time With the Moog? *The New York Times*.

Schutz, M., and Lipscomb, S., 2007: Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception*, **36**(6), 888.

Small, C., 1998: *Musicking: The Meanings of Performing and Listening*. Wesleyan University Press. ISBN 0819522570.

Smalley, D., 1986: Spectro-Morphology and Structuring Processes. *The Language of Electroacoustic Music*, 61–93.

Smalley, D., 1994: Defining Timbre - Refining Timbre. *Contemporary Music Review*, **10**(2), 35–48.

Smalley, D., 1996: The Listening Imagination: Listening in the Electroacoustic Era. *Contemporary Music Review*, **13**(2), 77–107.

Smalley, D., 2001: Spectromorphology: Explaining Sound-Shapes. *Organised Sound*, **2**(02), 107–126.

Smalley, D., 2007: Space-Form and the Acousmatic Image. *Organised Sound*, **12**(01), 35–58.

Stockhausen, K., and Barkin, E., 1962: The Concept of Unity in Electronic Music. *Perspectives of New Music*, **1**(1), 39–48.

Tenney, J., 1963: Sound-Generation by Means of a Digital Computer. *Journal of Music Theory*, **7**(1), 24–70. ISSN 0022-2909.

Tenney, J., 1965: Die Physikalischen Korrelate der Klangfarbe (The Physical Correlates of Timbre). *Gravesaner Blätter*, **26**, 103–105/106–109.

Tenney, J., 1969: Computer Music Experiences, 1961-1964. *Electronic Music Reports*, **1**, 23–60.

Theremin, L., and Petrishev, O., 1996: The Design of a Musical Instrument Based on Cathode Relays. *Leonardo Music Journal*, **6**, 49–50.

Thompson, W., Graham, P., and Russo, F., 2005: Seeing Music Performance: Visual Influences on Perception and Experience. *Semiotica*, **156**(1/4), 203–227.

Toop, R., 1979: Stockhausen and the Sine-Wave: The Story of an Ambiguous Relationship. *The Musical Quarterly*, **65**(3), 379.

Vail, M., 2003: Vintage Gear: SalMar Construction - Hybrid Performance Synthesizer. *Keyboard*, **29**(7), 118.

Varela, F., Thompson, E., and Rosch, E., 1992: *The Embodied Mind*. The MIT Press.

Waisvisz, M., 1985: The Hands, A Set of Remote Midi-Controllers. In *Proceedings of the International Computer Music Conference*, 313–318. International Computer Music Association.

Waisvisz, M., 1999: Riding the sphinx - lines about 'live'. *Contemporary Music Review*, **18**(3), 119–126.

Walker, W., Hebel, K., Martirano, S., and Scaletti, C., 1992: Improvisation Builder: Improvisation as Conversation. In *Proceedings of the International Computer Music Conference*. Interational Computer Music Association.

Wessel, D., 1979: Timbre space as a musical control structure. *Computer Music Journal*, **3**(2), 45–52.

Wessel, D., 1991: Instruments that learn, refined controllers, and source model loudspeakers. *Computer Music Journal*, **15**(4), 82–86.

Wessel, D., 2006: An enactive approach to computer music performance. *Actes des Recontres Musicales Pluridisciplinaires, Lyon, Grame*.

Wessel, D., and Wright, M., 2002: Problems and Prospects for Intimate Musical Control of Computers. *Computer Music Journal*, **26**(3), 11–22.

Wiener, N., 1948: *Cybernetics*. J. Wiley.

Windsor, W., 2001: Frequency Structure in Electroacoustic Music: Ideology, Function and Perception. *Organised Sound*, **2**(02), 77–82.

Wishart, T., 1993: From Architecture to Chemistry. *Journal of New Music Research*, **22**(4), 301–315.

Wishart, T., 1994: *Audible design: a plain and easy introduction to practical sound composition*. Orpheus the Pantomime.

Wishart, T., and Emmerson, S., 1996: *On Sonic Art*. Routledge. ISBN 371865847X.