Multi-Sensory Perception, Embodied Experience,

and the Development of an Interactive Multimedia Performance Environment

By Brian M. Knoth

B.A., State University of New York at Geneseo, 1998

M.A., Emerson College, 2003

A.M., Brown University, 2007

A Dissertation Submitted in Partial Fulfillment of the

Requirements for the Degree of Doctor of Philosophy

in the Computer Music and Multimedia Program

Providence, Rhode Island

May 2011

This dissertation by Brian M. Knoth is accepted in its present form

by the Department of Music as satisfying the

dissertation requirement for the degree of Doctor of Philosophy.

Date________________ ______________________________________

Dr. Todd Winkler, Advisor

Recommended to the Graduate Council

Date________________ ______________________________________

Dr. Joseph "Butch" Rovan, Reader

Date________________ ______________________________________

Dr. Chris Salter, Reader

Approved by the Graduate Council

Date________________ ______________________________________

Dr. Peter Weber, Dean of the Graduate School

ACKNOWLEDGMENTS

This dissertation project would not have been possible without the generous insight and expertise of my advisor Dr. Todd Winkler. In addition, the inspiration provided by my readers Dr. Joseph "Butch" Rovan and Dr. Chris Salter will fuel a lifetime of continued learning and creation.

The professionalism and calm demeanor of my collaborator Emily Beattie will always be remembered and greatly appreciated. I offer a warm thanks to her for allowing me the privilege to work with such an amazing artist.

Last but not least, I especially thank my wife, Sharon LaFrenaye, who throughout the entire process of pursuing this degree has been a constant source of love, support, and, most importantly, a grounding sense of humor.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

This document, the written portion of my dissertation project, is a companion to the interactive multimedia performance work *Unless* (2009), which featured movement artist Emily Beattie and was performed twice in May of 2009 and November of 2009 respectively. It details the inspirations and foundations for the creative project from a theoretical, technical and aesthetic perspective. In addition, it contains critical commentary and analysis, but should never be read solely as a review or summary of the work completed. Instead, it should also be considered an exploration of the underlying fundamentals that form the basis of an original interdisciplinary approach to aesthetic production, research and general inquiry.

This document comprises five chapters. Chapter 1 offers justification for the study of multi-sensory interaction with an interactive multimedia arts focus. This provides the theoretical and philosophical groundwork for the entire project, including discussions on the subjects of audiovisual interaction, music and movement, synaesthesia, multi-sensory integration and embodied cognition. My aim is to organize current thought in a way that is useful for artists, such as myself, exploring the interaction between physical motion, our perception of the body in space and technologically mediated music and imagery. In addition, this chapter introduces a central aspect of the research concerning the audience experience of kinesthetic empathy and embodied sensory simulation.

Chapter 2 is a discussion of the technologically oriented artistic concepts that helped to provide a foundation for the creative work and research. The ideas of immersion and interaction are highlighted throughout with the goal of defining these concepts in a way that clarifies their relationship to the development of dynamic digital multi-media performance environments. In the process I also survey artistic implementations of such ideas with a focus on physical interfaces and sensor based systems that translate human movement and expressive gesture into parameters that transform digital media elements in aesthetic contexts.

In Chapter 3, I survey three specific works of art that directly inspired the development of my dissertation project. Just as science and philosophy have brought us closer to a more complete understanding of multi-sensory experience and dynamic human-environment interaction, artistic practice has developed in a similar fashion. In light of these parallel movements I survey three pieces, *Messa Di Voce* (Levin and Lieberman 2003), *16 R(e)volutions* (Troika Ranch 2005) and *Schwelle* (Chris Salter et al. 2007), which represent contemporary examples of an holistic approach to multi-sensory aesthetics within the context of interactive digital performance.

Chapter 4 is a discussion of the aesthetic and technical development of *Unless* (2009) which represents the creative portion and main focus of my dissertation project. This process resulted in an original digital performance work featuring a movement artist and interactive multimedia environment. Special emphasis is placed on the interactive elements of the piece in this chapter as the specifics of artistic form are highlighted in

Chapter 5 along with an analysis and critique of each respective version of the piece as realized.

In Chapter 5 I discuss form, content and the specific performance contexts that gave rise to the two respective versions of *Unless* (2009). In the process I offer a thorough account of the work's formal structure, in general, and as related to each specific performance situation realized. In addition, through personal reflection and detailed accounts of performer feedback and audience experience, I offer a critical analysis of the work as a whole. The chapter ends with some concluding thoughts on the entire process from production through post-performance reflection.

Throughout this document there is an attempt to provide equal space for the theoretical, technical and aesthetic aspects of the work. This is a direct result related to my process, which engaged with the aesthetic concerns of interactive multimedia composition in addition to those aspects dealing with technical systems creation and conceptual underpinnings. Therefore, both the document and the work produced rise out of a convergence of interdisciplinary theory and practice, as well as an understanding of the close relationship between materials and content.

This document references the creative work in a variety of ways as the work is always discussed alongside the theoretical, technical and aesthetic ideas that gave rise to it. Thus, it is important that a reader fully engages with the work by periodically reviewing the documented performances on the supplied Data-DVD (See Appendix A for DVD contents). In addition, the accompanying Data-DVD includes Max/MSP/Jitter patches (customized software) that have been created by the author specifically for the work. These are discussed as compositional material, similar to the way in which one might discuss a musical score or a storyboard.

Most importantly, the main idea of this dissertation is the development of an original approach to interactive multimedia composition. As such, the work should be viewed both in its own right, and also as part of a larger perspective concerning the emerging discipline of interactive multimedia composition. In addition, the research component should offer a perspective that will help in developing a new interdisciplinary domain of knowledge relevant to both academic circles and the culture at large.

CHAPTER 1: INTERACTIVE PERCEPTION

Recently, the scientific community has exhibited a growing interest in understanding how the senses combine to create a unified percept. On a parallel track, the technologically mediated arts have been combining media, and the senses, in a rapidly evolving succession of forms resulting in a multi-media arts community of unprecedented proportions. This chapter surveys the most current research on multi-sensory interaction.

Along the way, I will frame the research in a way that creates a foundation for this dissertation project and the advancement of knowledge for multi-sensory arts practitioners (in general). My aim is to organize current thought in a way that is useful for artists, such as myself, exploring the interaction between physical motion, our perception of the body in space and technologically mediated music and imagery. Michel Chion would have to be credited as one of the first artists to thoroughly connect at least two of the senses, audition and vision, through the lens of film study. Therefore, it is with this perspective that I begin the investigation.

Audio-Vision

Composer and film theorist Michel Chion's ideas regarding the importance of sound in the audio-visual experience of narrative film provides a concrete foundation for experiential investigation into the perceptual dynamics of multi-modal art forms. Chion argues that film sound cannot be studied separately from its image and vice-versa. The sound and image combination inherently forms a feedback system of meaning in a film. Sound and picture, together, modify the perception of what is both seen and heard. However, Chion feels that most people are under the illusion that cinema is an "art of the image" which can be best explained through his concept of "added value" (Chion 1994, 5).

> Added value is what gives the (eminently incorrect) impression that sound is unnecessary, that sound merely duplicates a meaning, which in reality it brings about, either all on its own or by discrepancies between it and the image. (Chion 1994, 5)

"Synchresis", one such principle illustrating the phenomenon of "added value", is described by Chion as a reflexive perceptual phenomenon that forges an immediate link between the sonic and visual elements in a film (Chion 1994, 5). The perceiver assumes a tight network of interdependence, often unrelated to reality, such that sound is free to express something beyond the true nature of its origin. Chion "forged" the word by combining "synchronism" and "synthesis" (Chion, 1994, 63). It is a principle that clearly demonstrates the dynamics of a film experience or any other multi-sensory artistic form.

> It is also a function of meaning, and is organized according to gestaltist laws and contextual determinations. Play a stream of random audio and visual events, and you will find that certain ones will come together through synchresis and other combinations will not. The sequence takes on its phrasing all on its own, getting caught up in patterns of mutual

> reinforcement and phenomena of "good form" that do not operate by any simple rules. (Chion 1994, 63)

In addition to the forced relationships that Chion observes through sound/image synchronism in film experience, there are also scientific studies that provide evidence for inherent analogies between the senses, hinting at the existence of a human perceptual system finely attuned to structural similarities between the senses. So, we must leave Chion's realm of film research and enter the more abstract realm of cognitive science and perceptual research.

Audio-Visual Analogy

Lawrence Marks' research (1987) provides some of the seminal foundations for work exploring structural similarities between vision and audition.

> …the present study makes it clear that a wide range of sensory qualities of different modalities are interlinked and that correspondences between attributes or dimensions of visual and auditory experiences permeate functional responses to sensory stimulation. (Marks 1987, 393)

In this research cross-modal interactions were apparent in perceptual judgments related to light intensity/color brightness accompanied by pitch; pitch accompanied by light intensity/color brightness; light intensity accompanied by sound intensity; and perceptual judgments of form when accompanied by pitch. The research could not conclude, however, whether or not the interactions took place at a perceptual level of processing or after semantic encoding.

In a more current study, Marks et al. (2003) found a pattern of results suggesting that such cross-modal interactions result primarily from post-perceptual "decisional processes". Their research, however, did not completely deny the presence of cross-modal interactions operating at the sensory or perceptual level.

> One of the major challenges of research in multi-sensory processes is to develop a theoretical framework that is adequate to explain when, why and how cross-modal interactions may arise at different levels of sensory, perceptual and decisional processing in all of the sense modalities. (Marks et al. 2003, 143)

In response, Robert Walker's study addressed the experiential influences on audio-visual pairings. The study focused on previously validated cross-modal matches such as frequency with vertical placement, waveform with visual pattern, amplitude with size, and duration with horizontal length. A very large subject pool of over 800 people of varying ages, cultural, environmental and musical backgrounds was studied. The results suggested the following pattern of influence:

> Consistency of choice was determined predominantly by musical training. Age was also a factor, although to a lesser extent, and cultural and environmental effects were observed in subjects with less exposure to Western life-styles. (Walker 1987, 491)

Scott Lipscomb and Eugene Kim (2005) also investigated perceptual analogy in audio-visual pairings of simple stimuli and their results suggested differences in "degree of match" between audio-visual components of an audio-visual composite.

> The highest ratings of perceived match suggest the following pairings: pitch-location, loudness-size, and timbre-shape. Also emerging, however, is evidence of non-unitary relationships. For example, the visual attribute of color matched equally well with both pitch and loudness in the auditory domain...though the primary pairing receives a higher mean score, secondary relationships (such as loudness-location or pitch-size) may provide a source of variation that can be integrated into the audio-to-visual algorithm. Also, the auditory parameters of timbre, pitch, and loudness may all provide an acceptable matched pairing for the visual attribute of shape. (Lipscomb and Kim 2005)

Interestingly, their findings did not vary significantly for participants with different degrees of musical or visual training. Overall, the analogy-based research suggests the existence of inherent, albeit culturally/experientially influenced, relationships between seeing and hearing.

Music and Perceived Motion

In the systematic musicological research of Zohar Eitan and Roni Granot (2006), such evidence of perceptual analogy is expanded upon to include the visual representation of physical motion in relation to musical parameters. *How Music Moves: Musical Parameters and Listeners' Images of Motion* provides new evidence along with a succinct review of existing literature suggesting the existence of consistently perceived analogies for music and motion imagery (Eitan and Granot 2006).

The results as summarized (also see Table 1.1), suggest the following analogous relationships between musical parameters and motion imagery (some more obvious than others) (Eitan and Granot 2006, 231-7):

- ***Dynamics*** - Sonic dynamics were linked to motion imagery with regard to distance from an observer, verticality, speed and type of motion. For instance, crescendi evoked images of approaching and accelerating motion, while diminuendi were associated with images moving away, descending and slowing down. In addition, it is suggested that crescendi are associated with "running type movements" and diminuendi with the "falling or sliding type."

- ***Pitch*** - Pitch rises were correlated with images spatially ascending, moving away (contrary to the Doppler effect), accelerating, and exhibiting “higher energy”. Conversely, pitch falls were associated with spatial descents, slowing down, and "lower energy". Pitch falls were also associated with motion to the left, although pitch rise was not significantly linked with motion to the right.   In addition, pitch rises were linked to "running or walking type movements" and pitch descents with "falling types". Lastly, pitch intervals (dyads specifically) were loosely associated with distance change, whereby decreasing intervals were correlated with motion away.

- ***Tempo*** - Perceived tempo, a more obviously analogous parameter, showed positive correlations with regard to the sonic and image realms. Decreasing inter-onset intervals (IOIs), or faster tempi, in the musical stimuli were correlated with speedier motion images and vice-versa.

- ***Articulation*** - Changes from staccato to tenuto were associated with perceived changes in distance between the viewer and the imagery, whereby staccato events were associated with moving away and tenuto events with moving closer.

Musical Parameters	**Motion Imagery Visualizations**
Dynamics	
Crescendi	Approaching, Accelerated Motion, and Running Movements
Diminuendi	Moving Away, Descending, Slowing Down, Falling or Sliding Movements
Pitch	
Pitch Rises	Ascending, Moving Away, Accelerating, Increased Energy, Running or Walking Movements
Pitch Falls	Descending, Slowing Down, Decreased Energy, Motion to the Left, Falling Type Movements
Pitch Intervals	Distance Change
Tempo	
Fast Tempos	Fast Motion
Slow Tempos	Slow Motion
Articulation	
Staccato	Moving Away
Tenuto	Moving Closer

Table 1.1. A summary of Musical/Motion-Imagery relationships based on the results of Eitan and Granot (2006)

Some of the more interesting relationships uncovered from this research have an asymmetrical nature. The researchers articulate this point by stating, "Opposite 'directions' in this space do not present symmetrical contrasts (pitch rise is not the opposite of pitch fall, and crescendo is not the opposite of diminuendo), but rather, each stresses different spatial and kinetic associations" (Eitan and Granot 2006, 238-9). These less obvious pairings seem most exciting for future research exploring the implications for interactive art forms and composition.

Overall, the research reveals that the cognitive mappings between music, space, and kinetic association are much more complex than first assumed. Though some of the widely shared assumptions concerning music and motion are supported by these studies - most importantly, the notion that people do consistently map musical features into spatio-kinetic ones - the process is dynamic. As Eitan and Granot conclude:

> Musical parameters associate simultaneously with several aspects of motion, not only those traditionally associated with them...and musical space seems to be skewed in many different ways...Discovering such complexity while using extremely simple stimuli suggests that an even more entangled, more challenging musico-kinetic web shapes actual musical materials and listening contexts. (Eitan and Granot 2006, 242)

As will be discussed later, such research, with its close connection to mirror-neuron theories, offers added inspiration when considering an "embodied cognition" approach to the study and design of physical controllers for technologically mediated music and new media art forms.

Perceptual Coherence

Evidence also exists for "abstract" structural congruence (or multi-sensory metaphor) when considering relationships between general sonic "quality" and the surface characteristics/shapes of objects in the world (Kohler 1929, Ramachandran 2004, Handel 2006). The phonesthetic principle was first illustrated in Wolfgang Köhler's pioneering psychology experiment, in which he asked subjects, "which of the figures below represents the sound maluma, and which one represents the sound takete?" (Kohler 1929). Nearly all viewers responded with the same answer, such that the jagged shape was associated with the harsher sound of takete and the more gently undulating nature of the sound maluma was metaphorically linked to the curvy amoeba-like figure (See Figure 1.1).

V.S. Ramachandran has replicated the study numerous times and 98% of his research participants have responded in the same manner (Ramachandran 2004).   In this case, one must conclude that there is indeed some structural congruency between the senses when regarding shape and surface characteristics for image in relation to sonic "quality". It seems likely that the sound's envelope and our kinesthetic perception of creating those sounds with our mouths are both fundamental to these perceptual relationships, thus influencing the correlation.

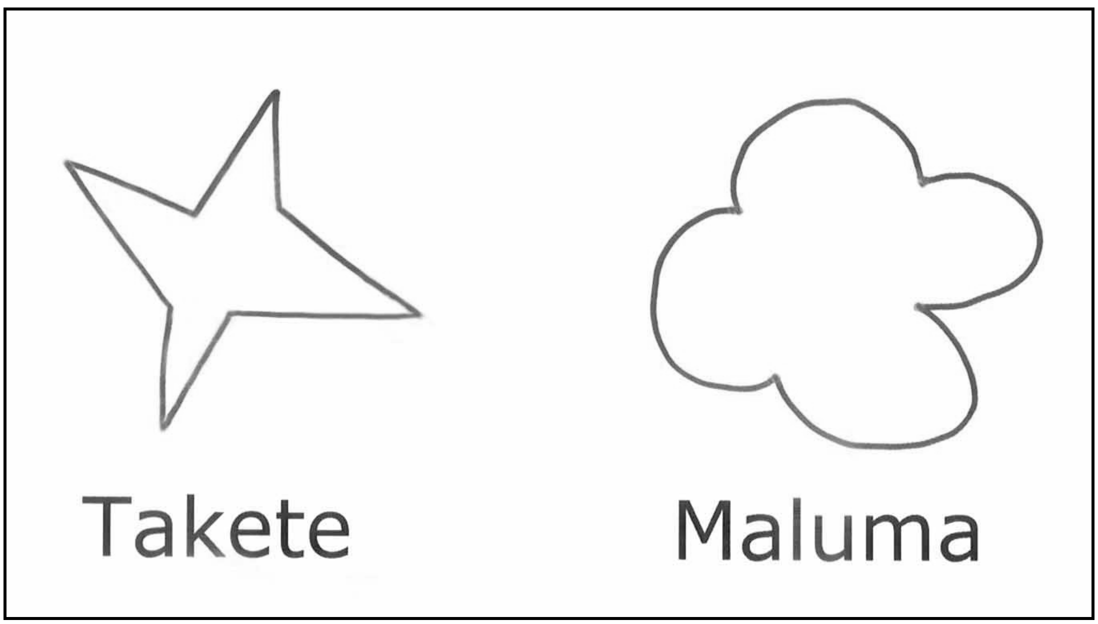

Figure 1.1. An artist's rendering of Kohler's (1929) stimuli by Manuel Krueger-Krusche (released to the public domain in wikipedia commons)

Obvious limitations exist in all of this research with regard to its relevancy to larger populations. Further studies could assess whether the results are applicable to wider and more complex musical/sonic contexts. In addition, it would be worthwhile to examine the effects of culture and native language on perceptions of sound/music and spatio-kinetic image congruence. Different cultures relate spatial concepts and musical dimensions in many varied ways. Thus, broadening the scope of such research into non-Western musical cultures could, "shed light on the interplay between cultural phenomena and possible inherent or acquired perceptual biases" (Eitan and Granot 2006, 241).

That being said, some researchers feel that the organizational principles in the brain are identical for all sensory modalities. This line of thinking is very inspiring for the interactive artist as such consistency could offer a concrete basis for creating abstract forms of art through the interplay of perceptual coherences and incoherences. Stephen Handel's book, *Perceptual Coherence*, articulates this idea from a cognitive science perspective:

> Perceiving is the construction of meaningful objects in the world. All such objects exist in a common spatio-temporal framework. All perceptual acts must also occur in the same framework, and that implies that there will be fundamental equivalences...(Handel 2006, 421)

The integration of information from multiple senses must conform on some level to create consistent predictions for perception.

All of the research combined thus far elucidates aspects of a core framework for interactive multi-sensory art forms such as *Unless* (2009), the project created in fulfillment of this dissertation. The following discussions on synaesthesia and illusions (or after effects) will continue this theoretical investigation by demonstrating the ways in which multi-sensory information can contradict and alter perception. Such information also provides insight into the mechanisms behind multi-sensory integration.

Synaesthesia

Perceptual abnormalities are another avenue to explore when investigating multi-sensory interactions. Synaesthesia could be considered an abnormal phenomenon that results from "cross-wiring gone haywire". By definition it is a neurologically-based phenomenon whereby stimulation of one sensory pathway leads to an involuntary experience in a second sensory pathway. While cross-sensory metaphors (e.g., "a loud color") are sometimes described as "synaesthetic", true neurological synaesthesia is reflexive, meaning that it occurs involuntarily and without any thought.

Ramachandran and Hubbard (2001) found that synaesthesia was a genuine perceptual phenomenon and not an effect based on memory associations (from childhood) or metaphor. Their research suggested that extensive cross-wiring between brain regions was the main factor. They thought that such a theory could explain the link between synaesthesia and creativity/metaphor (and, thus, the higher incidence of synaesthesia among artists and poets).

> Thus, our experiments on synaesthesia and our theoretical framework attempt to link several seemingly unrelated facts about the human mind. Far from being a mere curiosity, synaesthesia may provide a window into perception, thought and language. (Ramachandran and Hubbard 2001, 3)

In a study by Ward et al. (2006) research on a specific type of synaesthesia offered a window into more general cross-modal mechanisms. For the individuals that they studied auditory sensations automatically elicited a definite and reliable visual experience. For instance, a particular note on the piano "might be red but the note three octaves higher might be green" (Ward et al. 2006, 264). Such individuals are known to have “chromaesthesia”, also known as sound-color synaesthesia (or music-color synaesthesia).

Ward et al. (2006) studied these synaesthetes in relation to a “normal functioning” control group in order to look at how cross-modal processes may compare. They showed that synaesthetes chose more precise colors and were more internally consistent in their choice of colors when given a set of sounds varying in pitch, timbre and composition (single notes or dyads). But in spite of the differences, both the control group and synaesthetes appeared to use the same methods of evaluation for matching between auditory and visual domains (e.g., pitch to “lightness”). Overall, they concluded that such results provided evidence to support the notion that synaesthesia may recruit “...some of the same mechanisms used in normal cross-modal perception...” (Ward et al. 2006, 264).

Audio-Visual Illusion

Beyond that of synaesthesia, investigations into perceptual peculiarities involving illusions and after effects have offered even more insight into the structural interplay of auditory and visual perception. For instance, some audio-visual experiences are shown to be so interactive that one modality actually alters the perception of the other. One can’t help but see the connection to Chion’s ideas, but now with a scientific bent.

> Vision is believed to dominate our multisensory perception of the world. Here we overturn this established view by showing that auditory information can qualitatively alter the perception of an unambiguous visual stimulus to create a striking visual illusion. Our findings indicate that visual perception can be manipulated by other sensory modalities. (Shams et al. 2000, 788)

A subject was given a stimulus which was one white dot appearing on screen for one frame. However, when this stimulus was accompanied by two quick "beeps", the subject saw multiple dots, one after the other (Shams et al. 2000, 788). Thus, a type of visual fission occurred. “The simplicity and robustness of the illusory flashing phenomenon indicate that it reflects a fundamental and widespread property of poly-sensory mechanisms in the brain” (Shams et al. 2000, 788). Additional research has revealed fusion illusions of a similar kind along with an equivalent visually induced auditory illusion (Andersen et al. 2004).

Another audio-visual illusion, called the McGurk effect, shows that the perception of mouth movement is integrated into our auditory perception of speech (McGurk and MacDonald 1976). The syllables that we perceive depend on both auditory and visual information as a dynamic entity. It seems that our speech processing system makes use of all types of relevant information regardless of the modality.

> Participants are shown a person saying simple consonant-vowel syllables coupled with an acoustic recording of the same type of syllable. In some cases, the acoustic and visual information conflicted, yet few adults perceived the conflict. The participants perceived one of two syllables that were actually presented, or they perceived a syllable that was not presented at all, or they perceived a composite sound based on the two syllables. For example, if an acoustic (ba) was synchronized with a visual (ga), the majority of listeners reported (da). (Handel 2006, 417-18)

Interestingly, with the aforementioned fission illusion audition “dominates” vision, which is rare. More often, our “primary” sense, vision, dominates audition in creating multi-sensory illusions as in the McGurk effect (Andersen et al. 2004, 301-2).

Another illusion exists whereby auditory source position is influenced by visual cues. For example, if static auditory and visual stimuli are presented synchronously but from incongruent positions, the auditory source is pulled towards the location of the visual stimulus. The "ventriloquist effect", as it has been dubbed, has been reported to be very robust and is observed even when participants are explicitly trained to ignore the visual distracter (Vroomen and DeGelder 2003, 357). One can notice a direct relationship to some of Chion's ideas here as well, especially with respect to his principle of spatial magnetization (Chion 1994, 70).

Vroomen and DeGelder (2003) wondered if the same kind of effect could be witnessed for more dynamic stimuli.

> Investigations of cross-modal spatial interactions with dynamic events are, however, far less frequent, and, at present, generalizations from static to dynamic events may not be warranted...(Vroomen and DeGelder 2003, 357)

In response to the lack of evidence for dynamic event based spatial interaction Vroomen and DeGelder (2003) show an auditory motion after effect that is influenced by visual motion information. Participants listened to rightward-moving sounds with falling pitch alternated with leftward-moving sounds with rising pitch (and vice versa). Auditory after effects were shown to be bigger when a visual stimulus moved in the same direction as the sound than when no visual stimulus was presented.

However, when the visual stimulus moved in the opposite direction, after effects were reversed and became contingent upon visual motion. They also found that when visual motion was combined with a stationary sound, no after effect was observed (Vroomen and DeGelder 2003, 357).

> ...there is a direct perceptual link between the visual and auditory motion-processing systems. In this view, visual motion affects auditory motion processing, but visual motion does not affect motion processing of a stationary sound, as a stationary sound does not activate the auditory motion-processing system. (Vroomen and DeGelder 2003, 360)

Beyond the realm of spatial interactions between the senses, some researchers have found interesting temporal effects as well. Morein-Zamir et al. (2003) investigated whether sounds could influence the perception of vision over time. In their research, participants judged which of two lights appeared first. Their results found that presenting a sound before the first light and after the second light improved performance for making such determinations, as if the sounds pulled the perception of lights further apart in time. They also found that sounds intervening between the two lights led to a decline in performance, as if the sounds pulled the lights closer together. Overall, their results suggested a "temporal ventriloquism" phenomenon analogous to spatial ventriloquism (Morein-Zamir et al. 2003, 154).

Wassenhove et al. (2008) found another example of temporal distortion whereby vision affected the perception of auditory durations. When participants were presented with incongruent audio-visual stimuli, the perceived duration of auditory events could be shortened or lengthened by the presence of conflicting visual information (Wassenhove et al. 2008, 1). Overall, the findings support the existence of multi-sensory interactions in the perception of duration, and in this case vision also can modify auditory temporal perception. Such temporal studies may raise more questions than give answers, however:

> The ability to estimate the passage of time is of fundamental importance for perceptual and cognitive processes. One experience of time is the perception of duration, which is not isomorphic to physical duration and can be distorted by a number of factors. Yet, the critical features

> generating these perceptual shifts in subjective duration are not understood. (Wassenhove et al. 2008, 1)

In conclusion, these studies that create awareness about some of the anomalous features of our multi-sensory integration systems raise important questions about perceptual consciousness and our relationship to physical reality. Artists who want to exploit and/or experiment with sensory interactions are testing hypotheses in much the same way. It is in this pursuit that audio-visual and/or multi-sensory artists share many of the same aims as cognitive scientists.

On the one hand, the multi-sensory integration system that has evolved in humans to provide information on objects and events in space and time consistently organizes features of our environment to create order and detect form. Cross-sensory congruencies, for instance, add weight to the perception of true objects and events in our world.

> As an example, speech perception is facilitated by viewing a congruent articulating face. Also, detection of weak auditory and visual stimuli improves drastically when they occur in unison. In contrast, illusory percepts often occur when incongruent information is presented to two modalities. (Andersen et al. 2004, 301)

Interestingly, it’s these illusions, however, that can help uncover some of the mechanisms underlying sensory integration (Andersen et al. 2004, 301-2).

In artist’s terms, one could say that congruent structures would be integral to experiments at the threshold of consciousness, while incongruent structures help to raise questions about our relationship to reality. The exploration of congruent vs. incongruent relationships should be a core aspect of any multi-sensory art form. While many questions still remain about the true mechanisms underlying multi-sensory integration, some work in the audio-visual realm offers a little more insight.

Audio-Visual Integration

Ecker and Heller (2005) present a study that attempts to clarify the processes underlying sensory integration. With depth and path perception as a focus, they highlight the fact that such a process is accomplished through the interaction of multiple cues. Beyond the interactions existent between visual cues, such as perspective, relative size, texture, shadows, etc., visual and auditory cues also combine and modulate in order to formulate a unified percept (Ecker and Heller 2005, 75). As they remark:

> Even an event that seems as intuitively `visual' as the perceived path of a ball can be modulated by auditory cues. Thus, a complete account of event perception, even for nominally visual events, needs to consider auditory information as part of the system. (Ecker and Heller 2005, 75)

Although it seems obvious that the perception of objects and events are influenced by input from multiple senses, the jury is still out on how such information is combined, "especially where sound is integral to the event" (Ecker and Heller 2005, 74). Ecker and Heller suggest that such a combination occurs at a level before conscious inference at a cortical site supporting multi-modal perceptual processing of objects and motion (Ecker and Heller 2005, 74). As Hein et al. (2007) point out, another piece of the puzzle is determining how "semantic" congruency can affect audiovisual integration.

> The cortical integration of auditory and visual features is crucial for efficient object recognition. Previous studies have shown that audiovisual (AV) integration is affected by where and when auditory and visual features occur. However, because relatively little is known about the

> impact of what is integrated, we here investigated the impact of semantic congruency and object familiarity on the neural correlates of AV integration. (Hein et al. 2007, 7881)

In a study utilizing functional magnetic resonance imaging they identified brain regions involved in the integration of both familiar animal sounds and images (of a congruent and incongruent nature) and of unfamiliar (artificial) sounds and images in arbitrary combinations.

> Unfamiliar object images and sounds were integrated in the inferior frontal cortex (IFC), possibly reflecting learning of novel AV associations. Integration of familiar, but semantically incongruent combinations also correlated with IFC activation and additionally involved the posterior superior temporal sulcus (pSTS). For highly familiar semantically congruent AV pairings, we again found AV integration effects in pSTS and additionally in superior temporal gyrus. These findings demonstrate that the neural correlates of object related AV integration reflect both semantic congruency and familiarity of the integrated sounds and images. (Hein et al. 2007, 7881)

These findings relate to Susan Broadhurst's ideas on neuroesthetics, especially regarding the use of abstracted imagery in art. According to her, works without representational foundations seem to activate fewer areas of the brain than those with recognizable forms. However, such abstract realms, like the digital art practices of Troika Ranch and Palindrome for instance, offer new links between multi-sensory components, as they conflict with most people's prior experience, activating a specific area of the frontal lobe which appears to deal with the resolution of experiential conflict (Broadhurst 2007).

Such multi-modal forms that also incorporate the body in integral ways inspire the following research. Much of the current literature on multi-sensory integration focuses on audio-visual relationships as shown above - and thus many artists influenced by these perspectives do the same, including Chion and Broadhurst. It is obvious, however, that somatosensory integration processes (including touch and proprioception) are undeniable factors that must be addressed as well. In fact, it will ultimately be suggested that "embodied" approaches to perception and cognition may actually be the future of research in this domain. Artists may be ahead of the curve in this regard.

Somatosensory Interactions (and Illusions)

It is reasonable to assume that precise knowledge of bodily orientation is critical for movement because the brain needs to compare actual and intended positions of the body for any "goal-directed" movement (Wolfe et al. 2007, 879). This awareness arises from an interaction of spatial vision, touch, and proprioception. Most studies, which have focused on position sense for the arm or hand, have found that spatial vision of the limb generally dominates the interaction resulting in the perception of the arm's position (Wolfe et al. 2007, 879). When vision and somatosensory inputs are incongruent, position sense is primarily determined by the visual information and largely disposes proprioceptive input (Wolfe et al. 2007, 879).

It has also been reported that subjects seeing a fake hand brushed synchronously with their own (concealed hand) report an illusion that the tactile sensations they feel actually emanate from the fake one (Botvinick and Cohen 1998, 756). The perceived orientation of their own hand is distorted to match the fake hand so that, in effect, the fake hand feels as if it is their own. The illusion is negated, however, when the visual sense shows that the fake hand is in a posture inconsistent with the orientation of the subject's body (Pavani et al. 2000, 353). Congruent visual and somatosensory sensations appear to be so firmly coupled that if the two sources of information are in conflict, the experience is driven by vision (Wolfe et al. 2007, 879).

Another aspect of somatosensory dynamics not yet highlighted concerns touch. The awareness of tactile information seems inherently correlated with that of spatial information. Conscious tactile sensations seem to relate directly to positions on the body in external space rather than on the skin surface where the tactile receptors are actually located. This is not surprising when considering consciousness for visual events, as it seems obvious that we locate the visual percepts in external space rather than on the surface of our retina (Gallace and Spence 2008, 24).

> However, what is particularly unexpected and striking here is not only that we do not perceive the stimulation at the location of the receptors on the skin, but that we can consciously perceive tactile stimuli in spatial positions where our body is not physically present! (Gallace and Spence 2008, 25)

This is not to say that visual consciousness ‘‘dominates’’ over tactile consciousness, but that multi-sensory integration may occur, “before access to consciousness is provided” (Gallace and Spence 2008, 25).

As was discussed before, phenomena related to multi-sensory integration, such as the visual capture of tactile sensation highlighted in the fake hand illusion, or the ‘‘ventriloquist effect’’ occurring between vision and audition are truly pre-conscious. These observations suggest, “that while the consciousness of a stimulus presented in a given sensory modality can be enhanced by multi-sensory integration, this latter step of information processing does not necessarily enter our awareness” (Gallace and Spence 2008, 26). Such phenomena raise another important question.

> Should consciousness (and its neurological substrates) be considered as being unisensory, multisensory, or even amodal in nature? In an attempt to give a speculative answer to this question, we suggest here that only one sensory modality enters consciousness (or that only one unimodal system is active at a level of neural activation able to support consciousness) at any given time, but that information presented from other sensory modalities can speed-up or slow-down, and more generally affect (or modulate), this process as a function of their spatial and temporal characteristics. (Gallace and Spence 2008, 26)

Multi-Sensory Integration

Contrary to even recent belief, sensory interactions are the rule as opposed to the exception in brain function (Shimojo et al. 2001, 505). According to Holmes and Spence (2005), the superior colliculus is a well-known site of multi-sensory integration and recent research on this structure enhances understanding of the mechanisms underlying multi-sensory integration in the brain. The function of the superior colliculus seems to be focused on the generation and control of eye and head movements based on input from different senses. While cells in the upper layers are visual, the majority of cells in the deeper layers respond to combinations of visual, tactile, and auditory information.

> Multisensory integration in the superior colliculi has been understood for many years in terms of several simple rules. The neural responses to multisensory stimuli tend to be enhanced when the stimuli in different senses come from approximately the same location (the 'spatial rule'), when they occur at approximately the same time (the 'temporal rule'), and/or when at least one of the two stimuli is by itself only weakly effective in exciting a neuron (the 'inverse effectiveness rule'). (Holmes and Spence 2005, 762)

Although a seemingly simple framework, the mechanisms giving rise to these integrative rules are only now being uncovered. For instance, the combination of visual and auditory stimulation presented together sometimes results in more "activity" than the sum when the stimuli are presented individually. This phenomenon is called 'superadditivity,' as the multi-sensory whole is greater than the sum of its unisensory parts. It is felt that superadditivity, however, is just one of several computations through which multisensory integration enhances the neural representation of sensory signals and the behaviors that depend on them. In summary, the neurons in the superior colliculus represent salient visual, auditory, and tactile stimuli and contribute to the formation of motor commands to orient (eyes, head, ears, body) to these stimuli (Stanford and Stein 2007, 792).

It seems that the superior colliculus is an important mechanism to consider for digital artists who experiment with the creation of new digital music instruments and immersive environments (topics to be covered in Chapter 2). In addition, the simple rules that guide multi-sensory integration are important factors to consider when thinking about the audience's experience of novel art forms.

Embodied Cognition

The brain is often described metaphorically as the computer and the body as its "all-purpose tool" (Bennet 2008). However, an expanding body of research suggests that something more interactive is going on.

> …children can solve math problems better if they are told to use their hands while thinking. Another recent study suggested that stage actors remember their lines better when they are moving. And in one study published last year, subjects asked to move their eyes in a specific pattern while puzzling through a brainteaser were twice as likely to solve it. (Bennett 2008)

This new model of mind is often called "embodied cognition”. In the embodied view, perception and cognition are much more than isolated brain functions. What's going on inside the brain integrally depends on what's happening with the body and how it is interacting with its environment. Rolf Pfeifer, director of the artificial intelligence lab at the University of Zurich articulates:

> If you want to teach a computer to play chess, or if you want to design a search engine, the old model is OK...but if you're interested in understanding real intelligence, you have to deal with the body. (Bennett 2008)

As The University of Wisconsin's Laboratory of Embodied Cognition puts it, "Ago ergo cogito”: “I act, therefore I think.” (Bennett 2008)

Building from this foundation, scientists now argue that human traits such as empathy, concepts such as space-time and systems such as language, are ultimately traceable to the idiosyncrasies of the human body (Bennett 2008). This emerging paradigm shifts the focus of any debate on consciousness and behavior towards a more dynamic understanding of the mind and human-environmental interaction. Moving one step further, it is also necessary to frame cultural manifestations such as musical experience, with its close links to language, within this scope.

The embodied cognition framework counters several centuries of thought on the human mind. Descartes, for instance, argued that the mind was an incorporeal entity that interacted with the body. This concept of mind-body separation remained the dominant model up through the 20th century (Bennett 2008). In opposition, Maurice Merleau-Ponty (1962) offered a more holistic perspective on the mind, and specifically perception, which thoroughly considered the body and its actions as part of the process. In addition, Merleau-Ponty considered the co-dependency of the human body/mind system with its environment.

Ultimately, the discovery of "mirror neurons" revolutionized thought on the topic as it was found that a system of neurons could respond similarly while one acted on their own or observed someone else performing the same action. Mirror neurons suggest that acting and thinking are not so separate after all (Bennett 2008). Researchers have also found that individuals who scored higher on an “empathy” scale activated this system more strongly, adding evidence to claims linking the mirror system with human traits such as compassion (Gazzola et al. 2006, 1824).

> We do not need to reproduce the behavior of others in full detail in order to understand its emotive meaning, just as action understanding does not require the actions to be replicated. Even if they involve different cortical circuits, our perceptions of the motor acts and emotive reactions of others appear to be united by a mirror mechanism that permits our brain to immediately understand what we are seeing, feeling, or imagining others to be doing, as it triggers the same neural structures (motor or visceromotor respectively) that are responsible for our own actions and emotions. (Rizzolatti et al. 2008, 190)

Mirror neurons may not only be triggered by visual observation. An auditory-motor mirror system has also been suggested (Lahav et al. 2007, 308). Gazzola et al. (2006) found that auditory mirror neurons responded while monkeys both performed mouth actions and while they listened to the sounds of such actions. They suggested that such a system might be critical for auditory action understanding and language evolution. In addition, preliminary FMRI evidence suggested that a similar multi-modal system could exist in humans (Gazzola et al. 2006, 1824).

These ideas fit with Rolf Godøy's model of music perception and cognition. His concept of "motor-mimesis" is the basis for this theory.

> I believe this points in the direction of what I would like to call a *motor-mimetic* element in music perception and cognition, meaning that we mentally imitate sound-producing actions when we listen attentively to music, or that we may imagine actively tracing or drawing the contours of the music as it unfolds. (Godøy 2003, 318)

Gødoy's ideas make sense when considering emerging views on the motor-system from the neurosciences.

> …the motor system has a number of other functions, over and above organizing our motor behaviour. These include certain processes that are normally considered as being higher order and therefore attributed to cognitive systems; for example, the perception and recognition of actions carried out by others, imitation, and gestural and vocal communication; it may well be that the primary neural substrate of these processes lies in the motor system. (Rizolatti et al. 2008, 20)

These mirror-neuron inspired ideas and the more encompassing theoretical framework of "embodied cognition" relates to the ecological approach to perception brought forth by J. J. Gibson. Gibson (1979) fundamentally stressed the importance of looking at perception as an active process. In addition, he was very sensitive to the larger system of human-environment interaction that guided sensory experience. His "Theory of Affordances" is instrumental to the argument:

> The *affordances* of the environment are what it offers the animal, what it provides or furnishes, either for good or ill. The verb *to afford* is found in the dictionary, but the noun *affordance* is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment. (Gibson 1979, 127)

Gibson's views could be considered post-cognitivist. The action component, especially, in his theory of perception counters the perspective of representationalism, which is a key assumption behind the cognitivist point of view. Representationalists purport that we do not perceive the external world directly. Rather, we only know our interpretations of objects in the world. Francisco Varela, Evan Thompson, and Eleanor Rosch (1992) have articulated a similar post-cognitivist perspective on perception within their enactive approach. "Enactivists" fundamentally criticize representational views of the mind, emphasizing the importance of embodiment and action in relation to perception and cognition (Varela et al. 1992, 205). Daniel Dennett (1993) concisely clarifies the enactivist/cognitivst divide in his review of Varela et al.'s (1992) book:

> Question 1: What is cognition?
>
> Cognitivist Answer: Information processing as symbolic computation--rule-based manipulation of symbols.
>
> Enactivist Answer: Enaction: A history of structural coupling that brings forth a world.
>
> Question 2: How does it work?
>
> Cognitivist Answer: Through any device that can support and manipulate discrete functional elements -- the symbols. The system interacts only with the form of the symbols (their physical attributes), not their meaning.
>
> Enactivist Answer: Through a network consisting of multiple levels of interconnected, sensorimotor subnetworks. (Dennett 1993, 122)

Although the ecological and enactive approaches to cognition and perception are very similar, there are still some fundamental differences that the enactivists' feel are important to delineate. Specifically, the enactivists feel that the environment does not exist a priori and is thus dependent on histories of structural coupling between human and environment.

> In a nutshell, then, whereas Gibson claims that the environment is independent, we claim that it is enacted (by histories of coupling). Whereas Gibson claims that perception is direct detection, we claim that it is sensorimotor enactment…by specifying the sensorimotor patterns that enable action to be perceptually guided, and so we build up the theory of perception from the structural coupling of the animal. (Varela et al., 204)

It is not my intention to define the philosophical battlegrounds of the cognitive science discipline, but making these distinctions is important. I admire the humor of Daniel Dennett when approaching this dangerous territory.

> When you get right down to it, many people just do not like the idea of their minds as computers, and hence are strongly inclined to endorse any champion who will stand fast against this deplorable juggernaut. (Dennett 1993, 121)

Getting back to basics, this is the crux of any position that attempts to bring the human body and human-environment interaction back into the mix of cognition and perception. This, in a nutshell, is my main intention and the focus of my inspiration for the creative portion of this dissertation project.

That being said, the most exciting idea to materialize from this scholarship (for my purpose at least) may be the conceptualization of human-environment interaction as a self-organizing dynamic system (Kelso 1995). William Warren (2006) has developed a theoretical framework called behavioral dynamics. Humans and their environments are treated as two dynamical systems coupled mechanically and “informationally”. In doing so, Warren argues that organized behavior need not be attributed to a centralized control system, but an agent-environment system (with particular reverence to Gibson).

> Interactions between the agent and environment under physical, informational, and task constraints give rise to the behavioral dynamics, an evolving vector field that reciprocally captures the agent’s behavior. The

> agent modulates the dynamics of the system in which it is embedded via the lever at its disposal—a control law—to enact a stable and adaptive task solution. The exhibited behavior is thus genuinely emergent. It is in this sense, as Gibson (1979) foresaw, that behavior can be regular without being regulated. (Warren 2006, 385)

Such ideas offer great foundations for the systematic investigation of interactive music and new media art realms. Marc Leman, for instance, has explicitly championed the embodied cognition perspective in his book, *Embodied Music Cognition and Mediation Technology* (2007). He provides the first thorough look at the relationships between active perception, mirror neuron theory, ecological psychology, enaction, and interactive music and new media technologies. The connections are clear as interactive approaches to music and new media art inherently embrace the dynamics of human behavior and mediated environments.

CHAPTER 2: IMMERSION, INTERACTION AND MULTI-SENSORY DIGITAL PERFORMANCE

This chapter investigates the technical foundations for my project in relation to the ideas of immersion and interaction. My goal is to define and explore these ideas in a way that clarifies their relationship to the development of dynamic digital multi-media performance environments. In the process I will also survey artistic implementations of such ideas with a focus on physical interfaces and sensor based systems that translate human movement and expressive gesture into parameters that transform digital media elements in immersive aesthetic contexts.

Some questions that have guided the investigation include:

- What is the meaning of immersion in relation to this dissertation project?
- How should interaction be defined?
- What techniques have been used to create immersive digital music and image environments?
- How does a participant/performer's physical movement relate to expressive interaction with sound and image?

- How can techniques related to modality-specific sound and image interaction be combined for dynamic multi-modal experiences?

Arts based implementations of Virtual Reality (VR) offer an interesting starting point.

Immersion and Interaction

For years virtual reality was bogged down by "cumbersome gear and prohibitive costs " (Hansen 2006, 1). Recently, an economically and artistically motivated movement has emerged that offers an alternative to the old paradigm of virtual reality. Dubbed "mixed reality", it does away with virtual reality's disembodied approach using head-mounted displays (HMDs), instead, creating experiences that merge the virtual with existing physical spaces. As Mark Hansen states in his book *Bodies in Code*, "today's artists and engineers envision a fluid interpenetration of realms" (Hansen 2006, 2). The virtual is no longer a distinct space with rules of its own, but a "space full of information...added to and transformed...as the user navigates...real space" (Grau 2003, 247).

The more organic and interpenetrative paradigm of "mixed reality" will serve as a focal point for this investigation as I survey the approaches available to artists exploring human interaction with immersive sound and imagery. Essentially, the body becomes central to the investigation of the interface.

> In what amounts to a position statement for the mixed reality movement as a whole, Fleischmann and Strauss speak of "turning the theory on its head that man is losing his body to technology"; as they see it, "the interactive media are supporting the multi-sensory mechanisms of the body and are thus extending man's space for play and action" (cited in Grau 2003). (Hansen 2006, 2-3)

Hansen cites interactive arts pioneer Myron Krueger as a father figure to the mixed reality movement. For Krueger, the development of immersive environments needs to put us in touch with our most primitive perceptual capacities. “Whereas the HMD folks thought that 3D scenery was the essence of reality, I felt that the degree of physical involvement was the measure of immersion” (Krueger 1993, 147). For Krueger’s notion of “artificial reality” to succeed (which could be considered a direct precursor to mixed reality), the body’s involvement must be unencumbered (Krueger 1993).

> The theoretical underpinnings of Krueger’s alternate vision of “artificial reality” have found their practical fruition in today’s mixed reality environments, and the “first generation” model of VR as a disembodied hyperspace free of all material constraints simply no longer has any purchase in the world. (Hansen 2006, 4)

Accordingly, my process has and will favor technologies that encourage free physical involvement with media environments rather than highly constraining physical interfaces such as head mounted displays and constrictive body-attached sensor interfaces.

Krueger’s vision contrasts with other notions of the “natural interface” as a more visually immersive data-rich media environment, as offered in HMD based VR (Hansen 2006, 4). Although I am not totally against exploring the more “data-rich” environment per se (as offered with these technologies), I feel that physical freedom and unencumbered interactions offer experiences that more than make up for the lack of full 3D visual immersion in traditional VR.

Glowflow (1969), a computer art project by Krueger and others at the University of Wisconsin, offered a glimpse into the future of mixed reality art making. Through this work Krueger began to define his own “responsive environment” paradigm (Krueger 1977, 423). Although responses were delayed, and the authors were ultimately uninterested in the audience’s ability to perceive their effect on the work, people still tried to make sense of their relationship to the environment.

> What proved compelling about the environment and most captivated Krueger’s aesthetic interest was the way that viewers inevitably, one might even want to say “naturally”, attempted to make sense of the environment by coupling intentional gestures (e.g., speaking or moving with the (mostly arbitrary) patterns of light in the tubes. (Hansen 2006, 29)

This realization was important in the sense that participant’s do not physically have to have an effect on a mediated environment to feel connected to it. Thus, artist’s can harness the natural human desire for connectivity to create situations that simulate real interaction and play with the conceptual spaces between the real and the perceived.

> This emergent and unexpected phenomenon alerted Krueger to the possibility and value of exploiting the margin of indetermination afforded by experience with new technologies and media environments. (Hansen 2006, 30)

More importantly, this ingrained desire for interaction inspired Krueger to think of the mechanisms of interaction and response as artistic mediums in and of themselves (Krueger 1977).

After a series of artistic experiments, including *Metaplay, Psychic Space and Maze,* Krueger got funding for a project called *Videoplace* (1970-1984).

> *Videoplace* works by capturing an image of the visitor’s movement, only in this case the image presents the outline of the visitor’s body processed (and distorted in various ways) by the computer. Several different programs facilitate viewer interaction with the schematic traces of her bodily movement; for example, they allow her to fill in the space within the lines with colored images of body parts or to interact with temporally

> divergent and continually reverberating captures of her movement… (Hansen 2006, 35)

Kruger thought of *Videoplace* not as a single work, "but as a general facility for exploring all the possibilities of the medium…" (Krueger 1977, 430).

Krueger's new medium was "based on a commitment to real-time interaction between men and machines" (Krueger 1977, 430).

> What made this possible and served to distinguish Krueger's environments from the majority of art environments from the 1960s was Krueger's insistence on clarifying for the visitor precisely how his or her actions called forth reactions from the environment. (Hansen 2006, 31)

Krueger's notion of "response as the medium" introduced something in stark contrast to the artistic concepts of the time. In addition, his ideas, to this day, still counter some contemporary paradigms of interactive art. As Hansen states, they were "more codified than the radically open-ended environments of process art and happenings and yet more open and flexible than the preprogrammed response repertoires that one finds in so-called interactive cinema and hypertext" (Hansen 2006, 31).

What is key to Krueger's notion of interactivity is the privilege afforded the human body throughout the experience.

> Differences notwithstanding, the point of all of his environments is to facilitate new kinds of world-construction and inter-subjective communication. If these environments do still serve as interfaces to the computer, they do not do so, as do all technicist conceptions of the "Human-Computer Interface" (including the mouse and the graphical-user interface), by instrumentalizing – and thus reducing – embodied enaction; here, rather, it is the technology that remains instrumentalized and human action that gets privileged. (Hansen 2006, 31)

Philosophically, Krueger's environments fall in line with the phenomenology of Merleau-Ponty, the enactivist ideas of *The Embodied Mind* (Varela et al. 1992), and embodied cognition frameworks. What seems evident from a historical perspective is the introduction of a new trend in art away from passive ideas about perception towards notions of active world construction and ongoing feedback systems between humans and their environments.

In *Transforming Mirrors: Subjectivity and Control in Interactive Media,* David Rokeby explicitly breaks down these feedback system components:

> The interactive system responds to the interactor, who in turn responds to that response. A feedback system is created in which the implications of an action are multiplied, much as we are reflected into infinity by the two facing mirrors in a barbershop…Whatever the differences, like Cage, interactive artists are looking for ways to give away some of the control over the final actualizations of their works. (Rokeby 1996)

The feedback systems of interactive multimedia art environments have something in common with the compositional processes of John Cage, when one considers the idea of control in relation to art making. Interactive artists allow for a certain amount of indeterminacy in the final realization of a work in order to create the potential for audience/performer-activated results.

> The immediate precedent for this is found in John Cage's chance compositions. In each of these works, Cage defined a set of rules and then used the tossing of coins to choose a specific composition from the range of possibilities allowed by these rules… In later works, Cage further removed himself from the compositional process through what he called 'indeterminacy'. In these works, the rules themselves were left intentionally ambiguous, leaving them open to subjective interpretation by the performers as well. (Rokeby 1996)

Thus, the work for an interactive artist is never finished. The work itself is ultimately a means for the realization of potentialities and relationships between an "interactor" and a mediated environment. These relationships are, in Rokeby's words, "simplified symbols that are used to refer to the more complex operations of what we call 'real' life" (Rokeby 1996).

Today, Krueger's seminal ideas of the 1970's and Rokeby's extensions still have great influence over the work of interactive artists. However, as the technologies involved in creating relational experiences have become more ubiquitous, artists strive to find more compelling dynamics between the interactive mechanisms involved and the overall experience of the work. Thus, the philosophical implications of such practice have become an even more important topic of contemporary digital arts theory. Passages from the book *Interact or Die* (2007), an Ars Electronica media research award winner, represent some current waves of thought on the subject. Arjen Mulder, one of the editors of this influential book, offers a particularly interesting perspective on the power of interactivity in today's culture.

> Interactivity liberates us from the hundred years of solitude of twentieth century art. Interactive art makes the realization flash up that there is actually no "I." No one is an island; you are produced by others just as much as you produce them. You are created by objects as much as you create them. Since interaction means changing each other, and only that which interacts with us is alive for us, we are changing everything around us as much as we are being changed by it – nature and technology, loved ones and strangers, room and city, bed and world. (Mulder 2007, 69)

Even more pronounced in his perspective is the idea that consciousness, experience, life in general is a dynamic flow of relations, never static, always changing. To be alive is to interact. One could argue that artistic practices that do not take this "reality" into account deny the emergent nature of existence. Brian Massumi expands on this idea in an interview with Mulder:

> Art brings back out the fact that all form is necessarily dynamic form…Art is the technique for making that necessary but normally unperceived fact perceptible…(Brian Massumi in Massumi and Mulder 2007, 77)

Again, one can see the relationship between current ideas of action-perception dynamics, embodied cognition and contemporary interactive arts practice. For my own purposes, these ideas have opened me up to considering all facets of an experience in the creation of dynamic multi-media/multi-sensory work.

> What interactive art can do – what its strength is, in my opinion – is to take the situation as its "object." Not a function, not a use, not a behavior, not an action-reaction. But a situation, in all its complexity. It can take a situation and potentially "open" the interaction it affords, so as to make it truly an event…It's essentially a question of framing in a way that doesn't just elicit behaviors but opens the situation out onto the qualitative level of thinking-feeling where expanding potential appears, so that being appears in the act of becoming. (Brian Massumi in Massumi and Mulder 2007, 79)

Massumi offers an example possessing this strength as he discusses *Artaudian Lights* (2006), a movement and responsive architecture experiment created by Michael Montanero and Harry Smoak at Sha Xin Wei's Topological Media Lab. In the process, Mulder and Massumi's discussion turned to the realm of digital performance in relation to the larger context of interactive art.

Referring to the piece, Mulder remarked, "In the dance example, the interaction is staged. There's the traditional theatrical separation between the performers and the audience. The interaction is only between the performers and the technology" (Mulder in Massumi and Mulder 2007, 89). To this, Massumi responded:

> While it is true that the audience was not in on the interaction, they were in on the relation. You couldn't not see the relation between movement and vision being recomposed before your eyes. You felt the dancers making an actual sight of their bodies' imperceptible movement talents. Kinesthesia was not only fused with movement, it was like vision itself was emerging from it. Body-vision. Kinesthesia was making vision appear in the bubble, and the bubble was making bodily qualities of movement appear – a double capture of vision by movement and movement by vision, in a unique composition. (Brian Massumi in Massumi and Mulder 2007, 89)

Here, Massumi refers to the human capacity to simulate kinesthetic experience via the mirror neuron system. One could argue that even though an audience may not be "in on" the interaction between a performer and digital performance system, the mirror neuron system allows for a kind of simulated experience of the relationship. This notion of kinesthetic empathy is an important concept to consider in the creation of interactive arts based "situations" involving a privileged performer and audience.

Such ideas are particularly important in relation to the design of interactive music systems or digital music instruments, where more often than not the mode of presentation is performance based and an audience must make sense of the relationship between a technologically mediated sound world and a human performer's physical gesture. As we know from audio related mirror-neuron research, sound alone can also elicit such simulated experience and induce empathetic responses in much the same way observations of physical movement can.

Immersion and Interaction: The Digital Music Instrument

This dissertation project has been especially influenced by research regarding the *digital music instrument* (DMI), a term coined from the computer music literature. In fact, any artist who combines physical interaction with media elements, in general, owes a debt to the researchers working in this area. It is important to clarify the difference between an ordinary instrument and one that qualifies as a DMI.

> …the term digital musical instrument is used to represent an instrument that includes a separate gestural interface (or gestural controller unit) from a sound generation unit. (Wanderley and Depalle 2004, 633)

A traditional instrument inherently combines the physical interface with sound production. The expressive qualities of a guitar, for instance, have everything to do with the way one's fingers interact with the string on the fret board. With a DMI, however, these relationships are more abstract as the sonic result is actually removed from the physical interface. There are actually three components that dynamically combine to produce a sonic event.

> The term *gestural controller* can be defined here as the input part of the DMI, where physical interaction with the player takes place. Conversely, the sound generation unit can be seen as the synthesis algorithm and its input parameters. The mapping layer refers to the liaison strategies between the outputs of the gestural controller and the input controls of the synthesis algorithm. (Wanderley and Depalle 2004, 633)

The gestural control unit must provide a believable interface between physical movement qualities and the sound produced. It is important for DMI designers to consider their own experience of traditional instrument performance gesture in addition to the intuitive physics and psycho-physics we accumulate through experience with the world.

> In order to devise strategies concerning the design of new DMIs for gestural control of sound synthesis, it is essential to analyze the characteristics of actions produced by expert instrumentalists during performance. These actions are commonly referred to as *gestures* in the

> musical domain. In order to avoid discussing all nuances of the meaning of gesture, let us initially consider *performer gestures as performer actions produced by the instrumentalist during a performance...* (Wanderley and Depalle 2004, 633)

Using the term gesture can be confusing, as it can be associated with many different meanings. From here on any mention of the word gesture refers to the DMI performer's actions while performing with the instrument.

In discussing approaches to gestural acquisition, Wanderley and Depalle point to two methods that are specifically relevant to my own project's development. They are direct acquisition and indirect acquisition.

> *Direct Acquisition*: Direct acquisition is performed by the use of different sensors to capture performer actions. Depending on the type of sensors and on the combination of different technologies in various systems, different movements may be tracked. (Wanderley and Depalle 2004, 634)
>
> *Indirect Acquisition*: As opposed to direct acquisition, *indirect acquisition* provides information about performer actions from the evolution of structural properties of the sound being produced by an instrument. In this case, the only sensor is a microphone, i.e., a sensor measuring pressure or gradient of pressure. (Wanderley and Depalle 2004, 634-5)

As will be seen, one particular DMI that I have developed for the project utilizes both modes of gestural acquisition in order to create more dynamic variation in the sonic textures resulting from performer interaction. In a sense, using both modes of acquisition, simultaneously, allows for a certain amount of "built-in" indeterminacy for both the performer and the audience experiencing the performer's interaction.

The specific DMI paradigm that I have followed falls closely in line with the *alternate controller*, as defined by Wanderley and Depalle (2004). Alternate controllers, such as *the Hands* by Michel Waiswisz (1985) for instance, "allow the use of other gesture vocabularies than those of traditional instrument manipulation, thus being in principle less demanding for nonexpert performers" (Wanderley and Depalle 2004, 636). Yet, performers still have to develop specific skill sets for interacting with such new instruments (Waisvisz 1985). Some developers believe that more intuitive and user-friendly approaches to new interfaces could usher in a new wave of generalized tools for musical performance while others claim that ease of use may actually impede the expressive capabilities of a new instrument (Wanderley and Depalle 2004, 636; Mulder 2000). In other words, effort may be the key to the expression (Ryan 1991; Ryan 1992). Effort may only need to be perceived by an audience, however. A performer using an alternate controller may just need to "sell" the interaction through dramatic expression. The role of dramatic effect, or ancillary performance gesture, is even important to the expression of a traditional instrumentalist.

Once the gesture is acquired, the next stage of interaction occurs between the data generated and the sound generation module. The design of DMI sound generation modules has benefited from the field of sound synthesis. Sound synthesis, which has been studied for at least a half century is generally considered to be the study of sound representation "that leads to appropriate implementation of sound generation devices" (Wanderley and Depalle 2004, 637).

Sound synthesis and processing can be classified into two main categories: physical models and signal models. Physical models describe the mechanical and acoustic behavior of traditional musical instruments in order to simulate them (Wanderley and Depalle 2004, 637; Välimäki and Takala 1996). Contrastingly, signal models are based on a more experiential paradigm, utilizing creative mathematical abstractions for the realization of certain qualitative properties of sounds.

> There is no explicit reference to the instrument which produces the sound. Among the more popular signal models are the phase vocoder, additive synthesis, the source-filter model, and FM synthesis. (Wanderley and Depalle 2004, 637)

As I will demonstrate, my alternate controller approach to DMI design works naturally with a signal model for sound synthesis. As opposed to simulating an existing instrumental design, I am interested in what Wanderley and Depalle refer to as extrapolation.

> Extrapolation consists in using synthesis models, outside the normal range of parameters, or conditions other than those usually met in natural instrumental situations. The goal here is to create new sounds. (Wanderley and Depalle 2004, 638)

An important intermediary step, however, between gestural acquisition and the creation of new sounds is the mapping layer. This is especially true for the alternate controller paradigm, as there are no standard models to be referenced.

For instance, when considering general human movement as the gestural input for a digital musical instrument, Winkler (1995) offers this interesting insight:

> Each part of the body has its unique limitation in terms of direction, weight, range of motion, speed and force. In addition, actions can be characterized by ease of execution, accuracy, repeatability, fatigue, and response. The underlying physics of movement lends insight into the selection of musical material. Thus, a delicate curling of the fingers should produce a very different sonic result than a violent and dramatic leg kick, since the size, weight and momentum alone would have different physical ramifications. To do this, physical parameters can be appropriately mapped to musical parameters, such as weight to density or register, tension to dissonance, or physical space to simulated acoustical space, although such simple one-to-one correspondences are not always

> musically successful. The composer's job then, is not only to map movement data to musical parameters, but to interpret these numbers to produce musically satisfying results.

As Winkler suggested, what is important is the way in which a designer interprets the parameters involved to produce interesting results. Thus, the intensive study of parameter mapping in digital music instrument design has grown in scholarly significance over the years.

It would be argued now that mapping is the key factor in determining the overall expressivity of a digital music instrument (Rovan et al. 1997; Hunt et al. 2002). Rovan et al. (1997) proposed a helpful classification system with three different groups of mapping strategies:

- *One-to-One Mapping:* Each independent gestural output is assigned to one musical parameter, usually via a MIDI control message. This is the simplest mapping scheme, but usually the least expressive. It takes direct advantage of the MIDI controller architecture.
- *Divergent Mapping:* One gestural output is used to control more than one simultaneous musical parameter. Although it may initially provide a macro-level expressivity control, this approach nevertheless may prove limited when applied alone, as it does not allow access to internal (micro) features of the sound object.
- *Convergent Mapping:* In this case many gestures are coupled to produce one musical parameter. This scheme requires previous experience with the system in order to achieve effective control. Although harder to master, it proves far more expressive than the simpler unity mapping.

For this project's purposes, I have equally considered divergent and convergent paradigms. It seems clear from the literature and my own experience that simple one-to-one mappings, although clear in their intelligibility, lack the dynamics to be expressive over extended periods of time. Another aspect to consider, however, is the degree to which indeterminacy affects the mapping relationships and overall sonic result. This, for me, was the most exciting aspect of parameter mapping, and proved to be an interesting component influencing expressivity.

Hunt et al. (2002) extended the convergent mapping approach suggested by Rovan et al. (1997) to create an argument for a strategy that they called the "multi-layered" mapping scheme. They offered a scenario utilizing an interface consisting of 2 buttons, 2 sliders, and a mouse controlling FM synthesis to illustrate the approach.

> ...let us imagine that a mapping layer has already been devised to abstract the inputs to parameters such as 'energy', 'distance between sliders', 'wobble' etc. Also let us imagine that there is a mapping layer before the FM synthesis unit, providing higher-level control inputs such as 'brightness', 'pitch', 'sharpness' etc. Now we can picture the designer making a relationship such as "energy controls brightness". On the surface this may appear to be yet another one-to-one mapping. Indeed it is – at the *conceptual* level. However, when you consider how 'energy' is calculated from the given inputs, and how 'brightness' has to be converted into the FM synthesis primitives, you will notice how many of the lower-level parameters have been cross-coupled. (Hunt et al. 2002, 5)

This approach helps the DMI designer focus on the conceptual level of mapping, thus creating higher-level interactions between the perceptual and physical qualities involved in gesture controlled sound synthesis.

Although these strategies and frameworks are useful for guiding process, mapping and DMI design (in general) often requires a more idiosyncratic kind of exploration influenced by the uniqueness of the human performer(s) involved and the artistic needs of the piece.

> …the user discovers how to control a device by exploring different input control positions and combinations, thus gaining an immediate response from the system. The user may appear to be 'playing around' with the

> control, but they are actually discovering hidden relationships between parameters within the system. (Hunt and Kirk 2000, 233)

Immersive Music

In a paper titled *A Framework for Immersive Spatial Audio Performance*, Mike Wozniewski, Zack Settel and Jeremy R. Cooperstock (2006) consider natural human body motion as an input mechanism for an immersive digital music instrument. From their perspective, such control makes direct use of what humans have already learned through daily practice. In particular, kinesthetic feedback provides information about the body's orientation in space (Wozniewski et al. 2006).

They feel that the current computer-vision based body tracking solutions being investigated exhibit poor performance due to a lack of temporal and spatial precision. They note that musical systems are particularly demanding of low-latency control. In response, they use input devices with lower latency and greater fidelity. Such examples include data gloves and accelerometers. Their input devices provide information formatted into a standard message in the range of 0 and 1 for each parameter. Messages are then available for control of more complicated mappings.

> Many of today's so-called computer music interfaces either end up being too simple to achieve virtuosic playability in the long term, or too complicated to learn in the short term. (Wozniewski et al. 2006)

In response to these claims, users (in their system) are perceptually immersed in the instrument and have the ability to explore complicated parameter configurations in a holistic manner. In addition, multi-modal feedback is provided so that users can learn how the system reacts.

For another perspective, *On the Development of a System for Gesture Control of Spatialization*, by M. T. Marshall, N. Peters, A. R. Jensenius, J. Boissinot, M. M. Wanderley and J. Braasch (2006) offers an in depth look at the development of a system for real-time gestural control of spatialization by performers themselves. Thus, performers control the spatialization of their own instrument's output. The core technology behind their implementation is called the ViMiC (Virtual Microphone Control), which includes a gestural control system that attempts to provide a standardized means of communicating data to the spatialization system (Marshall et al. 2006).

They feel that the ViMiC offers an alternative to conventional spatialization techniques such as Wavefield Synthesis (WFS) or Ambisonics.

> ViMiC is a computer-generated virtual environment, where gains and delays between a virtual sound source and virtual microphones are calculated according to their distances and the axis orientations of the microphone directivity patterns. Besides the direct sound wave, a virtual microphone signal also contains early reflections and an adequate reverberation tail. This depends both on the sound absorbing and reflecting properties of the virtual surfaces and on the geometry of the virtual enclosed space. Sound sources and microphones can be spatially placed and moved in 3D as desired. (Marshall et al. 2006)

The authors' gesture tracking system utilized a Polhemus Liberty electromagnetic tracker with an update rate of 240 Hz per sensor for 8 sensors and an accuracy of 0.03 mm for X, Y, Z position and 0.15 mm for orientation (Marshall et al. 2006). The user wears wired sensors on the body points that are to be tracked. It should be noted that such systems are also used for tracking within immersive visual contexts.

Using the Gesture Description Interchange Format (GDIF), their proposed framework for standardizing gesture related information, they separate the movement related data into three categories (Jensenius, Kvifte, and Godoy 2006). Raw Data, the unprocessed data coming from sensing devices (e.g. electromagnetic trackers, accelerometers and floor pressure sensors etc.); Body Data, the information about orientation and motion of the body/limbs; and, Meta Data, the information about general motion qualities such as "intensity" (Marshall et al. 2006).

Intensity, the only meta-data descriptor used in the system, looked at the overall quantity of motion in the hands as related to the reference point. In the future, "other meta level descriptors will be added during the testing and evaluation sessions with performers." (Marshall et. al. 2006)

> Careful analysis of the performers gestures will be required to allow us to determine the optimum mappings for the system. Possible control gestures might include measurement of the performer's weight balance (or centre of gravity) to allow them to control parameters through slight swaying of the body. Other possibilities might include the movement of the feet, or for certain instruments, the movement of the performer's head. (Marshall et al. 2006)

In a subsequent paper, entitled *Non-Conscious Control of Sound Spatialization*, Mark T. Marshall, Joseph Malloch, and Marcelo M. Wanderley (2007) set out to answer these questions (at least with regard to instrumentalists/performers). Through research with multiple composers, performers, systems and venues they suggest a set of guidelines which may minimize performer distraction and maximize mapping success when making use of non-conscious performer control of spatialization.

> Map resting body states to neutral spatialization states. If resting body states are mapped to dramatic or obvious spatialization effects, the

> performer is likely to become distressed, as they will be unable to "stop" performing...
>
> Use non-conscious control for high-level spatialization control. High-level spatialization parameters, such as system changeability, the flocking behaviours of clouds of sound sources, or control over other algorithmic behaviours, may be successfully controlled using ancilliary performer movements...
>
> Use non-conscious control for subtle spatialization control. Similarly, subtle effects, such as sound spread or diffusion in the space or boundary reflectivity, may be less distracting for the performer than direct source position or volume...
>
> Avoid changing mapping relationships dynamically. If the audience is intended or permitted to perceive the nature of gesture control, this effect may be confused or destroyed by changing scaling or transfer functions... (Marshall et al. 2007)

These specific explorations of gesture-controlled spatialization offer some interesting ways forward for digital music instrument design in general. The ideas generated from these very recent research projects have directly inspired my process, especially with regard to sensor implementation and conceptualizing relationships between gestural quality and sonic quality.

However, I have been less influenced by the specific implementations regarding performer-controlled spatialization. Performer orientation in respect to the sound field seems to be a practical deterrent to effective implementation. In fact, the movement towards non-conscious control in the research speaks to this inherent difficulty. Simon Emmerson articulates this best.

> I have suggested several areas that control may be returned to the performer but the direct influence of the performer on sound diffusion is the most difficult. The very idea seems impossible – it is orthodoxy to assume that only a person at the optimum listening location can judge the acoustic balance. (Emmerson 2007, 97)

Immersion and Interaction: Moving Image Systems

In moving towards a total multi-sensory/multi-media aesthetic framework, one must also consider the technical/aesthetic precedents regarding immersive and interactive moving image and/or graphics systems as well. It is my choice to cover unitary subject areas, before considering the whole, as there are ideas to consider from focused examples dealing with one modality at a time.

As mentioned before, Myron Krueger's work with the *VideoPlace* environment offers a seminal example of the perceptually immersive and interactive aesthetic paradigm realized in the form of a dynamic moving image system. At the time, Krueger's work realized an abstract relationship between the body and media elements that was unprecedented in the arts. The physical interaction did not result in a graphical imitation of movement, but an abstracted extension.

> More exactly, by deploying sensors to capture the movement of the body through a concrete space and in conjunction with a virtual element, it displaces the body image – that is, a predominantly visual representation of the body, a primary resemblance – and offers in its place a new, representative but not representational "image", a graphic symbol that attains its force only through its inclusion within the body schema, through its motor coupling with embodied enaction. (Hansen 2006, 37)

Following the lead of Krueger's artistic revolution Simon Penny offered an artistic virtual reality environment called *Traces* (1998-1999):

> In Traces, virtually all sound and visual experiences are generated in real time based on the users behavior. Unlike other VR projects, I have no interest here in illusionistic texture mapped models, the illusion of infinite virtual space or building "virtual worlds". All attention is focused on the ongoing bodily behavior of the user. (Penny)

Penny's project was originally designed for the CAVE (computer assisted virtual environments). *Traces* essentially offered a space for sensori-motor interaction driven by three-dimensional representations of bodily motion picked up by four video cameras mounted in four corners of the CAVE. In this way, it privileged physical interaction over the "illusionism" of most VR as the driver of the immersive experience (Hansen 2006, 45).

> In "Traces," the interactive experience is not premised on hypertextual paradigms such as interactive storytelling. It is based on bodily, temporal, kinesthetic sensibilities. *Traces* most closely resembles a combination of sculpture-making and dancing. The paths of bodies will persist in the space, each body 'painting' as a three-dimensional brush. (Penny)

Raphael Lozano-Hemmer's *Relational Architecture Series* also deserves recognition when considering the power of physicality in mediated arts experience. Relational architecture can be defined as the, "technological actualization of buildings and public spaces with alien memory" (Lozano-Hemmer). His work in this series attempts to transform the "master narratives" of a specific public structure by manipulating media elements to "affect it, effect it and re-contextualize it" (Lozano-Hemmer).

> Accordingly, Lozano-Hemmer's relational architecture installations deploy media interventions into existing architectural space precisely as a means of triggering embodied reactions (reactions which, as we shall see, tap into the disembodiment constitutive of embodiment). When these reactions subsequently enter into resonance with the media transformations triggering them, they establish feedback loops in which embodiment and information mutually catalyze one another's ongoing evolution…(Hansen 2006, 95)

Lozano-Hemmer's embodied mediation environments tap into some very interesting tendencies for humans regarding physical expression and public interaction. Sometimes his work elicited physical interactions between participants that were not even intended showing the inherent proclivity humans exhibit with regard to non-verbal physical expression. For instance, in some of his works, the shadows of participants were intended to serve as an interface mechanism to the mediated large-scale interaction. However, many participants found the big shadows they would cast in public space to be interesting in and of themselves, resulting in large-scale impromptu shadow theatre exhibitions by complete strangers.

On a side note, Lozano-Hemmer's work offers a great example of the navigable structure paradigm for interactivity as articulated by David Rokeby. Rokeby's ideas conceptually inspired this dissertation project. Although my specific approach resulted in the navigation of conceptual sonic space, the ideas nonetheless deserve recognition.

> The navigable structure can be thought of as an articulation of a space, either real, virtual or conceptual. The artist structures this space with a sort of architecture, and provides a method of navigation. Each position within the conceptual space provides a point-of view, defined and limited by the surrounding architectural structure…In Virtual Reality systems, the architectural metaphor can be taken literally. In other works, the architecture is more like a conceptual paradigm, a method of organization of intellectual perspectives, opinions or emotions. (Rokeby 1996)

The Topological Media Lab's (TML) work offers a contemporary example of immersive interaction practices such as the navigable structure paradigm within the context of digitally mediated performance environments. Over the years, this group has invented "novel forms of gestural media, expressive instruments and compositional systems." Their work has traditionally drawn inspiration from studies of embodiment, performance and music, "as well as the poetic uses of dynamical systems" (Topological Media Lab). In addition, their work often fuses sensory modalities such as "visual dynamics and kinesthesia" (Massumi in Massumi and Mulder 2007, 86). One such work is exemplified in *Artaudian Lights* (2006):

> The concept was very simple. There were two dancers, going through a choreographed routine on stage in front of a large screen. A motion-sensing camera analyzed their movement. When the movement reached a certain qualitative threshold – a certain speed and density of gesture – a video window opened up on the screen. But it wasn't at all like a Windows window, thankfully. It was like a visual bubble that grew from nothing and expanded. It was like vision was flowering out of the screen, expressing a quality of movement, its speed and density, purely visually, in a sight that doubled the actual movement. (Massumi in Massumi and Mulder 2007, 86-7)

Referencing Antonin Artaud directly in the work is an interesting homage to the French drama theorist who foresaw the type of work that the TML started and the kind of work that this dissertation project aims to create as a humble continuation of efforts; another example of the movement towards interactivity and "total" multi-sensory immersion in the creation of dynamic digital performance experiences. "Artaud conceived the ritual impact of dramatic performance as a visceral experience that restores the integrity of the self through the involvement of both body and mind" (Ryan 2001, 303).

Immersion and Interaction: Dynamic Multi-Sensory Performance Practices

Outside of the fine arts, issues regarding sensory immersion and interaction are big business. It's interesting for artists to consider the utilitarian aimed endeavors that also focus on the need for more sensory immersed, physically connected, interactive relationships between humans and their technologies.

> WE ALL HAVE five senses; how sad that our connection to computers is "sensory deprived and physically limited," as Nicholas Negroponte so aptly phrases it. (Moggridge 2007, 515)

Designing Interactions

In his book *Designing Interactions*, world-class interaction designer Bill Moggridge (he designed what was arguably the first lap-top computer) clarifies the revolutionary ideas that the cognitive psychologist J.J. Gibson put forth with regard to human perception (Gibson 1979).

> He asserted that perception is designed for action, claiming that our whole evolution has been geared toward perceiving useful possibilities for action, and that rather than recognizing separate properties like length or mass, we perceive complicated combinations of these things that have validity for us to do something with or about. We see surfaces for walking, handles for pulling, space for navigation, tools for manipulating, and so on. He called the perceivable possibilities for action "affordances". (Moggridge 2007, 579)

In order to expand on these ideas, Moggridge interviewed Bill Gavers, who has held many titles including consultant for Apple, Inc. (he led their first foray into sound design for the Mac OS), professor of interaction research at the Royal College of Art (London), and, professor of design at Goldsmiths College (London). Gavers elucidates what Gibson's notion of the affordance means in relation to interaction design.

> ...you need to think more about the biophysics of a situation and combine that with fairly low-level perceptual information – in other words, information that can be conveyed through light or sound about the physics of a situation, without a lot of mediation by culture or learning. (Gavers in Moggridge 2007, 579-80)

Issues regarding the interface's immediacy are always central to the task of designing a believable interaction, whether it is for an OS or an art project. Often, our ability to convey a believable interaction in the digital domain must be based on our intuitive knowledge of physics, as garnered throughout a lifetime of "real life" experience. Physical extensions of the body into digital and/or conceptual spaces must be immediate and quickly intelligible, yet complex enough to be engaging over relatively long time frames. Hiroshi Ishii's development of *PingPongPlus* (within MIT's Tangible Media Group) illustrates this kind of interaction design. *PingPongPlus* is a digitally augmented version of the classic sport that features a "reactive table" incorporating sensing, sound and image projection technologies.

> Ping Pong is one of my favorite sports; that's why we designed PingPongPlus. Ping Pong really allows very engaging interactions, using full body motion, speed, vision, and other kinesthesia. We can learn a lot from this Ping Pong paddle about how interaction design should be really invisible and transparent, so that the player can concentrate on the main task, playing Ping Pong, trying to hit the ball. A good fit to the body is really critical to make the interface transparent. To make a good fit, you have to choose the right form, size, and weight…we can design the interface to become transparent and also to become extensions to the human body, for both digital and physical domains. (Hiroshi Ishii in Moggridge 2007, 531)

Ishii's ideas about the ping pong paddle echo, in a way, Merleau-Ponty's description of the blind man's stick, which Hansen points to as an immediate and practical expansion of the body (Hansen 2006, 43):

> Once the stick has become a familiar instrument, the world of feelable things receded and now begins, not at the outer skin of the hand, but at the end of the stick. One is tempted to say that through the sensations produced by the pressure of the stick on the hand, the blind man builds up the stick along with its various positions, and that the latter then mediate a second order object, the external thing…the stick is no longer an object perceived by the blind man, but an instrument with which he perceives. It

> is a bodily auxiliary, an extension of the bodily synthesis (Merleau-Ponty 1962, 175-6)

Nintendo's Wiimote could be considered one of these interaction design examples, where the immediacy of the interface is married with dynamic interaction possibilities, allowing it to be believable as a tennis racket, baseball bat, steering wheel, fishing rod, etc. As I have found out, it can also be a fairly believable interactive multimedia performance environment interface when in the capable hands of a talented movement artist.

Multi-Modal Music Research

In tandem with the movement towards full multi-sensory engagement in the more industrial ranks of human-computer interaction design, researchers in arts related disciplines are making their own strides.

For instance, studies of expressive gesture in music (and dance for that matter) are no longer focused on uni-modal feedback, but more complete multi-modal experience. The overall expressivity of a musical performance cannot be studied on the basis of auditory perception and musical cognition alone. Aspects of the visual performance gesture, which offer cues regarding the performer's physical relationship to the music (and expressive intention) must also be accounted for when probing an audience's experience (Camurri et al. 2004, 11).

This makes even more sense when considering the literature on mirror-neuron theory which would suggest that our visual and auditory perceptual mechanisms can also give rise to simulated physical responses that attempt to mimic the performer's musically motivated actions. This new approach to music research will only gain in importance.

> This approach is grounded in multi-sensory integration, the coupling of perception and action, the study of motor imitation, and issues that relate to affect, emotions and subjectivity. (Leman and Camurri 2006, 1)

As technology advances to accommodate these more holistic research aims, artists benefit from these activities as the research technologies developed also lend themselves to direct artistic application, such as the *EyesWeb* platform etc. (developed by Camurri et al.). *EyesWeb* has the capacity to map expressive gesture data from multiple modalities (including music, full-body movement etc.) via various sensor technologies (including multiple video camera feeds and wireless on-body accelerometers) into/onto multimedia output such as sound, music and visual media. Art forms, such as the creative result of this dissertation project, using these types of systems harness this technology to influence music and sound with a dancer's movement, or graphics projections with musical sound.

> The whole approach should be seen as an attempt to extend the limitations of symbolic interactions with machines towards gesture-based interactions. The platforms also extend the concept of hyper-instruments in that they offer an environment for measurement and modeling, in addition to a platform for interactive artistic expression. (Leman and Camurri 2006, 4-5)

Some of the expressive motion cues that seem most salient in the research regarding this type of interaction design are quantity of motion and contraction index (Castellano et al. 2007, 391). Contraction index (CI) is related to the overall shape of the body and the extension of its limbs while quantity of motion (QoM) is related to the speed of movement over specified time periods. Utilizing the *EyesWeb* platform, Castellano et al. (2007) focused on “full-body movement” interactions while relating QoM and CI to the tempo and sound level of music.

> QoM and CI are mapped onto tempo and sound level respectively. In this way the user’s performance gets faster for faster movements (i.e. higher QoM values) and louder for the user is expanding her body (i.e. lower CI). Articulation is controlled by mapping the emotion detected in the user’s movements to the articulation parameters in pDM…more *legato* articulation for emotions with lower activity and more *staccato* articulation for those characterized by higher energy. (Castellano et al. 2007, 391)

In addition, The *EyeCon* computer vision system, developed by Frieder Weiss for the experimental inter-media performance group Palindrome, has also been utilized in multi-modal performance research and practice. Their process has resulted in award winning works combining full body movement, interactive music and moving image.

Seine hohle Form

Palindrome’s award winning piece *Seine hohle Form* took the goal of creating music based solely on bodily movement very seriously. In this collaboration with Butch Rovan, the choreography was influenced by the sound just as much as the movement created the music. Thus, a blurring of roles took place between composer/choreographer and musician/dancer (Broadhurst 2007, 114).

As in all of their work, Palindrome considered issues of “gestural coherence” paramount to the experience. Specifically, they were interested in exploring the perceptual coherence between sound and the movement that gave rise to it. Underscoring their efforts is a belief in the premise that, “an emergent integrity arises when the relationship between dance and music is believable" (Rovan et al. 2001, 3). For them, this credibility is founded on a perceptual coherence between sound and movement based on a layer of parameter mapping mediating the physical input and music (Broadhurst 2007, 114).

Bodily Instruments and Instrumental Bodies

In 2006, Contemporary Music Review dedicated an entire double issue to themes related to the body, technology and performance entitled, *Bodily Instruments and Instrumental Bodies: Critical Views on the Relation of Body and Instrument in Technologically Informed Performance Environments.* As the editor states, this collection of articles explored:

> …sound in the widest possible sense in relation to the technologically informed body. I was seeking stimulating ideas from writers who were prepared to define or critically examine the threshold conditions of instrument and performer, who wanted to challenge continuities or discontinuities of instrument and performer, and who were eager to question not merely the role of the instrument, but the nature of the instrument itself, as well as its relation to the performing body. (Schroeder 2006, 1)

Of most relevance to this dissertation project were ideas regarding the moving body. For instance, Julie Wilson-Bokowiec and Mark Alexander Bokowiec's article, entitled *Kinaesonics: The Intertwining Relationship of Body and Sound,* examined the relationship of the body with computer driven music.

Out of this examination, they discovered the notion that technologically mediated performance could reconfigure, "the experiences and sensations of the body" (Schroeder 2006, 3). As Schroeder states:

> The authors' Bodycoder performance system serves as a means for highlighting the adaptability and resourcefulness of the body as well as the body's capacity for change. (Schroeder 2006, 3)

The first version of the Bodycoder system was developed in 1995 – 1996, and featured a wearable sensor interface that sent movement data wirelessly to a DSP environment. The performer had real-time access to all of the processing parameters in addition to control over the sensitivity of sensors. The performer essentially had control over all aspects of the system in performance (Bokowiec and Bokowiec 2006, 49-50).

The system could utilize up to sixteen bend sensors that could be placed on any suitable part of the body, along with glove attached buttons that could be assigned different functions depending on the needs of the piece (Bokowiec and Bokowiec 2006, 50). As the authors articulate, their conception of the instrument-system is much more holistic than an instrument-object paradigm.

> The Bodycoder system itself comprises a number of elements: the body of the performer, the performance space and sound diffusion system, the hardware interface elements, as well as bend sensors, switches and a wireless radio transmitter, a wireless receiver hardwired to a computer and an MSP software environment that also comprises a variety of programming levels: patches and sub-patches containing a variety of programmable function objects. Unlike the design of a traditional instrument that is a thing (object) in itself such as a piano, the Bodycoder system is multidimensional and multimedia in nature. (Bokowiec and Bokowiec 2006, 50)

This holistic conception of the system/instrument works well with the basic assumption they ground their work on: that all sensory experience is fundamentally synaesthetic, or, in other words a dynamic network of cross-modal perceptions (Bokowiec and Bokowiec 2006, 56).

This basic idea fuels their design of the performance experience. The performer is cognizant of the fact that the sound comes from the speakers and not the body, even though the body is controlling and influencing that sound.

> This form of remote manipulation—the articulation of objects by the body, but at a distance from the body—that new technologies enable is a new experience for the mind/body. Even though the intellect can rationalize the concept of dislocation, the sense of the body still has to find ways of working with the experience. What we suppose is happening with kinaesonic gestures is that the mind/body draws from its pre-formed patterns of sense perception in order to remodel a form or proprioception in a way that is useful for the task we are asking the body/mind to undertake, which is to physically control the movement of sound. (Bokowiec and Bokowiec 2006, 56)

Thus, the mind/body system creates a palpable link between body movement and the resultant sound by virtually fusing the inherent spatial disassociation, effectively restructuring the, "proximal sensation of a texture in the arm based on what the mind/body knows about movement and touch from other familiar patterns of proprioception".

> The intervention of technology into the space between body and world as tool, as prosthesis and as a medium, a 'virtual reality' for various forms of interaction in close proximity or at remote distance, has prompted the mind/body to develop new patterns of proprioception. Kinaesonic gestures made possible by the development of interactive technologies are opening the door to the possibility of new sensations and experiences. (Bokowiec and Bokowiec 2006, 57)

Because such sensations and perceptions are new to the mind/body, this novel form of constructed proprioception is actually experienced consciously (Bokowiec and Bokowiec 2006, 56).

Because the perceptual events we most often encounter on a daily basis come and go via unconscious processes, such new sensory reconfigurations become meaningful. Thus, even the creation of an interactive performance work that elicits conscious meditation on sensory experience as its only aim is a significant artistic endeavor in and of itself. One might even speculate that such an event would also heighten our experience of the perceptual act, in general.

CHAPTER 3: MULTI-SENSORY DYNAMICS AND DIGITAL PERFORMANCE IN PRACTICE

As was touched upon in Chapter 1, Synaesthesia, as a neurological phenomenon, refers to an abnormal condition in human perception where the activation of one sensory modality gives rise to perceived sensations in another sensory modality. The word comes from the Greek *syn*(together)-*aesthesis*(sensation). "Synaesthesia means 'feeling together,' just as syn-chrony means at the same time or syn-thesis means different ideas joined into one or syn-opsis means to see all together" (Cytowic 2003, 5). For instance, a typical form of this phenomenon induces the perception of distinct and repeatable color sensations in relation to specific pitches. This is known as chromesthesia or "color hearing".

Throughout the late 1800s and early 1900s, psychologists researched synaesthesia as an inroad to more general investigations regarding the modular vs. unitary theories of sensory organization, which were intensely debated (Campen 2008). Painters, poets and composers such as Kandinsky, Mondrian, Baudelaire, Schoenberg, and Scriabin, experimented with the "idea" in order to explore connections between media. In addition, the Wagnerian concept of the *Gesamtkunstwerk* (the total work of art) inspired further aesthetic inquiries into synaesthesia. As social science and fine arts researcher Cretien van Campen states:

> It is remarkable that both the symbolist and the early abstract art 'movements' interest in the Gesamtkunstwerk was synchronic with a boom of psychological publications on synaesthesia in the years 1880-1930. (Campen 2008)

Moving forward into the mid to late 1900s, a survey of such artistic precedents would not be complete without mentioning the work of visual music pioneers such as Oskar Fischinger and John Whitney.

Many of the artistic and psychological experiments in the past have focused on the relationships between audition and vision. Examples such as color-organs, musical paintings and visual music all fall into this category. More recently, interest regarding the body's relationship to other sensory modalities has intensified within areas of technologically mediated performance. This movement has surged in tandem with advances in digital technology that allow for sensory/media connections beyond those imagined in the past. Theoretically speaking, this aesthetic movement, like the scientific movement of embodied cognition (mentioned in Chapter 1), can be traced to the philosophical thought of Maurice Merleau-Ponty, who championed the idea that the body was the source of perceiving and knowing the environment (Campen 2008, 98). Therefore, it seems that new forms of multi-sensory art involving the body are evolving in a similar fashion to the synaesthetic forms of the modernist age. Parallel movements in art, science, and philosophy have brought us closer to a more complete representation of multi-sensory experience and dynamic human-environment interaction.

In light of these developments, this chapter surveys three particular works of art, *Messa Di Voce* (by Golan Levin and Zachary Lieberman in 2003), *16 R(e)volutions* (by Troika Ranch in 2005) and *Schwelle* (by Chris Salter in 2007), which represent contemporary examples of the more complete multi-sensory aesthetic within the context of interactive digital performance.

Synaesthesia in *Messa Di Voce*

Synaesthesia has clearly been an influence in the aesthetic process of Golan Levin and Zachary Lieberman. One recent piece, titled *Messa Di Voce* (Ital., "placing the voice"), was based on the cross-sensory phenomenon of phonesthesia, which was mentioned in Chapter 1. In phonesthesia, sound triggers consistent associations with the perception of visual form.

In *Messa Di Voce*, the concept of phonesthesia was incorporated with interactive system design, abstract language, and extended vocal techniques, in a 35-minute live audio-visual performance (Levin and Lieberman 2004b, 4). As Levin and Lieberman state, "*Messa Di Voce* presents the fiction that the voice can be seen. Like our previous augmented-reality artworks *RE:MARK* and *Hidden Worlds of Noise and Voice*, it is concerned with the poetic implications of this proposition, in an attempt to address the question: if we could see our speech, what might it look like?" (Levin and Lieberman 2004b, 4).

As with the two previous augmented-reality works, their implementation of speech visualization techniques deviated from the more conventional approaches found in the scientific disciplines of phonology, psychoacoustics, and computational audition.

> Practitioners in those fields have developed or refined a range of widely adopted and useful visualization techniques, including waveform graphs, spectral and formant plots, autocorrelograms, cochlear and pharyngeotracheal diagrams, etc., which help solve the problems those practitioners encounter. (Levin and Lieberman 2004, 1)

The speech visualization techniques in *Messa* were considered purely for their artistic value, with the artists describing their techniques as, “(A) explicitly non-utilitarian and (B) designed solely to establish a perceptually and aesthetically plausible interactive fictional universe in which speech is somehow visible” (Levin and Lieberman 2004, 1).

The two aforementioned installations served as instrumental precursors to the realization of the *Messa Di Voce* performance system. Untrained lay people were targeted for the interactions, thus serving as reference for further development of the markedly more professional and nuanced interaction environment in *Messa Di Voce* (Levin and Lieberman 2004, 5).

In *Hidden Worlds*, for instance, participants saw each other’s voices made visible in the form of animated graphics that emerged from the participants’ mouths as they spoke. *RE:MARK*, on the other hand, was created as a companion piece to the larger scale *Hidden Worlds* project and was designed for only two participants. *RE:MARK*, like *Hidden Worlds*, presented an interactive visualization of its users’ speech. However, as *Hidden Worlds* primarily attended to the abstract qualities of vocal sounds, *RE:MARK* shifted the inquiry towards the symbolic domain of the spoken and written word (See Figure 3.1) (Levin and Lieberman 2004, 4).

> In RE:MARK, sounds spoken into a pair of microphones are analyzed and classified by a phoneme recognition system. When a phoneme is recognized with sufficient confidence, the written name of the phoneme (for example, oh, ee, ah, etc.) is projected on the installation's display. If the user's sound is not recognized by the system's classifier, then an abstract shape is generated instead, according to parameters derived from the timbral (spectral and formant) characteristics of the vocalization. Among other mappings, sounds with high-frequency spectral centroids are represented with pointier, more irregular forms. (Levin and Lieberman 2004, 4)

In turn, Levin and Lieberman sought to augment these audiovisual mapping techniques through the creation of a dynamic performance system allowing for more nuanced interaction with virtuosic vocalists. "Our foremost challenge in doing so, as artist-engineers, would be to create interactive systems that could be deeply instrumental for these vocalists, and commensurately expressive." (Levin and Lieberman 2004, 5)

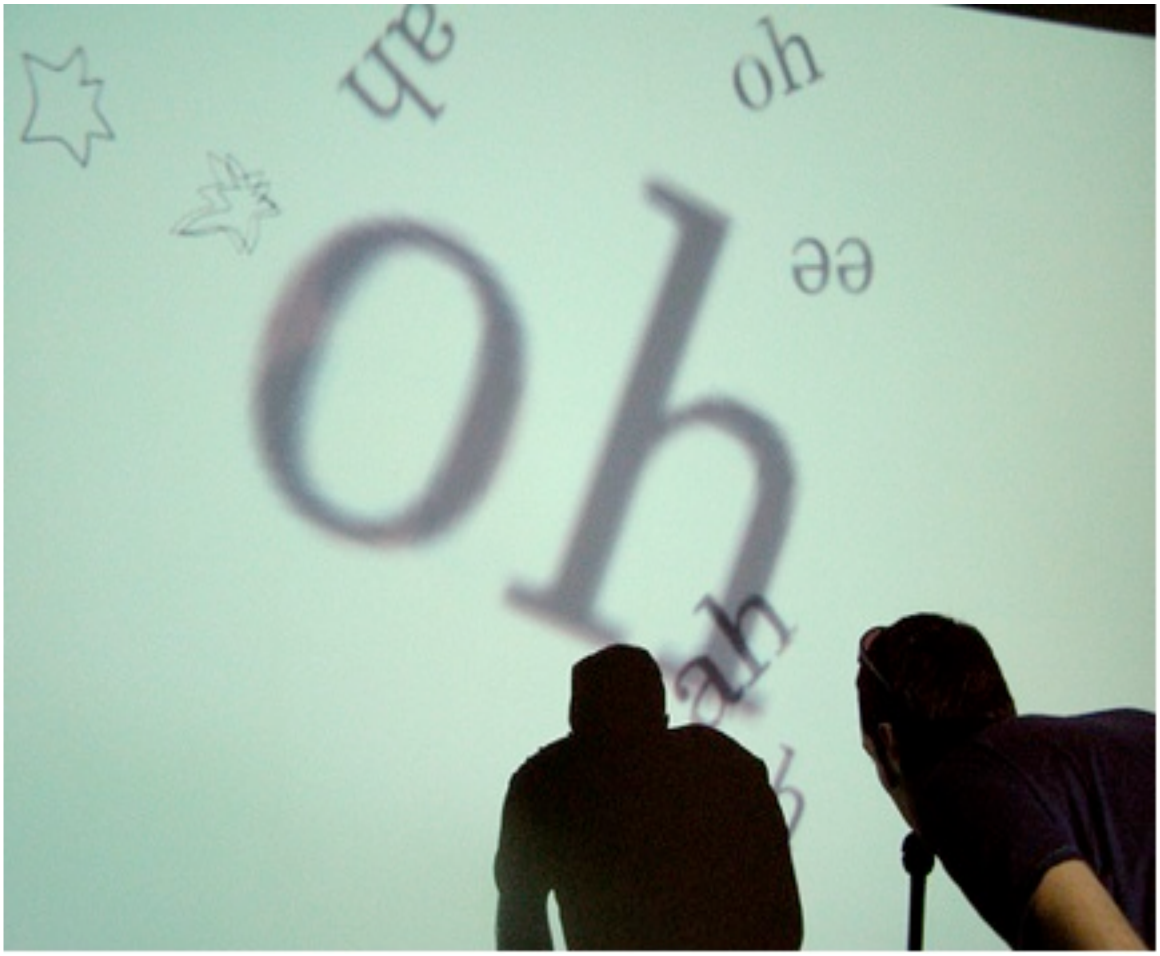

Figure 3.1. A participant interacting with the RE:MARK system. Permission granted by the artists.

The realization of this effort was *Messa di Voce*, a truly interactive performance piece. Speech, shouts and songs produced by two vocalists were augmented in real-time with custom visualization software. Appropriately, the piece utilized the virtuosic abilities of co-collaborators, Joan La Barbara and Jaap Blonk, who are known for their extended vocal techniques. The work touched on themes of primitive communication and synaesthesia within the framework of vignette style performance exhibition.

In real-time, computer vision techniques tracked the locations of the performers' heads. In addition, analysis of the sound emanating from the performers' wireless microphone systems extracted features such as pitch and spectral content. In turn, the computer/projection system displayed visuals behind the performers that were spatio-temporally aligned. The head-tracking system drove the orientation and placement of the projected material so that it appeared to emerge directly from the performers' mouths (Levin and Lieberman 2004, 6). In the process, Levin and Lieberman made use of an effective assortment of visualization techniques, including particle systems, elastic spring meshes, fluid simulations, cloud simulations, etc. (Levin and Lieberman 2004, 6).

Analysis

The performance score is guided by the linear progression of 12 separate software modules that support the vignette style of exhibition (with two being solos for each performer). The unique characteristics of each software module serve as structural guides determining performance dynamics. Over the course of the 12 vignettes, the vocal performances follow a trajectory from pure sound to speech to song with perceivable levels of complex audio-visual interaction also increasing. Actually, it seems that a "verbal language evolution" motif creates a conceptual bridge between successive modules/vignettes. Thus, two distinct progressions unfold over time. Firstly, the vocalists become more nuanced (breath to primitive speech to song) in their vocal expression, and, secondly, their interaction with the performance environment becomes more dynamic and deliberate. Based on the available documentation, I have developed detailed analyses for a select group of the performance modules.

In the fourth module, *Bounce*, Blonk performs alone and creates projected "bubbles" through extended (cheek based) vocal techniques that sound akin to duck-like utterances (See Figure 3.2). The overall effect is humorous, but some of the visuals seem to be out of sync with the location of his mouth. Sonic event density and bubble creation, however, are very well correlated. Feeling satisfied, Blonk's character stands proud and observes the creation. Suddenly, the bubbles start to fall one by one with increasing intensity. In response, he urgently attempts to kick the bubbles away. This chaotic display triggers a cacophonous barrage of sound; a replay of what was initially emitted over the course of their creation. Interestingly, his kicking gestures interact well with the bubbles' physical movement dynamics.

In *Ripple* (a duet), attention is drawn to a pair of "wetland" creatures (Blonk and LaBarbara) as they communicate in bizarre inflections that displace the fluid environment. In turn, they create beautiful patterns that match the complexity of the sonic disturbances. The spatio-temporal connectivity in this module was especially effective as performer movement also created "wakes" in the liquid environment behind them. As the performers are not movement artists by training, the vocal expressions are much more interesting to experience. Performer movement seems to primarily serve the purpose of connecting their expressions in a spatio-temporal fashion rather than highlight the spatio-kinetic interactions.

The sixth module/vignette, *Insect Nature Show*, increases a bit in narrative sophistication, while new sonic complexities are introduced in the form of speech-like sounds. The performers take on complementary dramatic roles: LaBarbara, as a bizarre creature; and Blonk, as the "nature show host" expounding in an abstract tongue about the creature he encounters (LaBarbara). LaBarbara's vocals morph the shape of her creature's projected outline as she moves around. Stronger upper partial densities and noise fragments contribute to the creation of extensions in her amorphous body that are correspondingly jagged and angular. The movements, visuals and sounds create an intriguing and humorous gestalt.

Vignette number seven, *Fluid*, introduces another synaesthetic device where the color of "plasma-like" substances are affected by phonemes. This development marks the last exhibition of a distinctly new audio-visual relationship. Like many of the vignettes (and the piece as a whole with its connections to phonesthesia and not complete multi-sensory interaction), priority is placed on the sonic and image based relationships rather than the spatio-kinetic interactions. However, they do engage in jousting like movements that add an interesting dynamic to the performance interaction. From this point onward it seems that the performer/environment interactions are integrated in a more nuanced fashion. In addition, it is the first exhibition of singing, which marks a complete progression in expressive capacity from primitive utterance to complex vocal gesture.

Pitchpaint (module 10) is the finale. The performers display what seem to be their most deliberate and dynamic interactive powers (See Figure 3.3). Thus, it could be argued that viewers witness a complete evolution from simple creatures to fully expressive beings. Full multi-sensory interplay results, as physical movement, sound and imagery all contribute to the performance dynamic. Although, as already stated, the spatio-kinetic interactions seem more utilitarian in function rather than aesthetically nuanced components of the experience. As the creators articulate (with emphasis on the audio-visual interactions):

> The performers take their most direct control of the canvas, painting bold gestures by singing. In this section, descending pitches curl clockwise, while rising notes curl counter-clockwise, and unchanging tones produce straight lines. (Stroke thickness is governed by the singer's loudness; the color of closed regions is linked to vowel quality. The performers erase their marks by making the sound, "Ssh!") From the hushed, simple tones of the previous section, the performers develop an expanded vocabulary of quickly-changing glissandos, melodic fragments, and abstract speech-songs. (Levin and Lieberman 2004b, 12)

Unfortunately, the experience seems to fall flat in some regards. As LaBarbara articulated in an interview, “I’d like a little more control…There’s also cross-talk between the two microphones. Sometimes I would be singing and doing a line, but because of what he was doing it would get jagged.” (Bersten 2003)

Conclusion

As Levin and Lieberman state:

> Messa di Voce lies at an intersection of human and technological performance extremes, melding the unpredictable spontaneity and extended vocal techniques of two master composer-improvisers with the latest in computer-vision and speech analysis technologies. Utterly wordless, yet profoundly verbal, Messa di Voce is designed to provoke questions about the meaning and effects of speech sounds, speech acts, and the immersive environment of language. (Levin and Lieberman 2004b, 3).

The work does indeed provoke these questions. In addition, the synaesthetic associations resonate viscerally and seem valid as aesthetic ends in and of them selves. As the piece’s vignettes evolve in complexity certain nuances reveal themselves over the course of the performance. The seemingly implicit conceptual device of “language evolution” is appropriate. In addition, within specific vignettes, complimentary integration of certain pseudo-narrative structures creates another level of engagement for the audience.

On the other hand, the entire piece might have been better served by a more compelling concept that could engage audiences beyond the module specific vignette style of performance organization. Overall, one witnesses a great string of digitally mediated performance experiments that could be more effective as elements of a more cohesive narrative structure. That being said, the piece will still serve as a great study for multi-sensory artists looking to expand upon the performer/interactive-media-system dynamic. Through the use of synaesthetic devices we see an emergent aesthetic connected to the roots of interactive perceptual research and multi-sensory experience.

Figure 3.2. The "Bounce" section of *Messa Di Voce.* Permission granted by the artists.

Figure 3.3. The "Pitchpaint" section of *Messa Di Voce*. Permission granted by the artists.

Embodied Mediation and *16 [R]evolutions*

Troika Ranch's aesthetic approach is a topic of interest for Susan Broadhurst in the development of a neuroesthetic paradigm regarding digital arts practice and criticism. For Broadhurst, the interaction of the physical and the virtual within their work creates a form of digitally driven multi-modality that, through novel cross-sensory activation, frustrates expectation and undermines any attempt at simple interpretation. This process results in a new kind of effect analogous to the perceptual experience of synaesthesia. Such confounding aesthetic experiences also create the conditions for other interesting neural firing patterns that warrant the specific inclusion of Troika Ranch's work within Broadhurst's wider neuroesthetic theory perspective.

> ...all works of art that conflict with our prior experience of visual reality or, in other words, frustrate our expectations of any clear resolution, such as the art of the Fauvists and the Surrealists - and by implication the digital practices, such as those of Palindrome and Troika Ranch, are likely to activate a specific area of the frontal lobe which appears to deal with the resolution of perceptual/experiential conflict. (Broadhurst 2007, 58)

A recent work by Troika Ranch, *16 [R]evolutions,* condenses 100,000 years of evolution into an evening-length performance featuring their signature integration of dance, theater and interactive media. The work traces a path that begins with the pure drives of the pre-human animal and ends with today's post-intellectual man.

> Using live camera tracking technology, the movements of the dancers generate interactive 3D imagery in real-time. At one moment suggesting delicate, living strands of DNA that envelope the performer, at another, becoming a gigantic digital rib cage that ripples and breathes in direct response to the movements of the performers. (Troika Ranch)

The theatrical performance begins with an abstract interpretation of "Adam and Eve", the moment at which, mythologically speaking, humans had shed their primitive nature (see Figure 3.4).

> Two performers are lit only by the aforementioned digital rib cage; the "ribs" move in direct response to the dancers' movement, creating a sublime animal triad as the duo performs within it. But, at the penultimate moment the scene cuts with a loud "crack" to a seated duo, each with a bowl of cereal before them, the hum of a transformer suggesting a location deep in the bowels of some institution. After methodically preparing utensils, milk, and napkins, they both lean in to take a bite. The sound of their crunching is amplified to a deafening level – the unexpected change of scale magnifying the banality of their activity. Through such contrasts we will emphasize the two polarities of human nature that are the focus of this work. (Troika Ranch)

In order to explore the transformation from pre-human animal to post-intellectual man they respectively use a mix of beautifully complex interactive video and absurd theatrical scenarios. With regard to the "animal" motion-triggered imagery, they suggest a tension between the "dormant seed of reason buried within the creature", and the seemingly "harsh actions of a creature whose mandate is to survive at any cost" (Troika Ranch).

Video cameras and software perform real-time motion capture of the dancers' movements allowing them to interactively manipulate video, sound and light as they dance. The motion capture is realized with EyesWeb, developed by Antonio Camurri and others. The movement trajectories generated in EyesWeb are then sent to Isadora (software developed by Mark Coniglio). Isadora handles the generation of imagery and the manipulation of other media.

> The shape and speed of the path are analyzed to sense the "quality" of the dancers movement: lyrical, sharp, slow, fast, simple, complex, etc. The relative intensity of each of these categories can influence the timbre of

> the music, the color of the image, and other parameters, so the feeling of the media reflects the quality of the movement. (Troika Ranch)

In addition, the movement of individually tracked points determines how certain images are produced. One particular section of the piece specifically tracks the position of the performer's head, which connects the shapes to the mouth. In the process, shape size and density are controlled by vocal amplitude through a wireless microphone (Troika Ranch).

Figure 3.4. The "Adam and Eve" section of *16 [R]evolutions* (photo credit - A.T. Schaefer). Permission granted by the artists.

Analysis

As Mark Coniglio, founder and technology director of Troika Ranch, articulates, the piece (and their work in general) is explicitly content driven. By that he means that the technological and interactive elements serve a narrative structure (albeit abstract) as opposed to being ends in and of themselves. Coniglio describes content driven work as essentially guided by narrative or abstract-narrative structures and materials driven work as attempts to exhaust the possibilities of "process" (in this case interactive technology) as an end in itself (Coniglio 2007, 79-80). In other words, Content Driven = Narrative and Materials Driven = Abstract. As I argued before, a more content driven paradigm could have been a more engaging structure for *Messa Di Voce.*

Troika Ranch's approach does not come without its critics, however. One colleague, after witnessing one of the first performances at the Digital Cultures Lab in 2005 (Nottinghan, England), felt that the decision to not fully explore the digital materials throughout the entire piece could be viewed as lacking rigor (Coniglio 2007, 80). A *New York Times* review stated, "Mr. Coniglio and Ms. Stopiello apparently don't have faith that their technology will provide enough variety or meaning to sustain interest for over an hour " (Rockwell 2006).

Coniglio responded.

> I have absolutely no faith that the technology (the visuals and the interaction) will provide enough meaning to communicate the themes in which we are interested. Instead I propose to rely on all of the elements to intensify each other in a complementary manner, so that the viewer's attention remains intensely focused on the content of the work. (Coniglio 2007, 82-83)

I believe the criticism above is actually warranted. Although I have not witnessed the work live, while viewing the piece on DVD, I did feel my attention drift during some of the "post-intellectual" sections that were presented as "pure" theatre without interactive technology. One could wonder why there was not an interactive component developed to appropriately support these theatrical scenarios and narrative structures.

At the same time, I understand that occasionally the interactive visual elements become so alluring that one loses sight of the underlying concept (see Figure 3.5). Thus, I understand Troika Ranch's decision to prioritize support of the narrative over materials and process. Utilizing interactive processes within narrative structures is always a tricky proposition, requiring endless give and take in order to find the right balance. Unfortunately, I believe that the decision to create such a polarizing contrast between the "animal" and "post-intellectual" segments brings attention to the use and avoidance of the technology. This fact may have motivated the more critical reactions.

Overall, the moments that did incorporate interactive elements worked very well as abstract representations of embodied mediation, while also supporting the narrative. Sonically, however, the piece did not satisfy as much as it did visually. Although the underscore and sound design elements did support the concepts appropriately, I felt that the sonic themes could have been more entangled with the movement and imagery in order to create a more unified experience. However, the visual elements are so stunning at times that the sonic elements might need to be more supportive in such a context.

Conclusion

In contrast to *Messa DiVoce*, *16[R]evolutions* uses narrative structure to drive the interactive components in a very effective way. The embodied imagery exhibits a very compelling representation of the narrative themes. In essence, the theatrical motives are appropriately intertwined with all aspects of the multi-sensory experience, including the sound design and underscore music, which consistently compliment the other elements in the realization of an effective gestalt. Although one should not compare apples and oranges and deride the work of Levin and Lieberman due to its lack of narrative structure and more limited corporeal engagement, it does seem appropriate to highlight the differences in these works. There are clear distinctions to be made between the more abstract audio-visually dominant experience of *Messa Di Voce* and the more embodied, content driven work of Troika Ranch.

Although the decisions made in *16 [R]evolutions* to contrast the narrative poles with the avoidance and utilization of interactive technology seems somewhat glaring, one cannot argue with the overall result. The themes are clearly presented and their process is ultimately successful. The piece represents a complete multi-sensory exhibition of aesthetically integrated elements, all supporting a successful abstract narrative structure. As an example of the content driven approach to technologically mediated dance theatre *16 [R]evolutions* offers a glimpse at the true *Gesamtkunstwerk* for generations of digital artists to follow.

Figure 3.5. The beautiful abstract imagery of Troika Ranch's *[16] Revolutions* (photo credit - Richard Termine). Permission granted by the artists.

Multi-Modal Dynamic Systems and *Schwelle*

Schwelle, created by Chris Salter, is a three-part project exploring the, "varying threshold states of consciousness that confront human beings in everyday life experience, such as the onset of sleep, or the moments before physical death" (Salter et al. 2007, 4). The work, with its structural underpinnings based on the *Tibetan Book of the Dead,* was in development starting August 2005 and premiered in February 2007 at the Transmediale Festival For Art and Digital Culture in Berlin. The first part of the work (*Schwelle Part I*) was an abstract high definition sound and image composition formatted for a very wide screen. In *Schwelle Part II* Salter (and collaborators) created a responsive media environment for a digital performance based theatre work exploring the same themes. Interactions between a solo performer, Michael Schumacher (movement artist), and the dynamic environment drove the experience of the piece (see Figure 3.6).

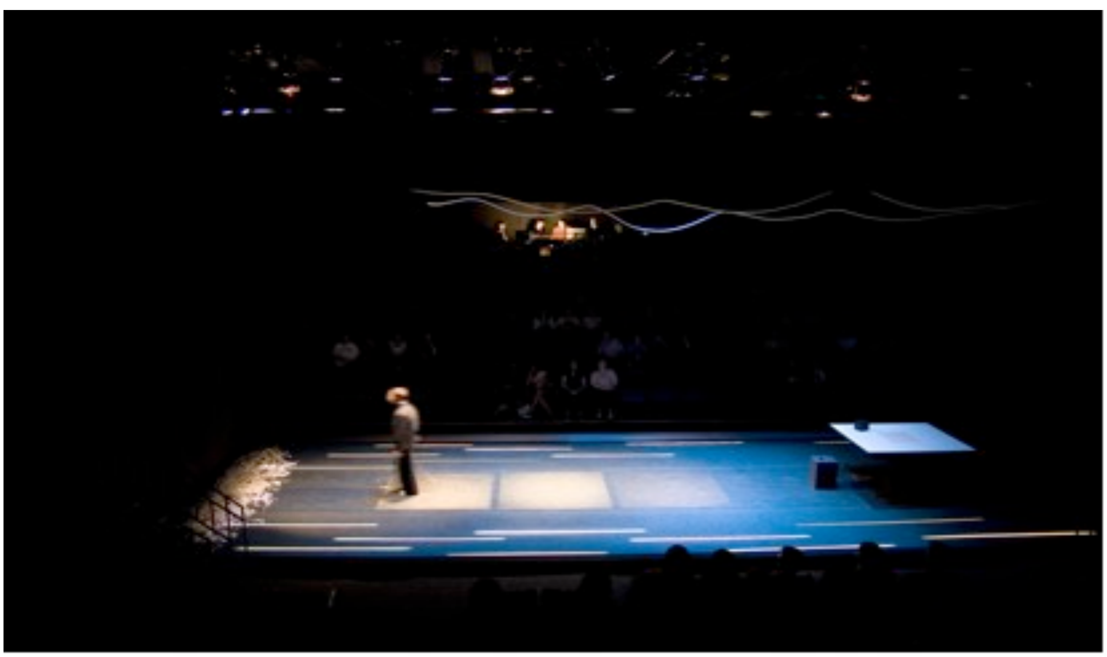

Figure 3.6. Michael Schumacher in *Schwelle's* responsive performance environment. Permission granted by the artist.

Salter found inspiration in the “responsive environment” paradigm credited to Myron Krueger in the 1970s. Krueger’s idea was generally defined as a physical space, “in which a computer perceives the actions of those who enter and responds intelligently through complex visual and auditory displays” (Krueger 1977, 423). Thus, Salter et al. (2007) created a multi-modal system constructed to generate dynamic sound and light patterns influenced by evolving environmental and physical gesture data sets over micro, “meso”, and macro time scales.

The system developed was a response to current human computer interaction (HCI) and algorithmic computer music approaches that usually feature playback-based composition. According to Salter, those methods of musical development do not address the complexities of composing for long-term, large-scale, multi-input, sensor-based environments. In response, Salter’s piece dealt specifically with the complex task of balancing legibility (listener understanding of mappings) with musical richness and complexity within the context of live performance, “where long term time evolution and duration are paramount to spectator/listener experience” (Salter et al. 2007, 2).

According to the authors, a particular set of technical and perceptual challenges presented themselves throughout the process. Because theater or dance performances take place over a specific amount of time and the audience is present for this entirety, “the perception of macro-structures (i.e., structural patterns, repetitions, sequences, motifs, themes and variations) becomes an essential element in the audience’s understanding of how the performance evolves over time” (Salter et al. 2007, 3). Because audiences are not involved in direct manipulation of instruments and controls, direct (one-to-one) physical interaction between the performer and the system has become the model for establishing perceivable relationships (Salter et al. 2007, 3). This seems unfortunate, according to the creators:

> The consequences of focusing almost exclusively on the short term level of mapping - sound generation - feedback is that most interactive performance works fail to exploit the thing that makes musical or dramatic performance compelling and affective to the listener: the unfolding of an event over a durational time frame...
>
> It is our contention that live performance involving sensor augmented responsiveness presents a particularly robust context in which to research new techniques for computer assisted composition in that explicit attention must be paid to the organization and structuring of sound over longer time scales. (Salter et al. 2007, 4)

So, their compositional challenge was in developing enough algorithmic complexity to generate compelling, yet conceptually driven patterns over long-term time scales, while also providing feedback in more medium (”meso”) and immediate time frames (Salter et al. 2007, 4). In a sense, *Schwelle* aimed to merge the successful aspects of Levin and Lieberman’s work, in terms of short-term interactions, with the overarching dramatic form of Troika Ranch’s aesthetic.

Analysis

The result in this instance was a theatrical performance that took place between a solo performer and a responsive space. The performer’s movements were captured with wireless accelerometers located on the two arms and chest. In addition, environmental light patterns (dynamically created by a lighting system) were picked up by photoelectric cells and wirelessly transmitted over radio frequencies to a central server. The continuously generated data from both the performer and the environment was then used to influence a dynamically changing multi-channel soundscape. In turn, it gave the impression of a "living, breathing room" for the audience (Baalman et al. 2007, 178). The audience's perception of responsiveness was the experiential key.

> ...the notion of responsiveness in the context of Schwelle as a theatrical event signifies two things: (1) the room is responsive in the sense that the environment provides information that is subsequently manifested (but not illustrated) through specific sonic structures having distinct timbral and rhythmic relations that behave differently depending on different "states" of the room and (2) the time evolution of this continually generating sound design has a certain set of identifiable patterns that emerge over the duration of the performance. (Salter et al. 2007, 4)

As the authors articulate, a certain tension always exists between musical structures that continuously evolve (based on new information from the performer/environment system) and the dramatic need to give those musical structures a clear sense of pattern.

It is in this sense that the piece truly succeeds. Like Troika Ranch's content driven process, Salter et al. clearly let the artistic concept drive the technological systems and aesthetic structures. The core dramatic foundation, taken from the *Tibetan Book of the Dead,* describes experiential states undergone by the body after death emphasizing the interactions that take place between the deceased and the environment he/she inhabits. The technology is clearly used in a way that fundamentally supports these abstract narrative underpinnings.

For instance, the choice of utilizing accelerometers to sense the performer's body movements had a specific narrative function. The authors feel that the most important expressive information in physical gesture is derived from force (Baalman et al. 2007, 180). Thus, there are clear moments in the piece where this force/affect relationship was exhibited. For instance, in sections that seemed to depict "the threshold states of death", the dancer's epileptic movements worked synchronously with frenetic sound design and pulsating light. In contrast, other moments featured more fluid movements with ambient lighting and atmospheric sounds. Overall, compelling congruencies were achieved between the performer and the responsive environment (see Figure 3.7).

The interactive relationships between the performer and sound design elements seemed to be particularly effective in support of the concept. The sonic system was organized around 16 layers of sound generated within the SuperCollider (SC3) programming environment. The layers were constructed with the following qualitative characteristics in mind (controllable parameters in parentheses): continuous background noise (frequency, amplitude, modulation speed); clouded events (density, frequency range, amplitude range, duration, amplitude); and regular, discrete events (frequency, tempo, amplitude, duration) (Baalman et al. 2007, 182-3).

> These sounds and their resulting musical parameters such as frequency, duration, rhythm, etc., were developed based on questions that emerged from Schwelle's overarching dramatic structure: what kind of range of affective or emotional behaviors would the room environment exhibit over the course of the performance and how could such behaviors, "attributes" or states be manifested to the audience/listener through qualitative phenomena like density and thickness, legato-like continuous structures versus transient "events", clustered versus particulate sounds and smooth versus jittery textures, among others. (Salter et al. 2007, 5)

An interesting choice was made in terms of not utilizing any sort of underscore material. The authors consciously abandoned "external" musical structures that would reinforce a particular mood or heighten dramatic tension (Salter et al. 2007, 5). I believe this worked to emphasize the internal experience of the performer as a dynamic and fluid part of the whole. This approach also gives the audience a true sense of "agency" in the reading of and influence over the work (Baalman et al. 2007, 179). The audience is meant to be part of the environment and an active part of the system.

To address the temporal evolution of the piece, the authors turned to work in the area of dynamical systems (Franklin et al. 1993). Their aim was to create a time-varying model of interaction where the compositional structure would depend on multiple feeds of real time data from the room and the performer. The resultant system was based on the early ideas of J.F. Herbart (1969), who developed a theory for the "strength of ideas" based on sensory impressions as a function of time and strength of impression (Salter et al. 2007, 5).

> It is this metaphor, which we also apply in *Schwelle*: the room's physical behavior makes certain sonic ideas stronger within the room over a longer time scale. Each type of sensor input has a separate scaling factor to determine the amount of influence it has on the system. Within the multiple layers of sound, each layer consists of one single sound object (object in the sense of a sound that stands on its own), whose intensity (amplitude) and amount of occurrence is dependent on how strong the "idea" within the system is at a certain moment in time. The Herbart system functions in such a way that there is no direct one-to-one mapping, but that its output also depends on the past of the system. This makes sense musically, as it reflects how we perceive and understand the temporal evolution of music. (Salter et al. 2007, 5)

After the first performance the authors decided to elicit some audience feedback. In response, they implemented a state-based system on top of the dynamic systems model to influence the patterning of more macro sonic structures. As a result, they were able to clearly define five qualitative sound states that occur during the performance, such as "meditation", "repressed anger", "transient anger", "sleep/dream" and "agitation."

> For each of these states, we have designed the parameter limits in which each sound layer would shift to reflect that state. Some sound layers are

> more present in some layers than in others, or have different frequency ranges, durations and so on within certain layers. (Salter et al. 2007, 6)

Essentially, the outcome is two compositional systems operating in tandem with one another: the dynamical system, which influences what and when certain sounds are strongest, and the state system, which controls the larger macro formations that involve "meso" structure repetitions over the course of the performance's evolution (Baalman et al. 2007, 183).

Conclusion

Schwelle Part II provides a complete multi-sensory aesthetic experience demonstrating the seamless integration of technological mediation within dramaturgical structures. It is obvious that the conceptual underpinnings of the piece, based on threshold states and human-environment interactions through death, are fundamentally tied to the responsive environment paradigm implemented.

In addition, the technological components compliment the composition over the long time frames required in such theatrical contexts. The performer's movement and theatrical gesture is clearly integrated with the exhibition of, and interaction with, the media elements. The fact that the audience also plays an active role in the environmental contribution to the work completes the loop of connected experience that is implicitly suggested in the conceptual framework.

Salter et al. clearly offer a bold and innovative example of technologically mediated composition that should become a model for other artists working within digital theatre and dance contexts. Specifically, they demonstrate an effective paradigm for complex performer-environment interactions in support of dramatic aims.

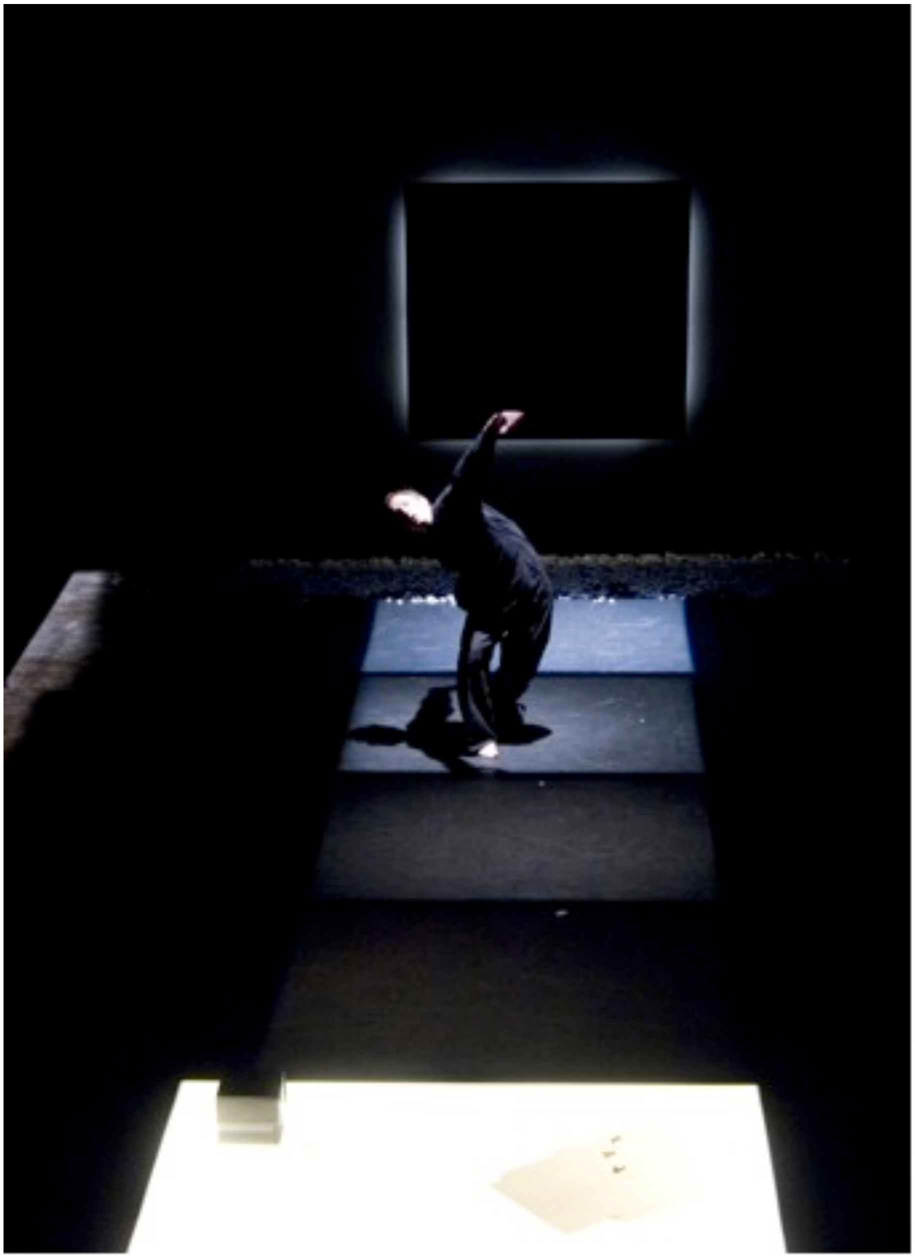

Figure 3.7. A still image of the performer in action. Permission granted by the artist.

CHAPTER 4: THE AESTHETIC AND TECHNICAL DEVELOPMENT OF *UNLESS* (2009)

The artistic portion and main focus of my dissertation project resulted in the creation of an original digital performance work featuring movement artist Emily Beattie and an interactive multimedia environment. The piece, entitled *Unless,* took the form of a structured improvisation in parts, and was presented twice in two different venues. The interactive multimedia environment was custom designed to respond to the dancer's physical gesture and offered compelling connections between specific aspects of the movement in relation to the sonic and visual elements of the piece.

Development of the interactive multimedia environment needed to satisfy two main objectives. One being the fulfillment of thematic and conceptual goals related to the abstract narrative structure of the piece (an emergent process to be further highlighted in Chapter 5), and, the other being the fashioning of kinesthetic/sonic/visual relationships that an audience could experience in a direct and visceral way. The piece consisted of three sections abstractly exploring a narrative trajectory examining humanity's inherent connectivity and capacity for empathy (via mirror neurons and the "social brain" etc.) in relation to the precarious situation we find ourselves faced with as we struggle to find a new balance between ourselves and with our environment(s). Therefore, an interactive environment aesthetic (echoing Krueger's responsive environment paradigm, but in a performance context) was deemed necessary to express these dynamic relationships.

On a theoretical note, it is speculated that spectators observing any physical action have the capacity to internally simulate movement sensations such as, "speed, effort, and changing body configuration" (Hagendoorn 2004). These phenomena describe the essence of the concept of kinesthetic empathy. Interestingly, audiences can tap into this experience even while sitting still, as they feel that they are participating in the movements they observe, and subsequently experience related feelings and ideas. As was alluded to in Chapter 1, these phenomena are at least partly related to mirror neuron activity in the brain.

Creating strong feelings of kinesthetic empathy in the audience (especially regarding their connection to the dancer) was extremely important to the overall experience of the piece. I hypothesized that a movement-activated multimedia environment might actually enhance feelings of kinesthetic empathy while also serving the conceptual and thematic aims of the performance piece.

Therefore, the technical development of the interactive multimedia environment and its modular system software components were of primary importance to the creation of the piece. Overall, the environment included two distinct interactive music systems (one used in the first section of the piece and the other used in the third section of the piece) and one graphics engine (used throughout the piece in various configurations). The second section of the piece featured pre-composed music and interactions between the sound, mover and visuals (the sound itself was not influenced by movement).

The physical interface for both interactive music systems consisted of two Nintendo Wiimotes (one held in each hand by the performer) that were used to detect motion in three axes (X, Y, Z) utilizing their built-in wireless accelerometer technology. In turn, custom designed software modules (one for each music system) could parse this raw data into useable motion parameters for further mapping within the two distinct interactive music modules. In addition, the separate visual graphics system could analyze the sound (which was mostly influenced by movement) and react accordingly. All of the software and media elements were original/customized creations of the author.

The first version of the piece was created over five months between December 2008 and April 2009. It was performed for over 60 people on May 1st of 2009. Three projections were used for a partially immersive (frontal horseshoe shaped) graphics display, while a 5.1 speaker system offered surround-sound music and sound design. In addition, a second version of the piece was created for a smaller gallery space in November of 2009 with two mirrored projections in a V shape configuration and an immersive 3-channel speaker array. The research and developmental processes regarding the interactive aspects of this work will now be detailed.

Observational Studies and Concept Development

The project commenced with an exploratory research phase in collaboration with Jeff Hutchison, a fellow graduate student from the cognitive sciences. I met Jeff, who is also a Butoh trained movement artist, in the spring semester of 2008 at a contact improvisation course in the Brown University dance department. As fate would have it, the course was taught by Emily Beattie, a Boston based choreographer and movement artist, who later emerged as my principal collaborator on the project.

Emily had put out a call to improvisational musicians at Brown to participate in the movement classes, providing spontaneous soundscapes for her students to move to. I got the gig and put together a duo with Bob Fish (a local drummer and percussionist with a world music background). In these sessions, Bob would create improvised rhythms while I added texture with an electric guitar and live electronics. This opportunity served as my first up close experience of movement art in process and helped to crystallize my goals for this dissertation project.

Looking back, it is now safe to say that these improvisational sessions served as my first bits of research for the work ahead. It was interesting for me to directly experience the affect of spontaneous music making on the movement quality of the students and vice versa. Fortuitously, I videotaped the sessions and upon review I was struck by how dynamic the connections were between our sonic textures and rhythms and the dancers' improvised movements (See Figure 4.1).

One moment, which reverberates in my mind to this day, was particularly memorable. Bob and I had been improvising without any real temporal pattern when suddenly Bob broke into a disturbingly iconic rock beat. Almost immediately, yet still artistically and fluidly, the movers transitioned in unison from smooth wave like motions into sharply jagged bouncing movements in lock step with the beat. It was a strikingly simple, yet elegant, and spontaneous example of organized sound and movement's inherent connectivity.

Figure 4.1. A still image extracted from video documentation of Emily Beattie's improvisational movement class

One of those students, as I mentioned before, happened to be Jeff Hutchison. One day after class Jeff approached Bob and me and expressed his gratitude for our volunteer work with the class. He mentioned that he was really interested in movement's connections with music. I mentioned that I was interested in creating an interactive multimedia environment (with sound/music and graphics) that would respond and be influenced by movement, and perhaps we could work together. In the meantime, I was finishing my qualifying exam work and starting to develop my dissertation proposal. This new partnership would help create the solid scholarly and artistic foundation that would shape the next two years of my academic life. The project was born.

Jeff and I started meeting in order to, first, brainstorm approaches to interactive musical systems (parts of a larger interactive multimedia environment) that could respond to movement and be effective in an artistic performance context. I was adamant from the start that the purpose of this process was not to create a demonstration module, but a useful artistic environment that could believably contribute to a performance work. This would be the real challenge.

Initially, Jeff offered his own insightful perspective on some of the scientific and philosophical foundations of the project. He let me borrow Francisco Varela's *The Embodied Mind* (mentioned in Chapter 1), a book that would become instrumental for me as I conceptually formulated the interactive media environment, and ultimately the artistic vision. We talked a lot about the mind/body system, human/environment dynamics and the emergence of consciousness from these mutually influential entities. Such ideas helped me conceptualize human movement, the computer, the physical interface and the sound world as a complete dynamic system of human/environment interaction, a feedback loop of sorts. It was helpful for me to think of the resulting sound world in terms that were just as influential on movement as the movement would be on the sound world.

The first stages of the project began with observational studies of Jeff's movement in different environmental contexts. Jeff and I decided that it would be interesting if he also abstractly vocalized his movements with pre-verbal phoneme-like utterances through a wireless lavaliere microphone while his movement experiments were recorded. We went to roadside construction sites, abandoned structures on the Providence River waterfront, the banks of the Blackstone River, the shore of the Atlantic Ocean and a stone quarry to name a few of the places (See Figure 4.2).

These studies offered some interesting insights, which would influence the technical and artistic conceptualization process moving forward. Generally speaking, they served as mindful thought experiments on human interaction with the environment. Given the nature of our times, this seamlessly led to thoughts about the current state of our relationship, as a species, to the environment. Although a loaded topic in itself, it would serve as artistic inspiration throughout the process, even if the connections were ultimately abstract and more representative of my internal process rather than explicitly tied to the realization of the work. That being said, the title of the work is secretly connected, as will be divulged later.

Specifically, these studies helped us to think more concretely about the movement/sound/image dynamics we wanted to explore with an interactive performance environment. For instance, the idea that the computer could create a semi-autonomous sound/graphics environment, with a "mind of its own", yet, still open to the influence of human interaction, became fundamental to the project. I was struck by the way environmental structures and moods could influence Jeff's movement. Therefore, I wanted to create a holistic media environment with its own influential power, dynamic in its interactivity, and not just an instrument/object under human control.

In addition, Jeff's vocalizations, also captured by the documentation, allowed us to think more deliberately about how sensory metaphor could contribute to the creation process. For instance, these early experiments helped me to clarify some of the natural ways in which we relate speech sounds, specifically, to movement. It was strikingly difficult for Jeff to move jaggedly while vocalizing a more fluid sound. Maybe we relate sound in general to movement based on these natural motor-language relationships.

Sharp and fast movements were largely correlated with angular vocalizations and abrupt sonic quality shifts while more fluid movements were characterized by smoothly textured vocalizations and more gradual qualitative state transitions. On a more fundamental basis, it seems that energy and effort are essentially the underlying factors that shape the quality of both sonic and physical gesture.

In tandem with these abstract movement studies, I started the development of custom software in Max/MSP that could parse and map movement data into parameters influencing sound. I also started experimenting with different physical interfaces that could reliably sense energy and effort without being invasive and encumbering to the artistic expression of movement.

Figure 4.2. A still image extracted from video documentation of a movement experiment at a roadside construction site.

The Conceptualization of the Graphics System

The exploratory research described above solidified the project's technical direction. The performance piece would be centered on dynamic interactions between a movement artist and interactive multimedia environment. I was not interested in the idea of a digital performance instrument under human control as much as the give and take dynamic flow of a human and its environment. That being said, I wanted the movement/sound relationships to be the most palpable of the human/environment interactions. I thought that an audience could very directly experience those elements existing together and mutually influencing each other in the same performance "world".

I also wanted to stress the idea that this performance "world" was located within another plane of existence. Therefore, I thought that an immersive projection-based graphics component could provide greater environmental context for this alternate reality, such as a constructed set provides context for a theatrical performance. Yet, in order provide more foreground focus on the movement/sound aspects of the human/environment interaction (at least for a large portion of the piece) some separation between the movement and the graphics system had to be considered. So, it was decided from an interaction design standpoint, that there would usually be some level of disconnect between the physical movement of the dancer and the moving images within the performance environment. I felt that it would be interesting to use the environment's sound as the mediating influence on the graphics rather than them being a direct reflection of some captured visual element (although that was not always the case).

Therefore, the graphics were conceptualized as an ambient force, emotional texture, or atmosphere in the "environment" of the piece rather than a foreground presence. The immersive visuals served as part of a dynamic “environment”, so it was necessary for them to influence, and be influenced by, the world they were a part of. In order to create something immediate, impacting and holistic, I decided that the most direct and visceral relationship between sound and image was the best for the piece. Therefore, I found the relationships between sonic amplitude and visual change, in general, to be the most effective parameters to work with. Theoretically, as well as experientially, this approach seemed well founded.

Additional Theoretical Justifications for the Graphics System

One of my influences derived from Michel Chion’s ideas (as detailed in Chapter 1), and research from the field of systematic musicology. For instance, as was mentioned, Chion defines "synchresis" as a spontaneous psycho-physiological phenomenon that contributes to a reflexive merging of sonic and visual events into one gestalt. Therefore, sound is free to express something beyond the nature of its origin. If sound had to mimic reality alone, the range of possibility for expression would decrease considerably (Chion 1994).

Research from the realm of systematic musicology provides empirical evidence aligned with Chion's observations while bridging his notions regarding film into the world of abstract graphics (Lipscomb 2005). This work confirms that synchronous relationships between sound and image inherently provide a basis for perceptual coherence.

In *The Perception of Audio-Visual Composites: Accent Structure Alignment of Simple Stimuli*, musicologist Scott Lipscomb's research suggests a correlation between the synchronization of “accent” in audiovisual composites and the perceived "effectiveness" of the material (2005). His generalized definition of an accent, based on David Huron’s ides, is described as follows:

> The term "accent" will be used to describe points of emphasis (i.e., salient moments) in both musical sound and visual images. David Huron (1994) defined "accent" as "an increased prominence, notice-ability, or salience ascribed to a given sound event." When generalized to visual images as well, it is possible to describe an A-V composite in terms of accent strata and their relationships to one another. (Lipscomb 2005)

The experimental stimuli utilized to create the audio-visual “accent” composites consisted of simple pitch sequences in equal durations in conjunction with simple computer generated animations of a circle moving on screen. When audio and visual accents were aligned, subjects' ratings of synchronization and effectiveness showed strong positive correlation. This research shows that audiences perceive closely synchronized material as being more “effective”.

The results of Lipscomb's experiments in concert with Chion's observations make it seem foolish to dismiss the impact of closely synchronized material. As Grierson (a computer arts researcher at Goldsmith’s College in London) states, "Audiovisual composition, in fact, may rely on an understanding of this 'effectiveness' and the complexity of its operation" (Grierson 2005). These complexities, however, need to be inspected further. For instance, Lipscomb did not feel that temporal synchrony alone determined the perception of an effective audio-visual composite. He speculated that a certain amount of perceptual coherence between the changing parameters of the components must be present for the perception of an effective composite experience.

Thus, in exploratory research (before performing the main experiment) using the same types of simple stimuli, Lipscomb studied musical variables such as pitch, loudness, and timbre in relation to visual variables such as location, shape, and color. His aim was to clarify the audio-visual relationships necessary for perceptual coherence, and ultimately, the creation of reliable accent composites. It was assumed that a noticeable change in a given variable could create the perception of an accent for a given audio or visual component. He also suggested that "perceptual congruence" between the changing audio-visual components would also be necessary for perceiving an effective combination (Lipscomb 2005).

In a related study, Lipscomb and Kim (2005) delved deeper into the notion of perceptual congruence. For that experiment, a total of 48 A-V composites were created in Flash. Again, extremely simple stimuli were used. Sounds were generated and manipulated (for timbre, pitch, loudness and duration) with a Yamaha DX-7 synthesizer and images comprised simple objects morphing in color, size, shape, and location. The results suggested that there are, indeed, differences in "degree of match" between the perceptual attributes of an A-V composite.

> The highest ratings of perceived match suggest the following pairings: pitch-location, loudness-size, and timbre-shape. Also emerging, however, is evidence of non-unitary relationships. For example, the visual attribute of color matched equally well with both pitch and loudness in the auditory domain...though the primary pairing receives a higher mean score, secondary relationships (such as loudness-location or pitch-size) may provide a source of variation that can be integrated into the audio-to-visual algorithm. Also, the auditory parameters of timbre, pitch, and loudness may all provide an acceptable matched pairing for the visual attribute of shape. (Lipscomb and Kim 2005)

Interestingly, the ratings did not vary significantly as a result of either musical or visual training.

Overall, the research suggested the existence of an inherent, albeit culturally based, coherence regarding the audio-visual relationships studied. Specifically, changes in loudness or amplitude seemed to be especially compatible with a wide range of visual attributes. Based on this knowledge and my own intuition, I decided to use real-time changes in amplitude to drive the primary audio-visual interactions in the graphics system.

Performer Research

Exploratory research was conducted throughout development of the interactive multimedia performance environment in different phases. As mentioned, my first collaborator, Jeff Hutchison offered a lot of insight on system conceptualization and development in what would be phase 1, through his environmental movement experiments (in the summer of 2009) and initial tests with prototype versions of both interactive music systems used within the larger "environment".

In an October 2008 email, after an exhaustive two months of research, Jeff summarized our status and offered some personalized wishes for the future:

> Hey Brian,
>
> My primary feedback is that I think it will be better to use data that will allow for larger over all gestures – with the subtler controls, such as tilt and acceleration, controlling the parameters of the situation, so if I do a large arm circle - this will have significant meaning for something like pitch etc. but the way in which that arm circle is executed - slow and

smooth for instance or fast and jerky or whatever - will define the effect - it will control the parameters of the situation.

I would also really like to see the soundscape changing in response to the sounds produced - so if the system could monitor itself so that getting into certain sonic ranges will send the system into a new state - bifurcations mediated by the movement itself instead of a button press – a relatively binary change.

Those are the biggest things that I think will really push us into a more free realm of movement – I am still a bit concerned with having these controllers on my hands. It is convenient, but my hands are so expressive and with so much concentrated on them, the movement feels more limited – I may need some time just to get used to the situation - and maybe the above suggestions will take care of it.

Peace,

Jeff

Jeff's need for greater system recognition (or sensitivity) regarding gesture qualities would be the focus of development long after his departure from the project. In addition, his insight regarding a need for some semi-autonomous system behaviors (as already mentioned) would be instrumental to the aesthetic development of the piece. His concern over the hand held nature of the physical interface would also continue to be discussed during later phases of the project. To say the least, Jeff's contribution to the project would have a lasting impact.

Shortly after this insightful email, Jeff decided to leave the Ph.D. program in Cognitive Science at Brown. Although this was always a possibility, it was still a disconcerting turn of events. In addition to leaving the program, he decided that he would move back to his native state of California, thereby marking his departure from the project.

Performer/Environment Symbiosis

With Jeff's departure I quickly started searching for a new movement artist who could commit to the project's performance deadline of May 1st 2009. Throughout the month of November (2009) I met with various student movers from Rhode Island College and Brown University. In addition, I solidified plans to use a black box theatre space at Brown for a week of movement experiments and "auditions" with the system.

Realizing that time was of the essence I also contacted Emily Beattie directly. Having worked with her and her contact improvisation class earlier that year, I knew she was a very experienced performer, and luckily she expressed interest in the project. After a couple of experimental sessions with her in November it was clear to me that she would be perfect for the project. Understanding that I needed a real commitment, she decided to dedicate some time in December to determine her true level of interest.

The week of auditions and experiments in December 2009 would prove to solidify the core elements of the project. The movement students from RIC and Brown offered many interesting suggestions and allowed me to get a feel for more varied perspectives on the interactive system components and the environment as a whole. Leading up to this week, I also had time to finalize certain aspects of the graphics system, which would lead to my first experiments with all interactive media systems functioning for the environment.

It was actually in one of the sessions with student movers when I realized that the second interactive music system needed to be greatly simplified. It became clear that I needed an interactive sound environment that complimented and contrasted with the more aggressive and chaotic nature of the first system developed (to be discussed in more detail later). This would help to realize the aesthetic goal of creating two uniquely textured sonic environments, allowing me to highlight a conceptual contrast between connectivity and struggle. The movers' sense of connectivity to the simplified system was instantaneous, marking a new stage in the creative process.

Throughout this week of work, it also became clear that Emily's skill level as a performer was perfectly suited to the task at hand. During a couple of very intensive sessions, Emily started to really connect with the music systems (now in close to finalized forms) and began the process of slowly relating to the graphics system. A symbiotic relationship emerged from this process. The separation between interactive media environment and human performer started to blur. Energized by the process, Emily decided to see the project through. Thus, the true artistic aspects of collaboration would begin in earnest from this point forward.

Overview of the Interactive Multi-Media Performance Environment

Before delving into complete aesthetic descriptions of the sections comprising the performance piece (to be highlighted in Chapter 5), I will describe the individual technical components contained within the larger interactive multimedia environment featured in sections 1 and 3 of the piece. The simpler section 2 environment, which included a playback system for the pre-composed musical element, will be discussed in Chapter 5. Figure 4.3 shows an overview of the system components within the "fully" interactive section 1 and 3 environment. The reader is also directed to the accompanying Data-DVD with (among other things) a folder of Max/MSP/Jitter Patches and other related software elements comprising all of the main software components in the environment (See Appendix A for a listing of the DVD contents).

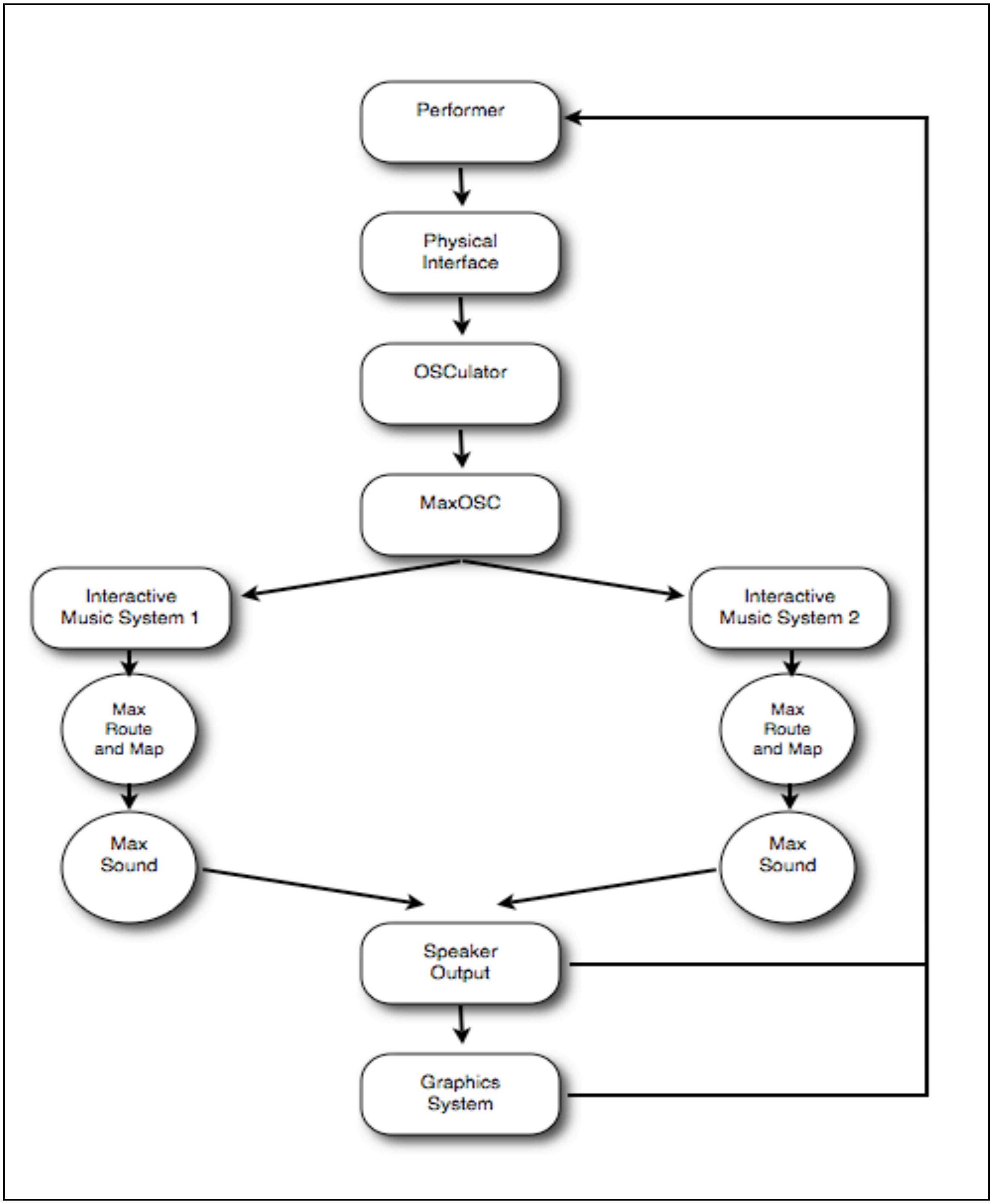

Figure 4.3 An overview diagram of the system components comprising the interactive multimedia environment.

The Physical Interface

The piece, from the start, was to be a structured improvisation and the interface needed to allow for spatial freedom and spontaneous movement. In addition, I wanted to have the liberty to make slight adjustments in response and to create smooth transitions during performance. Computer vision (CV) techniques were quickly ruled out because this sets up an unwanted spatial hierarchy on stage. Simplicity was a key force in the decision making process.

Based on my knowledge of accelerometers and their ability to sense force/affect relationships, I felt that a wireless device that could transmit this kind of data would be ideal. So, Jeff and I started experimenting with the WiTilt by Sparkfun and the Wiimote by Nintendo (See Figure 4.4).

With some help from NYU's ITP program website (article by Pinky Magbanua), I was able to configure the WiTilt in Max OSX's Terminal program to access its wireless data stream via Bluetooth before receiving with the Max serial object (Magbanua). It proved to be a fairly reliable device. However, with longer distances away from the receiving computer the connection sometimes failed. Within about 10 feet it performed well, but once outside of 15 feet, its reliability would drop considerably. This was not a very comforting experience for either Jeff or myself. It was hard to tell if this was a Bluetooth connectivity issue or a problem related to the Max serial object, which has always been a sore subject for my colleagues and me. From there our focus turned to the Wiimote.

In the beginning the Wiimote occasionally suffered from some of the same reliability issues regarding distance from the receiving computer. However, overall, it seemed to behave in a more consistent manner. After some initial experiments and tireless scouring of the Max discussion boards, I stumbled across a potential solution to the occasionally flaky behavior. At that time, I was still using the ***aka.wiimote*** object to receive data from the Wii. It became apparent, that there was a better solution with a great reputation for reliability. ***OSCulator***, by Camille Troillard, a standalone shareware program, could apparently run in the background, behind Max, and receive the data over Bluetooth and subsequently provide the capacity to format the data according to the OSC protocol. Not only was device pairing and synchronization with the computer efficiently and reliably handled, the program also took full advantage of the Wii's data resolution from each of the 3 axes of the accelerometer. The ***aka.wiimote*** only allowed for 8-bit data reception. In discovering this software, I have also wondered if Camille could program OSCulator to help with the WiTilt data, thus avoiding Max serial altogether. So far, he has not attempted this feat, but I plan to continue requesting this feature.

In addition to the 3-axis accelerometer, the Wii also has an infrared (IR) camera built in. Initially, it was interesting to explore that capability, but I determined that its low data resolution was not sufficient for the kind of robust expressivity we were aiming for. The accelerometers from a Wii held in each hand proved to be expressive enough for an interactive music system. ***OSCulator*** handled two Wiis at once with very acceptable data sampling rates. With the physical interface finalized I started working on the music modules.

Figure 4.4. Sparkfun's WiTilt and the Nintendo Wiimote side by side.

Wii Data (OSCulator)

The first layer of each of the music systems (identical for both) comprised the shareware software ***OSCulator (version 2.6.5)***. ***OSCulator*** functioned to pair/sync the 2 Wiimotes with the computer via Bluetooth (MacBook Pro, OSX 10.4.11) and provided the capability to transmit the Wiimote data over a local network utilizing the OSC data protocol to Max/MSP. As stated on the ***OSCulator*** forum in response to my question, Troillard reports that the software is able to sample the Wiimote accelerometer data with 10-bit precision on the X axis and 9-bit precision on the Y and Z axes. The sampling rate for the Wii data is roughly 100 hz. In turn, I was able to set the program up to transmit this data at full resolution and sampling rate to Max/MSP via OSC over localhost:9000. ***OSCulator*** also offers some simple data smoothing (filtering), which I applied minimally to the accelerometers.

The pitch, roll and yaw data, extracted from a portion of the raw accelerometer data, proved to be the most useful for expressivity. ***OSCulator*** offered a nice option for transmitting this data separately from the raw acceleration data. In aviation or nautical terms a vessel can pitch (with the nose moving up or down), roll (tip left or right, such that one may fall off the deck into the water), or yaw (with the nose shifting its forward angle left or right, such as to change direction) (See Figure 4.5). Such parameters were interesting to explore for movement, as they lent them selves to a navigational metaphor akin to the navigable system interaction design concept suggested by David Rokeby (as mentioned in Chapter 2).

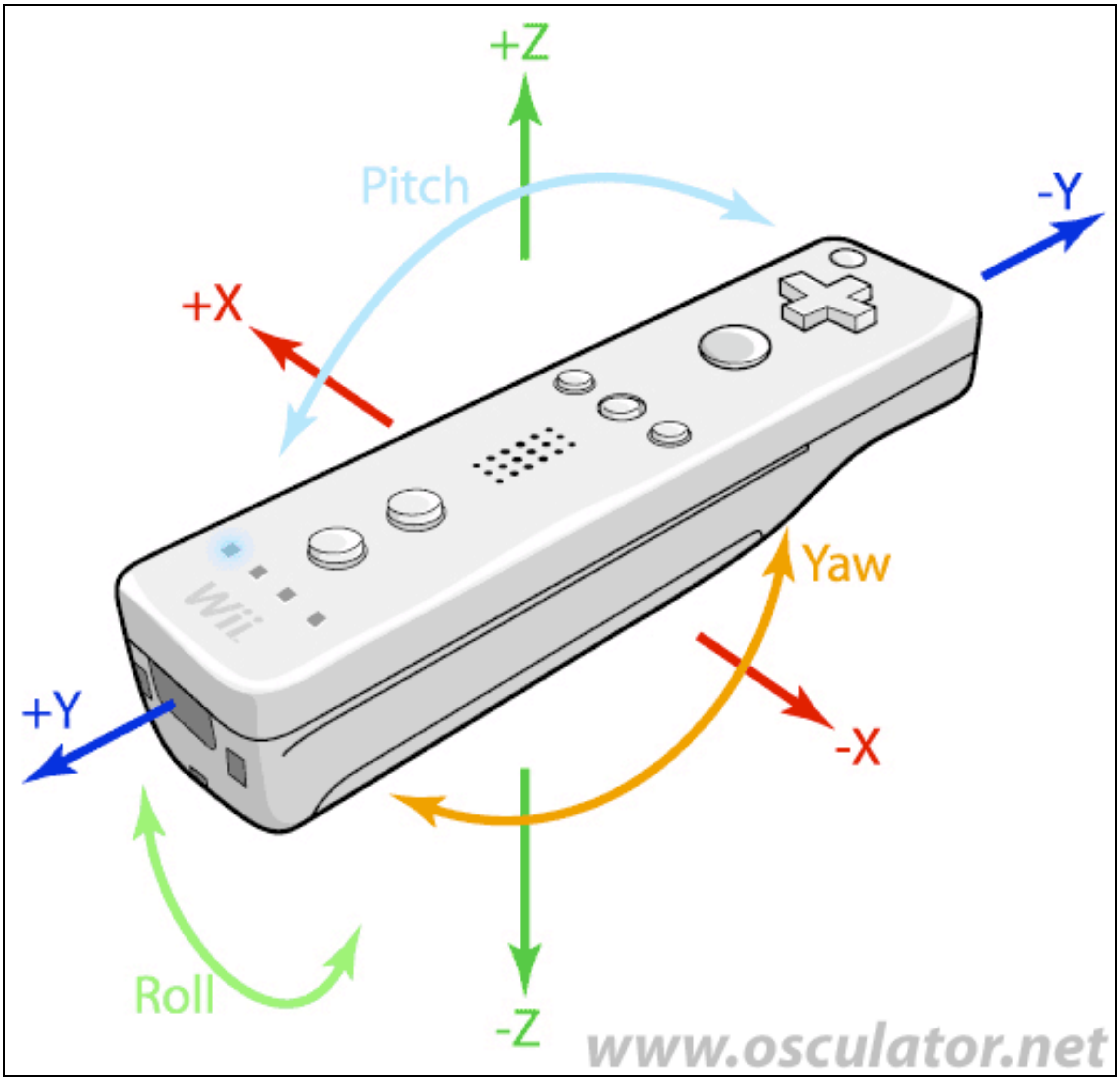

Figure 4.5. A diagram illustrating the pitch, roll and yaw capabilities of the Wiimote. Courtesy of www.osculator.net and Camille Troillard.

OSC in Max

The second layer of each music system (also identical for both) took advantage of external objects created by Matt Wright, Andy Schmeder, and Michael Zbyszynski of CNMAT for OSC data reception and routing in MAX/MSP (See Figure 4.6). After some experimentation, Jeff and I determined that out of all of the accelerometer data given, pitch, roll and overall acceleration were the most intuitive movement parameters to use for mapping within the system. This data was passed from OSCulator to the udpreceive object in Max over localhost:9000. Subsequently, the OpenSoundControl object in tandem with cascading OSC-route objects allowed me to use send objects to transmit the unpacked data to the next layer.

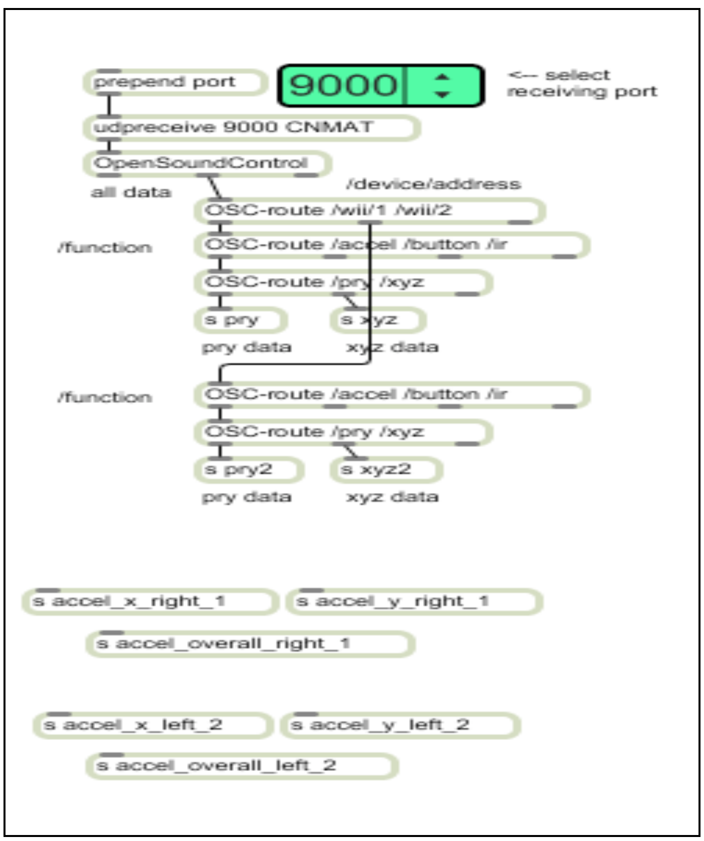

Figure 4.6. A screenshot of the OSC receiving/routing layer.

Details of Interactive Music System 1

Interactive Music System 1, the more aggressive and chaotic of the two, comprised a unique data routing/mapping module, along with a fairly complex sound generation engine. The sound generation engine consisted of a granular sampling engine, which was augmented through live processing in the form of delays, ring modulators, overdrive, and reverb. In addition, an analysis/resynthesis engine accompanied the granular texture in real-time. The resynthesis module utilized an FM enhanced additive synthesis sound generation engine before also being routed in parallel to the effects processors. The system will now be described in detail.

Max Data Routing and Mapping

This layer of the system allowed for efficient routing and creative mapping of the movement data in order to influence the core sound generation module (See Figure 4.7). Specified ***receive*** objects received roll, pitch and overall acceleration data from the OSC layer in float form between 0. and 1.0. ***Multislider*** objects allowed me to monitor the information stream in a visual form for real-time feedback and trouble shooting. ***Matrixctrl*** objects allowed me flexibility in routing the movement parameters to specified ***scale*** objects, which scaled the data into useable ranges for influence over the core sound module. This proved especially useful during the exploratory phase of system development while experimenting with different mapping techniques before deciding on final configurations.

During the exploratory phase I had the ability to map roll and pitch data to continuous control of grain duration, grain pitch, the resynthesis scaler, the modulation index (for FM resynthesis), and two different delays. I settled on using one Wii's roll and pitch data for grain duration and pitch respectively, while the other Wii's roll and pitch data was used for the modulation index and synthesizer scaler parameters of the additive/FM re-synthesis module. Grain duration was scaled to 0-300 ms, grain pitch to 0.2 to 10.0 (with 1.0 representing the original pitch of the sample), the re-synthesizer scaler ranged between 0.1 and 10. (with 1.0 representing the original frequencies extracted from analysis), and FM modulation index ranged between 0.0 and 5000.

In addition to these continuous controls I mapped overall acceleration into a switch of sorts using a ***past*** object as a threshold for signaling more global system state changes. This allowed the system to respond to fast and abrupt attack movements in a percussive fashion that created major shifts in the overall quality, dynamic and texture of the sound environment. One of the Wiis, for instance, could trigger global preset shifts for processor parameter configurations (affecting the delays, ring modulators, reverb and overdrive settings) in addition to changing the sound sample feeding the granular engine. The other Wii had the capacity to create dynamic shifts in the re-synthesized partial components in addition to amplitude shifts in the relative mix of effects processors. ***Counter*** objects stepped through these numerically determined state changes, rarely repeating the same arrangement of states.

Overall, this interaction offered a chaotic, and seemingly indeterminate system response (experientially speaking) to abrupt and forceful attack-like movements. This relationship between the mover and the sound environment spoke on many levels, appropriately serving the conceptual needs of the section, as will be discussed in more detail later in Chapter 5.

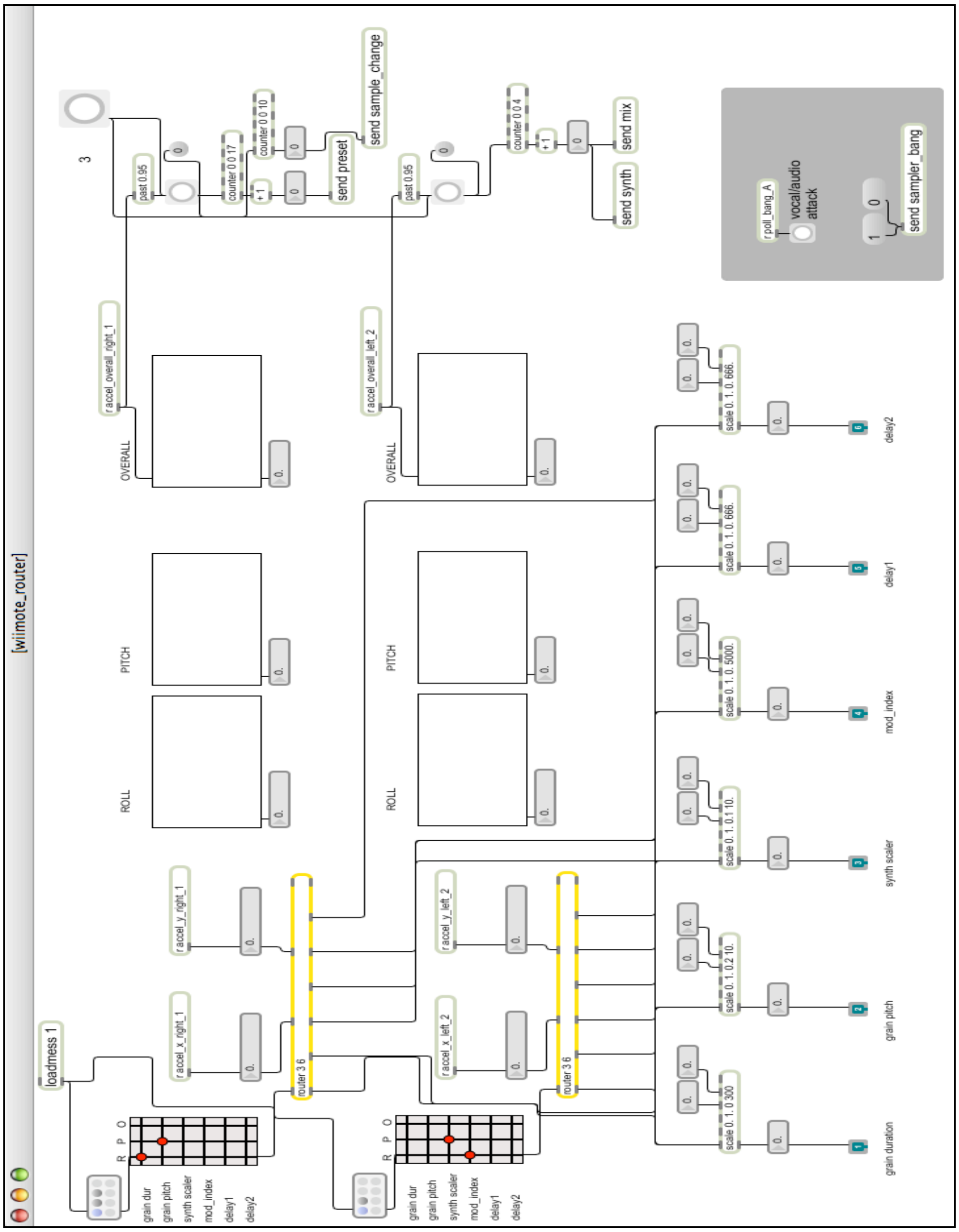

Figure 4.7. A screenshot of the Max routing/mapping layer of System 1.

Sound Generation

As already summarized, the sound generation engine of System 1 (See Figure 4.8) comprised a granular sampling engine augmented with live processing in the form of audio signal delays, ring modulation/AM synthesis, overdrive, and reverb. In addition, an analysis/resynthesis engine accompanied this amalgam of DSP, analyzing the resulting granular texture in real-time. The resynthesis module utilized an FM enhanced additive synthesizer before being routed in parallel to the effects processors.

Each of the processing components had their own fader for mixing before reaching the master output level of the patch. Therefore, each component effectively served as a pre-master sub-mix element in the composite texture. Both the granular engine (pre-analysis) and the resynthesis engine (post-analysis) were fed to the processors at all times.

The granular engine, which produced ten grains, was based on techniques learned from Butch Rovan. In addition to the continuously controlled parameters of pitch ratio and grain duration (in milliseconds) for all grains, the playback speed was set to the original timing of the sound file at all times. Thus, time compression and/or expansion was never a factor in the sonic result. In hindsight, I wish I had explored that dimension further. The grain levels (for all grains) were also set to an amplitude ratio of 0.8 in reference to the original signal level. The engine could receive one of eleven different sound files at a time. The files were cycled through with a ***counter*** object triggered by overall acceleration from one of the Wiis (as already mentioned). Obviously, the nature of the sound file content had an immense impact on the overall result. Thus, my approach to sound design for each of the files will also be discussed in detail later on in this chapter.

Ultimately, each of the ten grains generated are sent to the processing engine of the system, which will be discussed further. In addition, a composite dry stereo signal was sent to the master mix as well as to the analysis/resynthesis engine.

The analysis/resynthesis engine was based in part on ***fiddle~*** by Miller Puckette. When the composite granular texture reached a pre-defined amplitude level a real-time analysis was triggered producing 8 sinusoidal partial estimates with both frequency and amplitude information. This frequency and amplitude information then fed a resynthesis module that produced unique FM synthesizer signals for each partial. The dry, unprocessed, FM signals were not mixed to the master in contrast the granular engine texture. Harmonicity ratios for each of the FM units stayed constant at 1.0, while the overall synthesizer frequency scaler (a pitch shift of sorts utilizing a ratio relationship for each of the incoming partials) and modulation index (for all of the FM units) were physically controlled in real-time. An overall mix of the 8 FM synthesizer units was determined by 5 different preset fader configurations that were cycled through with a ***counter*** object triggered by overall acceleration. Each of the 8 unique partials derived from the FM synthesizer units then fed the processing engine before being mixed to the master output.

The processing module actually comprised two layers. As mentioned, both the granular engine's 10 unique grains (plus composite stereo feed) and the re-synthesis engine's 8 FM synthesizer units fed the first layer of processing. A second layer of processing received specific elements of the first layer.

Each of the live processing units within the module had a number of controllable parameters. Every one of the controllable parameters' UI objects were uniquely named so the ***pattr*** object could control major state shifts. A total of 18 ***pattr***-based presets could be cycled through to create global state changes as each controllable parameter had the capacity to be altered. None of these were continuously controlled with gesture input, although experiments with movement-influenced delay times were conducted early on in system development. The preset controlled parameters (in parentheses) for each processing unit were as follows: 2 delay units (delay time, modulation depth, modulation rate, feedback); 2 ring modulation units (frequency); overdrive (overdrive select – single or double, overdrive level 1, overdrive level 2, overdrive low pass filter); amplitude modulation (modulation rate, modulation depth, modulation amplitude); reverb (size, decay time, high frequency damping, diffusion). The second layer of processing consisted of two more delay and reverb units, also controlled by the above parameters, fed by the first layer ring modulators, amplitude modulator and overdrive units. In total, this layer produced 12 discrete channels of processing with separate mix levels feeding the master mix. All twelve processing units' mix levels were triggered by a ***counter***, which cycled through 5 states based on overall acceleration levels. The master mix layer received 14 channels total when counting the dry stereo composite signal of the granular engine.

Subsequently, the master output layer of the system had the ability to mix to stereo or multi-channel configurations depending on the performance situation. Before outputting the signal to the speakers, a master stereo mix was fed through the ***omx.peaklim~*** object, a peak-limiter used mainly for sound system protection. Then, depending on the situation, that stereo feed was diffused to the surround speakers as needed. Both left and right channels were fed as composites to the center and sub, while left and right front signals were crisscrossed to the right and left rear surround speakers respectively.

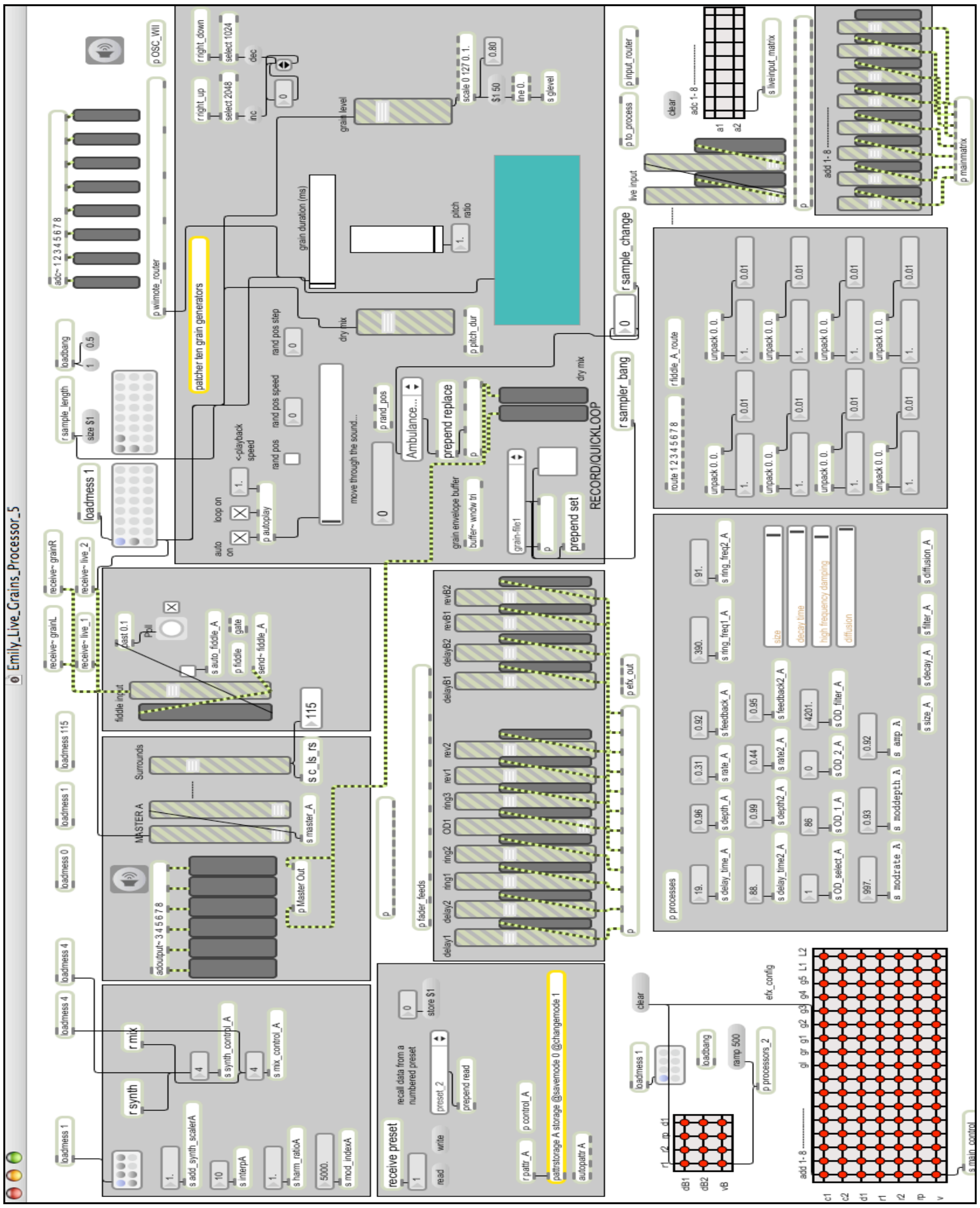

Figure 4.8. A screenshot depicting the main interface level of the sound generation engine of System 1.

Sound Design

Eleven audio files were designed specifically to feed the granular system prior to processing and analysis/resynthesis. Conceptually, I was interested in this interactive music system representing an "opposing force", in a sense, suggesting the human/environment dynamic we currently face as a species with regard to climate change. I wanted a meta-level of abstract representation characterized by the sound files even if their more explicit symbolism would be mostly veiled in processing.

The sounds all had interesting structural differences, which helped to create different effects through the system. For instance, some had interesting percussive envelopes that were also somewhat rhythmically irregular at times. Others tended to be more fluid and soothing. Specifically, those with irregular rhythmic structures and augmented low ends were especially visceral.

Details of Interactive Music System 2

This system served as a contrast and aesthetic compliment to the more chaotic and aggressive qualities of the first system detailed above. It offered a more stable counterpart to the seemingly indeterminate response character of System 1, appropriately serving the conceptual needs of the section it served. Under the hood, its sound generation engine was made up of a polyphonic FM/AM synthesizer with ADSR envelope control and some reverb.

Max Data Routing and Mapping (System 2)

As with its function in System 1, this layer of the system allowed for efficient routing and creative mapping of the movement data in order to influence the core sound generation module (See Figure 4.9). Much of the basic routing and data monitoring mechanisms were identical with System 1, however the specific parameters controlling sound generation were of course different for this system, so they were dealt with accordingly.

Specifically, one Wii's roll data was used for MIDI note generation while the pitch data was utilized for both the modulation index and harmonicity values of the FM aspect of the synthesizer. The second Wii's roll and pitch data were used for the modulation frequency and tremolo depth parameters of the AM aspect of the polyphonic synthesizer. MIDI notes were scaled to a range between 20-110, modulation index was scaled to a range between 0. and 24., harmonicity between .1 and 10., modulation frequency between 0. and 20., and tremolo depth between 0. and 2.0.

In addition to these continuous controls I once again mapped overall acceleration to create a switch of sorts that responded to fast and abrupt attack-like movements. This time both Wiis could trigger global preset shifts affecting wet/dry mix levels and subtle high frequency damping for the reverb component in addition to some random MIDI note generation between 30 and 90. ***Counter*** objects stepped through state changes in a linear fashion making the system a little more predictable than System 1.

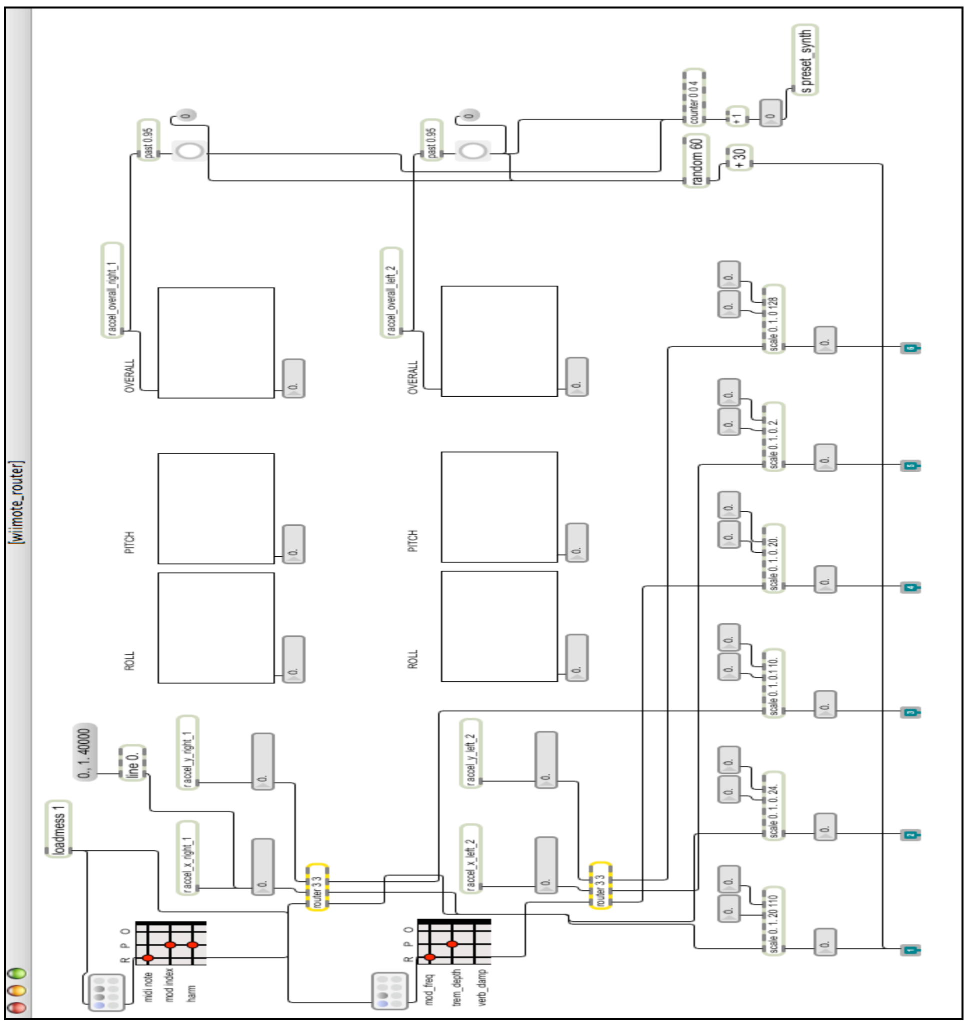

Figure 4.9. A screenshot of the Max routing/mapping layer of System 2.

Sound Generation

The sound generation module of System 2 was much simpler than System 1. As mentioned, its core feature was a polyphonic FM/AM synthesizer with ADSR envelope control. In addition, some basic reverb processing was employed to augment textural variation (See Figure 4.10). Although relatively simple in its design, it did provide a dynamic sonic environment for the performer to explore with an interesting mix of textures available. Its slightly more predictable nature also allowed it to be played more like a traditional instrument, creating an entirely different dynamic between the performer and the sound world (when contrasted with System 1).

The core sound generation module was created using ***poly~***. Housed within the ***poly~*** framework was an FM/AM synthesizer sub-patch and ***adsr~*** based envelope control module. MIDI note and velocity information was converted to frequency and amplitude data before feeding the FM/AM synthesizer module and ADSR envelope object respectively. Attack, decay, sustain and release values were programmed to be constant after many experiments with different configurations. Accordingly, the parameters were preset to: attack (1140ms), decay (190ms), sustain (.21 – as a factor of amplitude - or 21%), and release (960ms). A very gradual envelope texture resulted from this configuration. In addition, the ADSR legato function was enabled for the object, so that while the envelope was active it would not drop to zero after being retriggered. Instead, the envelope would ramp to the new amplitude over the attack period.

Other preset determined aspects of the patch could be varied in response to overall acceleration, or fast and forceful “attack” movements (as mapped). They consisted of 5 preset modes that influenced the reverb processing components in relation to the mix. These relationships included the amount of dry signal sent to the reverb processing sub patch in addition to overall dry signal mix proportions in relation to the “wet” processed signal as mixed to the master output faders. In addition, some subtle shifts in high frequency damping for the reverb unit could be influenced.

As already mentioned, the continuous movement-influenced aspects of the system were solely focused on control of the FM/AM synthesizer module itself. Movement activated note generation and FM/AM parameter control allowed physical gesture to be tightly correlated with sound generation. The “attack” movement controlled mix state changes enhanced this effect by viscerally connecting overall force and acceleration with global sound world texture variations.

As with System 1, the master output section of the system had the ability to mix to stereo or multi-channel configurations depending on the performance situation. Therefore, if necessary, as in the first version of the complete performance work, a stereo feed could be diffused to the surround speakers. Both left and right channels were fed as composites to the center and sub, while left and right front signals were crisscrossed to the right and left rear surround speakers respectively.

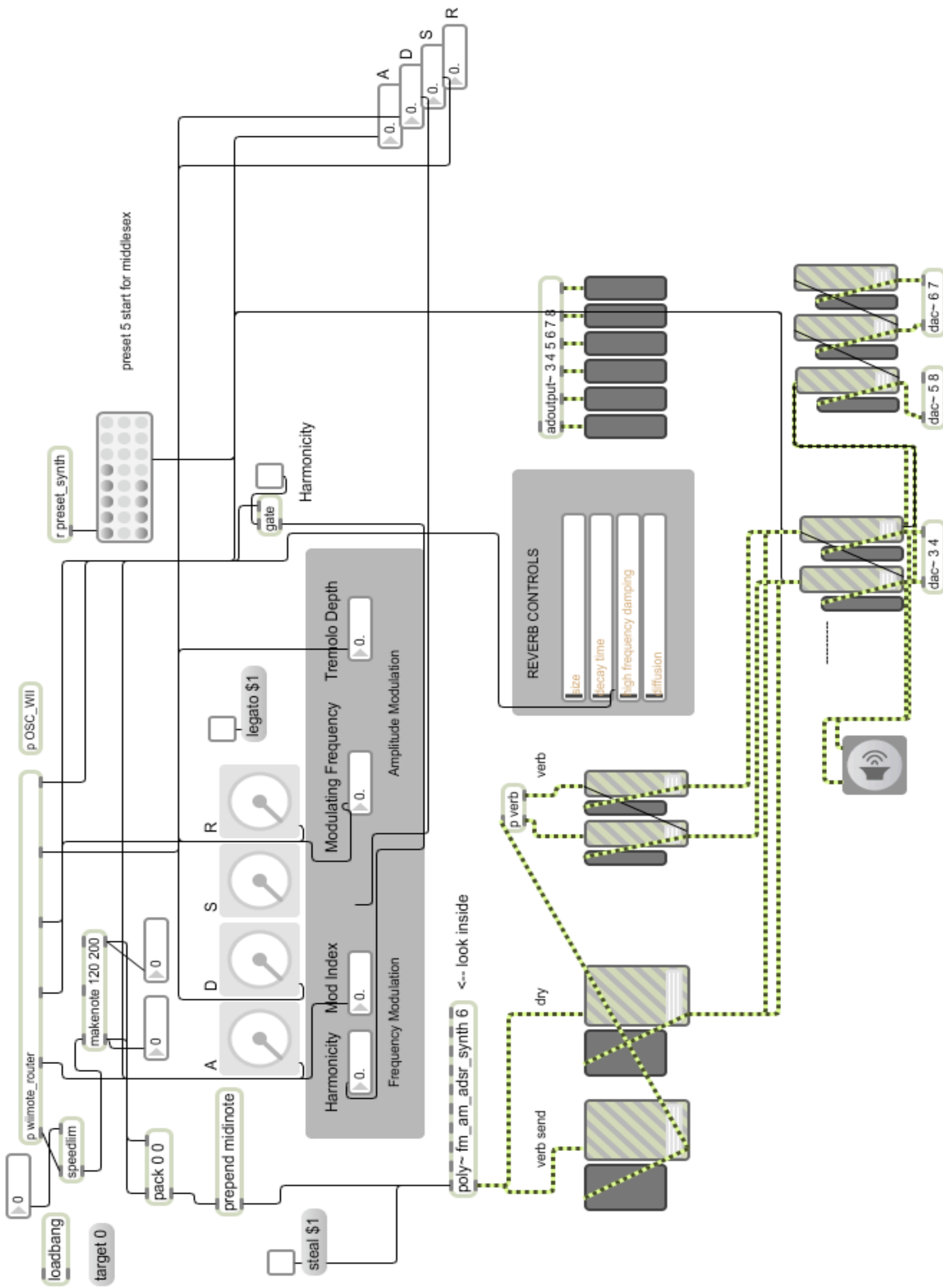

Figure 4.10. A screenshot of the main interface screen for System 2.

Details of the Graphics System

The graphics system, largely contributing to the overall atmosphere and mood of the performance world, was designed to be directly reactive to the sonic elements of the piece. Although its specific role was slightly altered for the two different versions of the piece (to be discussed in Chapter 5), the core elements of the systems stayed intact. The influence of sound over the motion graphics was in the form of amplitude related interactions. Changes in loudness seemed to be especially compatible with a wide range of visual attributes.

The graphics system, built with Max/MSP/Jitter, initially comprised three main visual generation modules for the first version of the piece with another being added for the second version of the piece. The system utilized, as much as possible, the Open GL architecture of Jitter to access the graphics card's robust processing power.

The development of a clear and manageable control/feedback panel was important to the system's interface design. During aesthetic development phases of the project, having quick access to important parameters and the ability to create dynamic presets involving combinations of these parameters was key. Subsequently, during rehearsal and performance phases of the project it was helpful to have real time feedback and efficient control over graphics scenes.

Open GL Model Morpher

The most prominent graphics generation module in the system was built in order to create reactive relationships between a live input audio stream's amplitude level and custom 3D models of abstract shapes. Although a relatively simple interaction design, the result was often very dynamic. A ***jit.gl.model*** object was used to load a 3D model (in the .obj file format). The matrix outlet of the Jitter object was enabled in order to output the data as a matrix, so that the .obj file's vertices could be parsed into separate planes that could then be manipulated with matrix math operations.

This matrix data was sent to a sub-patch called ***fft_warp***. Inside the ***fft_warp*** sub-patch the .obj file's matrix data was unpacked (with ***jit.unpack)*** in order to manipulate the data representing the x, y and z vertices of the 3D model. The rest of the model's data was sent back through to a ***jit.pack*** object unaltered. The unpacked matrix planes representing the x, y, and z data were sent to ***jit.op*** objects, which performed matrix math operations. One inlet of each ***jit.op*** object received the vertex data in matrix form while the other inlet received matrix information representing amplitude data from the live input audio stream. The live input audio stream underwent Fast Fourier Transform (FFT) analysis before feeding a ***jit.poke*** object, which converted the audio signal data into Jitter matrix data that was used to affect the 3D model. The ***jit.op*** object was programmed to multiply the vertex data with the amplitude data before being sent to another ***jit.op*** object that could scale everything. This was done in case I wanted control over the overall size of the 3D model and vertex displacement transformations that were occurring.

Each vertex was also sent through a ***jit.slide*** object that allowed for additional amplitude control (after scaling) over the temporal aspects of the amplitude based vertex transformations. For instance, increased amplitudes would create smoother interpolations in vertex displacement. The transformed vertex planes were then joined back together with the other 3D model matrix components through a ***jit.pack*** object that subsequently passed the data through an outlet back to the main level of the module.

Back in the main level of the module, the transformed model continued on its path in real-time to a ***jit.gl.mesh*** object that could reconstruct the matrix data into Open GL space again. Additional transformations took place here including amplitude influenced line width or point sizing (depending on how the GL model was instructed to draw) and amplitude influenced rotational speed. The ***jit.gl.mesh*** object was also programmed to capture its contents to a texture. This result, sent through a ***jit.gl.texture*** object, was passed on to a dual chain of ***jit.gl.slab*** objects that could utilize the Jitter Open GL shader architecture for further processing on the graphics card including texture displacement and temporal blending. Subsequently, the processed texture was sent to the mixer level of the Graphics System before being output.

Audio Matrix Texture Generator

Another graphics generation module utilized controlled randomness to fill a 2 dimensional 100 x 1 matrix with 4 separate sine wave generators corresponding to each of the 4 planes of the matrix. Thus, each randomized sine wave frequency filled the alpha, red, green and blue planes respectively, with new numbers generated every 2 seconds. The audio signals fed 4 separate ***jit.poke~*** objects, that were controlled by a saw tooth wave generating ***phasor~*** object, which was influenced by audio amplitude. The greater the level of the incoming audio stream the faster the matrix cycled through its 100 horizontal pixels. The audio signal generated matrix was then passed through a ***jit.slide*** object for temporal smoothing and another jit.matrix to fill the 2 dimensions for output (also dependent on the aspect ratio and size necessary for the piece).

Then, like the Open GL Morpher module, the properly sized matrix was passed on to a dual chain of ***jit.gl.slab*** objects that could utilize the Jitter shader architecture for further processing on the graphics card including texture displacement and temporal blending. Subsequently, the processed texture was sent to the mixer level of the Graphics System before being output.

Live Camera Feed

A third visual module was created very simply to capture a live digital video camera feed for mixing with the other visual elements. For aesthetic effect, I decided to down-sample and subsequently up-sample the captured DV stream to create an abstraction of the dancer's movement. Specifically, a 320 X 240 image was captured at around 30 fps and down-sampled to a 16 X 16 matrix before being up-sampled to a larger resolution with interpolation. The resultant texture was then sent to the mixer level of the system. No extra slab processing was used for this element.

Video Sampler

A fourth visual module was also implemented for version 2 of the piece (in November of 2009) that could play back a pre-composed Quicktime file and send the texture to the mixer level of the system.

Mixer

All modules, regardless of their origination on the CPU in matrix form (as in the Live Feed and Sampler modules) or in the Open GL world of the graphics card (as in the Open GL Morpher and Audio Matrix Texture) ended up at the mixer level of the system. This simple, yet dynamic, element was the most important piece influencing the composite imagery sent out to the performance world. As will be discussed, the mix combinations were ultimately the largest influence on aesthetic effect rather than the individual modules themselves, although the Open GL Morpher module was fairly prominent in its display in the first section of the piece.

This module used a ***jit.gl.slab*** object with 4 inputs that could process and mix the 4 separate graphics generation modules (3 for the first piece) based on shader code written specifically for handling the task. The nature of the processing and mixing was not simply based on opacity levels (alpha blending), but a fairly dynamic additive mix (taking a, r, g, and b data into account) articulated by the following equation: output = input 1 + input 2 + input 3 + input 4. Thus, both negative and positive values could influence the composite blend.

Control and Feedback

Last, but not least, a friend of mine, Ed Guild, helped design a nice user interface for the system. Because his contribution was initially focused on the Open GL Morpher as a standalone patch, I had to expand it (albeit, less artistically than he) with specific control and feedback elements necessary for further aesthetic development (See Figure 4.11). As

I added modules and augmented certain features, the system required greater control and feedback mechanisms in order to work in a complex rehearsal/performance situation.

An Audio Source panel allowed me to control input level while monitoring amplitude in real time. A 3D Object panel allowed me to select different abstract 3D models for input to the Open GL Morpher. This section also allowed me to select between two different textures for the 3D model, including a fractal generated color scheme and a texture generated from the Audio Matrix module. In addition, the 3D panel also allowed me to select between two different drawing modes for ***jit.gl.mesh*** (points and lines), while also offering a slider for overall size scaling. All of these parameters were under preset control. In addition, a Background Color panel allowed me to influence the color of the GL space (also preset controlled), while the Object Morph section gave me feedback on temporal smoothing, line and point size and rotational smoothing (all set to automatic control in order to be influenced by audio amplitude changes).

On the right hand side of the patch, a Render Control panel handled the rest. Specifically, this is where I programmed and created preset scenes influencing the aforementioned parameters in addition to texture displacement shader selection (for the Open GL Morpher and Audio Matrix modules) and mix levels (for all elements). Additionally, I was able to experiment with different texture displacement shaders (with a smart menu) and mix levels before creating presets. A master ***metro*** served as the main control for render context drawing, while an ***ezdac~*** on/off and master "kill switch" (bang = cut to black) gave me efficient control over global functions. Real-time frames per second feedback provided instantaneous updating on processing performance.

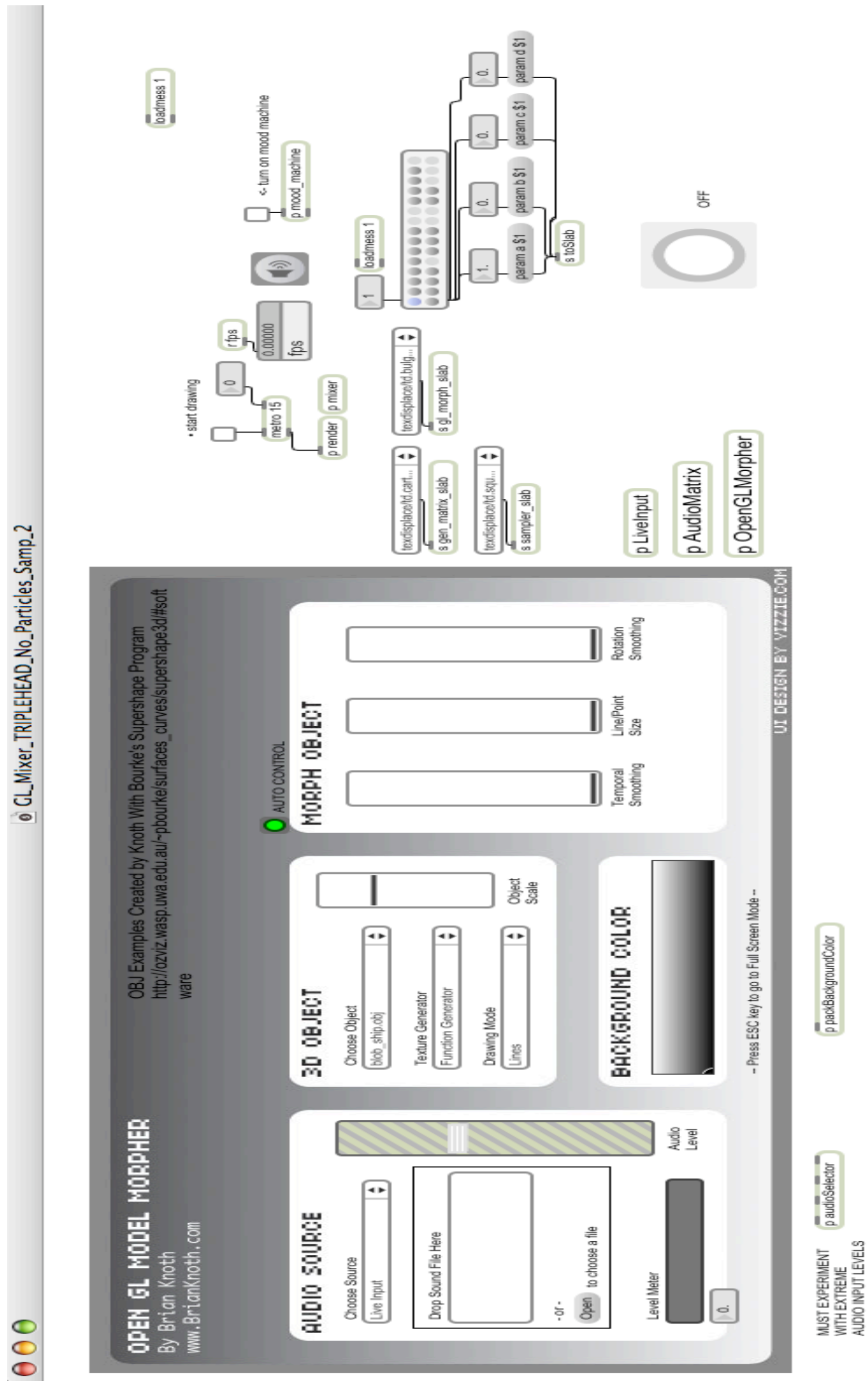

Figure 4.11. A screenshot of the *Control and Feedback* level of the *Graphics System.*

CHAPTER 5: PERFORMING *UNLESS* (2009)

> I felt alive. It was an exercise for my brain. It allowed me to escape and give thought to what I was seeing, hearing and feeling.
>
> – anonymous audience member (May 1st 2009)

Both versions of *Unless*, performed in May and November of 2009 respectively, consisted of three sections abstractly exploring contrasting themes related to interconnectivity (both human/human and human/environment), humans' seeming inability to fully embrace this nature, and by extension the conflicts we now face as a species. Throughout the process of defining the work's structure, all factors, including: materials (technology), content (themes/concepts), environment (the performance situation) and human input (myself – the artistic director, and the performer) interacted as tightly interwoven forces influencing form. Throughout this chapter I offer an account of the artistic results that emerged from this process. Specific differences between versions will be highlighted in relation to the influence of each unique performance situation. In addition, through personal reflection, performer feedback and audience research, I infuse a critical analysis of the project as a whole.

Emergent Structure

The dynamic process of developing the piece led to a resulting structure that could be categorized as emergent. The idea of emergent structure in relation to artistic form can be conceptually linked to the behavioral dynamics theory of William Warren (2006), mentioned in Chapter 1, which is a compelling attempt to apply dynamic systems theory to human behavior (in general). Since artistic processes could be considered one facet of human behavior, the conceptual leap is not all that radical. Warren considers human behavior an emergent phenomenon in that it cannot be attributed to a central control mechanism of any kind (i.e. the brain), but is the result of many levels of interaction between a human and their environment(s). One can think of the artistic process in much the same way.

The process of developing the structure of the performance was genuinely emergent. The human factors, Emily and myself, were in constant interaction with the technology and environment(s) throughout the process of creating the piece. I never actually envisioned my role as that of a central control agent strictly enforcing the way in which the piece should evolve (even though I was indeed the project's artistic director). The process was dynamic throughout, with my input to the "system" merely representing one factor influencing the result.

Emily and I initially allowed ourselves time to improvise with and explore the systems comprising the interactive multimedia environment of the piece. Much of these initial experiments carried out between December of 2008 and March of 2009 were geared towards finding moments of connectivity and disassociation between Emily's movement and the sonic environment. Through this process we quickly determined that each of the interactive music systems offered a contrasting relationship to Emily's movement.

> The relationship between my movement and the interactive media environment was complicated and connected. This seems to be some sort of definition of "interactive". (Emily Beattie reflecting on the environment as a whole)

Whereas the first system developed (the granular system with multi-layered DSP) created a more chaotic sonic environment that was harder to "control", the second system (the polyphonic fm/am synthesizer with reverb) offered a simpler interaction that exhibited more quickly discernible links between specific movements and sound. It became clear that they should be utilized in two distinct sections of the piece.

Conceptually, I related the more "connected" system with the human capacity for empathy, and the more disassociated/chaotic system with our failure to fully embrace this nature. Such perceptual and conceptual contrasts resulting from the performer-system dynamic would be considered structural material. These differences in system/performer interaction were consistent with my specific need for a conceptual structure. Although I was not concerned with the work, as a whole, conveying specific thematic elements to an audience, as a director I needed to think about form in terms of both raw perceptual contrasts and conceptually related themes.

My goal for the audience was to offer a dynamic and lively performance experience that could carry them through varying states of emotion. Although my process for developing structure was partly cerebral, I was not as interested in the audience experiencing the work in that way. I wanted to offer something more visceral. Just as the process of determining the work's structure was emergent, I felt that the audience's perception of the work and its meaning should also be emergent. I was not interested in forcing themes and concepts on the audience such as a traditional narrative approach might. I was more interested in offering a novel experience that would encourage active processes of perception, cognition and interpretation. In a way, one could relate my goal for this piece to the concept behind Rorschach testing. An audience's interpretation of the work is actually more reflective of their own personal histories and internal thought processes.

Presenting *Unless* v.1 (May 1st 2009)

> I felt that the work overall was very consonant, beautiful and aesthetically pleasing. The movement and gestures driving some of the musical and visual transitions helped me view the piece as a cohesive whole. It wasn't dance with projections and music. It was all one thing.
>
> – anonymous audience member (May 1st 2009)

In the following paragraphs I will now describe the first version of the piece as it was performed in conjunction with Boston Cyberarts on May 1st 2009 at Grant Recital Hall (Brown University) in Providence, Rhode Island.

As mentioned, the piece comprises three sections. Each of the sections seamlessly transitions from one to another. This iteration of the piece ran about 17 minutes and 40 seconds total, with section 1 at 7 minutes, section 2 at 5 minutes, and section 3 at about 5 minutes and 40 seconds. My intuition was to temporally balance the sections as much as possible in order to sustain interest over the entirety of the piece, although Emily and I never discussed timing in absolute terms due to the improvisational nature of the piece and the emergent nature of our process. Since each section offered something distinct in its overall effect, I felt confident that the audience would stay engaged throughout.

Section 1

The opening section of the piece emerged from Emily's movement experiments with the more "connected" interactive music system. This section served to abstractly represent our capacity (as humans), although sometimes fleeting, for interconnectivity and empathy. Emily articulated this connection in post performance reflection:

> In full improvisational mode a dancer must offer equal energy dancing as well as listening. This type of listening occurs in the full body, not just the ears. As I traveled through the first section, I was easily able to hear my impulses in sight and sound. (Emily Beattie on the May 1st performance)

These ideas were further augmented with the graphics system providing an immersive representation of such a connective "force" either within us or surrounding us. This force provided both an abstract representation of "environment" and an "other" being that she attempted to connect with.

The projection, at the start, was displayed far above Emily's head, so that it was out of reach and looming over her. Once she began moving it grew larger, as her movement would excite the sound system, which in turn would influence the scope of the projection. The movement-activated bursts of sound would cause the projection to radiate to the point of immersing her.

> Some points of synchronized movement-sound-projection interaction felt like communication with this other being. (Emily Beattie on the May 1st performance)

The forms used within the graphics engine were custom generated in Paul Bourke's *Supershape* software (Bourke is a scientific visualization expert and research professor in the Western Australian Supercomputer Program). Bourke's 3D model generator is based on Johan Gielis' superformula equation for modeling natural forms. The superformula is a geometric transformation equation that encompasses a wide range of forms found in nature (Gielis 2003). For instance, transformations applied to a circle produce archetypal shapes such as starfish, shells and flowers. The archetypal quality of the imagery was important for creating the mood of the section.

The interactive sound design for this section was meant to be both fluid and frenetic. I was pleased with how the sound of her movement in this "world" conveyed an almost "electro-magnetic" connection with her surroundings including the abstract "force" represented in the graphics. The held interfaces, I believe, sold the experience of interactivity as they provided a distinct locus of control, while the holistic impression of her connection to the environment remained a full-body movement experience. As Emily states:

> The most satisfying movements that altered sound were twisting and spiraling from the spine to the arm that held the remote…Although the remotes were located at the extremities of my body, I wanted to be sure that my most internal impetus activated them. Whether I was standing, leaning, curled up, or arched, the remotes were pulled along with my entire body in action. (Emily Beattie on the May 1st performance)

Emily's ability to navigate her connection to the media environment was compelling. Although she sometimes struggled to maintain her connection, as feelings of interconnectivity for us all are indeed fleeting, this constant yearning for a bond forced a more "in the moment" movement improvisation style. Her movement was less dependent on an automatic gesture vocabulary and more on a fluid reaction to the situational flux. Emily articulates:

> There was a pressure in the environment and a pressure in my body that were for a brief time synchronized. I think these moments were pinpoints of visual recognition of my role in the "performance world" for the audience as well. This control was short lived, as it seemed that the system would slip out of my control as quickly as I connected to it. (Emily Beattie on the May 1st performance)

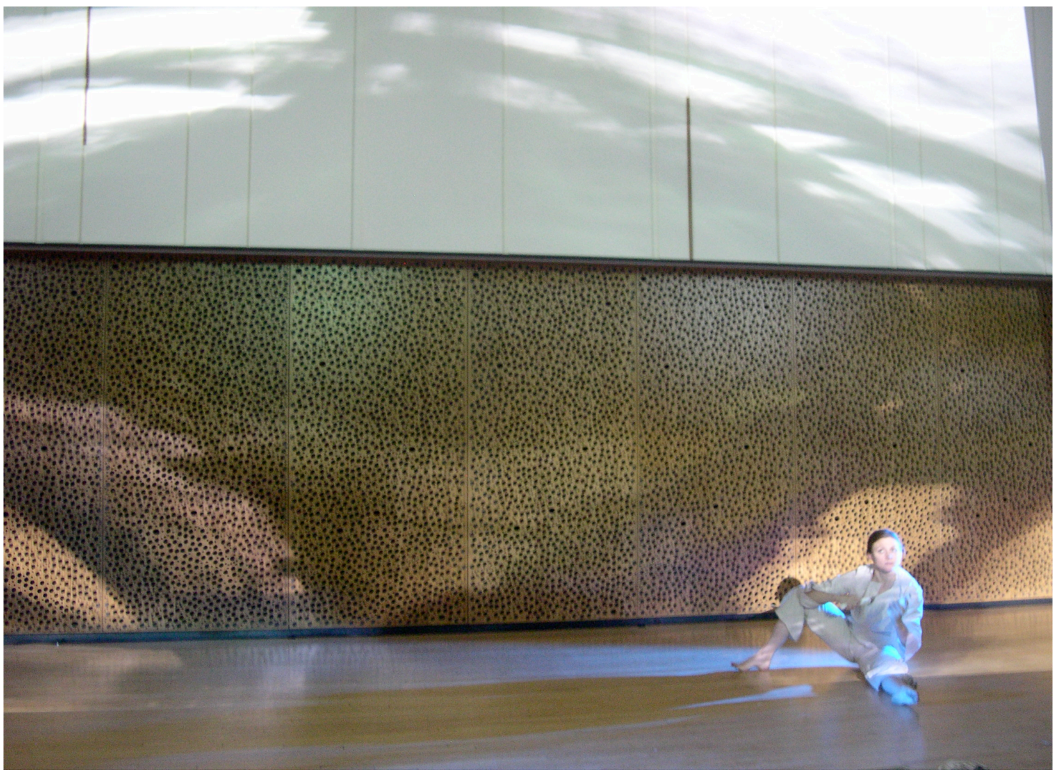

Figure 5.1. Emily Beattie performing during the first section of the piece.

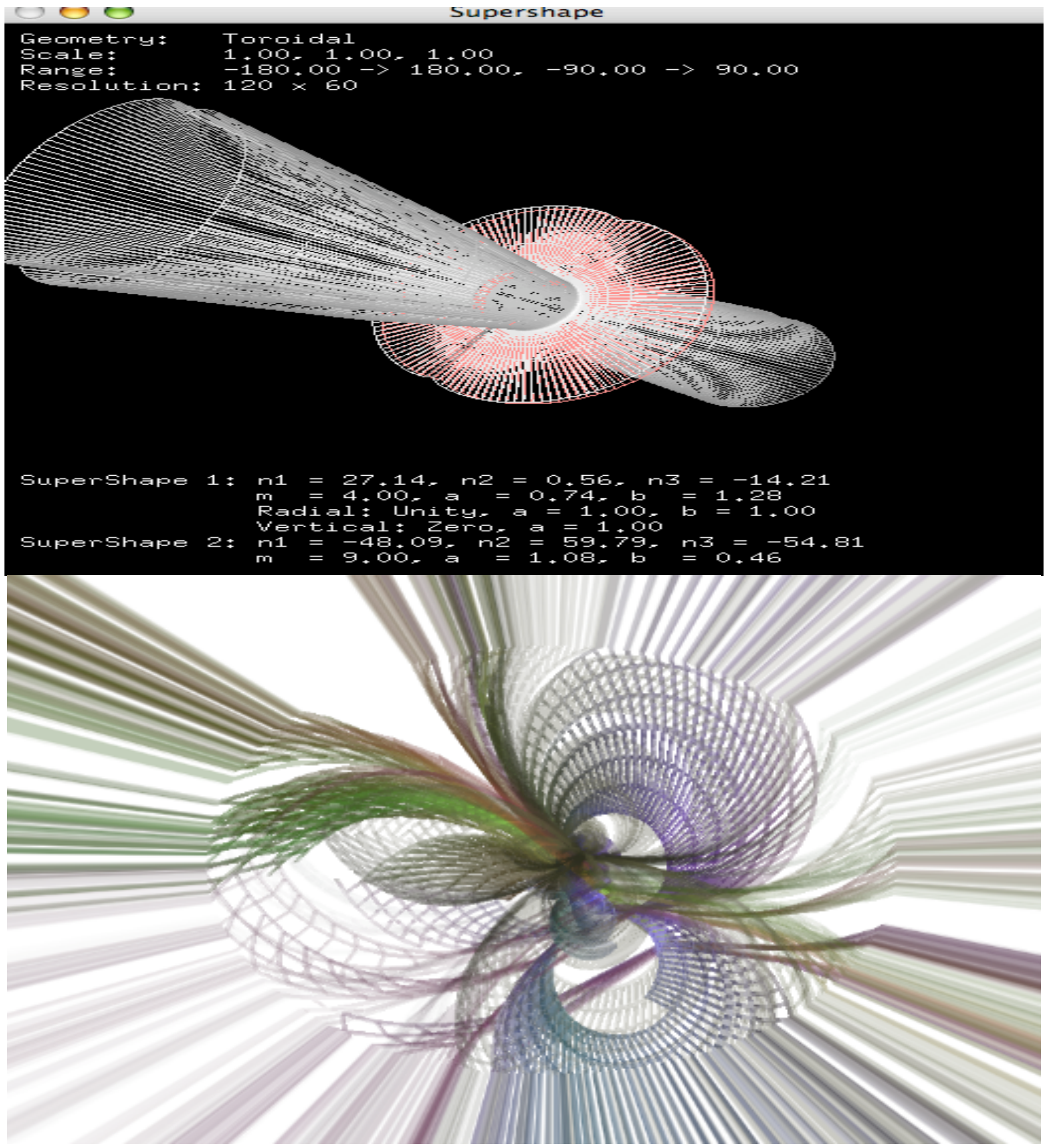

Figure 5.2. A representative 3D object used in the piece before texturing/processing (top) and after (bottom).

Section 2

The second section of the piece featured a pre-composed music piece, thus offering a respite, for both Emily and the audience, from the fully interactive media environment featured in section 1. In addition, this section allowed Emily to turn her focus fully on to the audience (without needing to mind the physical interface) and create an empathetic connection with them in a more direct way. The contrast to interactive sound was profound, as she was no longer attached to the remotes. Formally, it served as a theatrical bridge and palate cleanser.

Conceptually, I wanted the section to reflect a sense of hope. For instance, maybe the notion of realizing a more enlightened connection among humans, and between humans and their environment, is still possible. Thus, the music, now serving as an underscore element, rather than a diegetic source related to specific actions in the performance "world", created a more dream-like atmosphere in concert with the projected light patterns (now much more ambient in nature). The sound design was most fluid and wave-like in this section with very gradual crescendos and swells of spectral brightness further augmenting the meditational nature of the scene.

Some interactivity between Emily's movement and a live input camera to the graphics system was also implemented in a subtle way, creating ghostly blobs of abstracted movement in the projected atmosphere. In addition, I performed manual control of the graphics system's brightness intensity, which allowed me to follow the music and Emily's movement with measured accents. The result was akin to creating a portal through the use of increasing light as she moved closer to the audience and began to enter their space.

Whereas before I was connected to the reactive projection and sound system (with my eyes and ears), now I could explore the details of who was in the overall space with me, by moving all over the space, creating light. (Emily Beattie on the May 1st performance)

Figure 5.3. Emily performing during section 2.

Section 3

The final section of the piece could be considered in two parts. The first part seamlessly segued from the final decrescendo of the set music piece into the static, jittery and quiet initialized state of the “chaotic” interactive music system. The tense and ominous sound design of this moment underscored the transition to an audience participation segment. As Emily explains:

> I selected and spaced the audience members in an effort to engage them in the work. As the presentation mode shifted from proscenium stage to theater in the round, our hope was that the audience would perceive a connection to the work that was unfolding. (Emily Beattie on the May 1st performance)

After connecting to the audience with direct and focused body language, Emily’s relationship to them undergoes a shift from connected and empathic to controlling and “other”. Conceptually, I argue that this moment calls attention to the duality of inter subjective experience. Our ability to understand the experience of others in an empathic way also allows for manipulation. As Emily states, “My position as omnipotent selector switched the power dynamic.”

Thus, this section represents the potential for a more ominous future, *unless,* we decide to act with a collective wisdom at this juncture of our evolution, and fully embrace our capacity to connect with and understand “others” and our environment. Hence, the title of the work reflects the precarious situation we are now faced with as a species.

After Emily selects and sculpts the audience members on stage, the next transition unfolds with her returning to the physical interface and subsequently exploding into a more oppositional struggle with the “environment”. Her increased bursts of energy cause the sound to respond with equally opposed force. The projected world of the graphics engine is absent at this point with abstract and minimal light design creating looming shadows of the human statues (selected audience members) and Emily's frenetic movements.

> The sounds in this section caused me to collapse into the points of my body that would fold quickly, therefore, unfolding, subsequently, in explosive movements, was also supported by the sound palette. I felt overwhelmed by this section. Finally, I surrendered beneath its intensity. (Emily Beattie on the May 1st performance)

In the end, the projected atmosphere returns and Emily retreats back to the place where she started. The environment then fades out.

Figure 5.4. Emily and selected audience members in section 3.

Audience Research

The first version of *Unless* was performed for over 60 people. The audience was surveyed right after the performance with a questionnaire asking them to describe their impressions (See Appendix B). Specific questions were designed to elicit responses related to their embodied multi-sensory experience of the piece, such as:

- How do you think the sound and imagery may have impacted your perception of movement?
- What aspects of the performance did you connect to?
- In your experience, what was the relationship, if any, of the sound, mover, and visuals?
- How do you think your perception of the movement would have changed if there were no music and/or visual component?

Over 50 anonymous audience members responded. Detailed accounts of these experiences, in the form of excerpted quotes, are organized in relation to specific research related concepts.

Kinesthetic Empathy

As already mentioned, people observing any physical action have the capacity to internally simulate that movement. This fact supports the concept of kinesthetic empathy, which is most likely mediated by the mirror neuron system. I speculated that a movement-activated multimedia performance environment could enhance feelings of kinesthetic empathy in the audience. Although by no means irrefutable proof, the following representative quotes offer some insight into the audience's experience of kinesthetic empathy regarding the first version of the piece:

- "The performance was exciting. It made me feel sharp and awake…I felt movement in the room and I felt energized."
- "I felt small in a big place, shaking with Emily as I witnessed her movement. Wanted to move myself."
- "Felt connected to the dancer, consistently. It's the opposite of a "typical" dancer responding to music, the dancer seemed to be creating the sound."
- "I was aware of myself imitating and sympathizing with the movements…"
- "I liked the tension between the sound and movement – I felt like the sounds gave a special tense quality to Emily's movements, as if I could feel the tightness in her muscles, just watching."

Immersion and Interaction

The audience research questionnaire was also able to tap into some of the audience's thoughts on the multimedia environment and their experience of multi-sensory interaction and environmental immersion. For instance, the music/sound and movement relationships seemed to be the most palpable, as illustrated by some of the following quotes:

- "The music added a dynamic edge to her movement."
- "The movements would not have alerted me without the sound."
- "If the sounds were indeed generated live with the white remotes, than it is a new way of amplifying movement…like a human synthesizer."
- "The movement was dramatic and was intensified/amplified/focused by sound."
- "The dance/sound combination was most beautiful and the best (most synergetic) combination."

Some audience members struggled to find connections between the visual aspects of the environment and the mover, thinking that there should have been a more direct relationship. Interestingly, some did, however, implicitly perceive the built-in layer of disconnect between the movement and visuals, which was sonically mediated by design.

- "I had trouble picking up on connections between visuals and music, but noticed movement feeding off of visuals, and movement influencing music."
- "I connected the stronger sounds with the visual light rather than the performer. The sound and mover made more of a connection to me than the visual and mover."

Regarding the immersive qualities of the work, some audience members had strong reactions to the interactive graphics' contribution to this aspect of the experience. It seems that the imagery was effective in creating an atmospheric and sensually engaging performance world.

- "I was in awe. I liked the use of having the visuals travel over the room. It engulfed me into the experience with the dancer."
- "I liked the immersive environment with the projections spilling onto 3 walls, lighting from the sides and noticeable clear relationships between wii-motes and sound/image."
- "The opening wall scaling sequence helped expand the space and immersive feel."
- "The visuals heightened my sensitivity. The stuff in my peripheral vision, especially, kept me on edge."

Audience Participation

One aspect of the piece that created anxiety for Emily and me was the audience participation segment. We were not sure how the audience would react to being brought up on stage and sculpted (at least those we did not invite to participate beforehand!). In the end, this segment served as an integral structural element, which seemed to reinvigorate the room for the final moments of the piece. For some audience members it may have actually been one of the most engaging moments of the performance as evidenced by these remarks:

- "I connected especially with the audience section."
- "Really got into the piece at the end when people were brought up on stage…something in my spine opened up."
- "I really liked the audience participation. It helped me feel more connected to the performance."
- "My blood pressure went up when I watched her pull people from the audience."

Interpretation

Lastly, the audience's interpretation of meaning in the piece was an intriguing question for me. I was interested to see how such an abstract treatment of thematic elements would be perceived and how a sense of meaning might emerge in relation. Some described their interpretations rather succinctly, stating that the piece represented a, "Connection to others", or reflected a "Transcendental", "Other-wordly", or "Galactic" tone. Others offered more detailed interpretations such as:

- "I felt that the dancer was the "individual in life" and expressed how we as unique selves experience an almost overwhelming cascade of sensations – but at the same time we can't actually "feel" other people's feelings and experiences."
- "It represented the moods of life – from stress to joyful – the sound was very important to the flow of moods."
- "I sensed bewilderment, loss, death and rebirth."

Presenting Unless v.2 (November 21st 2009)

> Very synaesthetic.
>
> – anonymous audience member (November 26th 2009)

The second version of *Unless* was performed on November 21st 2009 at the Machines With Magnets gallery space in Pawtucket, RI. This iteration of the piece ran about 16 minutes and 4 seconds total, with section 1 at 4 minutes and 40 seconds, section 2 at 5 minutes and 20 seconds, and section 3 at about 6 minutes and 4 seconds.

By comparison to the first venue, this space was much more intimate and only offered floor seating. Grant Recital Hall is a fully functioning performance venue with a seating capacity of 135 while the gallery space at Machines With Magnets had space for a total of 30 audience members (mostly seated). The non-traditional nature of the performance setting in conjunction with its spatial limitations created new challenges while also inspiring some unique and interesting aesthetic possibilities. It became clear from the start that the piece would need to be re-evaluated in relation to this new context including Emily's interaction with certain aspects of the media environment.

That being said, many aspects of the piece in terms of basic structure and conceptual flow remained the same. The interactive systems and interface mechanisms were minimally altered, thus allowing Emily to continue developing her relationship with them. For instance, her insights into the movement/sound interactions after this performance seemed to expand in complexity.

> At certain points in the first and final sections, the impulses were coming so frequently, that it all coalesced into static noise, which resulted in a constant energy level for my movement. It was then that I would listen

> and feel for moments to stretch and extend time. (Emily Beattie on the November 21st performance)

By design, the sound palette offered moments of incessant activity and moments of rest. Emily's movement was directed by these contrasting behaviors that were inherent to the environment's responsiveness. Her conversation with the resulting sound always drove her forward, seeking the next connection.

The new space and its more intimate character did require some significant revisions in terms of media output, however. The audio system, for instance, was scaled back and reconfigured for 3 outputs in an irregular triangle shape immersing the entire audience and performer. In addition, the graphics system was rearranged for two mirrored projection outputs in a V shape behind the performer, facing (yet still partially immersing) the audience (See Figure 5.5).

Thus, the most experientially impacting modifications were due to the spatial limitations of the gallery space, which now required the media elements to exist on a more equal vertical plane with the performer and audience creating a more sensually immersive atmosphere, especially for Emily. The projections were placed on bisecting walls at about head height this time. The fact that Emily could touch the projection made her connection to it more intimate. Since the projections grew in size with increased sound intensity (by way of movement activation) she became immersed. They also splashed more light on to her body allowing her to feel "activated" more of the time. "I was more actively in conversation more of the time" (Emily Beattie on the November 21st performance). Therefore, the projection element emerged in the new space as a more present character. In turn, Emily's interaction style would be directed by this more tangible component.

Section 1

The time period between versions allowed us some critical distance, which helped to stimulate new thoughts about some of the conceptual relationships we were exploring in the work. For Emily, the idea of "control" in relation to the "environment" emerged more strongly in the second version, specifically with regard to the movement/sound interactions in the first section of the piece.

> There was some relationship that I could perceive but not control. I never got to display this understanding due to the system's nature, which changed every time I became close to owning it… I desired control the whole time, but was only strung along by minor access to that control. (Emily Beattie on the November 21st performance)

This kind of interactive performance dynamic was not so much a problem, as a challenge, and a satisfying one at that. I believe that the slight bit of indeterminacy inherent to all of the interactions (some more than others) did more to sustain interest than frustrate.

Interestingly, Emily's thoughts on this dynamic had as much to do with seeking control as feeling controlled.

> The conversation was not easy once I decided not to be controlled by the sound, but to retreat and return to it with my movement. Creating counterpoint gave me some autonomy and allowed me to make decisions in performance. I find that more difficult but rewarding. (Emily Beattie on the November 21st performance)

I had always imagined that her relationship to the media environment would reflect, in a sense, a human/environment interaction in the most holistic sense. Although humans seek to control their environments, they cannot, and doing so in many cases, leads to more conflict than resolution. At the same time, the environment is not really controlling the human. So, the most reasonable conclusion to draw would be that human-environment interactions are dynamic and feedback on one another. I believe Emily's relationship to the media environment reflected this concept. This is true interaction.

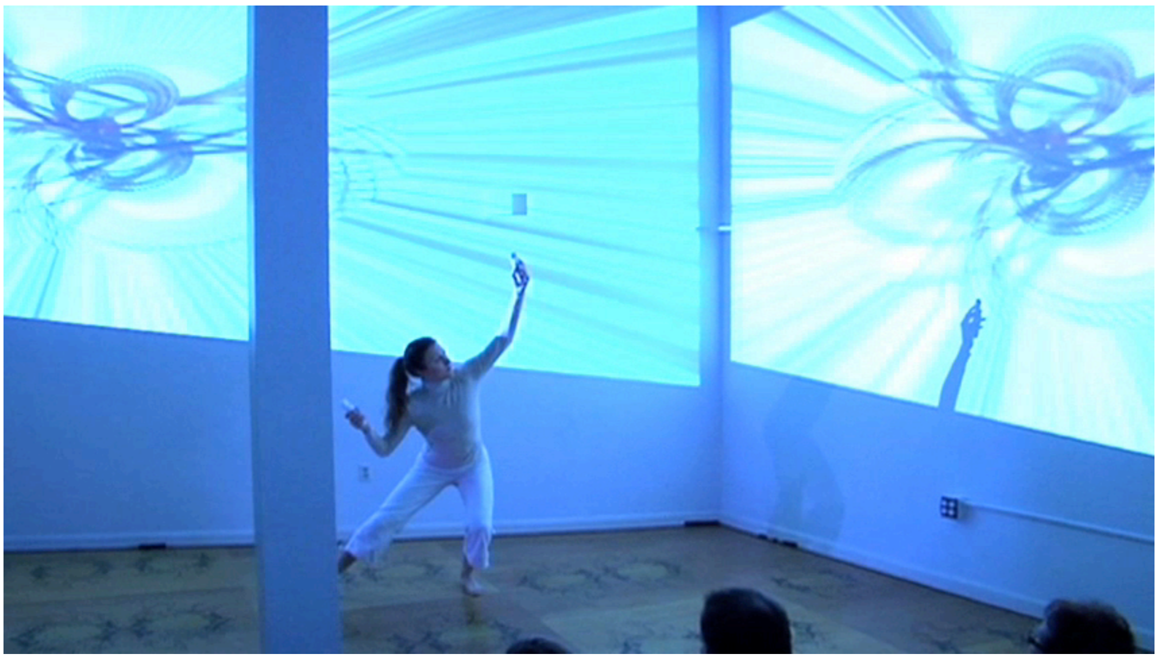

Figure 5.5. Emily performing during section 1 of the November 2009 performance.

Section 2

The intimate space of the gallery factored strongly in post-performance reflections on section 2 of this iteration. Since the music was unaltered from the first performance, it was revealing to see how the spatial constraints influenced Emily's choreography. In general, the gallery space offered less room for exploration. This required her to rethink her movement and expression in relation to what were the same temporal demands, musically speaking, as in May. What resulted was an increased focus on the emotional life of her character. With less room to move and decreased distance between her and the audience, expressive behaviors such as facial expressions and more nuanced body language could be more impacting. The spatial limitations actually augmented the potential for a more heightened dramatic experience.

In more intimate spaces there seems to be a connection between performer and audience that is more palpable and somewhat inevitable. The audience's occasional eye contact with the performer, or the dancer's momentary miscue, which would otherwise be lost on a large stage, captures the audience in a more immersive fashion. The spectators, who are really also participating in such spaces (if they know it or not), are able to empathize with the movement artist in a way that transcends the more traditional role of performer as objectified entity.

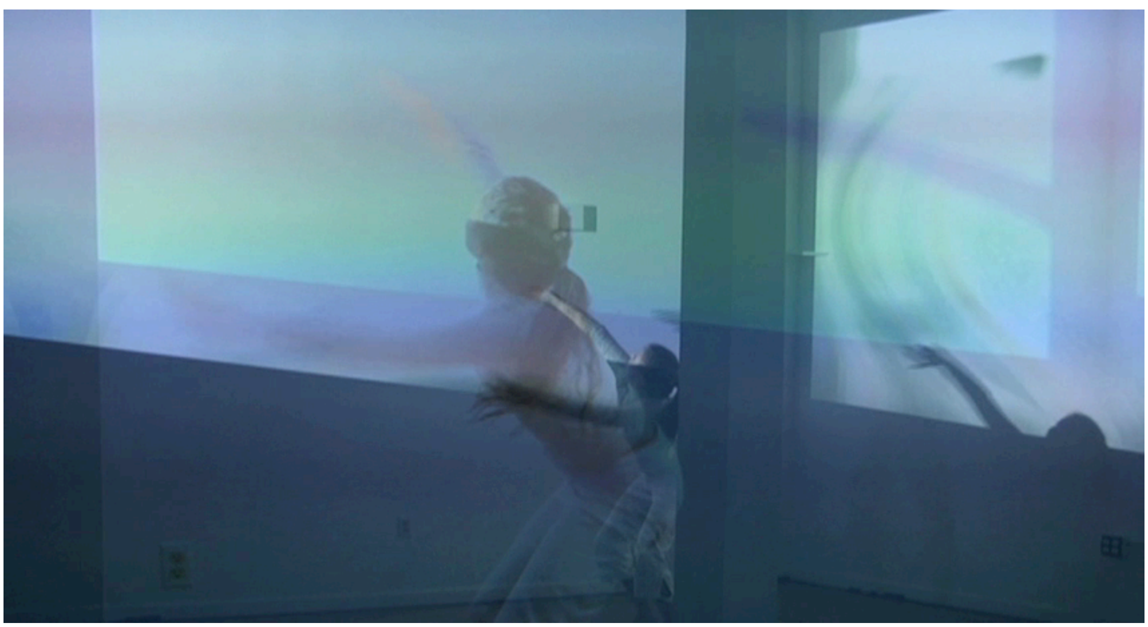

Figure 5.6. Emily Beattie reaching out during section 2 of November 2009.

Section 3

Section three exhibited some of the largest changes from the first version. In addition to the minimal lighting design of the May show (creating silhouettes of the human sculptures), I added a red and black crosshatched graphics component (projected) that also reacted to the amplitude of the interactive soundscape. Escalating amplitudes in the sound would create faster strobe effects with crosshatched line density also increasing in unison. Also, the agitated "redness" of the lighting design was more pronounced this time and offered an even more distinct contrast to the tranquil "blue" hues of the preceding section. This augmented the intensity in the atmosphere.

As Emily reflected, the confined space of the gallery also had an impact on the audience participation segment:

> In comparison to May, I found that the audience members that I recruited were more available, observant of my actions on them, and open to the environment. Being on the same level and in a more intimate space could have influenced the ease that I sensed. (Emily Beattie on the November 21st performance)

The nature of the non-traditional performance venue seemed to establish a different dynamic between the performer and audience, which led Emily to perceive a more relaxed disposition from the audience. This is an interesting insight, and warrants more in depth investigation. It would be interesting to systematically research the effects of the size and style of the performance venue on the nature of audience interaction.

In fact, the audience was more or less "on stage" the entire time in November. Emily was able to dance very close to them, effectively "invading" their personal space on a consistent basis. Since it could be argued that the audience already felt as if they were participating in the piece, as they were becoming desensitized to their proximity to the performer, the barriers to inclusion were eroding.

Interestingly, just as the spatial limitations would limit certain possibilities for movement exploration, they seemed to also provide artistic challenges that encouraged deeper performance practices. In terms of Emily's improvisational aesthetic, it seems that the constraints actually demanded even more mindfulness.

> It was more difficult to respond as freely to the sound with such spatial limitations. I felt that not only was the sound bouncing off walls, but I too was being bounced around (from person to person and wall to person). This served to make me more tuned into my improvisation because my usual movement vocabulary was clipped. (Emily Beattie on the November 21st performance)

In addition to demanding an increased awareness from Emily, the venue's dimensions also added to the power of the sound design. The larger the sound in relation to this more compressed space, the more the tension grew.

The structure of the space clearly affected perception of the piece in more ways than one. Even in its closing moments, Emily was left with a lasting impression of the room's effect.

> At the close of the work I backed into the corner that bisected the mirrored projections. The architecture allowed me to absorb the two projections and the audience's reaction in my peripheral vision. It connected the audience to me and to the system because it was all in focus. (Emily Beattie on the November 21st performance)

Her ability in this space to attend (at least peripherally) to every aspect of the experience allowed for an intense feeling of closure.

Figure 5.7. Emily Beattie and selected audience members during section 3 of the November 2009 performance.

Audience Research

The second version of *Unless* was performed for about 30 people. Instead of being surveyed right after the performance, I gathered email addresses and invited audience members to respond to an online survey (via surveymonkey.com), after the fact, using the same questionnaire format as in May.

The contrasting post-performance research methods in May and November educed some interesting variations in audience feedback. It seems that the more immediate style of surveying in May elicited a greater percentage of responses overall. However, in November, although much smaller in number, those that did decide to respond were able to express their feelings in a more detailed manner.

9 anonymous audience members responded within 2 weeks of the performance. Again, detailed accounts of these experiences, in the form of excerpted quotes, are organized in relation to research related concepts.

Kinesthetic Empathy

Impressions of kinesthetic empathy in the November audience seemed less explicit, yet more connected to other aspects of the performance. For instance, the awareness of a sensory gestalt emerged in some feedback, evoking commentary on how raw sensations could give way to higher level thoughts about form and vice versa.

- "I felt very surrounded. It was an entire sensation, yet was focused in direction. Sometimes this could be overwhelming…when this occurred the continuity of the piece was stressed…"
- "The sound and imagery added to the movement of the dancer. I connected to the part when Emily was dancing around the 4 people, it seemed like they had negative energy and she didn't want to absorb that into herself. The sound, Emily, and visuals were interconnected. As Emily moved, the sound and visuals changed. Without the music or visuals, the piece would not have made as much of an emotional impact. Without the dancer it would have appeared to be technologically interesting, but more abstract."

Other impressions that implicitly referred to experiences of kinesthetic empathy implied an intuitive knowledge of an auditory-motor mirror neuron system. Connections between the mover, sound and visuals were perceived with ease.

- "Very interesting. I definitely felt movement while viewing the performance. The sound and movement created a very visceral experience."
- "There was a sense of movement in the music, which was hard to pinpoint. The dance component was a great visual way to experience what we were hearing and to become involved with the music...they made sense together."

- “The dancer was obviously hugely talented and in control of her performance. She evoked an emotion that took the music and visuals to a different level of personal connection.”

Immersion and Interaction

I was particularly pleased with one audience member’s comment about the immersive nature of the experience. I believe that this clear and concise reaction to the piece was not only a reflection of the intimate space of the gallery, but the quality of the work. As they remarked:

- “I felt very much immersed in and transplanted into a work of art...surrounded by it.”

The piece was not without its critics, however, as one audience member speculated that prior knowledge of the physical interface’s relation to the multimedia environment was necessary for comprehension.

- “I really think that some explanation is needed for the audience, as I heard a lot of feedback that if one hadn't known previously about how the Wiis worked, they would have been frustrated and lost.”

Based on the feedback in May and my knowledge of the audience in November, I know that many in attendance were not privy to the mechanics of the interaction and still reacted very positively to the experience. Although I do not have statistical data to back this up, my general conclusion is that the piece did not require a verbal explanation of the technologies involved in order to be appreciated. Most feedback regarding the interactive components of the piece was very encouraging.

- “The machine and the dancer were remarkably synchronous in intention. She moved in a way that was totally believable, the lines blurred.”
- “The balance between the dance, visual, and auditory components was key. The three elements synthesized well and while one or the other might have been highlighted in certain parts, they smoothly transitioned in and out of the spotlight.”
- “The performance was great! I think all 3 components (music, dance, video) complemented each other very well. If one element were missing it would have felt less intense. I really enjoyed how the movement affected the sound.”
- “If you removed any one of the elements (movement/music/visual) it would not stand as well. Perhaps the dancer and music could stand-alone. But, the visual really binds everything together.”
- “The music and movement and visuals were very tightly integrated.”

The most negative reaction concerned the visual component. Although the graphics were mainly intended to help establish the atmosphere of the performance world, I should not have assumed that all in attendance would understand that. The comment is fair and will be expanded upon in the Reflections section.

- “My only criticism would be that I would like to see perhaps more variety in the visuals. The circular 3D imagery got a bit monotonous and it would be nice to see some more organic textures.”

In contrast, the insight exhibited by one audience member regarding the more holistic experience of the interactive multimedia environment was particularly uplifting. It’s comments like these that inspire one’s next project.

- “The human dancer element made for a more captivating experience. Connections between movement/dancer-expression, projected images and sonic elements were very strong. Abrupt movements translated to abrupt sonic/imagery changes. Subtle-to-subtle as well. But there was more! Timbre, dynamics, complexity, it was all there…more than just strict binary operations. You have touched upon usability and expression in a way I have seldom or never seen before...”

Audience Participation

Both Emily and I were less concerned with the audience participation segment in November having witnessed its success in May. However, a different group of people in attendance will always change the dynamic of such a situation, so some anxiety will always be warranted. Once again, this section provided an interesting combination of tension (anticipation), release and renewal to the space as evidenced by these remarks.

- "The incorporation of the audience in the last part of the piece was interesting...it could have become cliché, but avoided it. I noticed that then, the piece for me became about fear...internally...I felt fear that I would be chosen."
- "I enjoyed the interaction with the audience by dragging and posing people. It brought a levity and spirit to what would have been a passive experience."

Interpretation

The audience's interpretation of the piece was once again intriguing. It seems that the intimate space of the gallery grounded the experience more than in May. In other words, since engagement was increased through perceptions of being more immersively located inside the performance world, the work provided more accessible points of connection for the audience with both Emily and the piece as a whole. Whereas the first version of the piece took on a more "Other-worldly" character, this time spectators seemed to create more associations metaphorically between the performance world and the real world. Emily's performance, although still transcendental in many ways, seemed to symbolize something very basic, earthly, and human.

- "The abstract theme left me thinking about phases of life and I was lead to think this because of the direct interaction with the audience which I enjoyed. The space was a nice choice because you felt like a part of the art as it evolved."

- “Thematically I would say it was as if she was "finding her way" in life. She was searching, then interacted with others, and then found peace.”
- “At the close of the performance, I felt refreshed, also somewhat dazed. Conceptually, it seemed to be about taking control or controlling one's surroundings. Really enjoyed the experience...thank you!”

Refreshingly, many audience members appreciated the performance for its raw experiential nature and chose not to read into it anymore than necessary. In many ways these comments reflect my intention as an artist, to offer a dynamic multi-sensory experience that can be appreciated for its visceral qualities alone.

- “Mostly, I felt hopeful, impressed, happy that artists are making important work that make us think about the world around us in a creative, dynamic way. Beautiful piece. Bravo.”
- “Overall, I really enjoyed it. Made my feelings varied over the course of the performance from calm to anxious and back to calm at the end.”
- “I suppose art in general isn't always supposed to be about any intended meaning so much as the individually interrupted moment.”

Reflections

Overall, I believe that both versions of the piece were successful in offering a dynamic sensory experience while eliciting a wide range of raw emotions from the audience. The situation stimulated those in attendance to give thought to their perceptions and actively engage in processes of interpretation. Refreshingly, many audience members appreciated the performance for its raw experiential nature and chose not to read into it anymore than necessary. In many ways, that reflects my intention as an artist, to offer an experience that can be appreciated for its visceral qualities alone. That being said, I appreciated those who offered imaginative thoughts about the possible narrative impulses driving the work.

Whether or not the work actually reconfigured the senses, or enhanced perceptions of kinesthetic empathy for those in attendance is debatable. This is one area where this kind of qualitative artistic research falls short. Future collaborations with scientific researchers could reveal additional insights. Also, it would be interesting to attempt a more structured approach to the audience research that might allow for deeper analyses of the relationships between artistic experience with the various media types involved and perceptions of the performance.

It does seem evident that people's lasting impressions of the dancer's interaction with the interactive media environment were reasonably vivid. This is an accomplishment in and of itself. It is safe to say that many in attendance walked away from the performance feeling as if their mind was exercised in a new and exciting way. In addition, sensory impressions clearly related to the notion of kinesthetic empathy were also evident in the feedback. Therefore, one might conclude that the audience was truly "feeling" the interaction.

In terms of my impression of the artistic result as a whole, I was very satisfied with the interaction between movement and sound. This relationship provided an intensely visceral experience for both the performer and the audience, which was clearly reflected in the feedback. The graphics component, however, seemed to fall a little short. In hindsight, it's possible that this aspect was not given the full attention it deserved. I completely understand the critique of that element in terms of it seeming somewhat monotonous. In terms of raw size, the visual would have seemed to represent a substantial part of the whole, yet it lacked the full interconnectivity and variety that was perceived in the sound.

In future work, I would like to spend more time balancing the impact of each element even further. At any given time, a person may experience one sensory modality more intensely than another, but one of my goals as an interactive multimedia artist is to achieve a lasting impression that reflects a fully balanced treatment of all the senses involved.

Lastly, I have realized through this entire process from its initial conceptual foundations through production and exhibition that a certain discord existed between some of the scientific foundations and the artistic development. Ultimately, I found that my creative process could not rigidly follow the multi-sensory perception theory that originally inspired the work. I quickly found that a constant tension existed between pursuits. That being said, I still believe that those theoretical foundations were drawn upon loosely, even if sometimes unconsciously, through a deeper experience of the artistic process and product as it emerged. I also feel that the audience research methods and analyses were thoroughly informed by those same scientific foundations.

Conclusions and Implications

> I was a bit skeptical at first, wondering about the real-world application/usefulness of such an exercise, but I found myself drawn in, probably because of the use of several mediums at once, addressing multiple senses…the integration of stimuli was compelling.
>
> – anonymous audience member (May 1st 2009)

The implications of this interdisciplinary dissertation project are multi-dimensional in scope. First, this project extends and develops the vocabulary of current multi-sensory research and multimedia arts practice. It integrates elements from the sonic, visual and physical computing arts in conjunction with philosophical and scientific thought experiments inspired by the multi-sensory perception and embodied cognition research disciplines.

Secondly, the project has successfully delivered on the task of developing some unique technical tools, which were necessary to make the artistic extensions possible. Through a process of experimentation, an interactive multimedia environment was realized for the precise purpose of exploring multi-sensory and multimedia relationships in the context of a contemporary performance work.

Finally, the project has attempted to demonstrate some links between scientific and philosophical discourses and a unique work of art. This project has successfully produced an artistic experience that is symbolic of a synthesis between multiple disciplines. Research surrounding multi-sensory perception and embodied experience has been integral to the process of creating the artistic work.

The focus on and utilization of coherences between our experiences of physical movement, sound, and moving image has resulted in a dynamic approach to interactive multi-media performance. Engagement with contemporary research regarding cognition, perception and embodied experience, is necessary for the continued development of multi-sensory and multimedia art forms.

Most importantly, this work suggests that understanding the way in which we perceive interactive multimedia art works is essential to understanding their value. There are indeed coherences between the way we perceive, create and experience, and that knowledge may be the most fundamental device of any artist.

REFERENCES

Andersen, Tobias S., Kaisa Tiippana, and Mikko Sams. 2004. Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research* 21, no. 3 (November): 301-308.

Baalman, M., D. Moody-Grigsby, and C. Salter. 2007. Schwelle: Sensor Augmented, Adaptive Sound Design for Live Theatrical Performance. *Proceedings of the 7th International Conference on New Interfaces for Musical Expression*.

Bennett, Drake. 2008. Don't just stand there, think. *The Boston Globe*, January. http://www.boston.com/bostonglobe/ideas/articles/2008/01/13/dont_just_stand_there_think/.

Bersten, Rosanne. 2003. Singing a stream of floating images. *Australian IT*. http://www.tmema.org/messa/press/australian_it.html (accessed September 16, 2009).

Bokowiec, Julie Wilson, and Mark Alexander Bokowiec. 2006. Kinaesonics: The intertwining relationship of body and sound. *Contemporary Music Review* 25, no. 1: 47-57.

Botvinick, Matthew, and Jonathan Cohen. 1998. Rubber hands `feel' touch that eyes see. *Nature* 391, no. 6669 (February 19): 756.

Broadhurst, Susan. 2007. *Digital Practices: Aesthetic and Neuroesthetic Approaches to Performance and Technology*. UK: Palgrave Macmillan.

Campen, Cretien van. 2007. *The Hidden Sense: Synesthesia in Art and Science*. Cambridge, MA: The MIT Press.

Camurri, Antonio, Barbara Mazzarino, Matteo Ricchetti, Renee Timmers, and Gualtiero Volpe. 2004. Multimodal Analysis of Expressive Gesture in Music and Dance Performances. In *Gesture-Based Communication in Human-Computer Interaction*, 357-358.

Castellano, Ginevra, Roberto Bresin, Antonio Camurri, and Gualtiero Volpe. 2007. Expressive Control of Music and Visual Media by Full-Body Movement. *Proceedings of the 2007 Conference on New Interfaces for Musical Expression.*

Chion, Michel, Claudia Gorbman, and Walter Murch. 1994. *Audio-Vision*. Columbia University Press.

Coniglio, M. 2007. Materials vs Content in Digitally Mediated Performance. In *Performance and Technology: Practices of Virtual Embodiment and Interactivity, ed. Susan Broadhurst and Josephine Machon*, 78-84. UK: Palgrave Macmillan.

Cytowic, Richard E. 2003. *The Man Who Tasted Shapes*. Cambridge, MA: MIT Press.

Dennett, Daniel C. 1993. Review of F. Varela, E. Thompson and E. Rosch, The Embodied Mind. *The American Journal of Psychology* 106, no. 1 (Spring): 121-126.

Ecker, Adam J., and Laurie M. Heller. 2005. Auditory - visual interactions in the perception of a ball's path. *Perception* 34, no. 1: 59-75.

Eitan, Zohar, and Roni Y Granot. 2006. How Music Moves: Musical Parameters and Listeners Images of Motion. *Music Perception* 23, no. 3: 221-248.

Emmerson, Simon. 2007. *Living Electronic Music*. England: Ashgate Publishing.

Franklin, Gene F., Abbas Emami-Naeini, and J. David Powell. 1993. *Feedback Control of Dynamic Systems*. Boston, MA: Addison-Wesley Longman Publishing Co., Inc.

Gallace, Alberto, and Charles Spence. 2008. The cognitive and neural correlates of "tactile consciousness": a multisensory perspective. *Consciousness and cognition* 17, no. 1 (March): 370-407.

Gazzola, Valeria, Lisa Aziz-Zadeh, and Christian Keysers. 2006. Empathy and the somatotopic auditory mirror system in humans. *Current biology : CB* 16, no. 18 (September 19): 1824-9.

Gibson, James J. 1979. *The Ecological Approach To Visual Perception*. Boston: Houghton Mifflin.

Gielis, Johan. 2003. A generic geometric transformation that unifies a wide range of natural and abstract shapes. *Am. J. Bot.* 90, no. 3 (March 1): 333-338.

Godøy, R.I. 2003. Motor-mimetic music cognition. *Leonardo* 36, no. 4: 317-319.

Grau, Oliver. 2003. *Virtual Art: From Illusion to Immersion*. Cambridge, MA: The MIT Press.

Grierson, M.S. 2005. *Audiovisual Composition*. University of Kent.

Hagendoorn, I. 2004. Some Speculative Hypotheses about the Nature and Perception of Dance and Choreography. *Journal of Consciousness Studies* 11:79-110.

Handel, Stephen. 2006. *Perceptual Coherence: Hearing and Seeing*. Oxford University Press.

Hansen, Mark B. N. 2006. *Bodies in Code: Interfaces with Digital Media.* New York, USA: Routledge.

Hein, Grit, Oliver Doehrmann, Notger G. Muller, Jochen Kaiser, Lars Muckli, Marcus J. Naumer, et al. 2007. Object Familiarity and Semantic Congruency Modulate Responses in Cortical Audiovisual Integration Areas. *J. Neurosci.* 27, no. 30 (July 25): 7881-7887.

Herbart, Johann. 1969. *Kleinere Abhandlungen zur Psychologie.* Amsterdam, The Netherlands: E.J. Bonset.

Holmes, Nicholas P., and Charles Spence. 2005. Multisensory Integration: Space, Time and Superadditivity. *Current Biology* 15, no. 18 (September 20): R762-R764.

Hunt, A., and R. Kirk. 2000. Mapping Strategies for Musical Performance. In *Trends in Gestural Control of Music*, ed. M. Wanderley and M. Battier, 231-258. Paris: IRCAM.

Hunt, A., M. Wanderley, and M. Paradis. 2002. The Importance of Parameter Mapping in Electronic Instrument Design. *Proceedings of the 2002 Conference on New Interfaces for Musical Expression*.

Huron, David. 1994. What is melodic accent? A computer-based study of the liber usualis. *Canadian University Music Society Theory Colloquium.*

Jensenius, A.R., T. Kvifte, and R.I. Godoy. 2006. Towards a gesture description interchange format. *Proceedings of NIME*.

Kelso, J. A. Scott. 1995. *Dynamic Patterns: The Self-Organization of Brain and Behavior*. Cambridge, MA: The MIT Press.

Kohler, Wolfgang. 1929. *Gestalt Psychology*. New York: H. Liveright.

Krueger, Myron. 1977. Responsive Environments. *Proceedings of the American Federation of Information Processing Societies*.

———. 1993. "An Easy Entry Artificial Reality". In *Virtual Reality: Applications and Explorations*, ed. Alex Wexelblat. Boston, MA: Academic Press Professional.

Lahav, Amir, Elliot Saltzman, and Gottfried Schlaug. 2007. Action Representation of Sound: Audiomotor Recognition Network While Listening to Newly Acquired Actions. *J. Neurosci.* 27, no. 2 (January 10): 308-314.

Leman, Marc. 2007. *Embodied Music Cognition and Mediation Technology*. MA: The MIT Press.

Leman, Marc, and Antonio Camurri. 2006. Understanding musical expressiveness using interactive multimedia platforms. *Musicae Scientiae*, Interdisciplinary Musicology.

Levin, Golan, and Zachary Lieberman. 2004a. In-Situ Speech Visualization in Real-Time Interactive Installation and Performance. *Proceedings of The 3rd International Symposium on Non-Photorealistic Animation and Rendering*.

———. 2004b. *Messa Di Voce Report*. http://tmema.org/messa/report/messa_report_600.pdf (accessed April 30, 2008).

Lipscomb, S.D. 2005. The perception of audio-visual composites: Accent structure alignment of simple stimuli. *Selected Reports in Ethnomusicology* 12:37-67.

Lipscomb, S.D., and E.M. Kim. 2005. Perceived match between visual parameters and auditory correlates: An experimental multimedia investigation. *Proceedings of the 8th International Conference on Music Perception and Cognition*.

Livingstone, Dan, and Eduardo Miranda. 2004. Composition for Ubiquitous Responsive Sound Environments. *Proceedings of the International Computer Music Conference*.

Lozano-Hemmer, Rafael. N.d. *Relational Architecture*. http://rhizome.org/artbase/2398/fear/relarc.html (accessed October 15, 2009).

Magbanua, Pinky. Wireless Tilt Sensing In An Occupational Therapy Application. http://itp.nyu.edu/physcomp/sensors/Reports/WirelessTiltSensingInAnOccupationalTherapyApplication (accessed July 30, 2008).

Marks, Lawrence E. 1987. On cross-modal similarity: Auditory-visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance* 13, no. 3: 384-394.

Marks, Lawrence E., Elisheva Ben-Artzi, and Stephen Lakatos. 2003. Cross-modal interactions in auditory and visual discrimination. *International Journal of Psychophysiology* 50, no. 1-2 (October): 125-145.

Marshall, M. T., Nils Peters, Alexander R. Jensenius, Julien Boissinot, Marcelo Wanderley, Jonas Braasch, et al. 2006. On the Development of a System for Gesture Control of Spatialization. *2006 International Computer Music Conference*.

Marshall, M.T., Joseph Malloch, and Marcelo Wanderley. 2007. Non-Conscious Control of Sound and Spatialization. *Proceedings of the 4th International Conference on Enactive Interfaces*.

Massumi, Brian, and A. Mulder. 2007. "The Thinking-Feeling of What Happens: An Interview with Brian Massumi". In *Interact or Die: There Is Drama In The Networks*, ed. J. Brouwer and A. Mulder, 70-91. Rotterdam: NAi Publishers.

McGurk, Harry, and John MacDonald. 1976. Hearing lips and seeing voices. *Nature* 264, no. 5588 (December 23): 746-748.

Merleau-Ponty, Maurice. 1962. *Phenomenology of Perception*. London: Routledge & Kegan Paul.

Moggridge, Bill. 2007. *Designing Interactions*. Cambridge, MA: The MIT Press.

Morein-Zamir, Sharon, Salvador Soto-Faraco, and Alan Kingstone. 2003. Auditory capture of vision: examining temporal ventriloquism. *Cognitive Brain Research* 17, no. 1 (June): 154-163.

Mulder, A. 2000. Toward a choice of gestural constraints for instrumental performers. In *Trends in Gestural Control of Music*, ed. M. Wanderley and M. Battier, 315-335. Paris: IRCAM.

———. 2007. "The Exercise of Interactive Art". In *Interact or Die: There Is Drama In The Networks*, ed. J. Brouwer and A. Mulder, 52-69. Rotterdam: NAi Publishers.

Noonan, L., and Golan Levin. 2003. Q&A: Golan Levin. *Huddersfield Metro*. http://www.tmema.org/messa/press/hud_metro.html (accessed April 30, 2008).

Pavani, Francesco, Charles Spence, and Jon Driver. 2000. Visual Capture of Touch: Out-of-the-Body Experiences With Rubber Gloves. *Psychological Science* 11, no. 5: 353-359. http://www.blackwell-synergy.com/doi/abs/10.1111/1467-9280.00270 (accessed April 17, 2008).

Penny, Simon. N.d. *Traces*. http://www.ace.uci.edu/penny/works/traces.html (accessed October 15, 2009).

Ramachandran V.S.[1], and Hubbard E.M. 2001. Synaesthesia -- A window into perception, thought and language. *Journal of Consciousness Studies* 8:3-34.

Ramachandran, V. S. 2004. *A Brief Tour of Human Consciousness: From Impostor Poodles to Purple Numbers*. New York: Pi Press.

Rizzolatti, Giacomo, Corrado Sinigaglia, and Frances Anderson. 2008. *Mirrors in the brain*. Oxford, UK: Oxford University Press.

Rockwell, John. 2006. Tracking Mankind's Rise, With the Help of Computers. *The New York Times*, January 20.

Rokeby, David. 1996. Transforming Mirrors: Subjectivity and Control in Interactive Media. http://homepage.mac.com/davidrokeby/mirrors.html (accessed September 3, 2009).

Rovan, J., M. Wanderley, P. Depalle, and S. Dubnov. 1997. Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. *Proceedings of The Kansei Technology of Emotion Workshop*.

Rovan, J., R. Wechsler, and F. Weiss. 2001. Seine hohle Form: Artistic Collaboration in an Interactive Dance and Music Environment. *1st International Conference on Computational Semiotics in Games and New Media*.

Ryan, J. 1991. *Some Remarks on Musical Instrument Design at STEIM*. http://www.steim.org/steim/texts.php?id=3 (accessed October 13, 2009).

———. 1992. Effort and expression. *Proceedings of the 1992 International Computer Music Conference* .

Ryan, Marie-Laure. 2001. *Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media*. Baltimore: Johns Hopkins University Press.

Salter, Chris, M. Baalman, and D. Moody-Grigsby. 2007. Between Mapping, Sonification and Composition: Responsive Audio Enviroments in Live Performance. *Proceedings of the ICMC 2007*.

Schroeder, Franziska. 2006. Bodily instruments and instrumental bodies: Critical views on the relation of body and instrument in technologically informed performance environments. *Contemporary Music Review* 25, no. 1: 1-5.

Shams, Ladan, Yukiyasu Kamitani, and Shinsuke Shimojo. 2000. Illusions: What you see is what you hear. *Nature* 408, no. 788.

Shimojo, Shinsuke, and Ladan Shams. 2001. Sensory modalities are not separate modalities: plasticity and interactions. *Current Opinion in Neurobiology* 11, no. 4 (August 1): 505-509.

Stanford, Terrence R., and Barry E. Stein. 2007. Superadditivity in multisensory integration: putting the computation in context. *Neuroreport* 18, no. 8 (May 28): 787-792.

Topological Media Lab. N.d. *Topological Media Lab*. http://www.topologicalmedialab.net/joomla/main/component/option,com_frontpage/Itemid,1/lang,en/ (accessed October 15, 2009).

Troika Ranch. Troika Ranch Website. http://www.troikaranch.org/ (accessed October 15, 2009).

van Wassenhove, Virginie, Dean V. Buonomano, Shinsuke Shimojo, and Ladan Shams. 2008. Distortions of Subjective Time Perception Within and Across Senses. *PLoS ONE* 3, no. 1 (January 16): e1437.

Varela, Francisco J., Evan T. Thompson, and Eleanor Rosch. 1992. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: The MIT Press.

Välimäki, Vesa, and Tapio Takala. 1996. Virtual musical instruments: natural sound using physical models. *Org. Sound* 1, no. 2: 75-86.

Vroomen, Jean, and Beatrice de Gelder. 2003. Visual motion influences the contingent auditory motion aftereffect. *Psychological Science* 14, no. 4 (July): 357-361.

Waisvisz, Michel. 1985. The Hands, a set of remote MIDI-controllers. *Proceedings of the 1985 International Computer Music Conference*.

Walker, Robert. 1987. The effects of culture, environment, age, and musical training on choices of visual metaphors for sound. *Perception & Psychophysics* 42, no. 5: 491-502.

Wanderley, M.M., and P. Depalle. 2004. Gestural control of sound synthesis. *Proceedings of the IEEE* 92, no. 4: 632-644.

Ward, Jamie, Brett Huckstep, and Elias Tsakanikos. 2006. Sound-Colour Synaesthesia: To What Extent Does it Use Cross-Modal Mechanisms Common to Us All. *Nature* 42, no. 2 (February): 264-280.

Warren, William H. 2006. The Dynamics of Perception and Action. *Psychological Review* 113, no. 2 (April): 358-389.

Winkler, Todd. 1995. Making Motion Musical: Gesture Mapping Strategies for Interactive Computer Music. *Proceedings of the 1995 International Computer Music Conference*.

Wolfe, Uta, Jacob A. Comee, and Bonnie S. Sherman. 2007. Feeling darkness: A visually induced somatosensory illusion. *Perception & Psychophysics* 69:879-886.

Wozniewski, Mike, Zack Settel, and Jeremy Coooperstock. 2006. A framework for immersive spatial audio performance. *Proceedings of New Interfaces for Musical Expression*.

APPENDIX A: DATA-DVD CONTENTS

***Unless* (2009) Documentation:**

Unless_May_1_2009.mov - May 1st 2009, Grant Recital Hall, Providence, RI

Unless_Nov._21_2009.mov - Nov. 21st 2009, Machines with Magnets, Pawtucket, RI

Max/MSP/Jitter Patches for *Unless* (2009):

Folders (with patches and most of the necessary externals):

May Graphics System

November Graphics System

OSCulator Patch (*helps the Wiimote talk to Max)

Section 1 Sound/Music System

Section 3 Sound/Music System

*Section 2 was a pre-composed music piece. See documentation above.

APPENDIX B: AUDIENCE RESEARCH SURVEY

Audience Research for Brian Knoth/Emily Beattie Collaboration

Unless (2009) -- Improvisational Movement with Interactive Media Environment

<u>Performance Experience (for Beattie/Knoth Collaboration):</u>

Please take a few minutes to think about your experience of the Beattie/Knoth movement, sound and image performance. Try to be conscious of how it made you feel and post your responses below and on the back of the sheet, using the questions as a guide. You can leave your responses with Emily or Brian. Or, feel free to email brian_knoth@brown.edu with your responses/feedback if maintaining anonymity is not important. Thanks!

What was your experience of the performance? How did it make you feel? Were you aware of any movement sensations while you were watching it? How do you think the sound and imagery may have impacted your perception of movement? What aspects of the performance did you connect to? In your experience, what was the relationship, if any, of the sound, mover, and visuals? How do you think your perception of the movement would have changed if there were no music and/or visual component? If you were to thematically/conceptually describe this piece how would you do so?