Directions in Sound/Image Synthesis -

Study of Bi-modality within Emerging Multimedia Practice

Michel-Charles Therrien

A Thesis

in

The Humanities Doctoral Program

Presented in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy at
Concordia University
Montreal, Quebec, Canada

December 2004

Your file Votre référence
ISBN: 0-494-09981-X
Our file Notre référence
ISBN: 0-494-09981-X

ABSTRACT

Directions in Sound/Image Synthesis -
Study of Bi-modal Emerging Multimedia Practice

Michel-Charles Therrien
Concordia University, 2004

Humanities Doctoral Program

In recent years, due to the high level of technological standardization of the home computer, entertainment systems and high-speed web paraphernalia, our access to, desire for, and perspectives on various technological elements have been taken for granted. We have become acclimatized to the latest super-product promotion, web surfing, or interaction with the latest object-oriented programming software. Although audio often enhances graphical user interfaces (GUIs) to deliver feedback to users, interfaces in which sound is the principal element, or even shares a parallel domain, are atypical.

Nonetheless, this practice is emerging as a viable junction within new technologies and digital media, encouraged by the ongoing development of mobile and all-pervading computing. Within emerging multimedia practice, spatial objects (visuals) require essentially to be structured in a two-dimensional space, whereas temporal objects (sound) require a sequential array. The goal of this dissertation is to develop a broader understanding of these two key elements, while providing designers and educators with useful information that is appropriate to familiar design discourse.

ACKNOWLEDGEMENTS

This manuscript is dedicated to my family: my wonderful wife Mayumi, whose unrelenting support, love and patience made this possible, and to my remarkable children, Ayumi, Akira and Araya, who inspire my work and my life. Special thanks to Okåsan who has encouraged my endeavor.

I would like to express my warmest and deepest gratitude to my principal supervisor, Mr. Allan Crossman, for his continuing encouragement and assistance throughout the years; his exceptional tolerance, assiduous advice and open-mindedness have been the cornerstone of my academic spirit. This work would not have seen the light without him.

I gratefully acknowledge the support and benefaction of my committee, Dr. Will Straw, Ms. Marielle Nitowslaska, and my appreciation to Dr. Matt Soar who accepted to step in late in the process. Also to Dr. Andra McCartney and Jonathan Stern for their insightful comments.

I would like to thank my good friend and colleague, Mr. David Wells, for his resolute support, and urging me to complete my work sooner than later; and Jessica Landry, who's work as a research assistant was invaluable.

I would like to thank my friends, students and colleagues at the Centre for the Study of Learning and Performance for their useful discussions and suggestions, and Dr Phillip C. Abrami for his support.

This work is dedicated to my late and dearly loved parents; Claire and Paul-Emile Therrien.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Orientation and Goals

In my dissertation, I shall investigate the practice of cross-modal association by means of two forms of art, non-verbal sound[1] and visuals, and how, when sound is constructively designed and appropriately integrated with the visual framework, it can foster apposite information acquisition, facilitate perception and enhance aesthetic concepts. Sound design[2] can support communication in a variety of ways: combined with visual information, this dual-channel process can increase and focus attention by reducing the distraction components of conflicting stimuli.

The study will investigate the relationship between sound design and visuals within a multimedia environment by examining the perceptual, socially preconceived notions of this relationship; as well, it will consider the interface between the written word and its sound. The dynamic assortment of innovative technological tools, which has shaped our expectations, will be reviewed to show how the use of sound design has permeated our contemporary technological environment.

---

[1] From this point on, when I refer to sound or sound design, it will be implied that it is non-verbal sound that is addressed, unless otherwise noted.

[2] "Sound design": this is any organized audio or 'sonic' integrant that is willfully included within a work; it is a more inclusive term than music and can encompass (though not be limited to) music, background and atmospheric sounds; it also concerns the modes of interaction between the various sound materials as well as those between auditory and visual space. I make no distinction between music, dialogue, and sound effects: these types of dialectical interaction can in fact involve any sort of sound, creating a coherent, organically structured sound design.

I have realized that sound design is progressively becoming a contingent fundamental, though often still disappearing from the effect (the memory) of the visual/sound complex, and it is obvious to me that even today's media users, students and instructors have relegated the sound component to a secondary aspect of their work and that designers rarely know how to use sound effectively, which usually has the misfortune of incorporating makeshift sound environments, leading to the audio constituent being judged as bothersome and inappropriate. Non-verbal or organized sound (sound design/music) is repeatedly treated as the "poor relation" of drama, images and visuals, and therefore perceived to be a deterrent to good instruction, since it is often associated with elements that deter from learning, as in Mayer's (2001) "Coherence Principle." One of this writer's objectives is to offer some essential details about variegated approaches that authors (sound designers, educators, producers, etc.) can use to incorporate sounds into multimedia venues; it is my belief that these multimedia authors in all probability want to create tools that are as "cutting edge" as possible (Lumsdem et al; 2002).

It seems that new digital communication media technology has often been accompanied by the argument that it is a novel way to represent reality, therefore asserting that it has surpassed older technologies and is able to focus on representation of the real in a more advanced and improved fashion. We need only look at the Analog vs Digital debate and the hi-definition TV (HDTV), 3D-animation and digital recording arguments promoting a more focused "realism." I would contend that this new refined "realism," made possible by digital technologies, is not bound to its representational

activities. We could study an aesthetico-social evolution of digital media, but for the purpose of this work, it would be fair to say that the major modifications in digital communication media, such as the current transfer and usage of computer and network-based technologies, not only direct us to the creation of new artistic venues but in addition can set into motion a positive tendency that existed in previous artistic ventures.

Technology is pressuring us to grasp our environment in a new way. I am concerned that the existing tendency to limit the role of sound design to that of an extraneous element will prevent authors in general from acknowledging its contribution to the valuable function that sound can play in facilitating various approaches to communication within mediated environments. We should validate the appropriate use and design of sound by emphasizing its role within these environments, and present a comprehensive impression of the elements within the multimedia environment - more specifically, the bi-modality concept of the sound/image model. I would contend that the appropriate treatment of sound design acts as a key functionality in the communication process: that competent design of sound elements should be considered on its own intrinsic merit; that we re-think the interdependence of the visual/sound construct; that authors be introduced to the various concepts underlying this sound/image synthesis, and acquaint themselves with examples of tools currently used by contemporary sound and visual artists.

My goals are:

First, to develop a theoretical formulation for the concepts of sound design and its relation to visual communication, beginning with the relationships acquired in early childhood - those of the alphabet effect (Logan, 1986) and various acquired schema, and to show how these can be conceptualized differently depending upon the sound/image model.[3] (Creativity is most often associated with 'seeing'; and when we come to the realization of an idea, we say: "Ah, I see!");

Second, to analyze sound design[4] and its multi-functional role in 'intermedia' and multimedia environments by focusing users' attention through reducing the distraction of conflicting stimuli.[5] One outlook is to argue that graphic user interfaces (GUI) could rely on sounds, so that a sound/vision correspondence could exemplify auditory strategies and symbolic sounds arbitrarily. For example, Earcons could be "formed" musical sounds created from any sound source;

Finally, to analyze some of the notable social implications of using new post-MIDI sound design tools. As the result of technological advancement and the accessibility of information and tools, society has become increasingly discriminating in its global outlook, and designers should re-think the use of sound - its development and its interdependence on other fields, whether academic and/or artistic in nature -

---

[3] Phonetically, anglophones "read" differently than the French or Spanish, even though the grapheme is the same.

[4] Sound never functions in a void: it travels through the air and is picked up by the listener's ears. Also, sound design never functions in a social and cultural vacuum. The perception of sound can be changed in subtle ways by the optical array of the visual environment. The flourishing of new audio technologies in the last few decades has had an impact on all forms of communication systems. Sound/music and visual arts have always offered a *space* in which aspects of a world view could be offered into possible consequences, as well as imagined into alternative possibilities. There are other new forms of space - web spaces, all of the internet spaces, CD-ROMs, and of course there is the entire area of virtual reality.

[5] Baddeley (1992) and Chandler & Sweller (1991) put forward the notion of 'Limited Capacity', i.e., that humans are limited in the amount of information they can process through each channel at one time.

and how the various innovative applications can accommodate and facilitate the design of sound within the different multimedia communication systems. The range of new 'techno-soft-tools'[6] that are available in this post-MIDI[7] era of sound-design tools has altered the design and development of multimedia environments, and a comprehensive analysis of the diverse social implications needs to be conducted. The use of 'sample' technology and object oriented programmed software has created a sort of cultural artifact. We[8] have conducted a series of interviews with academics, artists, technologists (Appendix B) that use the various current software available and makes reference to established and emerging technologies.

Moreover, as support for these goals, and a salient part of this dissertation, I have included a unit of instruction (Appendix A) that will illustrate an integrative contemporary approach to sound design within dynamic visual arts; also included is a DVD (Appendix D) with the various components for the instructional unit, as well as examples of the author's current work on the development of educational software, and some earlier work in sound design.

[6] This refers to the new variety of software that is replacing traditional sound equipment, such as synthesizers, recording stations (tape or other), and even traditional instruments.
[7] Musical Instrument Digital Interface: this is still being used today, but the new wave of musicians use software based instruments and USB/Firewire connectivity.
[8] Interviews were carried out largely by my research assistant throughout the 2003-04 academic year.

Design of Work

The dissertation begins by developing, through a literature review, a brief and focused chronological look at early encounters with the audio-visual model. From our first sound/visuals experience (sounds and their corresponding written symbols) to a more sophisticated use of its relationship, it is argued that early visual biases have shaped our perception and that the shift to the new digital technologies has changed our expectations.

In the second part, the various sound design tools, along with their popularization, wide accessibility and user-friendliness, will be looked at from a broad perspective. The digitization of these various tools has allowed for a wide user base, but has it also pressured users into a "Software Culture" directed by technology? The current use of discrete sound bytes (samples) as "objects" or "patterns" that the new wave of sound designers are applying to create complete compositions is a prime example. Through a few chosen interviews, a look will be taken at some of the current local practitioners and their work on sound design, and how some of these works seem to have initiated the use of sound elements as discrete objects (Chion, et al). The concepts of emancipation through new technologies and digital collective confinement are various aspects of our recent social environment.

Finally, I will acknowledge the current trends and pragmatic uses of various sound design notions that draw from music, communication studies, human-computer

interaction, visual arts, and instructional design. I will refer to contemporary artists, technophiles and educators[9]. The purpose will be to describe how these artists integrate the various concepts and examples of the sound design model. This should go a long way in helping other designers develop an up-to-date methodology for the use of sound within multimedia environments. I will also give examples of fieldwork, drawn from a design and development project in which the writer is currently participating as a media producer and creative director: *Alphie's Alley* and *The Animated Alphabet*[10]. The aim will be to address essential questions related to the significance and consequence of sound design within multimedia-based instructional designed projects.

To summarize:

- Chapter I: contains the literature review and will serve as liaison to many of the various topics included in this thesis.
- Chapters II & III: discusses the more social, perceptual and sequential aspects, from our early encounters with symbols to the more extensive bi-modal discourse.
- Chapter IV: a theoretical and pragmatic look at sound design and its use within pedagogical multimedia environments, and how information processing and dual-channel assumption is key to learning.
- Chapter V: will look at the 'current state of the art', some of the object oriented tools now in use and views of current sound designers and the role played by *drag & drop* technology within our communication systems.

[9] We have interviewed sound design artists and professionals and selected interviews can be found in Appendix B (list of all interviewees is in the footnote of the Appendix).
[10] See Appendix C - DVD,

Having taught and worked in the field of multimedia design for over a decade, I observed that the various tools (software, CPU, object-oriented programs, etc.) have evolved quite significantly over the past decade or so; media are active metaphors that extend the human senses. It is the desire of this author to put forth a document that is balanced and engaged in elucidating the concepts of the sound/image model. It is hoped that the resulting theoretical framework will provide useful information about the applicability and impact of selective sound design and the various methods of treatment, as well as show many of the principles of effective sound models and their application.

The latest developments in recording, storage and sharing influence multimedia production at various levels, and sound design is becoming more central to multimedia environments. It is not that the contiguous relationship between image and sound design in multimedia is a novel idea, but the use of sound design prior to multimedia was within the realm of recording, performing and communicating - music existed as a form of participation "for doing": musicians, sound artists, and others shared the experience with an audience, dancers, or other performers (*musica practica* - Barthes, 1985). With the ascendancy of technology, sound bites are something you possess, collect and replay.

In Appendix A, I will provide an example of guidelines for the formulation and evaluation of effective sound design intended for use within multimedia environments, guided by informative instructional system design tools and design patterns.[11]

[11] 'Design Patterns' or 'Re-useable Objects' comes from the computer science field and is based on 'Object Oriented Programming,' where software designers have developed discrete objects that other programmers could reuse in software design.

I anticipate that practitioners, teachers and students of multimedia, sound design (film & video), instructional design and software authoring will be responsive to the main issues, current concepts and terminology involved in the use of sound design in multimedia communication systems. It is also hoped that this study might serve to clarify some of the questions involved in researching the use of sound design within multimedia applications, such as how sonically-enhanced graphical human-computer interfaces can benefit by addressing some of the advantages of adding or even being initiated by sound to complement visual interfaces, and that it will encourage further research among new media artists, instructors and researchers alike. It is not the intention of this dissertation to address technical considerations such as delivery platforms, programming languages, compositional techniques, theory or production; moreover, it will not serve as empirical evidence for the effectiveness of sound design within multimedia environments. There are numerous reference sources on these subjects, and the bibliography contains a rather comprehensive list of such works.

Origin of Project

The idea for the dissertation is the combined result of my research in multimedia and instructional design (MA thesis), lecturer on film music and multimedia, my pragmatic experience as composer for film,[12] teacher/coordinator of the Digital Image & Sound program (Concordia University) and media producer/creative director for a major National Science Foundation research grant for the use of technology in enhancing

[12] as in-house composer for the French Drama section of the Canadian Broadcasting Corporation (CBC), with Director Jean-Yves Laforce.

the Success For All (SFA) reading program.[13] My research and experiences have revealed gaps in the process of how individuals, both academic and non-academic, deal with the sound schema, and will identify these gaps through an analysis of the effects of sound design in multimedia.

[13] I authored and developed the original computer-assisted tutoring software – The Reading CAT (Computer Assisted Tutoring), now called *Alphie's Alley*. There is a demo of this software on the included DVD (see Appendix C).

CHAPTER I:

REVIEW OF RELATED LITERATURE

Though most of the literature in this area has focused on the verbal element for the audio modality (Mayer, 2001; Pavio, 1986), save for the ground-breaking work on Earcons[14] by the Multimodal Interaction Group (MIG) led by Prof. Stephen Brewster,[15] I intend to direct my study to the less common area of non-verbal sound design and the new methods of production in use in our current digital industry.[16] As a teacher of multimedia,[17] I have noticed that many students are eager to incorporate a variety of sounds into their works, but fall short of finding the appropriate sources, elements, or techniques to integrate their sound ideas; furthermore, my colleagues (except for the musicians) generally tend to overlook the sound facet altogether, and those that try to venture into this area are mystified by the various resources available and feel confused regarding the possibilities and how to use them.

As presented in the works of Eisenstein (1947); Kracauer, 1960; Thomas, 1973; Evans, 1975; Prendergast, 1977; Weis & Belton, 1985; Gorbman, 1987; Zettl, 1990; Thomas, 1991, analysis of the relationship between music and the visual element in

---

[14] Earcons were first proposed by Meera Blattner in 1989. They are abstract musical tones that can be used in structured combinations to create auditory messages. Blattner defines earcons as *"non-verbal audio messages that are used in the computer/user interface to provide information to the user about some computer object, operation or interaction".*

[15] Web Site: http://www.dcs.gla.ac.uk/~stephen/

[16] I will include a variety of interviews with contemporary sound and visual artists, with comments on the use of the various technological tools available to them.

[17] Advanced Design, Development and Production of Computer-Based Multimedia (ETEC 568/668): a graduate course in the Educational Technology department at Concordia University (Montreal, Q.C.)

motion pictures has proven capricious at best, and moreover a search for related literature in the area of non-verbal sound in multimedia has shown a forbidding scarcity of empirical research. Therefore, I will outline a variety of audio-visual encounters, from childhood experience to professional sound design. Some themes to be addressed in the review of literature are:

- Get the Picture - Some of our first encounters with the audio/visual model are from early childhood. As small children, we are taught early on to associate images with sounds - "ma-ma" with mother's face, and various familiar objects that have embedded associated sounds (telephone ring, tea kettle, etc.). We then identify sound with the written word: as we learn to read, we decipher symbols and their associative sounds. (There are many models of cognitive strategies used in learning to read.[18])

- Auditory Information Design - Subsequently, I will look at some of the groundwork realized in Sound Design, giving a general overview[19] of the origins of sound design and its relation to visuals, most notably film music theory and its convergence with the multimedia environment; following that will be a critique of the various aspects of the more postmodern sound

[18] It is not in the scope of this study to explore this question comprehensively, but for the purpose of demonstration I will illustrate a sample of a linguistic sign model.

[19] I say "general overview" because much of this has been covered by various scholars, as well as some of my colleagues, and it would be redundant to re-invent the wheel. I will, however, give ample reference to those studies that I believe useful for the understanding of my text.

design literature, along with some aspects of its associated aesthetic domain.[20]

- Universal Genre – An inquiry into the area of the design of sound as it relates to visuals in multimedia has shown that this is quite an untried field, and that there are still very few scholars addressing this question. But with the advent of PCs and universal web technologies, we are quantifying and developing worldwide standards that direct our perspective, from basic semiology to sonification of Human Computer Interface (HCI).

Get the Picture

> The alphabet was the most significant of the boons conferred upon mankind by Phoenicia. It is generally considered the greatest invention ever made by man.
>
> P.Hitti[21]

In The Alphabet Effect, R.K.Logan (1986) posits that the alphabet is the gateway to learning and knowledge, and that which children first learn after speech. As children learn to decipher symbols and their associated sounds, and as phonetic perception is ascertained early on in our childhood, we become accustomed to hearing certain sounds associated to images, objects, and so on within our particular language;

[20] Once again, it is not within the scope of this dissertation to undertake the vast subject of perception and aesthetics, but I will refer to the relevant literature as necessary.
[21] Hitti P. (1961) The Near East in History, Princeton N.J., p.102

and as we develop in a particular culture, we lose plasticity in our hearing and identification of non-native phonemes. Iverson, Kuhl et al (2002) state that "exposure to speech during childhood alters the neural organization such that individuals, born capable of learning any language, develop perceptual and cognitive processes that are specialized for their native language. The changes in neural organization are evident when an individual tries to learn a second language as an adult."

The various approaches to reading instruction have been developed, from Basal reader to Direct Instruction. Though all these approaches follow their own philosophy, the ultimate goal is to have children identify and give meaning to various symbols, and how of course these symbols are associated with a sound. Nuckolls, J.B. (1999) looks at sound-symbolic processes and how linguistic significance is not natural but learned. I will therefore look at reading (phonetic identification of symbols) in the more conceptual sound/symbol relationship to phonemic awareness, as it is not in the scope of this study to analyze or criticize these various reading models. As Perfetti, Beck, Bell and Hughes (1987) put it, we might think of phonemic awareness as "a constellation of abilities that center around the child's emerging understanding of the segmental structure of the spoken language." Coding and Decoding as defined by Logan (1986) is "converting auditory signals or sounds into visual signs." For Gough & Tummer (1986), reading is many different things: they argue that decoding and general language comprehension are separable components of reading comprehension; skills are cognitive processes that are executed automatically, without the reader's conscious attention or choice. In contrast, strategies are deliberately chosen and applied to reading situations (Paris,Lipson &

Wilson 1983). A skill is a cognitive process and executed without conscious attention; a strategy is more deliberate in nature, as in puzzles, etc. Skills, strategies, and knowledge interact differently in decoding and comprehension: we use strategies and skills to recognize words and various symbols automatically. Learning to decode symbols, including words, needs established concepts, such as direction of decoding and knowledge of patterns.

In my current work as Creative Director for the CSLP and developer of a multimedia tool, *Alphie's Alley* (a computer-based learning environment designed as a performance support system for literacy tutors), it has become evident that the close association of audio and visuals in a mediated environment, with appropriate Graphic User Interface (GUI), is key to enhancing literacy skills in children. *Alphie's Alley* was created to increase the quality of implementation of tutoring in the Success For All[22] (SFA) (Slavin,& Madden,1999) curriculum by assisting the tutor to assess, plan, and carry out tutoring sessions that are customized for each student. Each reading goal (e.g., phonemic awareness) is described and objectives (e.g., auditory blending) are identified and explained. Eighteen different activities provide multimedia content, practice, and assessments for students focused on phonemic awareness, phonics, fluency, and comprehension. Using this multimedia environment has significantly increased students' ability to learn reading skills. *Alphie's Alley* was evaluated in 2003-2004 with four SFA schools, two assigned to use the computer and two assigned to use the traditional SFA tutoring. With a small group of students (N=27) in a matched experimental design,

[22] Success for All is a comprehensive reform program for elementary schools in the US. The Success for All elementary reading program is currently being implemented in more than 1300 schools in 46 states.

researchers were able to show a significant effect of the treatment on the Word Attack subtest of the Woodcock Reading Mastery Test (ES=+0.47). Currently, a large-scale study (N=1,000) is being conducted to determine further the effectiveness of this tool. Feedback from teachers and students is extremely positive.

It is my belief that the combination of visuals and sound can be designed within a multimedia environment and that such a venue can significantly increase and develop comprehension and meaning by organizing and determining a relevant point of view or focus for the user, whether it be educational (as the *Alphie's Alley* study above), cultural or purely entertainment.

Auditory Information Design

There is an historical lack of reference to non-musical sound in film criticism, while conjecture on and theory of music and visuals in film have been part of the literature for quite some time (Burch, 1973; Eisenstein, 1947; Kracauer, 1960; Prendergast, 1977; Thomas, 1973; Zettl, 1990). This may extend from the perception that music is seen as having separate characteristics and for some is not altogether emotive. Peter Kivy (1993) for example, a leading thinker in musical aesthetics, in his book The Fine Art of Repetition : Essays in the Philosophy of Music, argues that music is evidently meaningless, devoid of semantic substance. He declares that music is a "fine art of repetition" not meant to express ideas or concepts, and what emotions it can convey are limited to what he calls "the garden-variety emotions, joy, sorrow, fury, bliss." He

expresses the view that music is more an expression of the composer's emotional state of mind and not a manifestation of absolute demonstrative musical effectiveness. But of late, researchers have addressed both the more empirical studies dealing with the interaction of non-verbal sounds (Lipscomp & Kendall, 1994), the area of screen-based graphic user interfaces (GUIs), and the relationship between visual cues (icons) and sound (Brewster, 1994; Lumsden 2002, Crease 1998).

The human-computer interface (HCI) is the most prevalent area for maximizing the use of audio-icons. The Glasgow Interactive Systems Group (GIST) conducts research into the area of human-computer interaction and sonification. In their proposal for the Audio Toolkit Project, Brewster & Gray[23] suggest that "...there is a small but growing body of evidence which indicates that the addition of non-speech sound to human computer interfaces can significantly improve performance and increase usability."[24] The summary goes on to say:

[23] http://www.dcs.gla.ac.uk/research/audio_toolkit/index.shtml

[24] In the literacy software *Alphie's Alley* (mentioned above), we make use of sonic enhancements to balance 'on screen information'; we are also working on adapting our software to students with special needs, and these types of enhancements could greatly benefit those students. As with the Lipscomp & Kendall (1994) study on the relationship between musical and visual components in film, the graphical human-computer interfaces (HCI) is considered to be a system, and the outcome is not seen as individual modalities (audio vs visual) but as a cross-modal interaction. In our audio/visual tool, The *Animated Alphabet* (See DVD in Appendix D for examples), for example, we use audio and animation correlations to teach and reinforce sound/symbol relationships. There are animations for 58 different graphemes that comprise all of the phonemes used in the English language. Each animation is between 30 seconds and one minute in length.

> Sonically-enhanced graphical human-computer interfaces allow more natural communication between computer and user. Such multimedia interfaces allow users to employ two senses to solve a problem, rather than using vision to solve all problems. This leads to reductions in the time taken to complete tasks. However, this area is in its infancy and there is little systematic research to demonstrate the best ways of combining graphics and sound. This means sounds are often added in *ad hoc* and ineffective ways by designers.

Stephen Barrass (1997) provides a thorough examination of non-speech audio in multimedia, of past and more recent approaches, and of existing guidelines for the effective design and use of sounds in computer-based programs. He cites the need for "a linkage between the characteristics of the data and the auditory perception of the data" so that auditory information can be designed more appropriately. His review of the semiotic approach to understanding audio in multimedia includes three principles of analysis:

- **Semantic**: a sign composed of a *signifier* (the vehicle of representation) and a *signified* (the concept being represented). *Auditory icons* operate within this framework, developed first by William Gaver in 1989 and outlined further by Stephen Brewster.
- **Pragmatic**: concerned with the signifier or form of the sign, delineating psychoacoustic principles such as timbre, intensity, pitch and rhythm as that which gives the sound-sign it's meaning.

- **Syntactic**: involving the organization of sound signs into larger groups with more complex messages, either *paradigmatically* or *syntagmatically*[25].

One interesting aspect of the application of these principles of analysis is that one is guided towards an expressionist alignment - that is, the author can blend visual cues with audio elements to "express" meaning or impetus, placing emphasis on the organization between the aural and visual strata within an associational (i.e. referential) context of multimedia meaning (Semantic). For instance, in *The Animated Alphabet*,[26] the animation introducing the short /e/ sound features an elephant pushing a rock with an "e" on it up a hill, making an /e/ sound with each push. At the top, the rock rolls down, and the exhausted elephant says "ehhhh" in frustration. The pairing of the memorable images, the letter sound, and the letter shape gives students many mental pathways for linking the letter with its sound. On the other hand, Audio-Icons or *Earcons* are arbitrarily constructed auditory messages that do not possess an intuitive connection to the object, action or activity they represent; much like language, they consist of sound-images (words) that do not meaningfully relate to the concepts they signify (as put forth by linguist Ferdinand deSaussure). These signs are not *motivated*, and are thus more easily classified into *motifs*, "short, rhythmic sequences of pitches that can be combined in

---

[25] **paradigmatic**, or substitutional, similarity: two words that are paradigmatically similar may be substituted for one another in a particular context. For example, in the context *I read the book*, the word *book* can be replaced by *magazine* with no violation of the semantic form of the sentence, and therefore the two words can be said to be paradigmatically similar;

syntagmatic similarity: two words that are syntagmatically similar significantly occur together in the text. For instance, *cut* and *knife* are syntagmatically similar, since they typically co-occur within the same context.

[26] *The Animated Alphabet* is a series of 3D animations developed by the CSLP that support phonemic awarness (See included DVD for examples.

different ways," like Barrass's sign categories. There are many examples which illustrate the sound designer's use of semiotic approach to the musical construct as a means of enhancing the effect of enduring ideas in the visual image: the recurring sounds associated with a "wrong or good" answers when clicking on a "submit" button, the musical 'ID' theme when turning on one's computer, etc. Walker et al (2003), in their ACM proceedings, "The audio abacus: representing numerical values with non-speech sound for the visually impaired," gives a strong example of the use of non-verbal sound utilizing the above principles (Syntactic), by which sound could be used at a checkout counter of a store: when items are being scanned, they could be associated with a "sonic" code that would allow the visually impaired to hear the exact price of the item.

Universal Genre

With the advent of personal computers (PCs),[27] the World Wide Web and wireless technology, the social status of various works of art has been transported to a level of being independent from the conditions of material and physical locality. But sound functions within a physical space: it is propelled by waves in the environment, the sound signal is emitted by an instrument (the voice is included as an instrument), is then driven through a medium (air or water) and is picked up by the audience ear (receiver). Sound can have many instances as it travels unobtrusively and can be used as a communication device, or just purely aesthetic performance, or both. Pragmatic device-to-human communication is not limited to what we see on a screen. "Sound is a reliable

[27] Not to be confused with the common use of PC to indicate Wintel CPUs. I will use PC for Personal Computers and Mac or Wintel to identify those models.

way to alert someone about an event" (Gerasimov & Bender, 2000). Likewise, art never functions in a social void: the awareness of visuals and sound can be altered in many complex configurations by the impression the performer will have on the audience, sociological factors, fashion, the listener's associations of that day, and a multitude of other factors studied in social science, anthropology, folklore, ethnomusicology, etc.

Visual and literary depictions create fictional localities. Does sound create fictional environments? One should draw on the fact that sound/music in and of itself can easily entice our imaginations, and without the aid of visual support can entertain, fascinate and create imagery. The aesthetic assumption that "art" is independent of "the real," or is to be treated as separate from everyday experience or uncharacteristic of pragmatism, is becoming more and more incongruous in our media-driven ethos. As Dewey proposes (1934),

> ...the forces that create the gulf between producer and consumer in modern society operate to create also a chasm between ordinary and esthetic experience. Finally we have as a record of this chasm, accepted as if it were normal, the philosophies of art that locate it in a region inhabited by no other creature and that emphasize beyond all reason the merely contemplative character of the esthetic.

Also linked to this aesthetics of sound design in multimedia, either as information design or just motivational assistance, are some of the theoretical models utilized in the earlier literature. In his Theory of Film Practice, Noël Burch (1973) writes:

> The fundamental dialectic in film, the one that at least empirically seems to contain every other, is that contrasting and joining sound with image. The necessary interrelationship of sound and image today appears to be definitely established fact, as even the most doubting critic must concede once he has examined the history of film.[28]

This statement is even more appropriate today - in our post-analog era, with the proliferation of computer-assisted productions, in which the new 'media artist' (direct agnates of the traditional film maker) is wearing more and more diverse hats: director, producer, editor and sound designer. Sound design aesthetics in the digital realm is a combination of the pragmatic (e.g., earcons and navigation aids) and the artful (e.g., sound-effects). In their book: The Art of Seeing: An Interpretation of the Aesthetic Encounter,[29] Csikszentmihalyi and Robinson are able to identify four major dimensions of the aesthetic experience: intellect, communication, perception and emotion (ICPE). According to the authors, the intellectual dimension of aesthetic experience consists of all aspects of the viewer's attempts to use knowledge in order to find meaning in the work of art. The communicative dimension of the aesthetic experience is characterized by the two modes by which an exchange is established via the work of art:

- communication across different time periods and cultures and
- communication between individuals (the artist and the viewer).

Perception is the third dimension of the aesthetic experience. It relates to the physical and formal qualities of a work of art and to the role that our senses play in

---

28 Burch, N. (1973). Theory of film practice / translated [from the French] by Helen R. Lane. London: Secker & Warburg.

[29] Csikszentmihalyi, M. and Robinson, R.E. (1990). The Art of Seeing: An Interpretation of the Aesthetic Encounter: Malibu, Calif. : J.P. Getty Museum ; Los Angeles, Calif. : Getty Center for Education in the Arts,

defining it. Finally, the emotional dimension of aesthetic experience consists not only of the positive emotions engendered by the work of art, such as joy and inspiration, but also the negative ones, such as fear and frustration.

We can easily substitute these dimensions of the aesthetic experience as paradigms for the aesthetic experience in sound design. In Notes sur le Cinématographe,[30] Bresson writes "A sound always evokes an image..." This statement concurs that music and sound could be designed as much as light and shadow compositions (Gessner, R. 1968.)[31] There has been much conjecture over the interplay of sound (more specifically music) and visuals in film (from Eisenstein in 1947[32] to Thomas in 1991.)[33] The literature within this domain, though, has shown very little research dealing with this interaction, or what I call the 'assemblage' of image and sound. Though psychologists have investigated quite thoroughly the perceptual relationship between visual and auditory sensory modalities (Radeau & Bertelson,[34] Staal & Donderi[35]), most of these studies deal with specific tonal or optic effects[36] and do not relate to the overall issues of perception in the aesthetic domain. Furthermore, given the availability and immediacy of various works of "art", in all its forms and on the net, art (both visual and auditive) has become more 'social'. Adorno[37] writes that:

---

[30] Bresson, R. (1977). *Notes sur le cinematographe*. New York: Urizen Books,.
[31] Gessner, R. (1968). *The moving image; a guide to cinematic literacy.* New York: Dutton.
[32] Eisenstein, S. (1947). *The film sense*. New York: Harcourt Brace Jovanovich, Publishers.
[33] Thomas, T. (1991). *Film score: The art & craft of movie music.* Burbank, CA: Riverwood Press.
[34] Radeau, M. & Bertelson, P. (1974). *The after-effects of ventriloquism.* Quarterly Journal of Experimental Psychology, 26, 63-71.
[35] Staal, H.E. & Donderi, D.C. (1983). *The effect of sound on visual apparent movement.* American Journal of Psychology, 96, 95-105.
[36] Radeau deals with tones, light and spatial seperation, or increasing loudness for more dramatic impact, etc.
[37] Adorno, W.T.(1997) *Aesthetic Theory,* University of Minnesota Press, Minneapolis, MN

> Under the pressure of an intensifying nominalism, the ever-present yet latent social character of art was made increasingly manifest.... Art, however, is social not only because of its mode of production, in which the dialectic of the forces and relations of production is concentrated, nor simply by its thematic material. Much more importantly, art becomes social by its opposition to society, and it occupies this position only as autonomous art.

Though I will not argue for or against Adorno's views expressed in his complex work - this could be a thesis in itself - I find that, from his point of view of the rejection of the dated definition of aesthetics as the theory of the beautiful, arts looks for a sanctuary in its own contradiction, and nothing concerning art remains self-evident. His emphasis on the connection between aesthetic theory and many other aspects of social theory is obviously significant in our contemporary digital media sphere: our creation of sound designs using Hip Hop or Bach loops and samples does not overwhelm cultural meaning or defeat a higher artistic purpose, but rather provides the culture industry with process.

Alternatively, aesthetic vision could be understood as the sense that gives us the best possible glimpse into the immediate cultural future as we look out on the sociological landscape that we are about to traverse, thus providing a sense of perspicacity about what may well come next. Broadly constructed, "artistic value" identifies a disposition or mode of being in which the artist comes into his full powers not through reason alone but through a process or activity in which reason and the senses, freedom and necessity, are in counterpoint. Visuals and sound are used by many societies as a means of establishing and mapping one's place in time and space. Simon Frith (1987) contends that we westerners employ popular music for this purpose and stresses

the importance of developing an aesthetic of popular music[38]. My view is that this is more critical today than ever, given the scientific, technical and cultural transformations affecting the structure in all domains of society, along with the growing social demand expressed through new and ever-changing resources that we aggregate under the term "technology." This so-called "process of innovation" has affected all areas of music activity, whether personal or at the 'industrial level. Actually, even the non-professional (amateur) musician is now capable of competing at the "industry" stratum, thanks to the development of new technologies (fast computers, WWW, CD burners, MP3's, and so forth). But this proliferation of technologies is not without social, political and economical problems,[39] thus implying that with technology comes responsibility. Responsibility in the area of human computer interface (HCI) is crucial, as metaphorical conventions are being developed worldwide on a daily basis.

Guidelines for an Auditory Interface Design

(Writer's note: In order to be more concise, I have put the instructional design guidelines into Appendix A rather than into the body of the text. The following section of this literature review is nonetheless essential to the study.)

> Design is that area of human experience, skill and knowledge
> which is concerned with man's ability to mould his environment
> to suit his material and spiritual needs.
>
> Archer (1973)

[38] Simon Frith "Towards an Aesthetic of Popular Music" in Music and Society: The Politics of Composition, Performance and Reception, ed. R.Lepperd & S. McClary, Cambridge University Press: 1987, pp.133-49.

[39] As seen with the latest court ruling on Napster. I will return to this later in my paper.

Instructional design has evolved both in philosophy and perspective. From self-instruction strategies (Keirns, J. 1999), multidimensional scaling (Schiffman et al. 1981), multimedia learning (Mayer, R.E. 2001, El Saddik, A. 2001) to modular architectures (Cholmsky, P. 2001, Wiley, D. 2002), instructional design has marked all facets of academia and industry. Learning theory has expanded dramatically in the past few years, from the 50's behaviorist standpoint to the more constructivist perspective (Behaviorist vs. Cognitivist vs. Constructivist).

The Constructivist approach supports the fact that learners create their own interpretation of the world. The Constructivist differs from the Cognitivist and Behaviorist approach in that the latter do not believe that the mind can be charted by instruction, but that learners need to adapt instruction to reflect their own experience. Researchers such as R. Spiro, P. Feltovitch & R. Coulson (1988) deal with transfer of knowledge and skills beyond the original learning circumstances - naming this learning of complex structures " Cognitive Flexibility Theory." Spiro & Jehng (1990, p. 165) state:

> By cognitive flexibility, we mean the ability to spontaneously restructure one's knowledge, in many ways, in adaptive response to radically changing situational demands...This is a function of both the way knowledge is represented (e.g., along multiple rather single conceptual dimensions) and the processes that operate on those mental representations (e.g., processes of schema assembly rather than intact schema retrieval).

Their principles are:

1. Learning activities must provide multiple representation of content;
2. Instructional materials should avoid oversimplifying the content domain and support context dependent knowledge;
3. Instruction should be case-based and emphasize knowledge construction, not transmission of information;
4. Knowledge sources should be highly interconnected rather than compartmentalized.

In Appendix A, I will put forth a concise instructional design model which integrates instructional exercises using sound design objects that need modification. It will also be expectantly recognized how one can, by closing the gap between the production goals, curriculum goals and design objectives, create a multimedia unit using the bi-modality concept. Reigeluth (1996; 1997) argues that, due to the transference to the Information Age, instruction needs to be adapted to a more learner-centered model. Gros *et al* (1997: 51-52) state that new instructional design models need to facilitate multimedia authoring. The designers need to recognize more adaptable design processes, including rapid prototyping, and create a transparent link between skill and knowledge acquisition. I am not suggesting one can design a sound environment in the same way that one chooses a brand name, but with sufficient consideration and planning, we can logically connect the dots using basic instructional devices, such as forecasting the desired effect and creating an efficient process to accommodate the various concepts.

CHAPTER II:

SOCIAL ORIENTATION OF THE STUDY

Foreword

I will elaborate on the overall direction of this study, linking the field of possible hypothetical and methodological approaches to the topic, and putting forward some of the basic conjectures of the dissertation: First, I will look at the initial and essential means of people's association with the audio-vision[40] model, i.e., our early encounters with the relationship between image and sound as the fundamental cornerstone of information organization, leading to mnemonic use of letter-sound associations or phonics. Second, I will highlight the post-analog tendencies of the multimedia field as well as the contemporary ideas of new media artists and educators, where interaction between sound and image is becoming more and more indigenous and highly influenced by information technology. Moreover, the use of multimedia has brought forth expanded applications and treatments of the sound/vision paradigm within cognitive theory assumptions. Mayer[41] proposes that three of these assumptions are key in multimedia environments: Dual-Channel, Limited Capacity and Active Processing. Third, I will explore the perspective that with the advent of digital sampling technology and *object oriented* music software[42], any "media artist" or other with any such interest can create concrete formal compositions, limited only by their imagination. "Post-analog"

[40] Audio-vision is a term first used by Michel Chion from his book L'Audio-Vision (1994)
[41] Multimedia Learning by R.E. Mayer, 2001
[42] Software programs such as Band-in-the Box, GarageBand, JamXmatrix, Orchestrator and the impressive array of fully-featured sequencers that allow an untrained person in the art of composition to assemble meaningful but sometimes ad hoc sound designs.

technology has provided an opportunity for media producers, educators and artists to download ready-made packets of musical composition and use sound austerely, allowing them to bring emphasis to their artistic conceptions by using this consequential pro tem media,[43] consequently augmenting the relevance of this "mix-media" analysis. As a composer, multimedia professor and new media artist, I have long observed the lack of convincing and compelling research into this unique area. Nevertheless, this "mix-media" context can, should and does provide a venue for various artists to apply a consistent environment to their work and needs to be addressed, as some of the interviews (Appendix B) with contemporary artists will attest. Fourth, I will determine that by illustrating the complex route through which technology infiltrates the arts community,[44] the 'audio-vision' aspect is filtered by our concepts of sound and image, it is my opinion that the causal inquiry of the aesthetics of sound design related to images has not caught up with this "techno" or post-analog development. Though it has been integral to the production of new artistic forms,[45] the aesthetic understanding of these newer healthier relations need to be re-evaluated in a purely autonomous method and not compared with traditional analytical formats. To underline the above, I will include a series of interviews with new media artists.

[43] The idea of having to integrate an element or craft that one is only moderately adept in is a notion that is becoming increasingly accepted in our contemporary technocratic art setting.
[44] The new media artist working with new technologies is drenched with information and streams of data; he can alter it, download it, re-assign it, email it, broadcast it, pitch change it,etc.
[45] The whole 'net' culture phenomenon is a prime example.

Resources for Creation

> We gestate in Sound, and are born into Sight...
> Walter Murch, 1994[46]

From birth our senses are inundated with information, and the ability to come to terms with it all is astonishing; most of us do however wittingly concentrate on honing and improving our sight and sound capabilities and rely on these for our development as "sapiens." Walter Murch writes[47]:

> So we all begin as hearing beings — our four and a half month baptism in a sea of sound must have a profound and everlasting effect on us — but from the moment of birth onward, hearing seems to recede into the background of our consciousness and function more as an accompaniment to what we see. Why this should be, rather than the reverse, is a mystery: why does not the first of our senses to be activated retain a lifelong dominance of all the others?
>
> Something of this same situation marks the relationship between what we see and hear in the cinema. Film sound is rarely appreciated for itself alone but functions largely as an enhancement of the visuals: by means of some mysterious perceptual alchemy, whatever virtues sound brings to film are largely perceived and appreciated by the audience in visual terms. The better the sound, the better the image.

One most important tool for the development of our identity is the use of codes as medium to organize and communicate; I am referring to the written symbols used for these conventions. Various cultures have developed an assortment of visual representation to articulate meaning, but we in the west have used the phonetic

---

[46] Forward to Audio-Vision by M. Chion (1994).
[47] From Stretching Sound to Help the Mind by W. Murch (October 1, 2000); See: http://www.filmsound.org/murch/stretching.htm

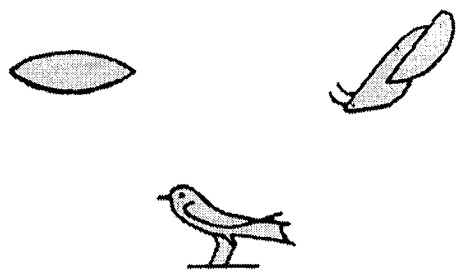
alphabet.[48] Previously, there was no link between the sounds of words people *spoke* and the signs they used to represent objects *visually* (Moorhouse, 1953).[49] When it was realized that a picture could represent both the object itself *and* the sound of the spoken word for the object, people began to use pictograms to represent the sound, even when the object was not involved (e.g., "mandrake" as being represented by the coupling of two separate symbols for "man" and "drake."). Future writing systems developed, pictograms became ideograms, and these were the evolution of purely representational pictures. Ideograms were now signs that represent ideas and abstractions and are able to communicate thoughts. *Phonograms* were used to signify the sound value of *whole words*. Examples appear in Egyptian and Chinese writings, with Chinese phonograms having only monosyllabic value and Egyptian phonograms having both mono and polysyllabic value, as Egyptian speech also consisted of polysyllabic words. Conversely, the Aztec language sometimes employed phonograms whose sound value came from only the *opening parts* of words (as with city names).

1. *Syllabaries* were used by the Sumerians where each phonetic sign represented a syllable of a word, so that a monosyllabic word would be written with one sign, a disyllabic word with two signs and so on. (Figure 1)

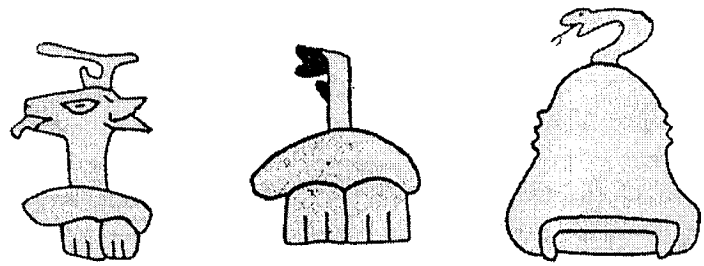
---

[48] In his book, The Alphabet Effect, R.K. Logan states, "the very first alphabet was invented over 3,500 years ago in the Near East by the Canaanites, a Semitic people, and contained only twenty-two letters. It became the model for hundreds of other phonetic alphabets, including English, French, Latin, Greek, Latvian, Russian, Romanian, Arabic, Turkish, Persian, Sanskrit, Korean, Hebrew, and Swahili. Each of these phonetic alphabets is descended from the first Canaanite alphabet through a complicated process of borrowing and adaptation. The English alphabet is derived from the Roman alphabet, which in turn can be traced back through the Etruscan, Greek, and Phoenician alphabets."

[49] Moorhouse, A.C. (1953) *The Triumph of the Alphabet.* New York: Henry Schuman, (p17)



Egyptian Phonograms.



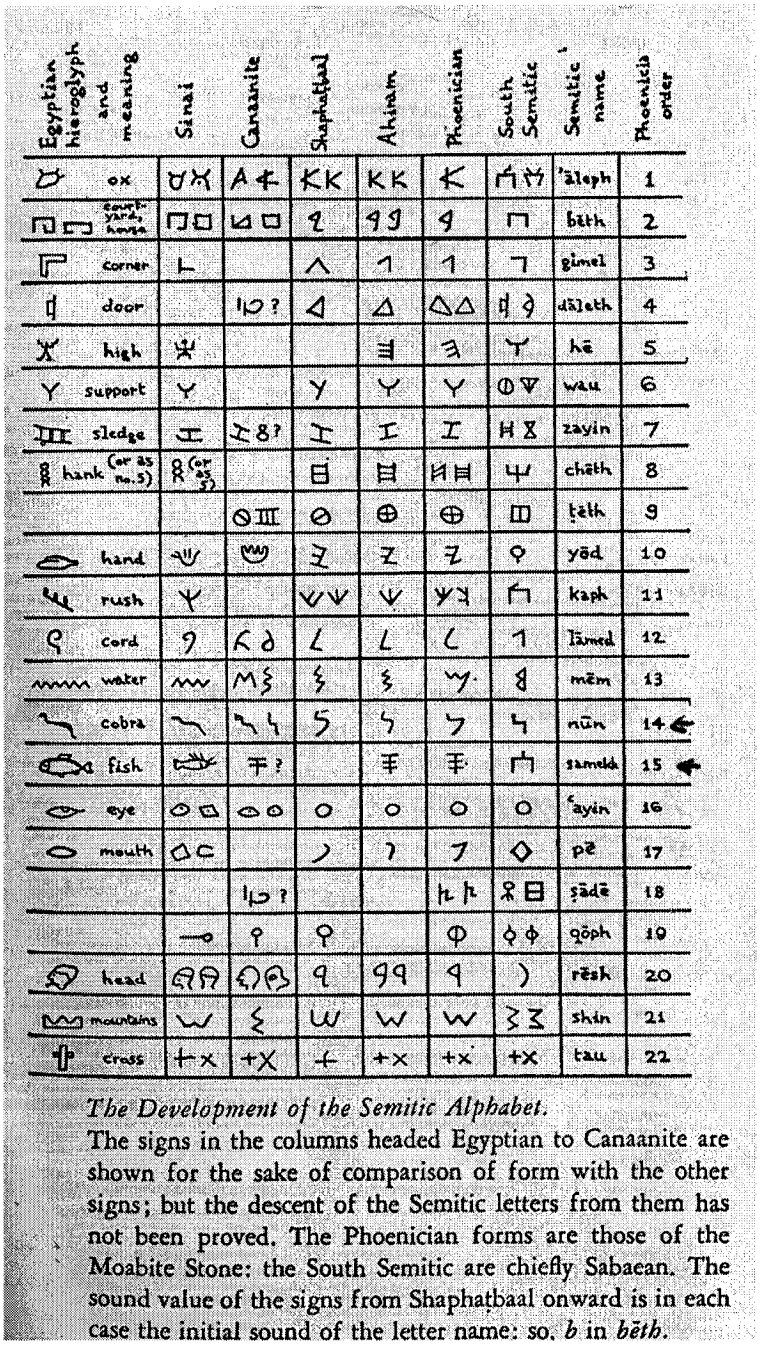
Aztec Phonograms of Town Names.

Figure 1 – Syllabaries or Phonograms[50]

2. *Alphabetic writing* consisted of a group of individual signs or letters that each stood, at first, for a single consonant and then eventually for single vowels as well.

    - The Semitic alphabet consisted of 22 consonant signs or letters that were named and arranged in regular order (Figure 2)

[50] Moorhouse, A.C. (1953); *The Triumph of the Alphabet.* New York: Henry Schuman,



The Development of the Semitic Alphabet.
The signs in the columns headed Egyptian to Canaanite are shown for the sake of comparison of form with the other signs; but the descent of the Semitic letters from them has not been proved. The Phoenician forms are those of the Moabite Stone: the South Semitic are chiefly Sabaean. The sound value of the signs from Shaphatbaal onward is in each case the initial sound of the letter name: so, *b* in *bēth*.

Figure 2 - Semitic Alphabet[51]

- Vowels were not considered essential to decoding the meanings of the word; they were not included as signs in the alphabet. Vowel sounds had to be considered instinctively during the reading of a text, leaving much of

[51] Ibid.

the decoding process up to interpretation. However, a desire to accurately read religious documents prompted more widespread use of vowel signs.

- The principle of **Acrophony**[52]: "The sound value of each letter is also the first sound in the name of the letter". These names have their own meaning (as opposed to English where our letters have been named with invented words.)

Were the letters of the Semitic alphabet conceived, then named, or were the pre-existing signs merely assigned the most suitable name from the Semitic language? If language is at the root of culture, do changes in language and communication shape the way we view our world and ourselves? Now that phonetic writing is a combination of sounds that no longer represent the whole word but uses letters to create a whole structure. Marshall McLuhan summarized this idea in a very concise way when he wrote: "By the meaningless sign linked to the meaningless sound we have built the shape and meaning of Western man"[53]. The twenty-six letters of our Roman alphabet are the first abstract sound/symbol genus that we strive to decipher and create the audio-visual link in order to decode meaning. This coherent and consistent perspective of codification has channeled our notions of knowledge, as Logan (2004) contends:

---

[52] Ibid; Moorhouse (p98).

[53] Marshall McLuhan, Gutenberg Galaxy (1962, p. 65)

> Not only has the alphabet performed admirably as a tool for literacy, it has also served as a model for classification. It has played an instrumental role in the development of the logical style of analysis that is characteristic of the Western way of thinking. Learning how to read and write with the alphabet has brought us more than literacy and a model for classification. It has provided us with a conceptual framework for analysis and has restructured our perception of reality.

As children, we first encounter the audio-visual world with the naming of pictures, and a generally accepted hypothesis is that this process has three major components: object recognition, semantic activation and lexical access (Ellis, Kay, & Franklin, 1992; Morton, 1985; Snodgrass & McCullough, 1986; Warren & Morton, 1982). Picture naming is semantically mediated, whereas word naming is not; Figure 3 shows three major processes underlying picture naming along with the basic processes proposed to be involved in the oral reading of words (Barry, Hirsh, et al., 2001).

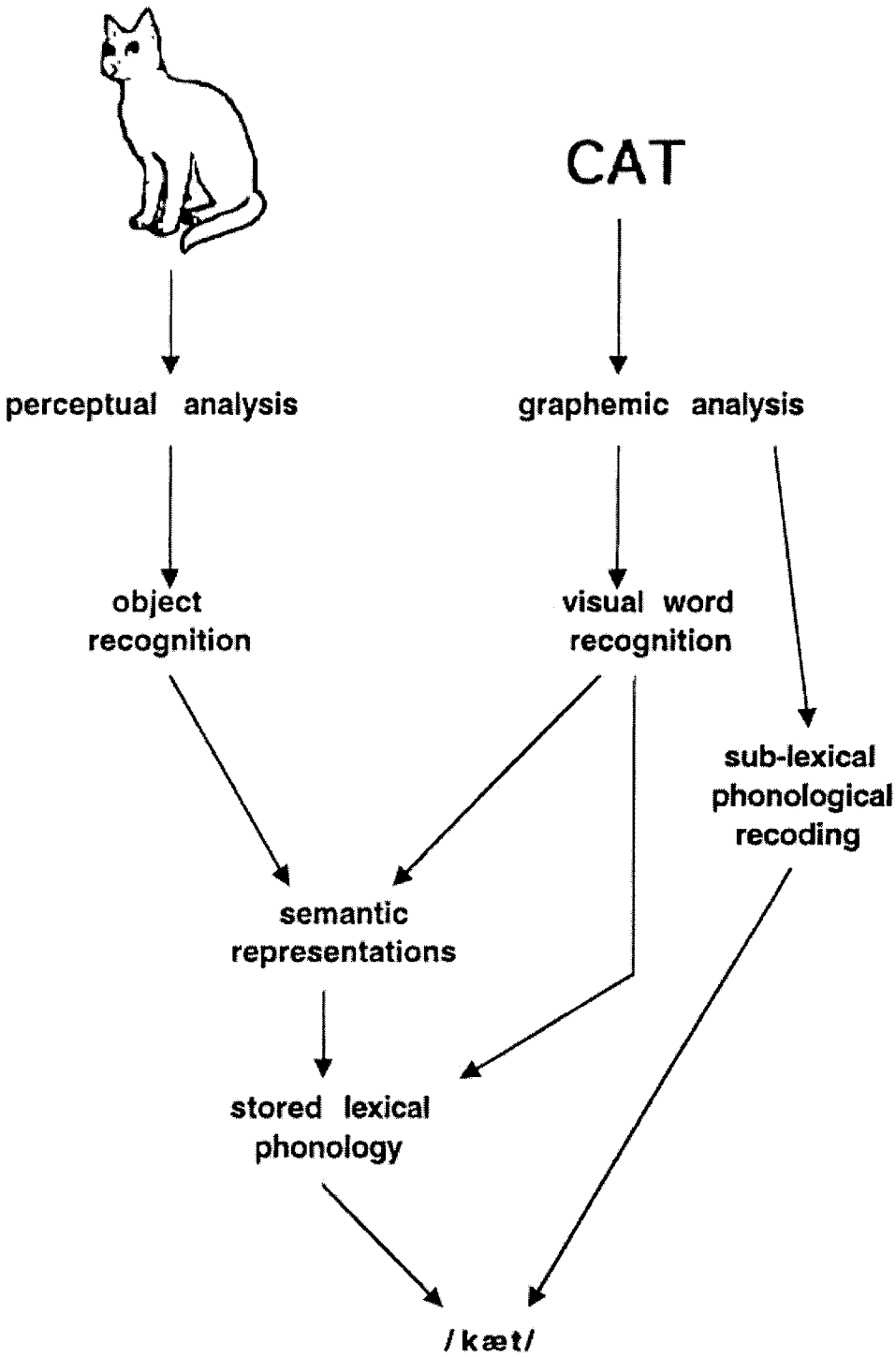


Figure 3 – Picture Naming Process

I will therefore look at reading (phonetic identification of symbols) in the more conceptual sound/symbol relationship to phonemic awareness, as it is not in the scope of this study to analyze or criticize these various reading models. As Perfetti, Beck, Bell and Hughes (1987) put it, we might think of phonemic awareness as "a constellation of abilities that center around the child's emerging understanding of the segmental structure of the spoken language." Coding and Decoding as defined by Logan (1986) is "converting auditory signals or sounds into visual signs." For Gough & Tummer (1986), reading is many different things: they argue that decoding and general language comprehension are separable components of reading comprehension; skills are cognitive processes that are executed automatically, without the reader's conscious attention or choice. In contrast, strategies are deliberately chosen and applied to reading situations (Paris,Lipson & Wilson 1983). A skill is a cognitive process and executed without conscious attention; a strategy is more deliberate in nature, as in puzzles, etc. Skills, strategies, and knowledge interact differently in decoding and comprehension: we use strategies and skills to recognize words and various symbols automatically. Learning to decode symbols, including words, needs established concepts, such as direction of decoding and knowledge of patterns. It is my belief that the combination of visuals and sound can be designed to build and organize meaning by establishing "feeling" though this meaning may not be the authentic[54] meaning.

Worthy of note is the fact that the phonetic alphabet has the ability to render fairly accurate pronunciation of unrelated languages to the Roman system just by using the phonetic code of the spoken word, so that if I wanted to say "thank you" in Japanese

[54] Authentic here refers to the author's intention.

though I could not read "ありがと" written in syllabic code. I could however read "Arigato," pronounce it phonetically and be understood by a Japanese native. Asian languages still use logographic coding to depict words using ideograms (symbolic) or pictograms. In our western culture we use the phonetic alphabet of meaningless symbols strung together to form meaningless assemblage from which we correlate significance. Furthermore, the writing process performs an essential function in the development of cultural identity and has performed a significant role in organizing the foundation of thought patterns. Also as we read, our mental processes call upon an audio cerebral word list, when we distinguish the various phonological parts of a word we perceive it in a dialectical mental vocabulary based on what we recognize using sounds embedded in our memory (the sound that particular letters make), and possibly a mixture of the shape of the equivalent letters. Also, Eddy & Glass (1981) found that hearing "high imagery" sentences produced better understanding than reading them.

A study by Frost, Repp & Katz (1988) contends that a language when written is directly and automatically mentally recoded into internal phonetic form, the implication here would be that as read a word, we most likely tend to search in a personal audio cerebral glossary based on our previous knowledge of what sounds the letters make, or possibly we "picture" a word mentally based on the familiarity of what we recognize as meaningful characters or graphemes. McKoon (1977) reports that research by Kintsch and Craik & Lockhart states that text is encoded and analyzed in several levels of processing: graphemic, phonemic, lexical, syntactic,[55] and semantic,[56] with each stage

leaving remnants in memory. The sound of a word is analyzed and reconstructed alphabetically. As mentioned earlier, my research and development work with the literacy tool *Alphie's Alley* shows that using pictograms or pictures associated with a letter has proven very effective in helping young learners memorize letter names and sounds. Hester & Francis (1995:85) states that: "Procedures for meaning-making," are learned during oral and social interactions in groups, and pictures are a valuable tool as they are likely to "invite children as co-participants" (Hester & Francis 1995:66).

Signs and Sound Symbols

> The proper use of images/symbols sanctions the generosity of information available in a sign, and it allows for an intense amount of information without the need of extraneous use of text. Multimedia is keen on being a significant vehicle, that is, its essential staying power is the fact that is can illustrate, inform, communicate and augment perception. There are communication tools, that by themselves have no intrinsic meaning but become effective when we invest them with consequential information, the sign is one of these elements we utilize for such input. Graeme Turner (1992) remarks that for something to succeed as a sign, it must have a physical form, refer to something other than itself, and be recognized as doing this by other users of the sign system. Chandler states that: "It is easy to slip into referring to Peirce's three forms as 'types of signs,' but they are not necessarily mutually exclusive: a sign can be an icon, a symbol and an index, or any combination." He

[55] Syntactic Knowledge deals with language units and rules for combining language units
[56] Semantic Knowledge involves the understanding of ones mental model of the major locations, objects and actions inside a given system.

continues: "Film and television use all three forms: icon (sound and image), symbol (speech and writing), and index (as the effect of what is filmed); at first sight iconic signs seem the dominant form, but some filmic signs are fairly arbitrary, such as 'dissolves' which signify that a scene from someone's memory is to follow." But sound and image can take on all three forms of sign, such scheme could be:

- **Symbolic**: an 'arbitrary' or purely conventional sound- where the relationship must be learnt or has been imbedded in our psyche – the high pitch strings in the shower scene of Hitchcock's *Psycho*, or
- **Iconic**: where the signifier sonically or sympathetically resembles or imitates the signified (recognizably sounding like it) - being similar in possessing some of its qualities – the wind through a forest, footsteps, etc…
- **Indexical**: where sound is directly connected in some way to the signified - this link can be observed or inferred – like a knock on a door or a baby crying.

More so than in graphic representation, the audio "sign forms" will vary with a substantial degree of skewed assumption. A sound unit (syllable, phoneme, effects, etc.) can go beyond its actual linguistic or absolute meaning as an idiosyncratic, non-meaning unit, as a fundamental expression of some kind of meaning. Wherein one expression of an audio sign is symbolic, iconic or indexical will be greatly influenced by the context, the same signifier may be used in two or more circumstances: for example, the sound of children laughing could indicate a dream state of someone remembering their past, or it can indicate a non-diegetic sound of children on the playground outside the window. Furthermore, Chandler (1994) mentions that "For a sign to be truly iconic, it would have to be transparent to someone who had never seen it before"; and it seems unlikely that

this is as much the case as is sometimes supposed. We see the resemblance when we already know the meaning. This is especially true with onomatopoeic words, which supposedly imitate the sound of their referent. Russian speakers for example regard the Russian words *puknut* and *pyornut* as onomatopoeic, but it is not possible for someone who does not speak Russian to work out their meaning from the sound alone (Cook, 1992).

In our contemporary setting, we are inundated with the daily use of now-familiar signs. With the habitual manipulation of the various tools that technology is spurting out by the second such as cell phones, personal computers, and web-based technologies, it is imperative that the designers of these ever-evolving technologies adhere to a 'iconic standard,' or else it is the author's belief that users could possibly be continuously misled into extensive instability. What the internet and personal computers show a propensity for is the manifestation of a semiotic medium to enhance information and communication. The personal computer and web-based tools are intrinsically designed to manage and organize signs, graphic, verbal or sound-based. As most computer users well know, we discriminate between the various signs and their reference: that is, what various buttons represent, what URLs refer to in terms of content once accessed, etc. The computer and/or web-based tools are a network of symbols, linking to other symbols steering the user to a content database. As multimedia application, they become the ideal agent for non-traditional means of communication, with the predisposition to use bi-modal approaches of firmly integrated graphics and sound.

Art and Technology

In his *Sketches of a New Aesthetic of Music* (1907), the Italian composer Ferruccio Busoni referred to an instrument invented by Dr Thaddeus Cahill, the *Dynamophone*[57], or *Telharmonium*, in such terms:

> Since pitch depends on the number of vibrations and since the apparatus may be set to any number desired, an infinite gradation of the octave can be achieved by the simple expedient of moving a lever corresponding to the pointer of a dial. Only after a long period of careful experimentation and training of the ear will this unfamiliar material become a useful tool of the art of the future.

The *Telharmonium'* can be considered the first significant electronic musical instrument, and was essentially a collection of 145 modified dynamos employing a number of specially geared shafts and associated inductors to produce alternating currents of different audio frequencies. These signals were controlled by a multiple set of polyphonic velocity sensitive keyboards (of seven octaves, 36 notes per octave tuneable to frequencies between 40-4000Hz) and associated banks of controls. In 1949, Dr. Meyer-Eppler said that the *Dynamophone* was used to broadcast concerts over telephone lines and used in communications systems; the instrument weighed over two hundred tons and looked like the engine room of a ship (Stuckenschmidt, H.H. 1969). This monstrous instrument occupied the entire floor of "Telharmonic Hall" on 39th Street and Broadway New York City for 20 years.

---

[57] Taken from Stuckenschmidt, H. H. (1969) Twentieth century music. New York: McGraw-Hill: World University Library

Obviously, we have come a long way since the *Dynamophone*. These days, in a world incessantly beset by transformation, where our techno-views can possibly determine the magnitude of delusion that our techno-imagination is willing to shadow, and where the progressive cyber-arts could translate into a postmodernism instance that is more than just a device "that is too complicated for us to understand,"[58] Marshall McLuhan declared 40 years ago that we are "...suddenly nomadic gatherers of knowledge, nomadic as never before, informed as never before, free from fragmented specialism as never before--but also involved in the total social process as never before."[59] On course with our ever-growing thirst for more resources and an original techno-viewpoint, information technology has destabilized our conventional significance of standards - such as space and time, sequential principles, telecommunications, etc. - and created a prevailing techno-culture genre of art. While it is indisputable that musical language is in most respects completely different from the visual arts rhetoric, my view is that the "digitalization" of the arts has merged our various art forms into a composite archetype that need not be simply re-interpreted in an analogue format, but a new generic language intrinsic to this matrix must be developed.

[58] Eveland J. D. and L. G. Tornatzky (1990) The Deployment of Technology: The Processes of Technological Innovation in L. G. Tornatzky and M. Fleischer. Lexington, Massachusetts, Lexington Books

[59] McLuhan, M. (1964) Understanding Media: The Extensions of Man: London:Routledge

Persona Lost

> Consciousness seems proportionate to the living being's power of choice. It lights up the zone of potentialities that surround the act. It fills the interval between what is done and what might be done.
> H. Bergson[60]

In our cybernetic culture, constructing or the act of creating works of art is not imitation or reproduction of our society, but a reference to a code of characteristic ambiguities. Our technological constructs seem to be pushing to assemble new views, new structures, and new occurrences, all with the underlying shadow of unspecificity. Many new media artists seem to be amenable to discarding unmitigated authority to a more accessible and flexible methodology. On the subject of electronic media, Glenn Gould[61] states:

> The most hopeful thing about this process—about the inevitable disregard for the identity factor in the creative situation—is that it will permit a climate in which biographical data and chronological assumption can no longer be the cornerstone for judgments about art as it relates to environment. In fact, this whole question of individuality in the creative situation—the process through which the creative act results from, absorbs, and re-forms individual opinion—will be subjected to a radical reconsideration.

Especially today, after the conjectural deficiency of social ideals and ventures, utopian promise is propelling its way into new spaces - more specifically, electronic technologies and new media. While the industrial generation overstated a myriad of variations on the myth of the machine as the universal instrument of production, the post-

[60] Henri Bergson, (1922) Creative Evolution, trans. Arthur Michell (New York: Holt,), 179.
[61] Gould, Glenn; 1984: *Prospects of Recording* in the Glenn Gould Reader, Ed. Tim Page,

modern age conveys the phantasm of the Universal Network: the aspiration toward a permissive mode of communication and the enhancement of the spaces of intellectual intercourse. As humans we have a very profound need to connect, I think that we are now starting to identify with communication technology, and realize that this technology offers extraordinary innovative possibilities for personal development, and shape a new variety of interactions, this new media can be an innovative touchstone for thinking critically. Social spaces are developed as open networks in which meaning flows out in diverse labyrinthine paths. Is this designed to tear the commodity makeup from the new media, so that they can be put to a communicative and gregarious use? I will scan through the possibilities in my work.

Since the configuration of this new media is essentially accessible, we are not faced with a technical problem but rather one of the networking of social relationships, artistic forms of production and recurring conventions. Centrality was key in our communication systems, where the process depended as much upon how the receiver decoded information as it did on the encoding practices of the sender, but the spatial links of centrality have been significantly transformed by the new technologies and by globalization. This generates complex theories around the definition and construct of centrality, where we now have transactions that develop through technologies which countervail distance and location, and this on a global scale; centrality has conventionally been represented in certain types of built-in environments and urban constructs, but economic globalization and the current information technologies have not only re-categorized centrality and its spatial junctures, they have also created new spaces for

centrality. The significance we attribute to accepted similes might change depending on cultural background in the future, as renewed inquiries lead to more adequate understanding of natural events (as was implied by Dewey's fallibilism[62]) does not require that our experience of the world at any given time may as a whole be errant. These advances in communication and information technology are changing the way people meet and converse, whole communities can meet, talk, and work together outside traditional meeting, home and/or workplace spaces. Dewey (1916) stresses the point that sensations, hypotheses, ideas, etc., come into play to mediate our encounter with the world only in the context of active inquiry. Once inquiry is successful in resolving a problematic situation, mediatory sensations and ideas; as Dewey says, "drop out; and things are present to the agent in the most naively realistic fashion." This new techno-space has materialized not merely as an information transmitter model, but as a predominant space for capital supply and the operations of global capital. This is an expression of statistical space comprehensively positioned in a significant domain within the dynamics of organized society, particularly productive power. Gould[63] adds:

---

[62] Belief that some or all claims to knowledge could be mistaken. Although Peirce limited the application of fallibilism to the empirical statements of natural science, Quine extended it by challenging the notion that any proposition can be genuinely analytic. Unlike a skeptic, the fallibilist may not demand suspension of belief in the absence of certainty.

[63] Ibid

> I believe the fact that music plays so extensive a part in the regulation of our environment suggests its eventual assumption of a role as immediate, as utilitarian, as colloquial as that which language now plays in the conduct of our daily lives. For music to achieve a comparable familiarity, the implications of its styles, its habits, its mannerisms, its tricks, its customary devices, its statistically most frequent occurrences—in other words, its clichés—must be familiar and recognized by everyone. A mass recognition of the cliché quotient of a vocabulary need not suggest our becoming saturated with the mundanities of those clichés. We do not value great works of literature less because we, as men in the street, speak the language in which they happen to be written. The fact that so much of our daily conversation is concerned with the tedious familiarities of common courtesy, the mandatory conversation openers about the weather and so on, does not for a moment dull our appreciation of the potential glories of the language we use. To the contrary, it sharpens it. It gives us background against which the foreground that is the habitat of the imaginative artist may stand in greater relief.

The flux of information structures found in our techno-space or new "mediascapes" does not simply connect various domains - by being connected they are also altered; it is their inherent characteristic to change their orientation and quality unexpectedly, a characteristic made very specifically by electronic media. Out of such changes new correlations are constructed that bring presumably autonomous components into an unexpected interdependency. For example, we can say that global spaces are hyper-concentrations of infrastructure, but also we can perceive them as both a geography of centrality and one of peripheral expanse. Out of these transformed correlations new interpretations emanate that conceivably re-construct the flux of information; therefore, these information constructs or techno-spaces are in constant mutation because this same mutation feeds upon itself, and this facilitates the flux of information structures to reorganize themselves according to the various circumstances. The questions and conflicts that these modes of thinking raise about the validity of

globalized media theory pre-suppose an insight into a "theory" of what they are, and where they expect to be redirected. Gould[64] goes on to say:

> It is my view that in the electronic age the art of music will become much more viably a part of our lives, much less an ornament to them, and that it will consequently change them much more profoundly. If these changes are profound enough, we may eventually be compelled to redefine the terminology with which we express our thoughts about art. Indeed, it may become increasingly inappropriate to apply to a description of environmental situations the word "art" itself—a word that, however venerable and honored, is necessarily replete with imprecise, if not in fact obsolete, connotations. In the best of all possible worlds, art would be unnecessary. Its offer of restorative, placative therapy would go begging a patient. The professional specialization involved in its making would be presumption. The generalities of its applicability would be an affront. The audience would be the artist and their life would be art.

Gould did not like words like "artist" or "public," as this implied a hierarchical relationship he was not comfortable with.[65] We are most certainly seeing a shift in the way we consume and create in the 21st century. The recording industry is desperately trying to come to grips with how to protect copywriting materials and the web has opened a whole new approach to publishing and distributing materials by offering resources that were not even conceivable ten years ago. So, in a way, Gould was raising issues that are now more or less common place, i.e., audience creativity enabled by digital technologies. With the synthesis of various art forms through the use of the multimedia model and the assimilation of computers, web based communications, and information technologies, we are formulating a new identity, converting, augmenting and adapting the various components of intellect. These innovations are happening at such a rapid pace that I hope the suggested use of an established bi-modal system - of spatial and

---

[64] Ibid

[65] Source: *32 Short Films About Glenn Gould directed by Francois Girard*

temporal objects – will foster and support the alignment of various multimedia resources and meet the evolving and various needs of this new form of digital media.

In the preceding chapter I speculate about early codification, art and technology, symbols, cyber cultures, etc. The following segment will elaborate more on representation and various aspects of contemporary concepts in multimedia and a broader discourse on bi-modality.

CHAPTER III:

SOUND DESIGN – Perception and Expression

Theory & Practice

At a time when more and more educators and students in the arts (in particular, students of film, animation, media arts and communications) are having to deal with up-to-the-minute technology and complex interdisciplinary concepts in their work, the need for solid decision-making regarding the sound environment is becoming essential. In my lectures and papers on the subject of sound design, I have taken issue with many theorists (mostly from the film sound domain) for stringently adhering to an associational (i.e., referential) meaning between the acoustic and the visual sphere. A brief summary follows:

For the most part, theorists conceptualize sound design as referential: referentialism[66] emphasizes the importance of the fact that musical sound can refer to extra-musical objects, ideas, and actions. This is due in part to the overwhelming concern of "seeing." I would be inclined to modify this vision to include sonic expressionism: this expressionism would suggest a musical meaning that is essentially intra-musical (i.e., areferential). Audio is a complex conduit that demands your attention as a constant - you cannot shut your ears down (like closing your eyelids) - and feeds the aural information system. Both referential and expressionist aspects of meaning are important to the media producer (educator, artist, etc). In contrast to the privileged status often given to

[66] Meyer, L.B. (1956). Emotion and Meaning in Music. Chicago, IL: University of Chicago Press.

associational (i.e., referential) aspects of sound design, I would propose a model of "intermedia" work that is organized fundamentally on the derivation of meaning within both the auditory and visual formulas. Hence, the relevance of the audio-visual experience would be solely under the control of the observer and allow for a gamut of experiences limited only by the observer's imagination. The audio-visual experience could be referential (i.e., pointing to something outside of the audio-visual experience) or areferential (i.e., inherent in the sound/optic relationship). Once again, these two types of dialectical interactions could allow for the apparent sound design to be created in the observer's imagination rather than imposed by a linear sound track.

Music is *expressive* (Hevner, 1935) in one form or another by virtue of its capacity to bring forth feelings in the listener: this implies its susceptibility to being operationally metaphorical. As mentioned above, the referential or areferential experience of music can be augmented by visuals, and therefore we can deduce that this expressiveness is appreciably related to emotion. So how is this expressiveness significant in exhibiting "extra-musical meaning"? By the skillful development of sound design as an information instrument, and the use of this instrument repositioning itself according to usage. In the traditional information design perspective, the "audience" is positioned for a specific intention; in the more recent perspective, one moves towards a process of "negotiation" or "construction." In traditional instructional design theories, meaning resides in objectivity or the instructional goal; in postmodern instructional design theories, meaning is a process of negotiation between informer and informed, sender and receiver. Along with negotiation, one must contend with the fact that the

informed may accept, modify, ignore or reject an idea according to their experience, attitudes and purposes. Thus, choosing a model that will benefit cross-modal association is a significant issue that affects the methodology of the current study. Non-verbal sound design and its relationship to images is an interaction of meanings whose consequences is an elevated appreciation of emotional conditions.

Sound Emotion

> Music is a moral law. It gives soul to the universe, wings to the mind, flight to the imagination, and charm and gaiety to life and to everything.
>
> Plato[67]

We all know that music can be suggestive and give focus within a collective perception, that it can influence how we react as a group, to the point of conditioning masses into a passionate commotion. We only need to look at the rock concert phenomenon or the discotheque fad. In ancient Greece, philosophers such as Plato, Socrates, and Pythagoras had an acute comprehension of the impressive influence that music and sound had on its audience. Plato stated: "In order to take the spiritual temperature of an individual or society, one must mark the music."[68] Pythagoras linked music to geometry, mathematics and even astronomy; he edified many scholars with instruction on how each of the seven planets produced a specific note within its own orbit; he used the distance between the earth and their bodies and the sub-division of the

[67] The Republic of Plato, translated by Benjamin Jowett. Oxford Clarendon Press, 1888, page 88
[68] Ibid

string as means to calculate the frequency data for a given planet - commonly referred to as the Music of the Spheres (Musica Mundana). Pythagoras elaborated a system that declared that various musical modes will have a particular effect on the individuals that hear them, and he supposedly could cure various illnesses by prescribing melodies in various modes. Socrates also felt that training in the art of music could influence character and play a part in manipulating the core of ones inner self, to the point of stating: "Musical training is a more potent instrument than any other, because rhythm and harmony find their way into the inward places of the soul, on which they mightily fasten, imparting grace, and making the soul of him who is rightly educated graceful, or of him who is ill-educated ungraceful."[69]

Sitting in a room facing an expensive apparatus, eyes closed, a human body is transported and transfixed while going through an assortment of emotions; a teenager gliding on the street on his skateboard swaying to a rhythm or a woman in a subway oblivious to her surroundings seems to be in a trance while others rush by. What are these people doing? They are listening to music, people listen to music for a variety of reasons - to forget, to be uplifted, to be reminded of a pleasant past, to feel a painful sensation through an imagined tragedy – all are possible scenarios, but the underlying characteristic here is that music can induce full-fledged emotions and it can evoke a variety of simulated experiences. Are these purely 'aesthetic emotions' different from 'real' experiential feelings? Although it is not in the scope of this study to delve into the psychosomatic regions of cognitive science, I will posit that individuals can experience music on a variety of levels whether through purely aesthetic understanding or pragmatic

[69] Ibid

conditions. Furthermore, does the experience come to a rapid standstill once we turn off the sound? Deryck Cooke (1959) argues that the response induced by a musical phrase will linger and develop after we have turned off the sound; he also observes that music's effect on us cannot be made up only of metaphors of past accounts, in view of the fact that, often times, music seems to generate feelings that are unique and different from anything we have felt before. When a passage of music upsets me or lifts my spirits, I am not directly sad or happy at the music but the generated link to that particular emotion that was awakened in me. According to Kivy (1993), we should develop objective principles to identify various emotive features of music, expanding the character of 'conventions' in musical descriptions; he feels that our understanding of emotional subject matter in music is culture-biased, but if a specific expression of emotion were "frequently associated" with it, objective reckoning could be managed. Kivy seems to discard much of the work done by film composers in the past 60 plus years.

In the next section, I will expand on how music can draw on articulate modes of expression. One factor remains coherent and observable: that music is susceptibly expressive, and has a substantial correlation to human sentiment, so that designing for representational or referential extra-musical meaning is from the onset plausible.

Music and Meaning

Accomplished sound design should generate a significant functional measure of meaning through the use of aural metaphors, and these aural metaphors can be used as a

device for a sound designer to communicate adeptly to an audience and engage them in a comprehensive exchange of ideas, atmosphere and emotions (Zaza, 1991). Kendall Walton[70] notes that "Literary and pictorial representation establish fictional worlds." He points to the fact that "story worlds" have an assemblage of characters, as is sometimes true in pictorial worlds: we can "enter" an allegorical condition and "monitor" and gauge our surroundings with the help of these fictional "communities," but does music have these metaphorical worlds?[71] Or is it just a collection of notes, harmonies and rhythm? I suppose that one can make up an imaginary world while listening to a musical composition, but this is even more subjective than the images fashioned by a novel. Walton continues by stating that "pictures and stories are representational by virtue of the fact that they call for such imaginings," he also admits that music "engages" our imagination but that these "are about elements of the music itself, about tones and harmonies and melodies." One interesting aspect of music's ability to "impress" upon the listener a variety of reactions is contextual significance: contrary to pictures and stories, where characters and scenes are mandatory to the creation of "emotive" response, music can make us feel joy or sadness without being joyful or sad about anything specifically.

Music, though, is dependent upon form and structure: Newcomb (1983)[72] states that " Formal processes themselves create expressive meaning." Moreover, he contends that interpretation can demonstrate "the perceived structural patterns of a piece into relation with other patterns of the listener's experience" and can "reveal new structural

[70] Walton, K. 1997; *Listening with Imagination: Is Music Representational?* (in Music and Meaning, Robinsion J. Editor) Pg 57-82
[71] Of course we are not referring to Opera, as this is closer to narrative fiction.
[72] Anthony Newcomb, "Those Images That Yet Fresh Images Beget," Journal of Musicology 2 (1983): 232

patterns in the music as well."[73] Ostensibly, we can infer that form and structure are a function of expressive composition, and the use of accurate and apposite formal sound design can create a representational locus of attention. This can be a key process in multimedia sound design, in view of the fact that the sound designer can reduce the need for schematic arrangements, that is sound designs that are based on representational or metaphoric material, leading to the effect of being able to prolong focus on multiple sources. The listener can imagine or experience the "feelings" without necessarily being involved in the idiosyncratic occurrence of that emotion, and therefore only extracting the meaning that is relevant for her/him at a given moment. Additionally, the corroboration between the actual sound design and our feelings is the result of the correlation linking various cultural associations, behaviors, and so on.

Karl and Robinson (1997) support the claim that musical expression can be evaluated as a "kind of gestural expression of emotional or other psychological states in a musical persona, whether it be the composer's or that of some indeterminate character or characters in the music."[74] This "gestural expression," or our capacity to have mental images that evolve from hearing certain sounds, is the nucleus of imagination - it gives us the potential to ostensibly experience things that are not an immediate occurrence of our various senses. We can, for example, have the impression of seeing, smelling or feeling particular things without actually going through the active processes of that particular sensual attribute. Skilled sound design can tap into our various emotions and expose various past experiences, dreams, fantasies, and so on.

---

[73] Anthony Newcomb, "Sound and Feeling" *Critical Inquiry* 10 (1984): 637

[74] Karl, G. & Robinson, J. (1997) "Shostakovitch's Tenth Symphony and the Musical Expression of Cognitively Complex Emotions," (in Music and Meaning, Robinsion J. Editor) Pg 154-178

Metaphorical Synesthesia

I introduce the phenomenon of Synesthesia at this point, not to examine this discipline within a cognitive scientific approach, as this would be a comprehensive dissertation in its own right, but as an element of the more contemporary concept of multimedia as a metaphorical imperative and connected medium, hence "Metaphorical Synesthesia." The Merriam-Webster dictionary defines Synesthesia as: "New Latin, from syn- + -esthesia (as in anesthesia): a concomitant sensation; especially: a subjective sensation or image of a sense (as of color) other than the one (as of sound) being stimulated." A more authoritative definition could be taken from Richard E. Cytowic's abstract to one of his many works on the subject, "Synesthesia: Phenomenology And Neuropsychology: A Review of Current Knowledge" (1995); he defines Synesthesia (Greek, syn = together + aisthesis = perception) as follows: "...the involuntary physical experience of a cross-modal association. That is, the stimulation of one sensory modality reliably causes a perception in one or more different senses."[75] Synesthetic metaphors are commonplace in our (and certainly other) cultures and are linguistically associated in terms belonging to one or more differing perceptual modes. For example, we will often hear "loud colors," "dark sounds," and "feeling blue." Of the different senses, hearing is generally agreed upon to be most consistently involved in the synesthetic experience. Within synesthetic metaphors in English, colored sounds are most widespread. (S. Day,

[75] Cytowic, R 1995; Synesthesia: Phenomenology And Neuropsychology
A Review of Current Knowledge; http://psyche.cs.monash.edu.au/v2/psyche-2-10-cytowic.html

1996).[76] Cytowic (1995) posits that: "Synesthesia is 'abnormal' only in being statistically rare. It is, in fact, a normal brain process that is prematurely displayed to consciousness in a minority of individuals."

Similarly, computers and the net allow us to discern various forms of data within a multimedia experience and are strategically positioned to create multi-sensory imagery within the receiver's mind; these mental images can be facilitated by efficiently crafted sound design and their industrious association with image, allowing them to foster audience perception and familiarity with their own experiences into a personalized "mental representation." Pertinent sound design tied to visuals should distinguish between the boundaries of sound and image and attempt to connect within the interface a link of the sound functionality and the audience's interpretive faculty. Speaking of these "new media," Dr Hugo Heyrman (1995)[77] states that they "are of a specifically telematic nature and for this reason effectively differ from the usual forms of communication." He speaks of 'Tele-Synesthesia' and adds that "by linking the concepts 'tele' and 'synaesthesia' to each other, we deal with the fact that the transmission of data creates a synaesthetic effect: 'tele-synaesthesia', synesthetes are in some sense, people of the future." The fact that we are living in a "digitized" society along with its implications of "cyber culture" and electronic archetypes influencing every aspect of our existence has significantly predisposed us to respond to multi-faceted input and interactive elements.

---

[76] Day, S. 1996; Synaesthesia and Synaesthetic Metaphors http://psyche.cs.monash.edu.au/v2/psyche-2-32-day.html

[77] Dr. Hugo Heyrman "Art & Computers: an exploratory investigation on the digital transformation of art. Tele-Synaesthesia, presentation of a hypothesis". Doctoral thesis, Universidad de La Laguna, Santa Cruz de Tenerife, Spain, 1995. See: http://www.doctorhugo.org/e-tsyn.htm

Dr Hugo adds: "The modalities of our sensorial perception become interactive by means of electronic mechanisms of control and selection. A 'tele-culture' is emerging, subjecting both the perceptual and the conceptual to —strictly speaking— continuous revision." We in effect add and fill in the details beyond the audio/visual information made available in one or other source. Audio is especially a strong link to this "tele-culture" as it absorbs and reflects information. What I mean here is that audio (music, sound design) is a vehicle, participatory in nature, that captivates an audience and is capable of yielding a multiplicity of information; on the other hand, the audience must differentiate, categorize and decode a variety of implicit messages and then create coherent meaning from the information received. I will call this "symbolic information," as it is transmitted solely through an auditory form, but as this "symbolic information" is processed with relevant accompanying elements, be it pictures, video, etc., the categorizing and decoding become easier to grasp due to the use mental imagery very similat to the dual-channel model. However, what of an extraneous association to various media elements? The channels of information risk being inundated with ambiguity and provoke indiscernible responses to the message design.

The multi-channel characteristics of audiovisual media can communicate information or emotional content appropriate to a given intention only if the relationship between sound and image are carefully composed; otherwise it can provoke made-up occurrences that were not the intention of the designer. Although no two people will ever have the same "fictional experience," the "symbolic information" and its associated

media will nevertheless dictate the experience, and can either be revealing and helpful or uncertain and confusing. I will discuss this point further in the next chapter.

Signal to Noise Ratio

> Noise is not only what we hear in the streets, in nature, in music. It is also a form of violence: noise hurts. Music is noise in order, noise with sense. Music is, therefore, a metaphor for the taming of violence, a metaphor for the role of religion and, later on, the role of civilization.
>
> Jacques Attali[78]

Communication theory acquaints us with the fact that anything that brings unpredictability or randomness to a signal is noise. In this age of total (or nearly total) accessibility, we seem to be excessively subjected to haphazard information even within a given systematic context - sensory overload is an everyday reality, as we walk in the streets we are swamped with noise, advertisements and microwave pollution. Technology has allowed for an array of possibilities regarding our access to information, we can surf the net and download music, video, text, graphics. But does the fact that we can access it give us the essential and desired outcome? Meandering in cyberspace, accessing a variety of interfaces on a multiplicity of digital systems can often appear sterile to the human senses, but this is our new reality and could allow for a viable participation in knowledge acquisition. If we look at a completely random array of pixels, what information will we gather from it? None! We will see only chaos, or snow - the moment order is imposed onto these pixels, a picture will begin to materialize. Attali (1985) speaks of organized

[78] During an interview with Ian Simons. See: http://www.nthposition.com/makingsenseofnoise.php (July 2002)

noise and music as a systematic means of regulating social order; he says: "Primordially, the production of music has as its function the creation, legitimization and maintenance of order. Its primary function is not to be sought in aesthetics, which is a modern invention, but in the effectiveness of its participation in social regulation..."

When I was a teenager I used to relish looking at the covers of jazz albums and reading the various liner notes. Buying an album meant saving, sometimes working long hours to purchase a particular album, which then was cherished and listened to religiously. Now with easy access to the various techno-tools, one can download hundreds of tracks within hours, while listening and/or doing other unrelated tasks, accumulating possibly thousands (if not tens of thousands) pieces of media, be it songs, videos, etc. Excessive sensorial redundancy, in our modern post-digital society, seems to take essence to a point outside of our conscious control - music is a good point of reference for this contention. In earlier times, music mostly reflected and required harmonious clarity, and then diverse and unusual sounds were added. Subsequently, the compulsory need and predilection for connecting and mixing different and unusual sounds, was increasingly showing popularity in the arrangement of our new ethos; as a transient departure from our safe consonant harmonies, chords were extended, taking pleasure in dissonance, modern music became liberated from and moved into continually more complicated forms, determined to incorporate the most dissonant, unfamiliar and discordant sounds, developing into a median of *musical-noise*, which is to some degree natural, but mostly socially expected or accepted. These musical innovations are analogous to the burgeoning of technology; things have moved swiftly forward since the

Industrial Revolution: from mass classroom education, which was a tedious symptom of industry, we have moved out of this industry-driven society into an information technology (IT) society. More and more IT is rendering our classrooms and workplace superfluous, as there are now hubs of communication - wherever there is a computer, a telephone line, a modem or wireless connection we can access our colleagues, clients, even a wide-ranging audience. One of the outcomes is that this new ethos is being pulled into a media marketplace and adjusted to more commercial characteristics.

What is more, we have moved from the agricultural age through the industrial age and into the information age in a span of just over one hundred years; with the introduction of the Internet we are moving into a second information revolution. (Dent; 1998). With this revolution comes the recognition that there is no one "fixed" method for doing things and very often we are likely to combine multiple alternatives in order to achieve a successful solution. We need just to look at our media: just a few years ago: those who wanted to be aware of the news would turn to the radio, newspaper or TV, and, as little as 20 years ago, television had 5 to 10 channels. We can now access over 600 channels with the possibility of setting up automated recordings of any one of the screenings, and even get instant replays. Multimedia has become the mediator whose task is to stimulate and contend with our sensibilities; just as with art music, which evolved towards the most intricate polyphony and the utmost variety, multimedia is evolving in style and form, adapting to allow for information overload, and it should be the contender to assuage media-noise. As the ear of the Baroque citizen would by no means be able to process and endure the discordant density of a variety of sounds produced by our

contemporary musical forms, humans have adapted depending on the culture, and within prevailing tendencies of our reality; the qualitative variety of information has created an evolution of the "media-centered" persona that was not possible before this time.

We are becoming accustomed to an infinite variety of "noise-sound" relationships, and quick sound bites are common ground in filmic gestures: as Chion (1990) mentions, "Goddard was one of the rare filmmakers to cut sounds as well as images, thereby accentuating jumps and discontinuities..." By today's standards of the music video and gaming generation, the limited specialty of pure sounds is undoubtedly shattered, and the infinite multiplicity of "noise-sound" conquered. It is recognized that most musical passages embrace an increase of feelings that are already recognizable and identifiable, and which predispose the listener to habitual response. Modernist composers began using industrial sounds in a variety of arrangements to signify to the modern day listener the various relations of audio array, while electro-acoustic artists are continually experimenting with a variety of new and every day sounds to create innovative forms of sound design, mixing a variety of noises, such as the rumble of thunder, the whistle of the wind, the roar of a waterfall, the gurgling of a brook, the rustling of leaves, with urban sounds like alarms, cars, and various mechanical tones. In addition, these sound artists, while creating these soundscapes, allow us to generate the missing visual details within the various settings; the metaphors engendered by the receiver usually come from an individualized psychological process. Imagery, by nature, is multi-sensory, and skillful audio sound design can expand human experience throughout the multi-sensory image-building capability of the mind. A common perspective that I made reference to in the

preceding section is that multimedia, through its multi-channel characteristics, can communicate information or emotional content appropriate to a given intention; but the relationship between sound and image should be carefully composed – otherwise, it can provoke false imaginary experiences. We have all experienced the mismatch of a sound with an image when we have read a book and then seen its screen representation at the cinema; those characters that are now too familiar in our imaginations occasionally take a good getting used to when we hear them interact in a not-too-convincing cast.

In the following chapter I will present how multimedia draws a good deal of its potential as a communication agent from two of our five senses (Mayer's multimedia principles and Dual Coding). With regard to the effectiveness of multimedia, our brains learn most from the bi-modal flow: as our brain retrieves information from two relatively independent systems, both of which are addressed by the multimedia schema, audio and visual representations are necessary to utilize fully the cognitive capabilities of the mind. In multimedia or intermedia, multiple perceptual signifiers and/or media such as speech, music, text, graphic, still, animation and video are often used in an integrated manner.

CHAPTER IV:

THEORETICAL ORIENTATION OF THE STUDY

Dual Channels

Mayer (2001) contends that multimedia designers often use single-channel and limited capacity processing systems[79] when designing materials, and do not take advantage of the auditory modes of presentation, basing their design on single-channel assumption - that is, that all information enters the cognitive system in the same way regardless of its modality. Also, by mixing, text, audio, graphics and words, the receiver has unlimited capacity for input, and by transmitting discrete pieces of information, the design is based on a passive-processing assumption. Mayer suggests that learners (or anyone on the receiving end of information transmission) have dual-channel coding resources, a limited capacity management system, and an active processing scheme. (see Table 1)

[79] in which people can pay attention only to a few pieces of information in each channel at a time - in contrast to active processing, where people understand the presented material, organize it into a coherent mental structure, and integrate it with their prior knowledge.

Assumption	Description	Related Citations
Dual Channels	Humans possess separate channels for processing visual and auditory information.	Paivio, 1986; Baddeley, 1992
Limited Capacity	Humans are limited in the amount of information that they can process in each channel at one time.	Baddeley, 1992; Chandler and Sweller, 1991
Active Processing	Humans engage in active learning by attending to relevant incoming information, organizing selected information into coherent mental representations, and integrating mental representations with other knowledge.	

Table 1 - Mayer's Dual Assumption Schema

Mayer's dual coding theory reveals that individuals understand and code information in two ways:

1. Imagery - translating from information to imagery, therefore allowing us to visualize and conceptualize.
2. Linguistic (audio) – translating from information to verbal semantic knowledge.

This obviously allows us to organize our knowledge through two conduits. Visual processing is more spatial, whereas audio processing is temporal. John Munro[80] (2002)

[80] Visual-Spatial Giftness by Dr J. Munro: See: http://www.edfac.unimelb.edu.au/LED/ELAGE/documents/GLT-Visualspatialgift.pdf

notes that visual/spatial learners are often preoccupied with space at the expense of time, and that knowledge acquired through visual modes involves:

- The synthesis of ideas,
- Intuitive grasp of complex systems (skipping "steps"),
- Simultaneous processing of concepts,
- Inductive reasoning, use of imagery and idea generation by combining disparate elements in new ways,

Ideas and concepts become more spatial and holistic, with discrete events linked and matched in space with or without relation to time. Logographic or pictorial style of communication is a usual construct of this type of thinking. In a study by van der Molen and van der Voort (1997), children remembered on-screen stories (visual+hearing) over stories they had read or even ones read out loud to them. Tests of recollection of information offered in a text versus one presented in movie format demonstrated that the text was forgotten more quickly, suggesting that film produces a "more emotional reaction," and this seems to influence our memory and recall ability (Baggett,1979). Auditory/sequential learners demonstrate responsiveness to time more than to space. Sequential learning involves:

- Analysis of information into smaller units,
- Orderly progression of knowledge from simple to complex,
- Skillful categorization and organization of information, and
- Linear, deductive reasoning.

The sequential point of view is more temporal: where cause and effect is stressed, this leads to transfer constructed from non-meaningful elements – individual letters of the alphabet as discrete elements.

The dual-coding theory proposed by Paivio endeavors to give equal substance to verbal and non-verbal processing. Human characteristic modes of learning have the particular ability to function concurrently with language/audio and with nonverbal/visual objects and procedures. Our communication system is distinctive in that it deals smoothly with audio input and output (as in talking and listening) while at the same time serving a representational activity with respect to visual objects, events, and activities.

Practical/Technical Implications

Mayer (2001) has developed a series of principles based on this dual channel theory designed for multimedia applications (see Table 2):

Empirical Results	Practical Implications
Multimedia Principle:	
Students learn better from **words and pictures** than from words alone.	On-screen animation, slide shows, and narratives should involve either written or oral text and still or moving pictures. Simple blocks of text or auditory-only links are less effective than when this text or narration is coupled with visual images.
Spatial Contiguity Principle:	
Students learn better when corresponding **words and pictures are presented near** rather than far from each other on the page or screen.	When presenting coupled text and images, the text should be close to or embedded within the images. Placing text under an image (i.e., a caption) is sufficient, but placing the text within the image is more effective.
Temporal Contiguity Principle:	
Students learn better when corresponding **words and pictures are presented simultaneously** rather than successively.	When presenting coupled text and images, the text and images should be presented simultaneously. When animation and narration are both used, the animation and narration should coincide meaningfully.
Coherence Principle:	
Students learn better when **extraneous words, pictures, and sounds are excluded** rather than included.	Multimedia presentations should focus on clear and concise presentations. Presentations that add "bells and whistles" or extraneous information (e.g., to increase interest) impede student learning.

Empirical Results	Practical Implications
Modality Principle:	
Students learn better from **animation and narration** than from animation and onscreen text.	Multimedia presentations involving both words and pictures should be created using auditory or spoken words, rather than written text to accompany the pictures.
Redundancy Principle:	
Students learn better from **animation and narration** than from animation, narration, and on-screen text.	Multimedia presentations involving both words and pictures should present text either in written form or in auditory form, but not in both.
Individual Differences Principles:	
Design effects are stronger for **low-knowledge learners** than for high-knowledge learners and for **high-spatial learners** rather than from low-spatial learners.	The aforementioned strategies are most effective for novices (e.g., low-knowledge learners) and visual learners (e.g., high-spatial learners). Well-structured multimedia presentations should be created for learners whom they are most likely to help.

Table 2 - Mayer's 7 Multimedia Principles

(Note: "students" here refers to the users of instructional multimedia)

Though I find Mayer's principles compelling, my view is that the use of multimedia tools, be it for learning or within artistic conditions, needs a broader venue of manifestation. Mayer has emphatically classified multimedia learning as "learning from words and pictures." Unfortunately, this rationalization does not take into account the vast array of new-media technology; with its sound design and audio icon metaphors, his approach is still tied to traditional re-use of the textbook syndrome. Though his theory has great value (this author uses it in his advanced multimedia course), in my view it needs to be updated to account for the array of new multimedia environments available to

the contemporary artist, teacher, etc. Since both pictures and words are in the visual information category, Mayer & Moreno (1998) suggest that it would be better to enhance visual information with audio instead of additional visuals. According to dual-coding theory, humans manage information through two generally independent channels: one channel processes verbal information such as text or audio, the other processes nonverbal images. (see Fig. 4)

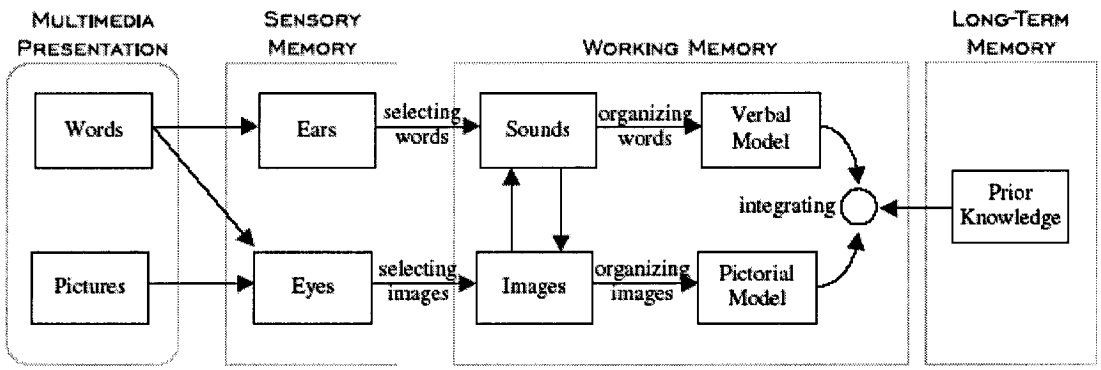


Figure 4- Dual-Coding Flow Chart

Multimedia, as with film and television, draws its full potential as a communication agent from two of our five senses, which are compelling when used in combination, that of sight and sound. When information is processed through both channels at once, it has an additive effect on recall (Mayer and Anderson, 1991; Paivio and Csapo, 1973). Research demonstrates that channel overload is less likely when information is distributed between the two separate channels and learning is more effective when the information is referentially processed (i.e., using both channel paths versus one) (Mayer & Moreno,1998). Referential processing will possibly construct a

complementary corollary because the learner produces additional cognitive paths that can be pursued to retrieve the information.

Bajo (1994) states that an image must be meaningfully recognized (illustrate semantic representation) before it can have an appellation (an audio representation), while a word may be identified before its significance is brought about. Visuals retrieve significance diametrically - that is, we process an image as a whole and in a more controlled fashion: we take the time to identify an image through a more consciously involved activity where, as we process audio more circuitously - though our reaction time is more direct and automatic - we still need to process the sound(s), which demands an allocation of attentional resources. Any representational theory must accommodate this dual functionality (Paivio, 1986). Visuals derive from a transient identification system that allows for a more proscribed conversion of accessible information. When we look at pictures and listen to audio, there are different processes that operate simultaneously, and in the dual coding approach, the audio conditions contribute to logic, direction and organization. The visual condition is comprised of holistic information, is non-sequential, and can better cope with spatial information (Gambrell & Brooks Jawitz, 1993). I believe that multi-modal or dual-channel information delivery - the perceptive association between audio and visuals so as to create simultaneous links between them - will be more effective when associations are clearly made between the information contained in each modality.

Bi-Modal Discourse

> With multimedia, the profound ethical and aesthetic challenge comes from the burden of responsibility that lies with the individual once the constraints on communications are so radically eased if not eliminated.
>
> Kenneth Dyson[81]

Today, "multimedia," or "new media," is a fashionable term in the present computer age, but the application of multimedia is basically using an assortment of media, whether visual or auditory, with the most significant purpose of communication. Computers have developed into a prevalent attraction because of their distinctive capability for helping us categorize information, as well as communicate more productively. The availability of low-cost, remarkably high-performance personal computers and the numerous easy-to-use object-oriented software programs, has equipped multimedia production with unprecedented accessibility. Sound and image technologies have evolved, though at varying pace, into a collective schema that allows the worldwide community to witness realism at its most authentic. We have come a long way since André Bazin (1967), talking of the camera, exclaimed: "For the first time, between the originating object and its reproduction there intervenes only the instrumentality of a non-living agent. For the first time an image of the world is formed automatically, without the creative intervention of man...The objective nature of photography confers on it a quality of credibility absent from all other picture making."

And so, from reality reproduction to reality re-design, technology enables autonomous processing, plasticity of treatment, and the possibility of assembling media

[81] See URL: http://en.thinkexist.com/quotes/kenneth_dyson/

into endless combinations. When a sound or image is digitized, it is stripped of all distinct identity and is hard-coded into a logical and consistent object of binary language; in addition, this object, acting as an information agent, becomes dependent upon a nucleus of opposition – ones and zeros. "Since bits are bits, it makes no difference if they represent video, sound or text, they are all the same" (Negroponte, 1995). Users of personal computers (PCs) now enjoy an enormous capacity for computational power as well as virtual communication, thus linking the creative process in ways never before imagined. This allows for integrated vision, and, for many designers, the key to creating noteworthy multimedia tools has been the ability to channel that power and create for the user a consistent and articulate system. The standard methodology inherent in digital design - that of high rate of occurrence, collage, and multi-sensorial stimulation - is intrinsic to multimedia and can have an overall homogenizing effect.

In a recent interview, Iain Cook,[82] part-time lecturer at Concordia University, speaks about an integrated creative project: "...but I think the most interesting experience I had was in my graduate studies: there was an online experiment of audio artists gathering together in these virtual spaces and with technology such as MIDI [Musical Instrument Digital Interface], building upon and creating sounds in this virtual, online space, collaborating from different countries..." He adds that, due to bandwidth considerations, everything was done using the MIDI standard, and that streaming audio files would have been impossible with the existing technology. Furthermore, regarding music education, the emphasis is to a large degree on playback proficiency - the ability to read music notation and reproduce it using a musical instrument. Composition is

[82] See complete interview in Appendix B.

frequently thought of as difficult and arid work, that it is complicated to convert musical ideas into musical notation. But technology has changed this concept, and, with the advent of the MIDI interface and MIDI-capable instruments, has unlocked the field of composition to a wider audience.

Obviously, using the digital realm of 1's and 0's was a key concept; he continues:

> ...exactly. As opposed to huge chunks of data. So somebody would have a basic MIDI track, which they would leave in this virtual studio, essentially with the doors unlocked and then anybody who wanted to could walk into that studio, or into this virtual studio, then add to that track, change it, modify it, then either save that new version or save another version, so you'd have the original left and then the enhanced version, and then somebody else could come in and you could either work on things alone like that or if by chance two or three people were online in that same virtual room then you would be typing away, 'okay what part of the song do we want to work on next, what do we want to do with the next part of the song.' So I thought that was the most interesting experience I've had with digital audio cause it tends to be sort of one person alone in a room, most of the time, working on things, as opposed to this really kind of exciting collaborative space where people from all parts of the world, from different backgrounds sort of sharing information and building stuff...

Peter Doolittle (2002) remarks that multimedia may be characterized in several ways, depending upon one's perspective. Here are a few that he mentions:

- Multimedia is the "use of multiple forms of media in a presentation" (Schwartz & Beichner, 1999, p. 8);
- Multimedia is the "combined use of several media, such as movies, slides, music, and lighting, especially for the purpose of education or entertainment" (Brooks, 1997, p. 17);
- Multimedia is "information in the form of graphics, audio, video, or movies. A multimedia document contains a media element other than plain text" (Greenlaw & Hepp, 1999, p. 44);
- Multimedia comprises a computer program that includes "text along with at least one of the following: audio or sophisticated sound, music, video, photographs, 3-D graphics, animation, or high-resolution graphics" (Maddux, Johnson, & Willis, 2001, p. 253).

In *Little Movies*,[83] Lev Manovich (1994), voices his speculation on digital minimalism or digital materialism. The following passage is revealing:

> I grew up in the U.S.S.R. where the material resources were quite scarce and I often had to travel from one end of Moscow to another because every art supply store would only carry a few colors at a time. So I would buy black paint in one store, get on the metro to travel to another end of the city, buy white and blue paints at another store, get on the metro again, and so on.
>
> Faced with the abundance of material and computation resources of the U.S., my reaction is to work against it. I don't need faster networks, more storage, more multimedia, more processing power. I want to figure out first what can be done with just a few pixels.

This basic property of digital media has a profound effect on the nature of visual and auditory realism. In a digital representation, all dimensions that affect the reality effect - detail, tone, color, shape, movement, pitch, timbre, etc. - are quantified. Quality, unlike in natural settings, has an immediate impact. Poor audio-visual fidelity is

[83] http://www.manovich.net/little-movies/

expressively alien. The distinctive background noise or diminished bandwidth on magnetic and digital media, for example, has no coincident or corresponding construct in an organic environment. The entire audio-visual domain that exists in this new technological medium typically extends from a defined artificial direction, specifically when the visual/symbolic representation suggests that it should be all around us. Of course, the new media artist or craftsperson has gone to great lengths to create more "natural" settings, using surround sound and IMAX-type technologies to create these manufactured environments.

As a consequence, the reality effect produced by the representation can itself be related to a set of numbers. In his essay "Aesthetics of Virtual Worlds"[84] (1995), Manovich discusses digitization - reducing all data to basic numbers - stating that "Digital media reduces everything to numbers." Manovich's minimalist approach compels him to assert that the computerization of culture leads to the spatialization of all information, narrative, and even time. His ideas on how we interact with computers, and that this, in all probability, will be our new way to work, communicate and play, can be viewed as a reaction to the potential control that the Internet (as digimachine/ digiclient) will have on our daily lives. Furthermore, by assuming that every viewer is a potential contributor, the Net gives us a misleading sense of creativity: Is the so-called creative and participatory nature of the Internet only number-deep? Are we just doing the paint-by-numbers shuffle? Is the fact that the Net and new-media phobia have swept the communication world an indication of a need for exploring new creative frontiers? Or is it that the facile nature of this new technology allows us to fit into a World Wide Cast?

[84] http://www.manovich.net/TEXT/virt-space.html

Evidently, the Internet can be instrumental in increasing access to education while likely improving the quality of users' learning experience; but if information technology is so enmeshed with existing programs – including a complete library of pre-assembled characters, 3D spaces and pre-fabricated sound-bytes - will we simply be working with templates and become reshaped avatars groping to keep up with the master charioteer? Very possible, since the complexity and difficulty of plans, assessments and decisions we make as unique individuals are explicitly tied to our personalities. Templates or canned procedures/models that are designed to answer all issues may lack a sense of perspective or creativity arising out of the tension between spontaneity and limitation; these pre-assembled templates may draw us into pre-determined arrangements and limit our ability to express our vision, which is essential to the work of art-making. On the other hand, templates used as direction or strategy that implement structure may actually help artistic constructs: the thoughtful and adaptable use of pre-assembled structures that reflect our individual experiences can be suggested by our usage of technology, and could very well be the door to a *gestalt opus*. Many of these new creative tools reflect this reality, (as indicated in Chapter 5): the ways in which media and technology are used by a diverse range of artists, and how this interaction will define society and culture, is to be seen; but the reality is that these tools permeate the neo-production industry of multimedia.

The common thread of these definitions "involves the integration of more than one medium into some form of communication...Most commonly, though, this term now refers to the integration of media such as text, sound, graphics, animation, video,

imaging, and spatial modeling into a computer system." (von Wodtke, 1993; Jonassen, 2000)

Multimedia, looked at through the lens of the above statement, is obviously time-dependent - that is, it contains at least one time-based element (audio or video); therefore, these structures include two or more media - at least one constant (time-dependent) and one distinct (time-independent) medium, and, consequently, information is both generated on a continuum and evolves in time. Today, multimedia designers have an assortment of various mechanisms for adding sound to their interfaces, from MIDI devices to software-based instruments and sounds (loops, sequences, and so on). Though MIDI is relatively unreliable in terms of predicting how the sounds will be heard on various individual computers (MIDI relies solely on the end-user's sound card for sound production), most of today's multimedia environments use actual wave files to generate sound, allowing designers to realize their auditory mappings with specific categories of timbre. Having control over the specificity of sound is also very useful as a communication channel for adaptive needs, and visually-impaired users can benefit greatly from this practice. Graphic User Interfaces (GUIs) will be enhanced by the use of audio icons: Kennel's (1996) AudioGraf technology,[85] using haptic devices[86] and audio indicators, enable visually-impaired users to identify graphic elements in a diagram;

[85] Kennel, A.R., 1996 AudioGraf: a diagram reader for the blind. *Second Annual ACM Conference on Assistive Technologies* pp. 51-56.

[86] Interestingly enough, the enhancements and progress of haptic devices (relating to or based on the sense of touch) may be due in part to the gaming industry, where the use of such devices is becoming commonplace – the gaming companies are using such devices to give more sensation to the game "experience."

furthermore, it was demonstrated that musical sequences could be used to communicate information about graphical objects. (Rigas & Alty, 1997)

Even though sound effects are a key component in most multimedia interface systems used today, music and sound design can be a central vehicle for the transfer of analogous information. The potential for well-designed audio to communicate information, whether computer-related or cross-cultural, needs to be attended to. The main instructional application of MIDI and multimedia software is as an excellent tool for sound design, providing students with a compelling way to capture their sound-design creations. Users can review their work, debug the errors, refine the timing and sequence, add supplementary parts, and create virtually unlimited arrangements.

Mediated Environments

> The world is for thousands a freak show; the images flicker past and vanish; the impressions remain flat and unconnected in the soul. Thus they are easily led by the opinions of others, are content to let their impressions be shuffled and rearranged and evaluated differently.
>
> J.W. von Goethe

The ability of the observer-participant to transfer knowledge between new contexts, through both direct and indirect experience, is greatly enhanced through familiarity with concepts. We can clearly predict that the use of the multimedia model, with its dual-channel coding and multi-modal system, can enhance the dominant idea that

adding multimedia functional designs to information constructs leads to improved transfer of information, such as:

- better retention, understanding,
- knowledge acquisition,
- impact of messages and entertainment.

The theoretical assumption of a virtual experience becomes essential as technological innovation in computers facilitates multi-sensory communication, since virtual occurrences and indirect experience are both part of a mediated environment (see Figure 5). In more creative and artistic conditions, virtual occurrences and indirect experiences can and should be contextualized, though at some point the framework must be abstracted in order for the observer/participant to become familiar with the concepts themselves, without the distraction from contextual details. These occurrences and experiences are desired in multimedia programs; in addition, bi-modality in creative systems allows for behavior modification: a way of responding to both the user's input and to other similar conditions. In other words, multimedia functional designs like these can result in exciting and unpredictable systems with a variety of outcomes, and assist the observer/participant in applying these ideas or concepts to new contexts.[87]

[87] Englemann, S. & Carnine, D. Theory of Instruction: Principles and Applications. NY: Irvington Publishers.

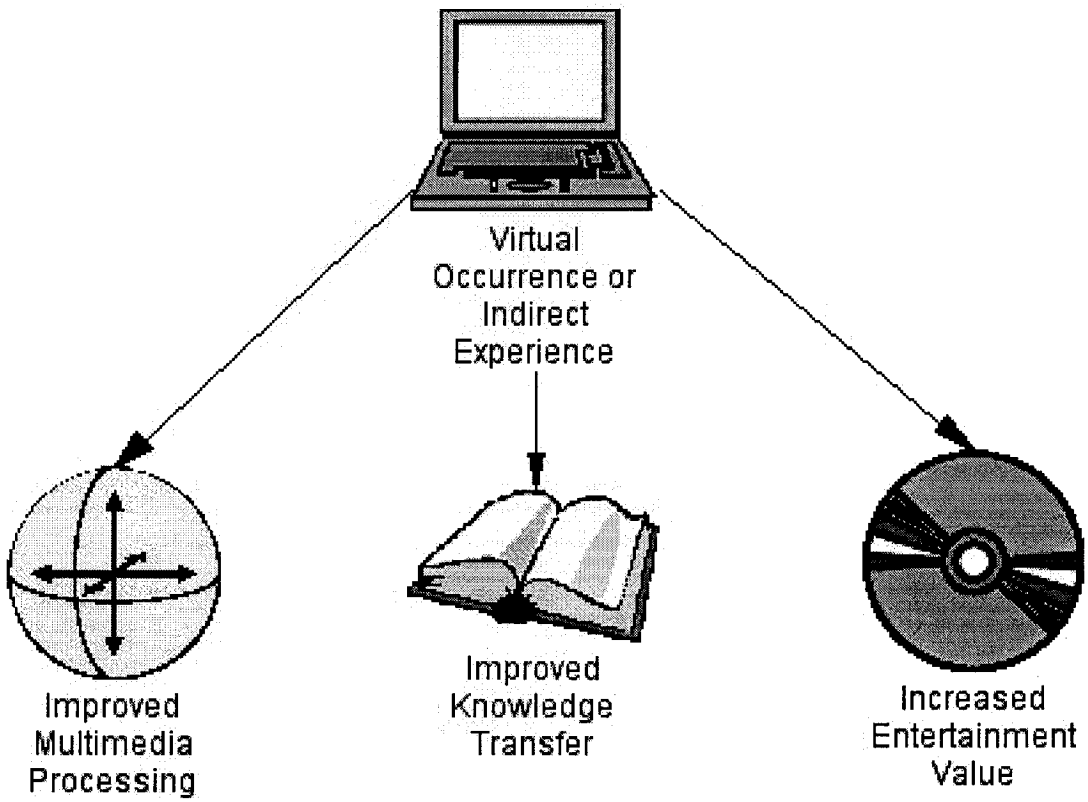


Figure 5 – Virtual Occurrence

In addition, we can articulate that each medium has its own unique qualities, and that, via the combination of these media, we may possibly give direction to a certain type of synergy. Audio is ephemeral and unrestricted - that is, once it is articulated, there is no record of it except in our mind's ear. We often associate words with intellect and music with sentiment;[88] graphics can be functional and practical or merely aesthetic: they can be used to communicate essential information (graphs, mind maps, etc.) or purely abstract representation; text can be descriptive, poetic, narrative, informative, and so on. Thus, the combination of elements stimulates our senses; if only parts of the possible

[88] For this discussion I will only compare music and text and leave sound effects for a later discussion.

combinations are in use, then we can use the free channel to augment or diminish a process (See Figure 6). This is the case, for instance, when we use audio warnings to identify an interruption or a risk during a given task.

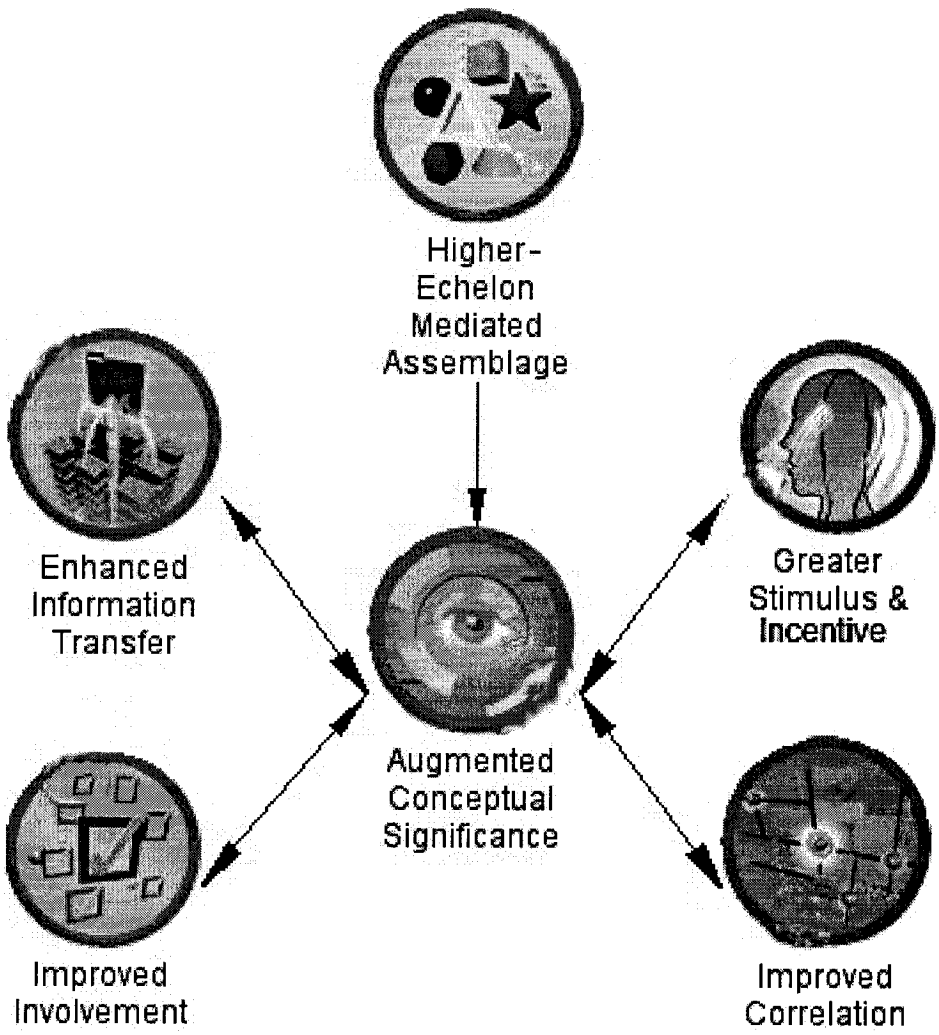


Figure 6 - ACS

Interface Crossroads

In multimedia, one must acknowledge the consequence of an active (interactive) user - most significantly, this activity must be meaningful for the user in order to enhance the disclosed process. Donna Haraway (1989) proposes that "Facts can be imagined as original irreducible nodes from which a reliable understanding of the world can be constructed. Facts ought to be discovered, not made or constructed. But the etymology of facts refers us to human action, performance - indeed, to human feats. Deeds, as opposed to words, are the parents of facts."[89] But in order for users to succeed in playing a meaningful, active and effective role in navigating within a specified set of parameters, they must be aware of what they are expected to achieve. Additionally, they must have a way of knowing whether or not they have accessed various echelons, especially when dealing with unfamiliar environments and/or subjects; this I believe can be resolved with the proper use of "metaphorical" interface and its complement, audio interface support. It is generally agreed that interfaces, for multimedia venues or other technologies, exhibit at least some of the following components:

[89] Haraway, Donna,1989; Primate Visions; New York: Routledge

- Metaphors: essential representations conveyed by means of graphics, text, audio, etc.; some of the more common models are taken from the analog idiom, such as bulletin boards, pages, video/film analogies;
- Mind maps: constructs of data, data bases, utility tasks, character assignments, role playing;
- Navigation: transaction by means of nodes, such as pull-down menus, dialogue boxes, tabs, control panels, icons, tool palettes, etc.;
- Interaction: input/output (I/O) and response tools, such as keyboards, mouse, pens, microphones, etc.;
- Graphic User Interface (GUI) or Desktop: using various analogies such as filing cabinets, folders, tree views, etc.

Auditory Interfaces

Interfaces for computers, PDAs (personal digital assistants), mobile phones and other machines dependent on monitors essentially manage information visually. Even though it is expedient to present information on monitors and 2D surfaces, icons and visual interfaces also have fundamental limitations. Screen size and area ratios are to be carefully considered, as the graphic user interface is size-dependent, and this is in direct correlation with the information that can be conveyed. With the current trend of always going towards the miniaturization of devices in our information technology society, space is a critical issue, and one way of reducing the visual footprint is to communicate miscellaneous information through the audio channel. This channel has the added advantage of not being predetermined as to direction - that is, the user does not need to use it as a focal point, and has the benefit of periphery. Mobile phones are used not only

for communication but also for information retrieval and transaction; the telephone interface is becoming vital to human-machine communication (Brewster, 1998). Visuals and sound are our two key senses of communication with the world at large, though sound is frequently relegated to a less significant role in order for us to direct our vision (Gaver 1997). But audio plays a much greater function in our understanding of the peripheral environment, and helps us adapt to it on a continual basis. Vision, an ambient array composed of rays from the immediate environment, scans the surface and is dependent on ambient light as a source of actualization. Hearing, on the other hand, tells us about the wider-range locality and the array of objects, and it communicates this information very quickly (Brewster, 1998). Sound is temporal in nature - that is, it has sequence; Gaver (1989) states: "Sound exists in time and over space, vision exists in space and over time." What we see are luminous and reflective objects: luminous objects are those that emit light—you can see them in the dark. Reflective objects are those that depend upon ambient light to be seen, and perception is affected by the amount of light present.

Sound has the advantage of not being bound by luminosity or locality; on the other hand, we cannot close our ears as we can our eyes. In our daily activities, audio and visuals interact naturally and supply complementary information for our everyday tasks. Most of our modes of communication have inherent limitations: emails cannot convey intonation, phone conversations cannot show expressions, and graphic user interfaces can harmonize this dual-channel modality. One past argument against using sound in the Graphic User Interface was that it potentially had the effect of annoyance and

intrusiveness; but with skillful design we can eliminate this incongruity. Current interfaces use audio only as cues, relegating the audio component to warning devices; also, audio feedback is restricted mostly to negligible beeps and short audio effects, unquestionably comprising the greater part of non-speech audio communication. We are used to hearing all sorts of alarms in our daily lives, from ambulances to fire detectors, and they all have a common goal of getting our attention at all costs, to supersede any other imminent ambient sound. However, good sound design has the potential to communicate, even with alert users, without disturbing their ongoing activity: it should be informative without being overly annoying.

What decisions should a sound designer make regarding the type of sound to use - high pitch, low pitch, long tones, short tones? Each one has a particularity that is inherent in its structure and will influence users in various ways. Our perception of timbre is elusive at best, though this field has been the subject of much research over the past four decades, including the use of multidimensional scaling (MDS) techniques, and other perception experiments, to determine that an altered signal could be distinguished from the original. Helmholtz (1877) pioneered research in this field, which to this day stands as the golden rule of acoustics. He hypothesized that timbral perception depends upon the spectral shape resulting from real-time frequency analysis on the basilar membrane. Clark, et al (1964) ascertained that timbre is determined by attack transients, modulation during the held state, and to some degree by one or more formants. Defining what exactly is or is not good sound design can be somewhat arbitrary: what one person

perceives as a pleasant sound is not what another perceives. Mitchell Akiyama[90], for example, has a modus operandi that makes use of a variety of subjective sound choices and manipulations; he speaks of his methodology and techniques in these terms:

> ...I'll do things like record myself playing guitar, I'll take a 2-minute excerpt of that, feed it into a patch in Reaktor[91] and do serious damage to it, and really explode it and deconstruct it and then take fragments from those sessions which I consider every bit as instrumental as the act of playing guitar because there are certain patches...I wish I was a programmer, and that I could program my own patches [as] a part of the gesture and that was a part of the process, but I don't have the time or the brain power to do that, so a lot of the patches I use I feel like I'm playing them like instruments, now so I know how I guide it to a certain level of expression. And then those improvisations on instrumental improvisations get recorded and re-imported into the sequencer, and then I might play along with the fractured guitar line. I'll play a new guitar line which might itself end up being really heavily treated, and it's just this constant layering of processes until I say that it's done.

Gaver puts forward two methods of listening to non-speech audio cues: "musical listening" and "everyday listening." He proposes that the "message" in musical listening is a consequence of the relationships of various acoustical constituents of a sound, such as pitch, rhythm, timbre, etc., whereas, in our everyday listening, we do not analyze what we hear in the same manner. We listen to the source of the sound and not its various acoustical characteristics. For instance, if we hear a motorcycle, we will most likely focus on its direction, the type of vehicle, or its velocity, whereas our listening to a

[90] Mitchell Akiyama,is a sound designer, Techno DJ, Producer, Videographer who has toured extensively world wide. See excerpts of the interview in Appendix B.
[91] Reaktor is a modular computer based sound studio. See http://www.nativeinstruments.de/index.php?reaktor_us for details.

similar sound in a modern sonic piece would result in a more "musical" analysis of the sound itself, such as, pitch, attack, loudness, duration. In today's perspective, we could as well speak of informational listening, active listening, evaluative listening - all modes of listening that vary greatly from "musical" listening and I will address some of these in the following section.

Auditory Ecologies

> Sick I am of idle words, past all reconciling,
> Words that weary and perplex and pander and conceal,
> Wake the sounds that cannot lie, for all their sweet beguiling;
> The language one need fathom not, but only hear and feel.
> - Georges du Maurier (1834-1896)

Gaver, Smith and O'Shea (1991) conducted an experiment (ARKola Simulation[92]) developing a simulated soft drink factory complex and creating a ground for experimenting with auditory icons, They dubbed the simulated soft drink ARKola, and named the plant accordingly. In this case, Gaver et al refers to "auditory icons" as

[92] Gaver et al (1991) designed an ecology of auditory icons which worked together to convey information about a complex and demanding simulation task, and observed users collaborating on it with and without sound. Results suggest that auditory icons can play a significant role in future multiprocessing and collaborative systems.

> ...everyday sounds designed to convey information from computer systems. They are based on the observation that environmental sounds are usually experienced in terms of the attributes of source events (e.g., size, material, and force) rather than in terms of physical attributes of the sounds themselves (e.g., pitch, volume, or duration). Most current uses of sound in the interface rely on arbitrary mappings between attributes of sound and the information to be conveyed. Closer, less arbitrary mappings may be created if attributes of sound producing events are used to represent attributes of computer events.

This is the strategy they used in developing their auditory icons, and also mention that, contrary to musical messages, auditory icons are often less annoying because they can be designed to be a counterpart to an existing ambient auditory environment. In describing the design of the ARKola experiment, Gaver et al formulate the following example:

> In understanding how to design an ecology of sounds for the ARKola factory, it is useful to consider the sound made by an automobile engine. On the one hand, people often experience automobile sounds as a unity, listening to whether the engine is running well or poorly. On the other hand, if something breaks the resulting noise is likely to be perceptually salient in the overall engine sound. Not only will the engine as a whole sound as if it is not running correctly, but people (especially automobile mechanics) can hear what has gone wrong. The sounds made by automobile engines are undesigned, yet related to their sources in lawful and perceivable ways. In designing auditory icons for computer systems, we have the freedom to choose only those sounds which will be functional, to retain the close mapping between sound and source, and to shape the sounds acoustically to be discriminable. In this way, an ecology of sounds can be designed that can be heard together as an overall plant noise or attended to separately to obtain information about individual machines.

This statement by Gaver et al demonstrates that sound can be designed to complement, even harmonize, a mostly graphic and/or tactile environment, but doing so obviously demands some forethought. Gaver and Smith (1990) give an informed

description of the makeup of *auditory icons* designed for auditory environments and targeting everyday listening. Some attributes of sound design based on musical listening – pitch, timbre, locality - are musical counterparts to graphic design elements - line, color, perspective.

One of the key attributes of multimedia audio production is that we can communicate and construct messages without using the visual channel. Evidently, written-word or graphical icons must be seen to be understood; audio communication is conveyed regardless of where one is looking. We all live in a visual world, and we tend to state our understanding of things by saying "I see," and acknowledge that "a picture is worth a thousand words." Thus, faced with a multi-modal situation, our foremost principle of evaluation will be that of the visual. However, sound dynamically models our understanding of the visual channel - it is a biological continuum necessity, capable of informing us of details that are far beyond the reach of our visual domain, and alerting other senses as well. We become skilled at correcting small differences in our various environments, but this instinctive process does not take effect when the sound is focused by choice: these small discrepancies become annoying when we focus on purposefully created sound settings, such as film soundtracks, theme songs, and so on. These sounds are produced by a specific source that envelops the addressees, whereas the optic array from a screen, for example, can be shut out by closing the eyes. Sound also gives a sense of time in a contiguous spatial environment; but a sound source is not always readily identifiable, and can be cause for concern, though it can also be treated to the advantage

of the designer. Therefore, audible material requires a greater recognition of time-frame than the visual object.

And so what does a proper use of the bi-modal element in a multimedia environment achieve? The interaction of these independent variables (sound and image) seem to give expected theoretical responses such as:

- possible enhanced recognition, as one sense can complement the other;
- possible enhanced level of participation, interest and concentration;
- possible enhanced level of stimulation, using bi-sensual modes (auditory and visual perception systems);
- possible enhanced emotional stimulation, as the combination of sound and image can stimulate various processes.

Figure 7 shows a hypothetical interaction between multimedia dual-channel prerequisites (left) and communication/entertainment value (right).

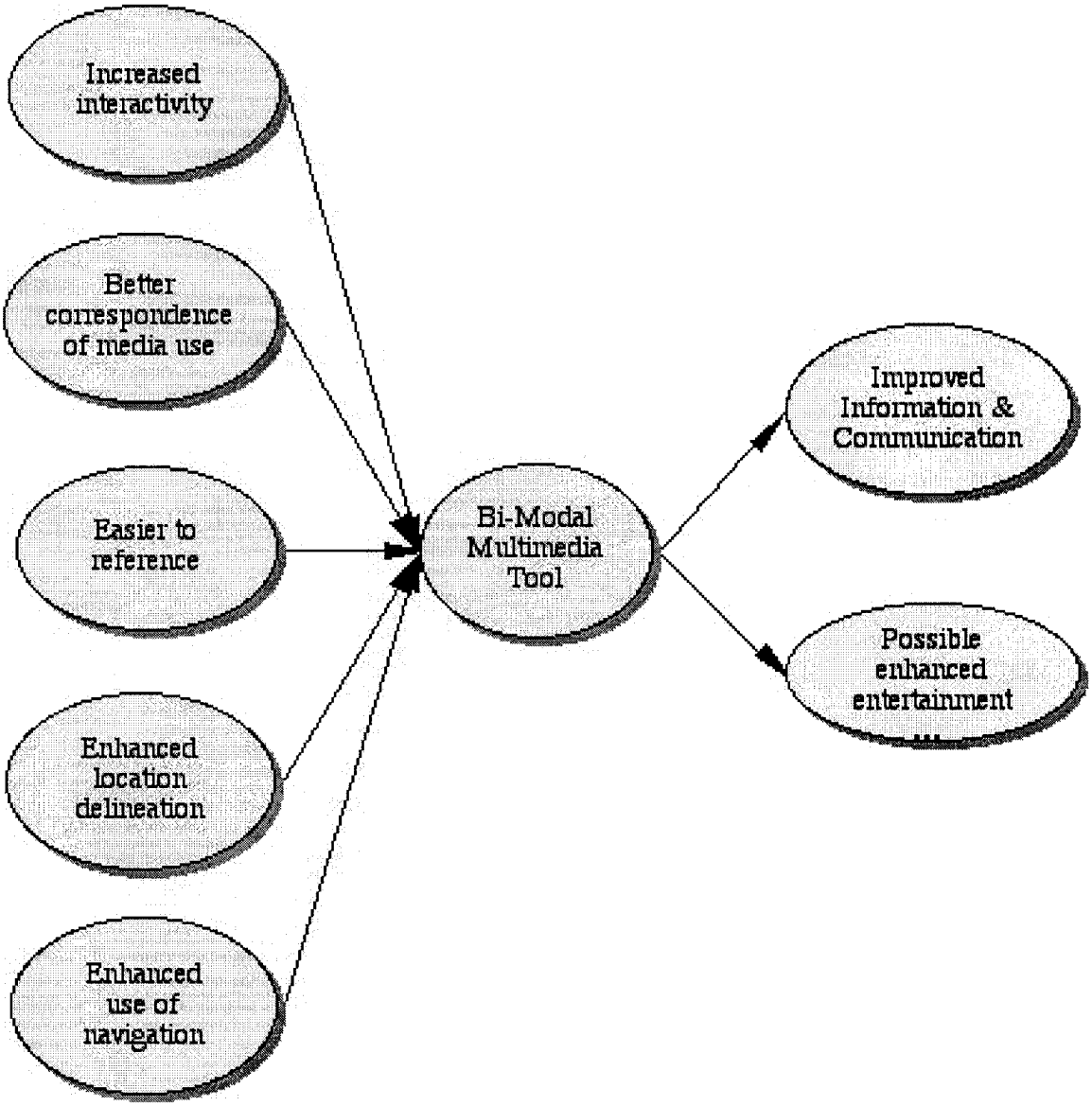


Figure 7 – Multimedia Communication Interaction Chart

In the previous sections, I looked at multimedia and the social implications of technology, as possible venues for allowing us to envision worlds fashioned from our dynamically charged recollections, ideas, and dreams. But what can make available to us that "modus operandi," translating these dreams into tangible effects? There are numerous options that are available and I will look at some of these in the following section.

CHAPTER V:

MULTIMEDIA PRACTICE - Current state of the art

"The best way to predict the future is to invent it."
- Alan Kay

Audio/Vision Convergence

It is only recently that computer processing power has reached a level enabling us even to consider synthesis of multi-channel sound and sound imaging - that is, software that can generate, sequence and sample sound on multiple channels and also use various algorithms to process images and translate them into sounds. Owing to the deep integration of television and computer-related technologies into society, we have transferred our textual mode of communication to a visual mode, and subsequently toward a more interactive genre of communication tradition. Authors have also transferred their locus of attention from purely visual or audio media to more network-oriented media such as television, video, web and computer applications, thus enabling a multimedia ethos to take hold.

In his online article, Andy Huntington (2002) looks at the development of music composition software and how the digital world and interactivity have changed the way people create and consume. This article, while not abundant in detail, gives an interesting insight that pulls together different ideas of multimedia, music, computer technology and creativity: "Where, before, musicians used one set of machines for their

task and painters used another, in the new media industry all use the same tool, the computer." Bi-modal programs, such as *MetaSynth*, *ArtMatic*, *Videodelic*, are all software-based programs that mix image production with audio synthesis; *MetaSynth*, for example, is a highly effective audio-vision tool, it allows one to "draw sound" or insert images which are then interpreted by the software into musical compositions; the developers (U&I Software) describe it thus: "*Metasynth* is an electronic music composition and sound design program [which] can create an astonishing variety of sound and music. Whether you are creating cutting-edge electronica or neo-classical masterpieces, *MetaSynth* will expand your sonic palette in ways you never imagined." (See Figures 8 & 9)[93]

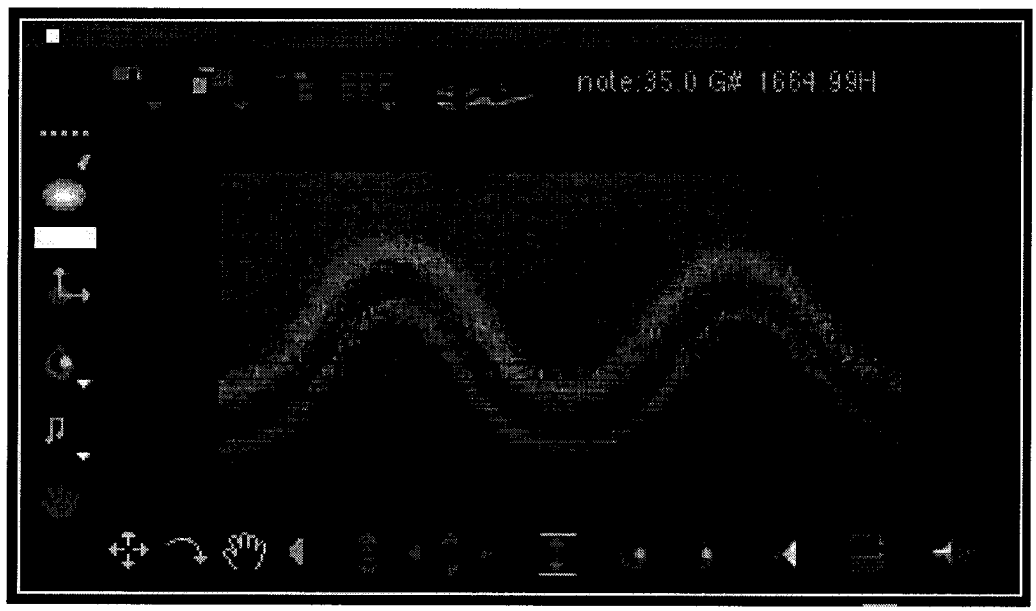


Figure 8 – *Metasynth* Interface[94]

[93] Examples taken from templates of the Metasynth Software tool.
[94] Inside the central area, you can create a drawing (or insert an existing one), and this in turn will be analysed and interpreted into sound.

Figure 9-Tree[95]

U&I has also created a software called *ArtMatic*, which "...lets you create incredible pictures, video effects, textures, animation and sound..." *ArtMatic* generates astounding graphics and sound using a tree of tiles that represent graphic functions and filters. While *ArtMatic* makes use of sophisticated mathematical components, no knowledge of mathematics or computer graphics is required to create a remarkable array of images. (See Figures 10 & 11)[96]

[95] A picture of a tree began this example. The sound is computed using a pure sine wave in 32 divisions of a tone, and with a rather slow tempo (about 134 ms for each pixel).
[96] Ibid

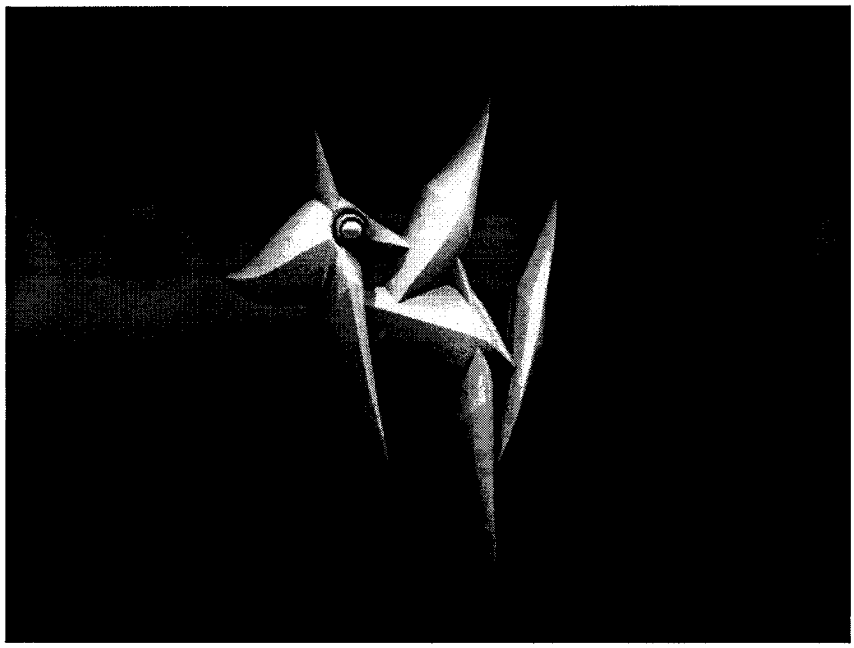


Figure 10 – *ArtMatic* Example 1

Figure 11 - *ArtMatic* Example 2

One interesting aspect of these intermedia tools is that, in contrast to the strictly visual arts or music, for which theories of representation are well established and more sophisticated, multimedia or intermedia works, as a form of cultural and social representation, do not have a locus of existence and lack the authority of a significant critical discourse. Authors whose works are exclusively created within a multimedia or intermedia model are conscious of the possibilities or limitations[97] brought about by the inherent properties of the medium. But contrary to the more established art forms such as painting and sculpture, multimedia works are not bound to a fixed space and can be disseminated and distributed, allowing the work to be experienced at numerous sites and at various times, or simultaneously via the web. Moreover, what the user experiences is not a representation of the work, but the work itself.

Logic, for example, is a composition and arranging tool with video input, allowing one to compose music as one watches the images unfold, and offers modern software synthesis; this tool also has various embedded synthesizers that generate any sound, that can be processed and shaped using plug-in effects enabling one to transform the audio in every way possible. Some new media artists like Fredo Viola, who has been experimenting with video and music since the 70's, uses *Logic* to record and compose at his home studio, and distributes his music online. Using the 15-second video function on his digital still camera and Adobe *AfterEffects*[98], Viola produced visual accompaniment

[97] I have added "limitations," as all things do have limitations, but it is the author's view that multimedia works have the potential of breaking through many of the limitations encountered by other forms.
[98] *AfterEffects* provides 2D and 3D compositing, animation, and visual effects tools: http://www.adobe.com/products/aftereffects/main.html

for his piece, *Sad Song.* This online[99] multimedia work is composed of five screens side by side in a panoramic view, each broadcasting different images simultaneously. The result is a captivating marriage and separation of both moving image and music, an example of rhythmic montage that illustrates Chion's idea of *synchrony,* to the point at which, as listener and viewer, we automatically and intuitively make connections between the emotional value of the music and the images of nature, of landscapes rushing by, and of the artist himself. *Logic*, as well as *Metasynth & ArtMatic*, are legitimate authoring tools that use the power of the modern CPU to assist in creating genuine works of art; they combine sound synthesis and sound processing into the same structural assemblage, and by altering the various control parameters, the synthesis design can offer a very sophisticated and accurate production of sound material. These tools do have a steep learning curve, while other software is designed to give a full spectrum of prosumer tools, be it sound or visuals, allowing the creation of broadcast quality audio-visuals with a minimum of technical or theoretical competence.

With the advent of digital sampling technology and *drag & drop* [100] music software,[101] any media artist with any such interest can create concrete formal compositions, limited only by their imagination. Post-analog technology has provided an opportunity for media producers, educators and artists to download ready-made packets

---

[99] See Viola's work at: http://www.zonezero.com/magazine/essays/viola/viola.html

[100] I use the term *drag & drop* throughout the work to exemplify object-oriented programming (OOP) software, designed to facilitate user interaction; using this type of software, the user need only drag various elements around to achieve simulated programming or virtual 'creation' effect.

[101] Software programs such as *Band-in-the Box*, *GarageBand*, *Live*, *JamXmatrix*, *Orchestrator* and the impressive array of fully-featured sequencers that allow an untrained person in the art of composition to assemble meaningful but sometimes "spontaneus" sound designs.

of musical composition and use sound austerely, allowing them to bring emphasis to their artistic conceptions by using this consequential pro tem media,[102] as a result augmenting the relevance of this "mix-media" analysis. As a composer, multimedia professor and new media artist, I have long observed the lack of convincing and compelling research into this unique area. Nevertheless, this "mix-media" context can, should and does provide a venue for various artists to apply a consistent environment to their work, and needs to be addressed, as some of the interviews (Appendix B) with artists currently using these venues will attest.

The increasing facility (user-friendliness) of production in this age of digital communication, when countless infrastructures are being cybernetized and musical form is burdened with the task of satisfying an array of diverse settings: Does technology and its facile modes of production show a detrimental or positive effect on artistic content, style or form? With regard to the various trends of *drag & drop* technology: the new multimedia artist working with new technologies is drenched in information and streams of data, altering, downloading, re-assigning, emailing, broadcasting, pitch-changing, and so on. However, with the advent of this *drag & drop* sound design software (*Band-in-the Box*, *Garage Band*, *Live*, *MP3* and MIDI files on the Web, or the fully-automated sequencers and self-contained PCs), any interested sound designer can create or download ready-made packets of musical compositions. By following the complex route which technology now takes within the arts community, we can reveal much about the distinctive interaction between the aspiration for detail, the need for knowledge and the

[102] The idea of having to integrate an element or craft that one is only moderately adept in is a notion that is becoming increasingly accepted in our contemporary technocratic art setting.

manipulation of techniques. This *drag & drop* technology, of course, like any phenomenon that centrally reflects technological development, can be regarded either as progress and added value or a step towards lethargic modernism. It is the opinion of this writer that individual artists will progress in whatever fashion suits their style; just as the addition of mechanical devices (keys) added to a tube full of holes allowed the modern flutist to expand his sonic horizons, so can *drag & drop* technology facilitate the development of brand-new musical structures for our contemporary sound designers.

I shall look at some instances in which *drag & drop* sound design tools and their integration are used to produce multimedia and/or intermedia[103] art, how these popular tools have made possible the integration of visual elements with sound, and how they can subsist fundamentally as a fused entity. These *drag & drop* software tools, along with other user-friendly tools, are shaping the modish multimedia environments. Both multimedia and intermedia use spatial and temporal objects within a two-dimensional space and are used within mostly computer-based (CD-ROM or Web) and video environments and their various interfaces.

An iota of meaning and content may help us surpass the rudimentary "push-button syndrome" and ceaseless input of data. However, the underlying analysis of music - analysis that has been central to the production of new musical forms - has not advanced as far as the techno-development: my view is that, as a result, our aesthetic understanding

[103] "Intermedia" (named and defined by D.Higgins: quoted in 'Color me Synesthesia', *Leonardo Journal* 32 #1, 1999, p7), is the combination of structural elements or syntax from different media into one. Contrary to "multimedia," where content is presented in more than one medium simultaneously, Intermedia can be included in a multimedia environment.

of sound design in general has possibly been weakened. I shall survey and analyze the assortment of some *drag & drop* programs and their evolution, considering how they can be adapted to assist novice sound designers in a positive fashion.

Garage Band and *Soundtrack* are excellent examples of this type of software. To understand this fast-track expansive evolution of audio software in our present-day setting, I will briefly examine these two, as they are representative of the more typical advanced computational models of sound synthesis environments. *Garage Band* is available as part of the new *OSX* on *Mac* computers,[104] and this is truly a *drag & drop* application with pre-recorded sounds and easy-to-use interface. Apple says about this software: "You don't have to play the piano. You don't have to read music. You don't even have to have rhythm. If you know what you like when you hear it, you can make your own kind of music."

Huntingdon (2002), mentions Sherry Turkle's A Tale of Two Aesthetics, and her assertion that new interfaces and *drag & drop* – type software enable

[104] What is of note here is that these software programs are now widely available at no or very little cost, wheras only 5 to 8 years ago they would have cost thousands of dollars (if available at all).

> ...a new breed of non-programmers to use computers without having to concern themselves with the minute processes at work: the "user." In her work, users are, for some reason, perceived only as a rather lackluster group, seduced by the computer's own capacity rather than as a means to express the user's capabilities. She contrasts this group with two others, the highly technical *hackers* and the homebrew *hobbyists* with their reductionist desires to penetrate the heart of the machine - groups that parallel those in music. The user is neither the amateur *hobbyist* at their piano or a virtuoso *hacker*, but rather one who seeks to "use" music rather than to understand its fundamental qualities and construction, whether at home or in the concert hall.

Though, I guess it is feasible to "use" music, as Turkle suggests, given that the software supplies many tools, such as embed pre-recorded sound bites, a variety of styles, etc., I would assume that a more creative use of these software programs are the norm for most resourceful users. Also, given that our powerful digital computational capability, the signal-processing functions of our modern-day computers allow sounds to be synthesized from concise mathematical specifications, and can render any musical combinations available at the click of a mouse, it can actually nurture and support a high-quality inventive process.

As an example, *Soundtrack* is a software application that allows film and video editors to develop their own musical scores with instrumental loops and to add sound effects (see Fig. 12). With a design similar to sound-editing programs such as *Peak*, *Logic* or *ProTools*, *Soundtrack* uses the conventional multi-track metaphor to visually represent the layering of different instruments in a composition. This software is truly user-friendly and selections include instrumental soloists, full orchestrations, common sound FX for film and video, and even some vocal loops. Each track is identified by an

easily recognizable icon to the left of the timeline (e.g., an acoustic guitar, a drum kit and so on). *Soundtrack* allows control of musical characteristics such as key, tempo (bpm) and positioning within a bar, and will automatically keep tracks in sync and in key, without concern for modulation or other aspects of music theory. It also allocates a track where one can see the video while scoring, in real time. This type of functionality was available only to very high-end state-of-the-art studios just a few years ago: the program was directed toward film and video makers without the time or money to employ a professional composer, or to secure the rights to copyrighted material. Sadly, this is still a reality in many production situations, where producers often do not give adequate consideration to audio until it is too late - at least this would appear to be the case in smaller productions. This is apparent in much of the *Soundtrack* tutorial material online, with articles introduced in such a way as to suggest that, "We know you probably haven't thought about audio yet, but look at this new tool that will make your life easy..." Titles include: "Music To Video Editor's Ears," "Let There Be Music" and "How to do the Blues," a tutorial that guides users through the composition of a basic blues soundtrack which, when looped, makes "a nice blues bed for that documentary about southern living you're working on!"[105]

[105] Reference URL: http://www.apple.com/soundtrack/

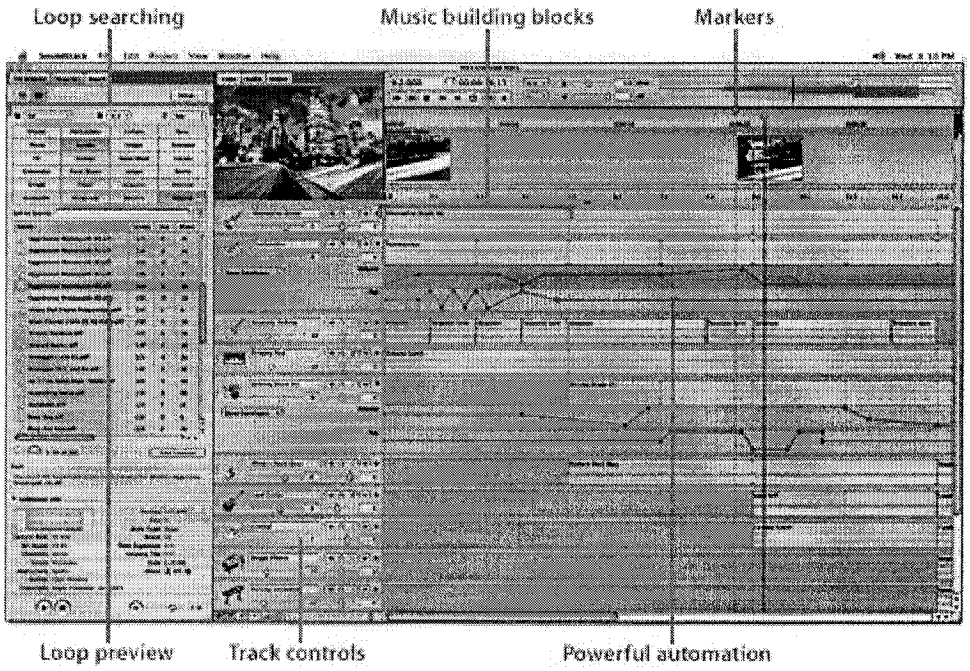


Figure 12– Sound Track

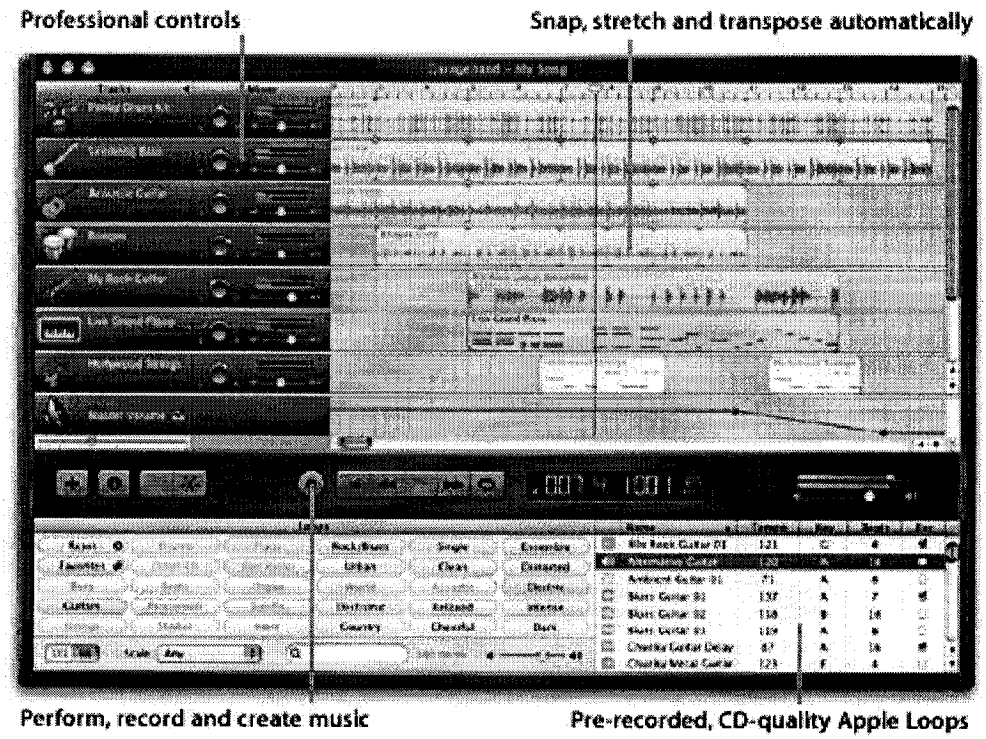


Figure 13 - Garage Band

As with *Garage Band* (see Figure 13), the idea behind *Soundtrack*, as well as the language used with online tutorial texts, suggests that composing music can be trouble-free and effortless, especially when there are other tasks that require attention (such as shooting or editing a film, for example). But are there elements of the soundtrack that are being sacrificed for speed and ease of use? With these types of software, users are restricted to pre-designed generic loops that must be matched with their images. If a user has no musical background, or has little concept of audio composition, it might be difficult for her/him to develop a complex or dynamic piece of audio to accompany the image. In a more ideal scenario, working with a composer or a sound designer, makers can begin with sounds and instruments that are better suited to their images, and create melodies that are inspired by and intimately connected with the picture. Collaborating with sound designers allows producers to refocus their energy on the visual, and to benefit from a fresh pair of creative ears. Moving beyond a *drag & drop* software also provides the possibility for more sound sources that are not necessarily musically motivated, such as field recordings. Granted, within these software programs, there is a reasonable selection of loops to choose from and, arguably, endless possibilities as to how they could be transformed and combined; but the danger lies in image makers' settling for less than ideal sound - in believing that the simple act of choosing a loop and drag and dropping it onto their work is enough to make a good soundtrack.

There is also the interesting issue of instant recognition, where audience members - or other audio and image artists - recognize a loop or particular sound and immediately attribute it to *Soundtrack* or other such software. Even though there are

different options for sounds (over 4000 loops in some cases), it is conceivable that two filmmakers would choose the same loop for their aforementioned "*southern living*" documentary, for example. In his online review at *creativecow.net*, Ron Lindeboom[106] actually claims the opposite, that "Because the songs you make are built up of bits and pieces and not lengthy patterns or progressions, your own use of these same bits and pieces used by others have little likelihood that your music will sound the same as anyone else's work." However, just like cliché images in film and video, there are audio clichés and image-sound relationship conventions that may unconsciously inform our design choices. In an informal interview,[107] Communications graduate Craig Desson (Concordia University) admitted rather ashamedly to using *Soundtrack* in several of his more recent video works. Desson was "underwhelmed" by the final product of his tinkering, noting the creative limitations of *Soundtrack*'s loop selection, and dismissing the "music" he had produced as a simple relative to what could have been composed by a musician. However, he explained that with his practical musical knowledge, his familiarity with similar programs - and given his temporal and financial constraints - it was just simpler for him to produce a soundtrack almost instantaneously with the application in question, rather than spend the time with a composer. In this digital era, the copy/paste syndrome is attractive to time-budgeted producers, and the maker must avoid the impulse to make too evident a mark, invisibility usually being the greater virtue.

106

http://www.creativecow.net/show.php?forumid=1&page=/articles/lindeboom_ron/soundtrack/index.html

[107] We have conducted a series of informal interviews with local artists and educators. I will cite parts of these in this chapter, but the full content of the chosen interviews can be read in Appendix B.

Nevertheless, *drag & drop* software offers the maker the capability of being intimately included in an overall design.[108]

Artistic works of this type can possibly have value by their ability to let one create without prior knowledge of the theory or tools; but undesired elements may appear that are not intended by the author, about which he/she can do nothing, due to lack of familiarity and understanding. Spatial or temporal depreciation and distortions can occur. It would seem that if we introduce discrete elements ad hoc, a continuum could be hard to achieve, and we would not be able to reproduce a pre-designed original idea completely from these pre-programmed snippets, therefore creating a string of artifacts. It seems that we could simulate the final idea by interpolating these pre-determined loops and sample data, but common sense - and experience - suggest that it would be only an approximation. I would call this phenomenon "process artifact" - the way the pre-constructed creation, using elements that are not altogether original, assumes a function that is inherently different from a truly original piece, in which all the compositional elements are structured from the ground up, or from source materials. In music, many such materials are often identified as familiar or unfamiliar, which means that the listener usually organizes the musical setting into sequences that carry musical events based on recognizable sound sources; if the "process artifact" is too predominant one may get the impression of *déja vu*, but in relation to disconnected and possibly incoherent audio-visual events. One element that a "process artifact" piece and the "original" work do have

[108] One aspect worth noting is that scoring work with this type of "prosumer-level" tool also gives users access to a wealth of technological support and peer advice via online email lists and forums, such as Creative Cow's *Soundtrack* forum.

in common is that both make a statement by illustrating and displaying ideas, albeit in very different types of authoring.

Artifacts and Reality

Between the idea
And the reality
Between the motion
And the act
Falls the shadow.

- T.S.Eliot ('The Hollow Men')

Artifacts demonstrate the ideology and customs of a particular culture. In this technology-laden age, time and space have become cultural artifacts. Artifacts alter practices, abilities, competencies, ways of thinking. In an interview, Dr. Andra McCartney, Professor of Sound in Media for the Communication Studies Department at Concordia University in Montreal (she is also a *Soundwalk* artist), discusses the dependence on formal training in order to create audio works and various other aspects that influence the contemporary artist.[109] Though Dr McCartney refers mostly to audio conditions, many of her comments can be applied to audio-visual settings. We first asked her about the importance of a musical background: "I don't think it's necessary at all...I think it is helpful because, for instance, *Studio Vision*[110] had a very heavy kind of musical metaphor that it was using for setting things up so there would be a score..." It became clear that certain elements of theory and music training were key, and she continued:

[109] See Appendix B for the complete interview. Interviews were conducted by my research assistant, Jessica Landry.
[110] Sound sequencing software.

"When I say no understanding, I mean no formal understanding of harmony, because I think people often have a very strong informal understanding that they may not realize they have." We then asked about the relationship between non-trained and trained artists still producing a similar product of comparable quality:

> I think it's great! I think often when people do start working with sound editing and sequencing software, it creates in them a desire to learn more about music, because they realize that they know a certain amount but they need to go further in certain areas; and I'm also aware that there are many really accomplished film music composers who will very openly say that they have no musical training.... it's very possible for people to do musical work without musical training and after doing those ten years, and more than ten years of piano performance training, I felt like in many ways, I was being restricted as to what I could do rather than opened up... Many people will pick up a pencil or a paintbrush and try painting a picture even though they don't have artistic experience; but then it has traditionally been much less widespread that people would pick up an instrument and say I'm going to compose something. And I think that this kind of software makes it fun to start composing and then people can listen to what they're doing and start to improve what they're doing...

On the subject of multimedia, Dr McCartney concurred that there is still a gap in the multimedia model, that there is no real multimedia aesthetics. She adds:

> ... in relation to multimedia, I think there's a lot less that's been done in terms of aesthetics; [there is] more in relation to film sound, radio or electronic music, for instance, than there is in terms of multimedia... partly because it's so new. The web for instance had really limited sound possibilities up until five years ago, so that's not enough time for a body of writing to start.... I think at the moment the best possibilities for learning about multimedia sound are mostly on the web itself.

Later in this study, I will point out how the role of independent variables that facilitate a multimedia effect needs to be reconsidered, and how digital technology has done much in the way of energizing multimedia projects and hopefully promote multimedia aesthetics. When we asked about the usefulness of "Earcons" on the web, Dr McCartney offered: "I do find them helpful. They'd be particularly helpful for somebody who had limited vision as well. It must get quite dull for somebody with limited vision who wants to use the Internet...I think there are so many sites that are [visually interesting], and yet the sounds have not been developed as much as they could be."

Dr McCartney examines the role of technology and human agency within the changing nature of musical composition. Performance and consumption, and in our day the "interactivity" of digital media, are often key factors that guide the various transformations and modes of communication in various artistic projects, possibly making us more aware of "representation," while pulling us in the direction of virtual reality and reproduction. I would suggest that, instead of exploring media arts in the form of imitation of nature, and its effects on conventional artistic genres, digital media can be utilized to put the locus of attention on the perimeter of the virtual and the

concrete, without trying to re-do what has been done before, but establishing new settings to probe the assets of a dynamic musical experience.

Interactivity Myopia

> A blind man cannot see how a timepiece is designed, or what distinguishes it from other clocks. Still, he may see that, if it is a clock at all, it will embody certain dynamical principles; and may explain the action to his young apprentice. The latter, however keen his vision, can only describe the perturbations of the clock...[111]
>
> - N.R. Hanson

One primary problem for designers of multimedia is: How are audio sensations rendered in a mainly visual environment? Every designer goes about this task using a variety of idiosyncratic approaches. In Chapter Three, I identified aspects of meaning of significant consequence to sound designers: *referentialism* - conveying an approach to extramusical objects, ideas, and actions; and *expressionism* - involving musical meaning as essentially intramusical (i.e., areferential). External to all functional considerations, this re-association of visual and audio elements is key in the sound-to-image relationship, generating skillful tensions and stimulating the inner psyche by what we see on the screen and what is activated in our mind's ear. Referential and expressionist aspects alike focus upon significance within both the auditory and visual modalities; significance is obtained by the recognition of experiences that represent varying degrees of abstraction. Chion (1994) calls this "rendered sound," causing a sound to appear more real than reality. Chion also states that "... audiovisual media do not just address the eye. They place their

[111] Hanson, N. R.,1985; Patterns of Discovery; Cambridge University Press, p.4

spectators – their audio-spectators – in a specific perceptual mode of reception…" which he calls "audio-vision." The ear functions in the time dimension, following (just slightly behind) a sound event, and hearing in slices across that event's timeline (like the sampling function on a CD or DVD). The eye also provides spatial information, and in a sense is limited by the nature of what it perceives: Chion points out that images in a film cannot be layered as successfully (or inconspicuously) as with sound. Intricate extensions of sound and image in motion pictures take advantage of the multimedia capabilities of this medium. In the larger context of everyday listening and viewing, observations about the seemingly natural relation between sound and the projected image are a direct result of how people interact: while it is easier to assume that film sound merely complements or accentuates meaning already present in the image, Chion stresses that the way we perceive and formulate filmic meaning depends entirely on the "construction" of image and sound as determined by the makers. This construction is at once informed by and contributes to a set of artistic and cultural conventions. Multimedia and emerging technologies take advantage of this relationship between image and sound: the personal computer is the instrument of choice as the predominant manipulation device for these new media and offers the potential of control and of so-called "interactivity"; these are often simply elements of playback and navigation in a predetermined structure.

Notwithstanding the predetermined nature of mediated environments, nothing is a given, and sound or sonic elements can offer *added value*. This concept of *added value*

makes clear how well-designed sound can create the illusion of its own superfluity, Chion[112] states:

> On one hand, music can directly express its participation in the feeling of the scene, by taking on the scene's rhythm, tone and phrasing; obviously such music participates in cultural codes for things like sadness, happiness and movement. In this case we can speak of empathetic music, from the word empathy, the ability to feel the feelings of others.

"Meaning" of a shot (or scene) is not found solely in the image or in the sound, but in the content created by the interaction of this bi-modality within the viewing experience. Is it possible to unite the two by first separating them? While considering sound as an individual entity and construction, we can still give it the consideration and analysis otherwise lacking in the study of film theory and aesthetic. There is still an enormous amount of potential to shape, manipulate and, in effect, destroy the expression of filmic time through image (e.g., *Memento, The Limey, Eternal Sunshine of the Spotless Mind*). Digital image editing has done much in the way of liberating the look of film or multimedia projects (since the time of Chion's publication), and digital sound editing could be considered a contributor or a result of the liberation. The role played by so-called interactive attempts at addressing the challenges of the mind's eye/ear correlation have shown that the ear analyzes, processes and synthesizes faster than the eye, and therefore abrupt sounds make more of an impress on our senses than quick movements. Because of sound's inherent temporality, it is more obvious when sound is played backwards, for example, than when an act such as playing piano occurs backwards in a film.

---

[112] Chion, M. (1994). *Audio-vision. French.* New York : Columbia University Press,. (P.8)

Sound is almost always imposed, in the sense that we are "always listening" – we are not as well-trained to block out unwanted aural stimulation. Consider the small number of people who actually wear earplugs to a bar or club vs. people who will wear sunglasses on bright days. Chion writes: "Sound more than image has the ability to saturate and short-circuit our perception." In film (unlike video art, installation work, etc.), the image occupies one place, *the frame*, much like a computer screen in more contemporary thinking. Sound has no container other than the listener's immediate hearing range, and this allows for the layering of tracks in film and various multimedia art forms; also, there is creation of narrative spaces, in the opposition between "on and off" screen sound, or, as Chion calls these two spaces, *visualized* (the sound's source is visually present or understood) and *acousmatic* (the source is hidden, abstracted or existing completely outside the diegesis of the film). Film is an example of the potential for "spatial magnetization of sound by image," i.e., when we hear a sound, we try to locate it, determine its "source."

Chion's audiovisual view is no longer unique to film, or even to television, but can be applied to the new media environments created by interactive installations, gaming, and the internet, many of which are designed around the technology of the personal computer. In some ways, when we consider our engagement with these types of media products, our real everyday experience takes the place of the narrative that would normally develop onscreen; our constant interaction with our surroundings, despite focusing our attention on the images at hand (whether they are videos, video games, or

the internet), creates a different listening/viewing experience than one might have in a movie theatre, for example. This is not to suggest that computer spaces are any less immersive, but only that other factors contribute to our immersion (e.g., the computer being dependent on our movements and action to function). As mentioned in Chapter 3, musical meaning can be highly referential (i.e., pointing to something outside of the music itself) or just as highly areferential (i.e., the meaning is inherent in the musical tones themselves and their relation to one another). As in the visual domain, other echelons of significance may be carefully planned as artifacts relevant to a multiplicity of areferential and referential attributes (see Figure 14). Considering multimedia as an audio-visual code, it is possible to place any specific arrangement in a field representing the dimension of the audio-visual construct; the effects of multimedia functionality on audio- vision processing would result in knowledge and information transfer. The degree of understanding is therefore high in both the visual and the audio domains, since the meaning results from a well-synchronized activity of sound and image.

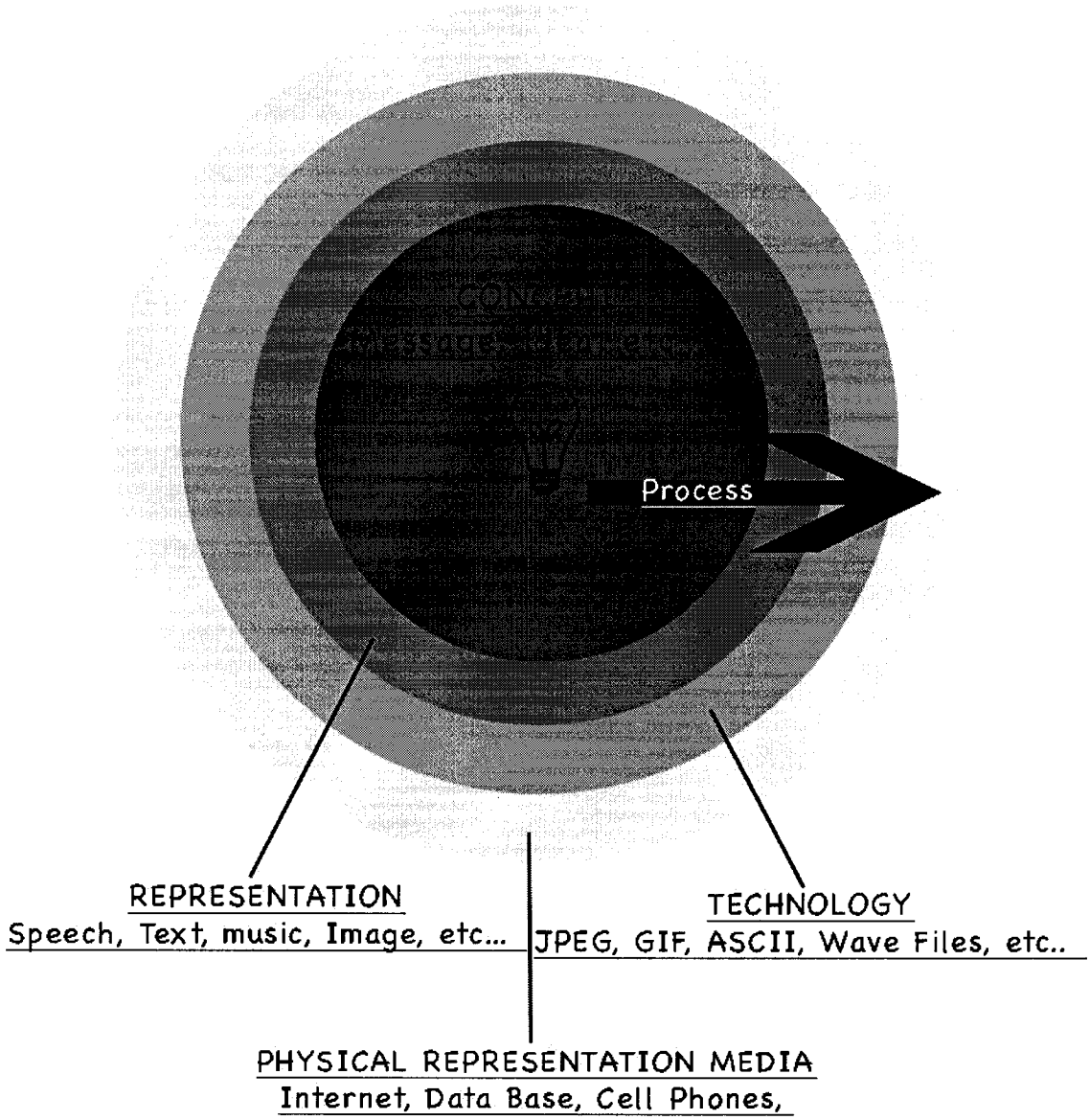


Figure 14 - Concept and Representational Media

CONCLUSION

In the introduction to one of his later works, Info-Aesthetics, Lev Manovich[113] declares: "Twenty-first century has arrived. What comes after modernism, postmodernism, and new media? Welcome to INFO-AESTHETICS." Manovich is one of our most prominent thinkers about new media theory and critic of the information society; yet, as many of his texts indicate, he rarely addresses the sound issue. He nevertheless intends to debate twenty-first-century hybrid media, but still shows a lack of the comprehensive multi-modal debate one would expect from an information-age thinker. For example, in the above mentioned "post new media" work, INFO-AESTHETICS, he declares:

> If the 1920s avant-garde came up with new forms for new media of their time (photography, film, new printing and architectural technologies), the new media avant-garde would introduce radically new ways of using already-accumulated media. Thus, the new avant-garde is the computer-based techniques of media access, manipulation and analysis. In other words, information society may not need new visual languages, new forms and new representational techniques because it can use computers to re-configure the old ones in radically new ways.

So it seems that even a modern thinker of new media technologies would still have a tendency to employ antiquated visual symbolism along with traditional textual information. However, while these traits are often coupled with the added function of interactivity, bringing to the forefront a multimedia model, one cannot comprehend why

[113] http://www.manovich.net/

the assimilation of the bi-modal functions of sound and image into a synthesized schema is not a standard process. It still emerges as and continues to be a rhetorical construct.

Is it that unfamiliarity, compounded by a long tradition of study in visual arts, makes the examination of bi-modality in new media potentially unattractive for many scholars? Nonetheless, as demonstrated in the previous chapters, the comprehensive synthesis of these two traditional extant art forms, sound design and graphics, to construct a joint experience has grown exponentially since the talkie films. But, from the increased appeal of home computers as communication and entertainment machines, they have combined into a multi-modal/multimedia effect. Multimedia relies on concept and its expression rather than on advanced audio-visual representation. It is embodied in the experience rather than the production. By illustrating the complex route through which technology infiltrates not only the public at large but also our social strata - even the learning community, we can reveal much about the distinctive interaction between the aspiration for specificity, the need for knowledge and the manipulation of process.

The contemporary author (teacher, instructional designer or media producer) working with new technologies is saturated with data that he can organize or regulate without limit, but an iota of meaning and content delimited within an audio environment may help us surpass the rudimentary pushing of buttons and ceaseless input of data. Sound can contribute to the structural outline and give identity to an otherwise amorphous context. The underling analysis of bi-modality within a multimedia environment - analysis that has been integral to the production of new media forms - has

not caught up with techno-development, and consequently the aesthetic understanding of newer forms of the audio/visual relationship has stalled. Conversely, it would be just as unrealistic to suggest that multimedia (of whatever form) could return us to a form of *media practica* - that is, doing away with mediated environments, human computer interfaces, or even mobile technologies - as it would be to say that the current media would remain impervious to the mode. New modes grow out of the old and may give direction, but would seldom eradicate them.

What remains to be seen is the change in the role of the production and the *creative practicum*, which multimedia and its use of bi-modality could bring. Contemporary communication and information delivery systems have improved over the years and the ability for users to access information has reached an unprecedented phase. As technology blends into a homogeneous, mediated system, multimedia, including the proper use of bi-modality, will unquestionably progress outside the periphery of what has been explored here. New structures, artistic concepts and multi-user experiences will be the norm, and the acceptance of these new concepts will bring achievements bridging the established arts and *media practica*. We tend to grasp concepts through experience; and we also make connections between the various concepts by categorizing - in audio-visual research, the "expectancy mechanism" (Borowsky & Besner, 1993). A "multimedia theory of learning" such as the Dual Coding theory, proposes that learning is more effective when the learner uses more than one sense-modality, such as verbal and visual processing, and when connections are clearly made between the information contained in each modality (Mayer, 2001).

We know that today our modes of accessing information can take the form of graphics, video, animation, multi-channel audio, web-based data and interactive CD-ROMs; the network possibilities, through hypermedia systems, have pushed the multimedia models to the vanguard of information system design. For the past few years, an array of sound design software to create multimedia tools, from the very complex (*Logic*, *Live*, etc.) to the basic (*Garage Band*, *Band in the Box*, etc.), have been used in a variety of projects many of us have developed and managed, with great success.

Multimedia authors have long recognized the strength of lifelike sounds and voices, even when linked to visuals of average quality. It would be my recommendation that designers produce a modus operandi for engaging environments that originate with audio. They should try to become aware of the various soundscapes, to preserve the familiar sounds, and to be certain that the aural information comes from the most effective sources, just as they would if the listeners were experiencing them in real time. For designers of multimedia, audio and its relation to visuals can have a significantly positive effect on the evolution of the construct. Audio can sometimes be easier to manipulate and alter than visuals, and there are simpler audio-fidelity solutions. But even if the yield of improved vision and sound were *at par*, the outlay suggests often starting with audio, as sound design can very often assist and contribute to the visual effect. Furthermore, even putting aside the technical quality of images, good sound design can cause users to think that the visuals actually look better. It is perhaps in the hybrid nature of multimedia that this synthesis becomes most appealing: each of these discrete traits is

present in the structure of their manifestation and with which we are all familiar - we understand and inhabit bi-modality. This amalgamation of the attributes of audio-vision is not synthetic - it would be more precise to say that the isolation of one modality from the other would be blatantly artificial.

This dissertation has suggestions concerning how to evaluate the theoretical status of seeing and hearing within multimedia systems and mediated environments. It is not my intention to promote either sense to a position of central or key role, or to give prescriptive directions on how multimedia should function. The foremost inquiry is not which is more important, but more accurately how each one works within a synthesized environment. In this study, we have looked at the theoretical concepts dominant in multimedia authoring using sound and image synthesis, and the importance of developing a conceptual approach to sound design as related to visuals. As there is a vast amount of material in a typical multimedia project, different users would have varying needs and should be able to position themselves at any number of standpoints in the sound-design process. The intricacy of the sound-designed mediated system lies in the fact that sound is not an object like a painting or a photograph, but an event - unlike these objects or the written word, it is a conceptual phenomenon that stays malleable within our imagination long after the event has taken place.

BIBLIOGRAPHY

(1997). *Handbook of human-computer interaction/ edited by Martin G. Helander, Thomas K. Landauer, Prasad V. Prabhu.* Amsterdam ; New York.

Alkin, E. G. M. (1973). *Sound with vision: sound techniques for television and film.* London [for] the British Broadcasting Corporation: Butterworths.

Appleton, B. (1997). Patterns and Software: Essential Concepts and Terminology. *Object Currents Magazine*, Vol. 3 No. 5, May 1997.

Appleton, Jon H. and Ronald C. Perera, *The development and Practice of Electronic Music* , Englewood Cliffs, N.J., Prentice-Hall, 1975.

Arenson, M. (1982). The Effect of a Competency-Based Computer Program on the Learning of Fundamental Skills in a Music Theory Course for Non-Majors. *Journal of Computer-Based Instruction,* 9 (2), 55-58.

Armbruster, B (1996) Schema Theory and the design of content-area textbooks. *Educational Psychologist, 21,* 253-276.

Attali, Jacques, (1985) *Noise: The Political Economy of Music*, Minneapolis, University of Minnesota Press.

Ausubel, D. P. (1967). *Learning theory and classroom practice* --. Toronto: Ontario Institute for Studies in Education.

Baddeley, A.D., (1992). Working Memory. *Science*, 255, 556-559.

Bajo, M. T., Burton, A., Burton, E., & Canas, J. J. (1994). Word recognition across orthographies: Another look at the interaction between context and degradation. *European Journal of Cognitive Psychology*, 6(2), 171-193.

Barrass, S. (1997) *Auditory Information Design.* PhD thesis submitted to The Australian National University.

Barry, C., Hirsh, K.W., Johnston, R. A., & Williams, C.L. (2001). Age of Acquisition, Word Frequency and the Locus of Repetition Priming of Picture Naming. *Journal of Memory and Language*, 44, 350-375

Barthes, R, (1985). "Musica Practica" in The Responsibility of Forms: Critical Essays on Music, Art, and Representation (Richard Howard -Trans.) Berkeley: University of California Press.

Bazin, A. (1971). *Aesthetic of Reality:Neorealism.* Berkeley: University of California Press.

Bergson, H. (1922*) Creative Evolution*, trans. Arthur Michell; New York: Holt.

Bordwell, D. (1989). *Making meaning : inference and rhetoric in the interpretation of cinema.* Cambridge, Mass: Harvard University Press.

Bordwell, D., & Thompson, K. (1986). *Film art : an introduction.* New York: Knopf: Distributed by Random House.

Borowsky, Ron & Besner, Derek. (1993). Visual Word Recognition: A Multistage Activation Model. *Journal of experimental psychology. Learning, memory & cognition,* 19 (4), 813-840.

Bransford, J. (2000). *How people learn : brain, mind, experience, and school /* John D. Bransford ... et al., editors ; Committee on Developments in the Science of Learning and Committee on Learning Research and Educational Practice, Commission on Behavioral and Social Sciences and Education, National Research Council. Washington, D.C.: National Academy Press,.

Bresson, R. (1977). *Notes sur le cinematographe.* New York: Urizen Books,.

Brewster, S. (2003) *Non-speech auditory output.* The Human-Interaction Handbook. Eds. Julie A. Jacko & Andrew Sears. New Jersey: Lawrence Erlbaum Associates Inc.

Brewster, Stephen A. 1998, Using Non Speech Sounds to Provide Navigation Cues. *ACM Transactions on Computer-Human Interaction* 5:3: 224 –259.

Brooks, D. W. (1997). *Web-teaching: A guide to designing interactive teaching for the World Wide Web.* New York: Plenum

Bruner, J. S. (1966). *Toward a theory of instruction.* Cambridge, Mass.: Belknap Press of Harvard University.

Burch, N. (1973). *Theory of film practice* / translated [from the French] by Helen R. Lane. London: Secker & Warburg.

Busoni, F. (1911) *Sketch of a New Aesthetic of Music, in Three Classics in the Aesthetic of Music;* New York: Dover Publications Inc. (1962)

Chandler, Daniel (1994): Semiotics for Beginners [WWW document] **URL** http://www.aber.ac.uk/media/Documents/S4B/ [*Accessed 04-07-09*]

Chandler, P. & Sweller, J. (1991) Cognitive Load theory and the Format of Instruction. *Cognition & Instruction*, 8, 293-332

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293-332.

Chion, M. (1985). *Le son au cinema*: Paris: Editions de l'Etoile

Chion, M. (1994). *Audio-vision. French.* New York : Columbia University Press.

Cholmsky, P. (2001). *Using Modular Architectures Within Distributed Learning Environments: A Means for Improving the Efficiency of Instructional Design & Development Processes.* Concordia University: M.A. thesis

Clark, J.M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review,* 53(2), 445- 459.

Cook, Guy (1992): *The Discourse of Advertising.* London: Routledge

Cooke, D., (1959); *The Language of Music*; London, Oxford University Press

Crease, M. G., & Brewster, S. A. (1998). *Making Progress With Sounds - The Design and Evaluation Of An Audio Progress Bar.* Paper presented at the ICAD 94: 2nd International Conference on Auditory Display

Cuban, L. (1986). *Teachers and machines: the classroom use of technology since 1920 / Larry Cuban.* --. New York: Teachers College Press,.

Cytowic, R. E. (1989). Synaesthesia: a union of the senses. New York: Springer-Verlag.

Cytowic, R (1995); *Synesthesia: Phenomenology And Neuropsychology A Review of Current Knowledge* http://psyche.cs.monash.edu.au/v2/psyche-2-10-cytowic.html

Day, S. (1996); *Synaesthesia and Synaesthetic Metaphors* http://psyche.cs.monash.edu.au/v2/psyche-2-32-day.html

Dent, H. Jr. (1998). *The roaring 2000s: Building the wealth and lifestyle you desire in the great boom ahead.* New York: Simon & Schuster.

Dewey, John (1916*) Democracy and Education*, The Free Press, NewYork, NY.

Dick, W., & Carey, L. (1990). *The systematic design of instruction.* Glenview, Ill: Harper Collins Pub.

Eisler, H. (1947). *Composing for the films. London*: Dennis Dobson.

El-Saddik, A. (2001). *Interactive multimedia learning: shared reusable visualization-based modules /* Abdulmotaleb El Saddik. Berlin; New York: Springer.

Ellis, A. W., Kay, J., & Franklin, S. (1992). *Anomia: Differentiating between semantic and phonological deficits.* In D. I. Margolin (Ed.), Cognitive neuropsychology in clinical practice (pp. 207P228). New York: Oxford Univ. Press.

Ertmer, P., & Newby, T. (1993). *Behaviourism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. PFI, 6 (4),* Essays on the sound pattern of English / edited by Didier L. Goyvaerts and Film sound : theory and practice / edited by Elisabeth Weis and John Belton. --. Oxford ; New York : New York: P. Lang, Columbia University Press,.

Eveland J. D. and L. G. Tornatzky (1990) *The Deployment of Technology: The Processes of Technological Innovation* in L. G. Tornatzky and M. Fleischer. Lexington, Massachusetts, Lexington Books

Forbus, K. D., & Feltovich, P. J., (1947)- (Eds.). (2001). *Smart machines in education : the coming revolution in educational technology /* edited by Kenneth D. Forbus and Paul J. Feltovich. Menlo Park, Calif.: Cambridge, Mass.: AAAI Press; MIT Press.

Gambrell, Linda B. & Brooks Jawitz, Paula. (1993). Mental imagery, text illustrations, and children's story comprehension and recall. *i* 28 (3), 264-276.

Gardner, H. (1982). *Art, mind, and brain: a cognitive approach to creativity.* New York: Basic Books.

Gardner, H. (1983). *Frames of mind: the theory of multiple intelligences.* New York: Basic Books.

Gardner, H. (1985). *The mind's new science: a history of the cognitive revolution.* New York: Basic Books.

Gaver, W. W., & Smith, O'Shea, T. (1991) *Effective sounds in complex systems:* the ARKOLA simulation Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology New Orleans, Louisiana, United States: 85 - 90 (1991)

Gaver, W. W., & Smith, R. B. (1990). *Auditory icons in large-scale collaborative Environments; Proceedings of* Interact'90, 735-740. D. Diaper et al. (eds.) Amterdam, The Netherlands: North-Holland

Gaver, William W. (1989). The Sonic Finder Human-Computer Interaction 4:1. *Elsevier Science.*

Gaver, William W. (1997). *Auditory interfaces.* http://www.billbuxton.com/AudioUI04interfaces.pdf

Gessner, R. (1968). T*he moving image; a guide to cinematic literacy.* New York: Dutton.

Getzels, J. W., & Csikszentmihalyi, M. (1976). *The creative vision : a longitudinal study of problem finding in art.* New York ; Toronto.

Gould, Glenn; 1984: *Prospects of Recording* in the Glenn Gould Reader, Ed. Tim Page,

Greenlaw, R., & Hepp, Eß. (1999). *In-line / On-line: Fundamentals of the Internet and the World Wide Web*. Boston: McGraw-Hill.

Gros, B., et al. (1997). Instructional Design and the Authoring of Multimedia and Hypermedia Systems: Does a Marriage make Sense? *Educational Technology, (*, 37)1, 48-56.

Hanlon, L. (1986). *Fragments : Bresson's film style*. London: Fairleigh Dickinson University Press.

Hawkes, Terence (1977): *Structuralism and Semiotics*. London: Routledge

Heyrman, H. (1995). "Art & Computers: an exploratory investigation on the digital transformation of art. Tele-Synaesthesia, presentation of a hypothesis". Doctoral thesis, Universidad de La Laguna, Santa Cruz de Tenerife, Spain,. See: http://www.doctorhugo.org/e-tsyn.htm

Helmholtz, Hermann L F (1877) *On the sensations of tone as a physiological basis for the theory of music.* Dover Publication, New York. 1954. Translated by Alexander J. Ellis from the 4th German Edition of 1877.

Hevner, K. (1935). Expression in music: A discussion of experimental studies and theories. *Psychological Review 42(2)*, 186-204.

Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology,*, *48*, 246-69.

Hoffman, S. (1997). Elaboration Theory and Hypermedia: Is there a Link? *Educational Technology,*, (37)1, 57-64.

Housen, A. (1983). *The eye of the beholder [microform] : measuring aesthetic development.* Doctoral Thesis. Boston, Mass.: Harvard Graduate School of Education.

Huntington, Andy (2002). *An Analysis of the Success of Interactive Music For The Home Computer.*
http://www.soundtoys.net/a/journal/texts/andyhuntington/inde.html

Innis, Robert E (Ed.) (1986): *Semiotics: An Introductory Reader.* London: Hutchinson

Jonassen, D. H. (1996). *Computers in the classroom: mindtools for critical thinking / David H. Jonassen.* Englewood Cliffs, N.J: Merrill.

Jonassen, D. H. e. a. (1997). Certainty, Determinism, and Predictability in Theories of Instructional Design: Lessons from Science. *Educational Technology, (37)1*, 27-34.

Jonassen, D. H., (1999). *Learning with technology: a constructivist perspective / David H. Jonassen, Kyle L. Peck, Brent G. Wilson.* Upper Saddle River, N.J. :Merrill.

Karl, G. & Robinson, J. (1997) "Shostakovitch's Tenth Symphony and the Musical Expression of Cognitively Complex Emotions," (in Music and Meaning, Robinsion J. Editor)

Keirns, J. L. (1999). *Designs for self-instruction : principles, processes and issues in developing self-directed learning / Johanna L. Keirns.* Boston: Allyn and Bacon.

Kemp, J. E., & Dayton, D. K. (1985). *Planning and producing instructional media.* New York: Harper & Row.

Kennel, A.R., (1996) AudioGraf: a diagram reader for the blind. Second Annual ACM Conference on Assistive Technologies pp. 51-56.

Kivy, Peter (1980). *The Corded Shell.* Chapters 1 & 12. New Jersey: Princeton University Press, 1980.

Kivy, Peter (1991); Is Music an Art? *Journal of Philosophy* 88 (October 1991; 553)

Kivy, Peter (1993); *The Fine Art of Repetition* (Cambridge University Press

Kivy, Peter. (1984). *Representation and Expression in Music.* Princeton University Press.

Kolb, D. A. (1994). *Experiential learning: experience as the source of learning and development.* Englewood Cliffs, N.: Prentice-Hall.

Kracauer, S. (1960). *Theory of film; the redemption of physical reality.* New York: Oxford University Press.

Kracauer, S. (1961). *Nature of film; the redemption of physical beauty.* London: Dobson.

Krathwohl, D. R. (1993). *Methods of educational and social science research: an integrated approach.* New York ; London: Longman.

Kubey, R. W., & Csikszentmihalyi, M. (1990). *Television and the quality of life: how viewing shapes everyday experience.* Hillsdale, N.J.: L. Erlbaum Associates.

Landauer, T. K. (1996). *The trouble with computers : usefulness, usability, and productivity / Thomas K. Landauer.* Cambridge, Mass.: MIT Press.

Lindsay, R. B. (1973). Acoustics: historical and philosophical development. Edited by R. Bruce Lindsay.

Logan, R.K (1986) *The Alphabet Effect,* Morrow, New York

Louis G. Tornatzky and Mitchell Fleischer, eds., *The Process of Technological Innovation* (Massachusetts/Toronto: Lexington Books, 1990), 7.

Lumsden, J., Brewster, S., Crease, M., & Gray, P. (2002). *Guidelines for Audio-Enhancement of Graphical User Interface Widgets.* Paper presented at the Procedings of the 16th British HCI Conference, London.

Maddux, C., Johnson, D., & Willis, J. (2001). *Educational computer: Learning with tomorrow's technologies.* Boston: Allyn and Bacon.

Main, R. G., (1997). Integrating Motivation into the Design Process. *Educational Technology, (33)12,* 37-41.

Manvell, R. (1975). *The technique of film music.* London [etc]: Focal Press, 5.

Marks, L. E. (1978). *The unity of the senses.* New York and London: Academic Press.

Marshall McLuhan, *Understanding Media: The Extensions of Man*, (New York: McGraw-Hill, 1964), 358.

Martin, B. L. & Briggs, L.J. (1986). *The affective and cognitive domains : integration for instruction and research.* Englewood Cliffs, N.J.: Educational Technology Publications.

Maslow, A. H. (1971). *The farther reaches of human nature*. New York: Viking Press.

Maslow, A. H., Cox, R., & Frager, R. (1987). *Motivation and personality*. New York: Harper and Row.

Mayer, R. E. (2001). *Multimedia learning*. Cambridge, UK: Cambridge University Press.

McLuhan, M. (1964) *Understanding Media: The Extensions of Man*: London:Routledge

McKoon, G. (1977). Organization of information in text memory, *Journal of Verbal Learning and Verbal Behavior*, 16, 247-260.

Meijer, P. B. L. (1992). An Experimental System for Auditory Image Representations. *IEEE Transactions on Biomedical Engineering, 39, No. 2*, 112-21.

Moore, B. (1992). Music, Technology, and an Evolving Curriculum. *NASSP Bulletin*, 76 (544), 42-46.

Moorhouse, A.C. (1953) *The Triumph of the Alphabet*. New York: Henry Schuman,

Munro, J. (2002): *Visual-Spatial Giftness* http://www.edfac.unimelb.edu.au/LED/ELAGE/documents/GLT-Visualspatialgift.pdf

Norman, D. A. (1976). *Memory and attention: an introduction to human information processing / Donald A. Norman.* --. New York; Toronto: Wiley.

Norman, Donald A. 1990. *The Design of Everyday Things*. Doubleday.

Nuckolls, J.B. (1999). The Case for Sound Symbolism. *Annual Review of Anthropolgy;* 28, Social Science Module

Pavio, A. (1986 Mayer, R. E. (2001). *Multimedia learning*. Cambridge, UK: Cambridge University Press.

Pavio, A. (1986) *Mental representations: A dual coding approach.* Oxford, England: Oxford University Press.

Paivio A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of psychology, 45(3), 255-287*

Peirce, Charles Sanders (1931-58): *Collected Writings* (8 Vols.). (Ed. Charles Hartshorne, Paul Weiss & Arthur W Burks). Cambridge, MA: Harvard University Press

Percival, F., & Ellington, H. (1984). *A handbook of educational technology*. New York, NY: Nichols Pub. Co.

Perfetti, C. A., Beck, I., Bell, L., & Hughes, C. (1987). *Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children.* Merrill-Palmer Quarterly, 33, 283-319.

Peters, D. G. (1993). Computer-Based Music Skills Assessment Project: A Portal to Artistic Innovation. *Bulletin of the Council of Research in Music Education*, 38-45.

Piaget, J. (1977). *L'equilibration des structures cognitives. English.*

Plato. *The Republic of Plato*, translated by Benjamin Jowett. Oxford Clarendon Press; (On line edition: http://www.fordham.edu/halsall/ancient/plato-republic.txt)

Prendergast, R. M. (1977). *A neglected art. Film music: a critical study of music in films.* New York: W. W. Norton.

Prendergast, R. M., & Ferguson, M. (1980). *Psychology of the TV image.* New York: Praeger.

Pudovkin, V. I. (1960). *Film technique, and Film acting.* New York: Grove Press.

Pullum. Geoffrey K. Chicoutimi Quebec: Ghent: G. Morin, E. Story-Scientia,. Gerasimov, V., Bender, W. (2000). Things that Talk: Using Sound for Device-to-Device and Device-to-Human communication. IBM Systems Journal. (39) 3/4

Ranzijn, F. J. A. (1990). *The instructional design for the acquisition of concepts / F.J.A. Ranzijn.* Amsterdam: Bloomington, Ind.: Thesis Publishers.

Raths, L. E. (1969). *Teaching for learning.* Columbus, Ohio: C. E. Merril Pub. Co..

Reenen, P. T. v. (1982). *Phonetic feature definitions: their integration into phonology and their relation to speech: a case study of the feature NASAL /* Pieter van Reenen. --. Dordrecht, Holland ; Cinnaminson, N.J.: Foris Publications,.

Reigeluth, C. M. (1996). A New Paradigm of ISD? *Educational Technology.* (36)3, 13-20.

Reigeluth, C. M. (1997). Instructional Theory, Practitioner Needs, and New Directions: Some Reflections. *Educational Technology.* (37)1, 42-47.

Reiguluth, C. M. (1979). In Search of a Better Way to Organize Instruction. *Journal of Instructional Development*, 2 (3), 8-15.

Reiguluth, C. M., & Stein, F. S. (1983). *The Elaboration Theory of Instruction.* In C. M. Reiguluth's (Ed.)., Instructional Design Theories and Models: An Overview of their Current Status.; Hillsdale, NJ: Lawrence Erlbaum.

René Berger, *Art and Space Age Communities*, in René Berger and Lloyd Eby, eds. *Art and Technology* (New York: Paragon House, 1986), 394.

Richey, R. C. (1997). *Agenda-Building and its Implications for Theory Construction in Instructional Technology.* Educational Technology. (37)1, 5-11. Rigas, D.I., Alty, J.L., 1997. The Use of Music in a Graphical Interface for the Visually Impaired. In: Howard, S., Hamond, J., Lingaard, G. (Eds.), Proceedings of Interact '97, Sydney. Chapman & Hall, London.

Rogers, P. L. (Ed.). (2002). *Designing instruction for technology-enhanced learning.* Hershey, PA: Idea Group Pub.: Information Science Pub.

Rowland, G., Parra, M.L., & Basnet, K. (1994). Educating Instructional Designers: Different Methods for Different Outcomes. *Educational Technology, (34)6*, 5-11.

Russcol, Herbert, *The Liberation of Sound* : *An Introduction to Electronic Music* , Englewood Cliffs, N.J., Prentice-Hall, 1972 1) p.32 2) p.32 3) p.38 4)pp.37-38

Schiffman, S. E., Reynolds, M. L., & Young, F. W. (1981). *Introduction to multidimensional scaling: theory, methods, and applications.* New York: Academic Press.

Schwartz, J.E., & Beichner, R.J. (1999). *Essentials of educational technology.* Boston: Allyn and Bacon.

Seels, B. (1997). Taxonomic Issues and the Development of Theory in Instructional Technology. *Educational Technology, (37)1*, 12-21.

Seels, B. (1998). *Making instructional design decisions* / Barbara Seels, Zita Glasgow

Simms, Bryan R. (1996) *Music of the Twentieth Century: Style and Structure*, Second Edition. New York: Schirmer Books; London : Collier Macmillan.

Smith, P. L., & Ragan, T. J. (1993). *Instructional design.* New York: Merrill,.

Snodgrass, J. G., & McCullough, B. (1986). *The role of visual similarity in picture categorization.* Journal of Experimental Psychology: Learning, Memory, and Cognition, 12, 147-154.

Spiro, R. J., Bruce, B. C., & Brewer, W. F. (Eds.). (1980). *Theoretical issues in reading comprehension : perspectives from cognitive psychology, linguistics, artificial intelligence, and education / edited by Rand J. Spiro, Bertram C. Bruce, William F. Brewer.* Hillsdale, N.J.: L. Erlbaum Associates,.

Spiro, R.J., Coulson, R.L., Feltovich, P.J., & Anderson, D. (1988). *Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains.* In V. Patel (ed.), Proceedings of the 10th Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Erlbaum. [available at http://www.ilt.columbia.edu/ilt/papers/Spiro.html]

Spiro, R.J., Feltovich, P.J., Jacobson, M.J., & Coulson, R.L. (1992). *Cognitive flexibility, constructivism and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains.* In T. Duffy & D. Jonassen (Eds.), Constructivism and the Technology of Instruction. Hillsdale, NJ: Erlbaum.

Spottiswoode, R. (1950). *A grammar of the film; an analysis of film technique.* Berkeley: University of California Press.

Sweller, J. (1999). *Instruction design in technical areas.* Camberwell, Australia: ACER. New York: Viking Press.

Thomas, T. (Ed.). (1979). *Film score : the view from the podium.* South Brunswick [N.J.]: A. S. Barnes.

Turner, Graeme (1992): *British Cultural Studies: An Introduction.* New York: Routledge

Vanechkina, Irina l. (1994) "Musical Graphics" as an Instrument for *Musicologists and Educators, Leonardo*, vol.27, no.5, pp.437-439.

von Wodtke, M. (1993). *Mind over media: Creative thinking skills for electronic media.* New York: McGraw-Hill.

Walton, K. 1997; *Listening with Imagination: Is Music Representational?* (in Music and Meaning, Robinsion J. Editor)

Webster, P. R. (1990). Creative Thinking, Technology, and Music Education. *Design for Arts in Education,* 91 (5), 35-41.

Weis, E. & Belton, J. (Eds.). (2002). *Look with the ears: Charles Tomlinson's poetry of sound* / New York: Columbia University Press.

Weis, E., & Belton, J. (Eds.). (1985). *Film sound: theory and practice*. New York: Columbia University Press.

Wiley, D. A. (Ed.). (2002). *The instructional use of learning objects*. Agency for Instructional Technology: Association for Educational Communications & Technology,.

Willman, F. (1992). New Solutions to Curricular Problems. *Music Educators Journal,* 79 (3), 33-35,68.

Wilson, B. G. (1997). Thoughts on Theory in Educational Technology. *Educational Technology, (37)1*, 22-27.

Winn, W. (1997). Advantages of a Theory-based Curriculum in Instructional Technology. *Educational Technology, (37)1*, 34-41.

Wishart, Trevor. 1996. *On Sonic Art*. Revised ed. Contemporary Studies Series vol.12. Ed. Simon Emmerson. Amsterdam: Harwood Academic.

Yeston, M. (1976). *The stratification of musical rhythm*. New Haven, CT.: Yale University Press.

Zaza, Tony. (1991) *Audio Design: Sound Recording Techniques For Film and Video.* Englewood Cliffs, New Jersey: Prentice Hall, Inc.

Zettl, H. (1990). Sight, sound, motion: applied media aesthetics. Belmont, Calif: Wadsworth Pub. Co.

Zook, K. (2001). *Instructional design for classroom teaching and learning,* Boston: Houghton Mifflin.

APPENDICES

APPENDIX A

This section demonstrates a practical application of a viable interactive rule-based sound design within a multifaceted system. This instructional module is a problem-solving approach to creating a sound constituent within a visual component (video, multimedia environment, etc.), which could be useful to anyone with little sound design or composition experience (producers, directors, students, etc.), and possibly to those sound designers or musicians that would like a systemic and systematic method of generating sound design for multimedia. I realize that composition, or any act of creation for that matter, is a personal undertaking, but the following is a suggested approach using ID principles and offers the possibility to work within a team setting.

The choice of an instructional design model is dependent on the model's prescriptivity (Martin & Briggs, 1986): the author derives an appropriate terminal objective directly from the implicit condition. The following is an example of an instructional design component in which the author would integrate the methodical sequencing of sound design objectives and close the gap between the media production goals, media curriculum goals, and sound design objectives to create a multimedia unit using the bi-modality concept. In the given instructional module, when I mention the use

of "a pre-recorded video segment," it is available on the included DVD[114] (See Appendix D); of course, a more personalized pre-recorded video segment could be used as an alternative for the suggested instructional treatment.

Suggested Sound Design Instructional Treatment

Objective One:

Topic: Brainstorming

Task: The author will be able to brainstorm original ideas.

Given: 1.- a pre-recorded video segment; 2.- time to look at it at home or work; 3.- opportunities to discuss ideas with others; 4.- use of intellectual skills and creative thinking strategies; 5.- storyboard with blank space to input sound ideas; 6.- definitions of sound design, sound structures, soundscapes; 7.- sound design tools[115]; 8.- access to sound bites (Web, CD-ROM, etc...) or pre-recorded music or sound effects.

Criterion 1: That brainstorming will generate thinking aloud, constructive experimentation with sequencing tools.

Criterion 2: That brainstorming will motivate the author on various uses of sound design.

Criterion 3: That brainstorming will generate original ideas, new thought patterns to enhance the visuals of the work.

[114] The videos without sound can be used in the creative process; the ones with sound can be looked upon as examples only and the ones with *ad hoc* sounds are experimental in nature and give an example of how sound can have or not have purpose and/or validation.

[115] whichever tool the user is comfortable with, suggested either in this dissertation or elsewhere.

Objective Two:

Topic: Preparation and expectancy

Task: The user should be able to initiate a sound design project successfully

Given: 1.- a pre-recorded video segment; 2.- sound design tools; 3.- opportunity to show draft ideas of sound design to others; 4.- storyboard with blank space to input sound ideas.

Criterion 1: Proposing several production ideas or constructs to peers or others.

Criterion 2: That the initial sound design ideas will generate interest or constructive criticism from others.

Objective Three:

Topic: Project definition and usage.

Task: The user will be able to discriminate the type of project and the best delivery platform (Digital Video, CD-ROM based media, Web based, etc...)

Given: 1.- a pre-recorded video segment; 2.- access to various multimedia and video examples; 3.- definitions of multimedia, sound design and dual-channel coding.

Criterion 1: Looking at the video will determine if it stands on its own or needs to be part of a larger work.

Criterion 2: If it stands on its own, then type of audio and sound design needs to be addressed for video format.

Criterion 3: If does not stand on its own, then type of platform needs to be chosen.

Objective Four:

Topic: Choice of Sound Design Style.

Task: The user will be able to determine the style of sound design best suited for the proposed work.

Given: 1.- a pre-recorded video segment; 2.- access to various multimedia and video examples; 3.- have a good notion of sound design styles available. 4.- access to sound bites (Web, CD-ROM, etc...) or pre-recorded music or sound effects.

Criterion 1: Determine style of sound design to be used means having a range of examples of different styles to chose from, and can compare styles by auditioning the various sound segments.

Criterion 2: Using the audio tools can vary the various elements of the sound design, such as tempo, key, instrumentation, etc....

Objective Five:

Topic: Discerning the value of silence vs sound.

Task: The user will be able to discriminate the value of sound design in relation to silence or segmented sound design.

Given: 1.- a pre-recorded video segment (one with no sound); 2.- access to various multimedia and video examples; 3.- opportunities to discuss the value of silence in a multimedia environment. 4.- access to sound bites (Web, CD-ROM, etc...) or pre-recorded music or sound effects.

Criterion 1: Using the audio tools can vary the input or not of elements of the sound design and make judgments on usefulness of silence or sound.

APPENDIX B: INTERVIEWS

As stated in the Introduction, this dissertation examines the relationship between sound design and visuals within a multimedia environment, first by examining the various theories relating to the written and the oral; this is followed by a review of the ways in which technological tools have shaped our expectations and channeled the use of sound design in our contemporary technological environment. Thus, it became appropriate that certain axioms be verified by asking scholars and professionals about their views regarding their work as it relates to sound design. Included here is the full text of three (out of many[116]) interviews, as I find them to be most pertinent to the present study: a scholar who teaches at university, an internationally known Techno DJ and recording artist, and a video & sound designer.

Interview with ANDRA McCARTNEY

My name is Andra McCartney and I teach Sound in Media for the Communication Studies department. I'm (also) a *Soundwalk* (explain in brackets or footnote) artist.

Do you have any musical training, professional musical background?

Yes, I did piano with Royal Schools of Music in England and then with the Royal Conservatory in Toronto, and I did up to grade ten and then I did pedagogy, history, theory, counterpoint and harmony. Then later on I did a Masters and a Ph.D. in music.

[116] The research assistant (Jessica Landry) and myself interviewed a variety of candidates: with some, we had casual talks and with others, more formal interviews. Among the various candidates were students (Lisa Gasior, Simon Abramovitch, Tom Mallah), radio and video producers (Craig Desson), DJs and techno-advisors (Colin de Laplante), sound artist and theorist Gerhard Eckel (Professor of Computer Music and Multimedia at the Institute of Electronic Music and Acoustics, University of Music and Dramatic Arts, Graz).

On a regular basis, do you use, either personally or professionally, sound editing software?

Yes, I use at the moment primarily *ProTools*, but then also a little bit of *Sound Edit 16*.

Do you have a preference of platform?

Macintosh. When I first started working with digital sound, there really was no contest. *Mac* was just so much better than IBM and so I got used to it. And I've just stuck with *Mac* ever since. I think now it doesn't really make a great deal of difference, certainly not the way it used to, but now that I'm used to *Mac* I don't really want to go learn IBM.

How did you learn how to operate within programs like ProTools and Sound Edit 16?

Well, initially, the first editing program I used was *Sound Designer*, which is like a precursor of *ProTools*, and I learned that informally through a number of different ways. One thing was that I knew a composer in Toronto, Wendy Bartley, and she gave me some lessons on *Sound Designer* and also on *Sample Cell*. At that time I didn't have my own equipment at home so I was using a place called InterAccess in Toronto, which is a kind of access for digital artists, so they have a studio set up in there with *Sound Designer* and *Sample Cell*. And the technician also taught me...about the programs. But then I just used them after that; I used them on my own and learned that way.

Have you ever dabbled in programs that are geared more toward looping and making music at home?

...the first way I worked with digital sound was with MIDI and so I had the original *Cakewalk* sequencer [1988], so I did basic sequencing and looping and entering note by note in *Cakewalk*, and then later I worked at Queens University and they had *Studio Vision*, so I did looping and sequencing in that. Then when I first came here we had *Qbase* for the first year that I was teaching here, so I did some work with *Qbase*, although I never really liked it.

What kind of work, when you say you produce pieces?

Just using a keyboard to perform on, different lines...using different MIDI instruments and making a piece that had several different instrumental parts, all of which I performed myself. And then sometimes...with *Qbase*, I combined that with some digital audio that I had previously recorded.

It seems that, in your case, having a musical background helped you in terms of using something like MIDI, where you can play the keyboard. Do you think, though, that it's absolutely necessary to have that background before you work with these...

I don't think it's necessary at all...I think it is helpful because, for instance, *Studio Vision* had a very heavy kind of musical metaphor that it was using for setting things up, you could enter a score or you could use a guitar, like something that resembled a guitar

tablature; so if you were aware of what those symbols mean it makes it easier; but it is possible for someone without any performing ability to enter note by note or someone with no understanding of harmony still to put things together...When I say no understanding I mean no formal understanding of harmony, because I think people often have a very strong informal understanding that they may not realize they have.

But as someone who did go through the formal training, how do you feel about people who don't necessarily have the same formation, still producing things that they then release, specifically music.

sequencing software, it creates in them a desire to learn more about music because they realize that they know a certain amount but they need to go further in certain areas and I'm also aware that there are many really accomplished film music composers who will very openly say that they have no musical training. Or Yanni, who's been wildly popular throughout the world as a performer, had no musical training. His training was in psychology... So, I think it's very possible for people to do musical work without musical training and after doing those ten years, more than ten years of piano performance training, I felt that, in many ways, I was being restricted as to what I could do rather than opened up. For instance, at the time that I studied that anyway, you were being a performer which meant you were supposed to interpret and you were supposed to be able to illuminate these works of the masters, but there was no training in actually putting your own music together, so often people who are highly musically trained don't necessarily have a background in composition. I think that's changed a bit lately. I think there are teachers around now who do teach composition right from the beginning with musical training but I think they're still probably the exception rather than the rule.

So it's the advancement in the technology and in the interaction, not just between the instrument and the musician, but with a more interactive base; it's allowing people to make things more their own...

I think so. To make things more their own and to be able to experiment with composition the way that...many people will pick up a pencil or a paintbrush and try painting a picture, even though they don't have artistic experience; but then it has traditionally been much less widespread that people would pick up an instrument and say I'm going to compose something. And I think that this kind of software makes it fun to start composing and then people can listen to what they're doing and start to improve what they're doing...

To take your example of painting: in my opinion, in the audio world, it's a lot more acceptable to be from that kind of background as opposed to a painter who would try to make it big without any kind of backing - whereas people doing this sort of thing on their own could more easily gain access to a wider audience because of the Internet...

Exactly.

In your experience as a professor do you find that students are now coming to you better prepared? Or with more experience or more prior knowledge in whatever field interests them?

In the last few years, it seems that students have a deep knowledge in certain areas in sound or music production or music aesthetics as well. I think that if I had a concern about it, it's that sometimes it seems very deep in one area but not so much in other areas, which is why it seems to work very well to get people talking to each other about "well

I'm doing this and I'm interested in this," and it sort of broadens out what everybody knows. Maybe people have less of a kind of broad musical knowledge of possibilities of rhythm and harmony, sense but have a very strong historical knowledge of particular kinds of genres or approaches.

Do you find that they have a good handle on the people of the past or even of the , such as composers, thinkers or performers, academics, people within the field?

Again, each student that I meet will know of certain people, but then maybe be completely unaware of other people that I think they should know about. It seems like there is more idiosyncratic knowledge.

In your own research and work, have you found it easy or difficult to gather information about work, sound work, specifically in relation to multimedia?...

...in relation to multimedia, I think there's a lot less that's been done in terms of aesthetics,...there's way more in relation to film sound or radio or electronic music for instance than there is in terms of multimedia, and partly because it's so new. The web, for instance, has had really limited sound possibilities up until five years ago, so that's not enough time for a body of writing to start about that. I think at the moment the best possibilities for learning about multimedia sound are mostly on the web itself.

When you're surfing the web do you notice sounds, do you think that there's an adequate sonic element to the Net?

No, I find a lot of it is way too much use of really short loops that get really annoying, with really low quality. I mean there are some wonderful exceptions: some sound work on the Net, even in terms of museum work - some of the virtual exhibits are really very interesting and have used sound in a clever way to get the most out of it. I'm thinking of Diane LeBoeuf's work, for instance, that virtual exhibit on Montreal Through Two Lenses with old photographs and then new photographs and then the sound...

What about software-wise - do you feel there could be more aurally motivated icons as opposed to being visual all the time? Do you find those things helpful?

I do find them helpful. They'd be particularly helpful for somebody who had limited vision as well. It must get quite dull for somebody with limited vision who wants to use the Internet.

The medium is actually so tactile with everything you have to do, but you just have to see it.

So if it were activated by sound, it could be very interesting...I think there are so many sites that are visually and in terms of touch really interesting, and yet the sounds have not been developed as much as they could be. Even that Montreal Through Two Lenses: it was interesting to note that the photographer was given a lot of space within that exhibit; they were talking about how the photographer had set up certain shots and what the challenges were for the photographer; and then the historians were given a lot of space,

obviously, but it didn't even mention who the sound designer was or what they had done, and I thought that was interesting because, even though the sound there was really well developed, they clearly have not gotten to the stage of thinking about it as an approach to working with history.

It potentially speaks to a larger problem of disregard or just general ignorance from people designing sound (for images)... Have you worked with image makers at all?

Only to a limited extent...I worked with a painter on a couple of exhibitions where I would give him a recording that I had done and then he would paint; not having been to the places where I had done the recordings, he would paint a response. And that was quite interesting. Other than that, I guess the way that I've worked with image-makers is I'm the person doing the installation and I get somebody that I want to do photography or manipulation of digital images...

Like the Lachine Canal exhibit with the photographs... Something I noticed in the Lachine Canal (exhibit) was that it was an interaction with the visual medium, not specifically with representation but with actual spaces, like the museum that you used for the installation...

That space was really important to me, partly because of its historical importance, partly because of its placement so close to the canal, and partly because of the sonic qualities of that space, which was a beautiful sounding space. And I also enjoyed being in it. I could sit on the floor of that room for hours and be perfectly happy.

Even if your vision isn't specifically directed toward something, such as in television or in film, you're still seeing when you're listening to things and that really impacts how you listen.

Yes, that's a good point.

Not as a maker, because your work doesn't specifically relate back to film and television, but as an audience member, what is your take on sound in those two media? Do you think it is under-utilized or embellished?

I think it depends on the program. I think in action films, for instance *The Matrix*, the second *Matrix*, which I went to see recently, I found the sound level to be almost painful and made it unpleasant for me...I went to the biggest one possible but then the sound was too much. I think sound has been handled really sensitively in some films, like *Smoke*, for instance. There are certain sequences of films often where, in the first 20 minutes, the sound is particularly well done, but then it slips through the rest of it. I think in television, science fiction shows are often really well done in terms of sound. (**short interlude about the X-Files**) Even early science fiction like *Dr. Who*, which I grew up on, had great sound - electronic composers working on the soundtrack all the time. And the *Star Trek* shows, not so much the first one but the later shows had really interesting sounds too. And real thought put into it because they weren't trying to be representational, because you can't represent something that's in the future - whereas a lot of sitcoms or dramas have really predictable sound.

You're perhaps more aware of sound than the "average" viewer because of what you do professionally?

I think so. Now even though I don't do film and television sound myself (I did sound for one film,); I think because I'm always reading in that area and because I always have

students who are doing tv and film sound, it makes me want to know more so that I can be critical of soundtracks that I'm listening to that are student works, and of course that affects what I hear in the public realm too.

More specifically, in relation to your work with soundwalking in the museum installation as the most recent project you're working on: in your interaction with the public, with your audience, do you find that there's an awakening or increased awareness of sound and of listening, this sense that we don't necessarily use, or notice that we use?

Definitely. I think that both with that show and the previous one which was based on sounds of my neighborhood and was shown in *Chicago* - in both of those I asked people to respond and many of the responses were about paying more attention to the sounds around them. And I remember, I think *Voir* had an online [piece] about the *Lachine Canal* installation and a woman had responded there and had said "It's an interesting idea all these everyday sounds that are recorded in this show, but I only wanted to listen to it for 15 minutes because then I wanted to get outside and be by the canal," and I thought right on! That's exactly what I wanted you to do! So it was interesting to talk to people like the guides, who had been through the installation several times, and they were commenting both on what other visitors said but also their own experience, and that they were really thinking a lot more about the sounds of the environment surrounding the museum and hearing sounds that they hadn't really been paying attention to until then. I'm really pleased when I get that reaction to exhibits.

Interview w/ Mitchell Akiyama (Excerpts)

These excerpts are included since they are quite typical of the new multimedia-user generation.

Musician, Artist, Techno DJ, Label Owner and Producer

Do you have any professional musical training?

I have a lot of amateur musical training...I studied piano, I studied classical piano for awhile, pretty seriously...I did my conservatory exams...only private musical training, no conservatory training I did some electroacoustics at Concordia, but aside from that I'm pretty much self-taught.

What kind of music are you making now?

There's no easy genre...I'd have to be a little bit more tangential...I won't tell you what kind of music I make, I'll tell you how I make it. I record instruments, and I transfer them to computer, I record them to computer and manipulate them and compose post-facto. So I don't always play along with what I'm playing - I sometimes take bits and pieces from other recordings and other improvisations and digitally mix them back together. So there are traces of my classical training, there are traces of jazz maybe, traces of post-rock and a lot of traces of different things all sort of crumpled up into one digital ball.

What kind of program are you using on the computer to compose and to construct pieces?

I mainly use *LogicAudio* which is a...sequencer like *Qbase*, but that's kind of just the shell...I could be using anything. It's mainly the plug-ins that have an effect on the sound in the end...There's a program called *Reaktor* (**http://www.nativeinstruments.de/index.php?reaktor_us**), which is a modular synth FX processor kind of software, but it's not linear; it's not editing software...so I'll do things like record myself playing guitar, I'll take a 2min excerpt of that, feed it into a patch in *Reaktor* and do serious damage to it, and really explode it and deconstruct it and then take fragments from those sessions which I consider every bit as instrumental as the act of playing guitar because there are certain patches; I wish I was a programmer and that I could program my own patches, and that was a part of the gesture and that was a part of the process, but I don't have the time or the brain power to do that, so a lot of the patches I use feel like I'm playing them like instruments now, so I know how I guide it to a certain level of expression. And then those improvisations on instrumental improvisations get recorded and re-imported into the sequencer and then I might play along with the fractured guitar line I'll play a new guitar line which might itself end up being really heavily treated, and it's just this constant layering of processes until I say that it's done.

What is mulching, exactly?

Mulching is just the program. There's...granulation...it does granulation really well, which is the process of cutting up sound into tiny, tiny little fragments and then having

control over how they play chronologically so you can shuffle them; it's like a cloud, it turns into really cloudy opaque sort of sound – you can do millions of things with granulation.

Have you ever had exposure to or experience with compositional software that "does it all for you," that provides you with loops that you can then assemble and mix together in order to make music, so to speak?

Actually I just did a contract where I used some old 70's drum loops, like soul breaks... that that's the first time I've ever done that...It was kind of fun, it was different...Seeing as though I don't do much commercial work...When I do commercial work, the important thing is to give the client what they want – but I do that maybe 1/20th of the time, the rest of the time...actually the commercial work I do do, people ask me to do because they want ME to do it, they want my sound and they don't want me to sound like anybody else...so when that happens, in my personal practice...one of my fundamental principles in producing is that all the sounds have to be either made by myself or by somebody I know. Because for me it's important to have a certain continuity of experience, and when I listen to my own stuff I want to feel and remember where I was when I recorded the rain outside my apartment or when I played that guitar piece where I was, whether it was Becky that played that cello part or whether that was me making this really awful attempt at playing the viola...it's important, it's important that there is a certain inherent narrativity there that might only exist for me, but, in the end who am I making music for, really?

And I guess it maintains that connection with what you're making, if you have a story for everything that you use...

...definitely, and I'd like to think that there is some element of that narrativity that sublimates out of [it]...a listener might not actually know the anecdotes and the stories that I know, or have the same kind of insight into that as I would, but I think that if there's something honest and coherent about it then it'll speak in a certain way, although not necessarily speak more eloquently or more interestingly than another genre, like when you have somebody who's a sample specialist like DJ Shadow or R2D2 or most hiphop producers...One of the amazing things about that is that you end up with this semiotic soup of all these different genres and eras of music all colliding with each other, and you could go off on all these tangents: DJ Spooky...he talks a lot about sample culture and the kinds of relationships that happen when you sample sources that don't traditionally have anything to do with each other.

... It's the same thing with video. Video didn't kill film...people make some really amazing videos...

...the piano that turns into the trumpet... Do you think that it speaks to a larger problem of people being generally unaware of listening on the whole and of hearing (as opposed to) vision, for example?

There are lots of things people take for granted, and one of the things I think people just never question...is that all music that you're not hearing live is electronic music - that's to say that it comes out of two speakers and not from the walls where it's reverberating: someone plays a saxophone in a room you're not hearing it out the horn, you're hearing

off the walls, you're hearing it off the couch, it's physical, it's tangible. It's also physical and tangible when you hear it from speakers, but you're hearing it from two points instead of one - the sax, you're hearing a representation of a spatialized room in a recording. Of course, it's still close enough to a sax that you can make a connection in the sort of semiotic chain between the sax and the sound. When you hear a piano dissolve into a trumpet, that process, even if you're really aware of how to do these things, isn't clear because what I do with one program, somebody might have done with another...and then to relate that to visuals...The conventions of music - music's a much more abstract medium than film is, or has been up to this point... I don't think that's really disputable...

Even just general listening doesn't have to be in a musical form; even just people often ignoring their sense of hearing, or privileging sight over hearing and not specifically for music but just in everyday living; and I wonder if people's capacity for listening is just not as finely tuned...

Oh for sure, it's even couched in the language in our expressions...I don't think this is a really unfair generalization: people will look around them and make compositions out of their surroundings with their eyes, and they say "that's a really nice yard" or "that's a nice flower." People don't compose with their environment; they don't say "listen to the way that plane plays with the traffic and the whirring of the car's engine." That's not something people do. So we definitely favour the visual sense. To go back to the recorded versions of these things and the organized versions of these things, they require a lot more organization in sound, because when sound is chaotic, even if it's not necessarily reverting back to a more naturalistic representation of the real world, it still

has that lack of structure that people require, the meter and the tempo and rhythm and harmonic consonance and all these things that are generally needed for people to say music - as opposed to beautiful visual sight.

Or maybe people are just less equipped to make sense of chaotic sound than they are of chaotic sight...

Well it's also a biological thing too because we're programmed...unpleasant sound means unpleasant experience...all pleasant sound tends to have a benign analog in nature: they say that the first inspiration for music and for singing would be birdsong. There are two schools, there's the birdsong school and there's the school that says that we have the tendency to, when mothers are nursing their children to raise their voice up an octave and to almost sing words to kids. So those are the two schools of thought as to where the origins of music come from. It's hard to talk about those things in the visual culture... Something that you know is unpleasant when you look at [is] not necessarily going to be aesthetically pleasant...it still just seems intuitive to me that you compose with your eyes a lot more than you do with your ears. Anyway, that's what all music is, structured noises, noises that are mediated...you can have so much control over that, whereas in the visual, there are always details and relationships that are going to come in there. In electronic music it's the ultimate form of that control because you can abstract it to any level you want.

Do you have any specific people that you refer back to, whether it be in your study or in your art...Any specific sound practitioners, not necessarily music-related but people who deal in sound art or soundscaping...?

It's funny you should mention that, because I've always left my music out of my master's program that I'm doing because I don't think I necessarily have anything to bring, I don't think that world has necessarily a lot to bring to my music world because I just think that [what] I do... is music. And the most recent record is called "If night is a weed and day grows less," which is really influenced, at least topically, by research that I've been doing on the lawn...it was partly influenced by the introduction Plateau by Deleuze & Guattari, so this idea of the rhizome as being something that spreads without middle or end... [with] a patch of grass, let's say, you can say this is the edge, but it's not like a tree, where you have this vertical rigid, hierarchical, definable delimitation... (http://www.jahsonic.com/Rhizome.html) So the way that that influenced that record was I had recorded piano compositions and improvisations and then deconstructed them and then put 5 minutes of one piece back with another and they kept infecting each other so that in the end, each [came] out of every other piece, and that was a theoretically influenced idea but it's still just music, just piano music...So that being said, I don't really pay a lot of attention to sound art so much...

Do you think that in general people are pretty aware of sound art, or at least within the academic realm?

In the art world, yes. It's installation art, sound art, video art, everything that falls under "not paintings on walls"...You get people that sculpt and say that they are very painterly with their sculptures, and everybody's always borrowing each other's metaphors - so in

the visual art world I think there is more of an attention to sound art than there would be in the music world, which isn't intuitive, but that's the way it is.

Interview w/ Iain Cook

(part-time lecturer in Communication Studies, experimental sound artist and film and video sound designer)

I'm interested specifically in your personal experience with any sound editing software, any kind of music generating software, anything you've come across in your studies or work.

Starting way back, even before I came into the department, I had an old, ancient Apple Plus and I had the first generation of sound editing software called *SoundEdit 8*, which was 8-bit audio, working with tiny 2MB of RAM, so you'd have maximum 20 sec of workable things. And then when I came into the department, I dispensed with working with the digital audio and learned analog audio and then back into the digital world, following the typical experiences of working with digital audio, with *ProTools*, *Digidesign*, all that. But I think the most interesting experience I had was in my graduate studies - there was sort of an online experiment of audio artists gathering together in these virtual spaces and with technology such as MIDI, building upon and creating sounds in this virtual, online space, collaborating from different countries...

All with MIDI?

All with MIDI, because at the time, the network bandwidth considerations were such that you weren't pumping digital audio across, you were pumping MIDI...

1's and 0's?

...exactly - as opposed to huge chunks of data. So somebody would have a basic MIDI track, which they would leave in this virtual studio, essentially with the doors unlocked, and then anybody who wanted to could walk into that studio, or into this virtual studio, then add to that track, change it, modify it, then either save that new version or save

another version, so you'd have the original left and then the enhanced version, and then somebody else could come in and you could either work on things alone or, if by chance two or three people were online in that same virtual room, then you would be typing away, "Okay, what part of the song do we want to work on next, what do we want to do with the next part of the song?" So I thought that was the most interesting experience I've had with digital audio because it tends to be one person alone in a room, most of the time, working on things, as opposed to this really kind of exciting collaborative space where people from all parts of the world, from different backgrounds, share information and build ...

That was going to be my next question: Who were the participants in this kind of project? Were they professionals, academics, musicians?

[Yes], running the whole gamut. I found myself working with people who were curious - they had read about this site in some magazine and they just happen to type that URL in their browser and were part of that - to academics who were more interested in networking technologies; so you had a whole convergence of expertise and people from different backgrounds. Then you had your 14-15-year-old musicians who wanted to jam, ...and it was "Wow this is great." I think some really interesting projects came out of that space, but I'm not sure if that's an ongoing kind of thing or whether it dissolved.

These other programs that allow people to profit from all the different sound sources and construct their own things, as you said, alone in a room - have you ever used any of those programs, music-in-a-box programs?

See, I'm not a musician by training, so anything that I tinker with comes from what was I playing with...*AudioMulch*, which takes audio and mulches it up, and it allows for a

certain amount of play. But what I also find interesting, all of these programs have visual metaphors that they're using - for example, *ProTools*, the visual metaphor is that of an analog recording studio, the icons are reel-to-reels, and there's tracks and whatnot. In *AudioMulch*, I'm not sure of the background [of] who designed it, but I think he's more of a hardcore engineer; as opposed to tracks that you're looking at, you have boxes, and you click on one box and then you draw a line to the next box. You're physically connecting the circuits.

That's like MAX ?

So that comes from a very different tradition. Obviously, *MAX* isn't being designed by people who worked in analog audio - they're coming from a very different world. So as a teacher...if you think back to someone who's moving from analog to digital, they see on screen the basic metaphors that they understand, multi-tracking, and they can think, "oh yeah, that's like an 8-track machine." But if students are coming in without that background, that 8track metaphor just doesn't make sense any more. The things that we take for granted, because there's a visual language that relates back to things we understand, is no longer there. So just explaining very simple concepts becomes harder without that analog background.

So do you think it's better or easier to teach students who have some kind of personal experience with sound editing, making, recording...

I found in my personal experience when I was working with *SoundEdit 16*, a lot of [what] they were talking about, I really had no idea what was happening, what I was supposed to do...[until] I got that basic fundamental training in working on a reel to reel: "Oh, that's

what they were talking about in *SoundEdit*. That's what those icons meant, that's what that language was talking about." And I don't think my experience is that special. I think that's a common experience, that it really is more helpful to have that analog understanding. As long as we have programs that use that visual language which is based on analog audio practice, then it makes it a lot easier if you have that experience. If they start designing programs with a different metaphor that draws upon different experiences, then the teaching obviously will change.

And you mean fundamental training at school, for example, in an environment like this, in Communications let's say?
Yes.

Because there are a lot of people right now that seem to just acquire the skills as they go, especially younger people - as you said, the 14-year-olds who want to jam.
Yes, I recognize [that] and I think that's a really important type of knowledge, that sort of informal knowledge. I think the basic ingredient that's really necessary is to be really curious about something, and not to expect to have that information pumped into you. Especially working with computers, it's a matter of, "okay, this is the program, how can I play with it and take what I want out of it?" I was just reading an interview with David Byrne, who's done a project using PowerPoint. And he had set up this sort of exhibit in the lobby of this big corporate company and the information technology guy came down and said "Well, you could have done this a lot better in *Flash*," and David Byrne's response was "Well, I know that, but there are limitations to the software; PowerPoint has so many glitches and it's so buggy that I wanted to explore those bugs to come up with something that was really interesting." So I think that the same level of curiosity, and

appreciating that the software has bugs - that it's not perfect - that you could then do something really interesting with that. But the key rests with the person who's sitting between the chair and the computer. That's where the power of the computer comes out.

You mentioned that you design for film and video. What is that usually more comprised of? Is it more musical, or is it more ambient, found sound?

I don't have musical training. Apparently I have a musical ear, because people who listen to it say, "Well, that's musical," and they recognize all these complex time signatures and..."oh yeah, cool, that 15/16ths, I wanted to do that..."[W]hen you think of musicality in terms of the whole world [being] musical, and just taking those elements... the process of abstraction I think is what I like to put into the work that I do for film and video. And I've been very lucky in that I work with video people or filmmakers who appreciate that sensibility. They're not looking for a straight ahead boxed-in perspective, so I've been really lucky to work with people like that.

Given your position in the chain of production of a video or film, what has your experience been with people and their response to that relationship between images and sound?

... I was working with people who [gave me their] very initial edits, so I knew the types of image that would be used and the quality of the images...the graininess, or the textures, then I would work on a sound design for that...They really appreciated the fact that they would have something [that they could] cut their images to, as opposed to the other way around. [This made me] feel that I was a real integral part of the process, as

opposed to, "Here's our finished edit, can you then match the sounds to that;" it would be more of an organic, holistic approach.

Do you think that people respond differently to those kinds of pieces versus pieces where sound is less...

I think so. I think on an emotive level I think it works because sound affects people and I think that really changes the quality; I would hope to think that it improves the quality of the final video or film. But the downside to that is that a really good film or video soundtrack is something that you don't really notice. So if it is too strong an element, if you draw too much attention to that, it's a fine balance. It's like in *Spinal Tap*, it's a fine balance between stupidity and brilliance.

When you work, what platform do you prefer to use?

My initial work was done on Macintosh. As I did a more realistic accounting of how much it costs for Macintosh versus Windows, the fact now that, on a technical level, the two platforms have their own strengths and weaknesses; it's a lot cheaper for me to invest in Mac than it is in Windows.

What about with the software, what do you use?

I actually find that I like free tools, and I find that there's a lot more freeware available for Windows than there are for Mac. So I may not look as glamorous cause it's a beige box that I'm sitting in front of, but I figure that the people who are listening to my work don't really care whether I've done it on a Mac or on Windows.

Is that going along with the philosophy of free stuff for everybody, giving everybody a fair shot at producing something?

Yes, and that's what I find is the most encouraging part of...what they're calling the digital revolution - that at not all levels, but at a given level, there is [a] kind of democratization: somebody who's curious enough and if they have access to the technology, they can do things that even five or ten years ago were unheard of, and to have access to such information and sources, and different types of information, it's really empowering.

Does that philosophy extend to the music industry, and to home studios?

Yes, for sure; it's really mind-boggling the types of work that you can listen to now, that are available to us, at least living in North America. I don't know if that question is related directly to the challenges posed to the record industry...

In the other interviews, this is the question that has come up that I hadn't even considered, that the record industry is being shaped now by the up-and-comers who don't even need to go through the record industry to get their music out. Is that giving the big "thumbs down" to people who perhaps "work harder" at what they do?

If I had come from that type of musical background, musical tradition of having to go through formal musical training, I would probably have a different take on it; but as somebody who hasn't taken a single class in traditional musical training, I think it's totally empowering. I can't imagine that it's the musicians that come from that background that would feel threatened; I think it's more of a dollars-and-cents kind of issue, and it's the people that are holding the moneybags that are a bit nervous right now.

What do you work on at home?

When I'm working with film or video I use *VidEdit*, which is a freeware for cutting up my sequences so I can then work on *SoundForge* and *CoolEdit*. Just the basics. Not the fancy *CoolEdit Pro*; I really like to limit myself. So, for example, in *CoolEdit* I'm working with severe limitations of two tracks; I'm kind of stubborn that way - I realize that if I worked with 16 tracks or 24 tracks, I could do things easier, but for me I really like constraints...perhaps it's because I come from a world where you have your *TASCAM* 4 track in your bedroom and you're limited and you're bouncing tracks over, and you come up with really amazing recordings [on which] I really like to pose that same kind of limitation. Maybe I'm just an old fart, but that's what I like doing...

In your experience as a teacher, do you find it easy or difficult to get information and sources on professionals in the sound world, either academics or people who could be inspirational (to students...)?

I wouldn't look necessarily at the academic world. But I would look at places like the SAT downtown, I would look at local institutions that have real people doing really interesting things. I think we're really lucky here - we have a really strong community of that kind of person; it doesn't mean going far afield, just check out what's happening in the local scene.

You're almost forced to do that, because it seems like this field is constantly developing and branching out into seemingly non-related fields, but they really are integrated...

I think the key or the trick is to try and be able to draw those connections, to make the web of connections and to interrelate, point out when something happening *here* has actually influenced something here. One example would be the software that was developed by hackers who met up with DJs at a hackers convention - *Virtual Mix*; instead of mixing vinyl, you're mixing these virtual albums. So that came out of the fact that there were these computer hackers having a conversation with these DJs, and the hackers couldn't understand why they were still using vinyl. So you have this conversation and out of that you have this amalgam of these two kinds of cultures; so if somebody is using that, they'll say that this is what hacker philosophy is all about, and this is what DJ culture is about, and this is what happens when these two groups get together - they come up with this interesting technology. And then you have people in town that are using that technology, and you can bring them into class, you can talk about them...

APPENDIX C: CONSENT FORMS

- Andra McCartney
- Iain Cook
- Mitchell Akiyama

APPENDIX D: DVD

The attached DVD contains both informative and functional materials. The demos, sound designs and compositions will give the readers of this thesis a good idea of the scope of a multimedia or sound design project. The video templates can be used with the Instructional Design module from Appendix A.

DVD INFO:

1. A demo of the literacy educational tool, *Alphie's Alley*, that I instigated 5 years ago and presently completing with the CSLP D&D team for scale production in over one thousand schools.

2. Video templates to be used with the Instructional Design module;
3. Examples of personal compositions (film music, sound designs, etc.);
4. Characteristic examples of some of the software mentioned in the thesis (*Metasynth* and *ArtMatic*)

This DVD functions as a regular viewing DVD with a table of contents that one can use to navigate through the various sections; as well, it operates as a DVD-ROM (if inserted into a computer), which contains the files that can be downloaded for use with the sound design ID module; these are in the User-File folder.

DVD CONTENT:

1. *Alphie's Alley* Demo
2. Making of *Alphie's Alley*
3. Animation with sound
4. Animation without sound
5. Animation with ad hoc sound
6. Two Metasynth & Artmatic examples (Patern Dance & Transformations)
7. Two Compositions & Sound Designs (Composed using Logic and Live)
8. Vrai Monde (Trailer of a Michel Tremblay film for which I composed the score)