Georgia State University
Digital Archive @ GSU

English Theses | Department of English

11-20-2008

The Morphic Orator: Transmogrified Delivery on the Audio-Enabled Web

Brian Johnson Snead

Follow this and additional works at: http://digitalarchive.gsu.edu/english_theses

Recommended Citation
Snead, Brian Johnson, "The Morphic Orator: Transmogrified Delivery on the Audio-Enabled Web" (2008). *English Theses.* Paper 49.

This Thesis is brought to you for free and open access by the Department of English at Digital Archive @ GSU. It has been accepted for inclusion in English Theses by an authorized administrator of Digital Archive @ GSU. For more information, please contact digitalarchive@gsu.edu.

THE MORPHIC ORATOR:

TRANSMOGRIFIED DELIVERY ON THE AUDIO-ENABLED WEB

by

BRIAN SNEAD

Under the Direction of Mary Hocks

ABSTRACT

Audio is an effective but often overlooked component of World Wide Web delivery. Of the nearly twenty billion web pages estimated to exist, statistically few use sound. Those few using sound often use it poorly and with hardly any regard to theoretical and rhetorical issues. This thesis is an examination of the uses of audio on the World Wide Web, specifically focusing on how those uses could be informed by current and historical rhetorical theory. A theoretical methodology is applied to suggest the concepts and disciplines required to make online audio more meaningful and useful. The thesis argues for the connection between the Web and the modern orator, its embodiment, its place in sound reproduction technology, and awareness of the limitations placed on it by design and convention.

INDEX WORDS: Online Audio, Sonic Literacy, Aurality, Web audio, Multimedia, Sound reproduction technology

THE MORPHIC ORATOR:

TRANSMOGRIFIED DELIVERY ON THE AUDIO-ENABLED WEB

by

BRIAN SNEAD

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2008

Copyright by

Brian Johnson Snead

2008

THE MORPHIC ORATOR:

TRANSMOGRIFIED DELIVERY ON THE AUDIO-ENABLED WEB

by

BRIAN SNEAD

Committee Chair: Mary Hocks

Committee: George Pullman
Elizabeth Lopez

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Sciences
Georgia State University
December 2008

DEDICATION

To my grandfather, who has generously made this degree possible

und an meine Frau, die nicht nur meiner ewigen wissenschaftlichen Grübelei zugehört, sondern sie auch überlebt hat.

ACKNOWLEDGEMENTS

The genesis of this thesis was the genuine interest my professors in the Rhetoric and Composition program showed in issues related to delivery. I was able to glean many ideas from my discussions with them, as well as from my fellow students. It is a pleasure to convey my gratitude to them *en masse* in this acknowledgement.

I would like to directly express my gratitude to the chair of my thesis committee, Dr. Mary Hocks, for her supervision, advice, and guidance over the long summer that it took to put this thesis together. Above all, she provided me encouragement and a structure of assignments and deadlines that saw me through to completion of this project.

I gratefully acknowledge Drs. George Pullman and Elizabeth Lopez who not only have served as my thesis committee, but have contributed more than they realize to many of the ideas and approaches imbedded in this work. Extra special thanks go to Dr. Lopez for the formatting wizardry.

Many thanks go to Dr. Jennifer Bowie for initially encouraging me to formally pursue a study of web audio, what I would have otherwise regarded as a hobby.

I also thank Dr. Karl Stolley at the Illinois Institute of Technology for his stimulating comments and questions between sessions at the ATTW 2008 conference in New Orleans.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS v

INTRODUCTION 1

CHAPTER

1. THE SOUNDSCAPE OF THE AUDIO-ENABLED WEB 9

The Peculiarities of Background Sound 12

The Variety of Signal 16

Introductory Pages 21

Interactive Audio 23

Discussion 25

2. SITUATED MECHANICS 29

Compression / Decompression Schemes and Psychoacoustics 31

Sound, Hearing, and Getting Sound to the Web 32

Online and Back Out Into the World 35

Discussion 36

3. TRANSMOGRIFIED DELIVERY 38

The Tacit Subordination of Delivery 40

Embodiment and Gesture 43

The Voice 45

A Negative Implication of Hypermediacy: The Barbarism 46

Discussion: The Limitations of Preservation and Transmittal 49

CONCLUSION 52

BIBLIOGRAPHY 58

APPENDIX: SURVEY OF 120 AUDIO-ENABLED WEBSITES 68

INTRODUCTION

The World Wide Web—an interactive, multimodal medium delivering content within a genre framework and a physical apparatus—is one modern incarnation of the ancient orator. In and of itself, the Web is informative and persuasive, imparting greater validity to the wide range of sights and sounds it projects into three-dimensional space. It works within local societies and societies abroad to inform, persuade, entertain, to obfuscate, and spur to action; it is a physical, electronic, and intellectual presence that over a billion people worldwide use and contribute to. In its visual space, values of invention, arrangement, and style are transmitted. Due to its electronic characteristics, an extremely high proportion of activity occurring in its physical space remains and is retrievable. If it were not for the problematic nature of delivery on the Web, it would be fully representative of the traditional Aristotelian rhetoric. Yet, as a communication system, the denial of the canon of delivery is as glaring an omission as delivery itself is difficult to contextualize: one might say that the Web paradoxically is delivered through a vehicle which is fundamentally mute and featureless, producing no more sound than the whir of very small fans and the movements of hard drive spindles and having no more visual presence than the machines containing those parts and the wiring connecting the machines together, an extreme of reductionism. Taken to an extreme of phenomenology, one might attribute the delivery of the Web to the experience of its rendering. The latter has generally been the verdict of web users, who have allowed it to assume an increasingly pervasive and vital place in society. Because of its power to persuade and the widespread conception that it is a thing with organic properties, the Web is considered in this paper as the "morphic orator."

Still, for the orator to fully immerge, the circumstances surrounding the definition of the Web must be addressed. The acceptance of the Web as phenomenological presence is not

without its limits, for it carries with it a set of values and prejudices: as the innovations of its premier began to settle into convention, the Web demonstrated every indication of having inherited the ancient distrust and disregard for delivery, the silence imposed on rhetoric since the early age of writing along with the attendant *de facto* notion that the visible word is superior to the aural one. Perpetuating the biases of sound inferiority are those sites that routinely make sound a noisy and ineffective banality; it is rare that online sound is ever held to the same exacting standards that online text is. This is bizarre and counterintuitive for several reasons. Sound is an extremely important component of life, pervading "every facet of our living" (White 3), "a primary means for us to acquire information about the world in which we live" (Yost 1). And as human life is occurring in part over the Internet and on the Web, sound would seem to be a necessary and desirable component. Theoretically, sound can be captured and digitized, hereafter, uploadable, storable, transmittable, and downloadable via the Internet just like any other digital file, giving every Internet user access to it. It can be implemented on web pages and can exist in the dimensions of offline space by projecting into it. It can also expand the dimensions offered by the Web, as the Web, constructed to operate over the infrastructure of the Internet, is metaphorically described, like the Internet, as space and place and thing—a superhighway, maze, library, and dimension (Ratzan)—where one can perform any number of three-dimensional activities, such as surfing, mining, or browsing. This online space is increasingly coming to represent and replace traditionally offline spaces. Commonly referred to as "Web 2.0," it is a network composed of responses to human social and aesthetic concerns, "linking people…people sharing, trading, and collaborating" (Wesch). It would seem that sound would play a more extensive role other than as a disconnected commodity. While a number of specialty sites use sound in at least a limited way, the "serious" sites (banks, universities, unions,

museums, news, governments) overwhelmingly do not use it at all. We return to a simple and vital observation: nearly all web activity occurs with the most incidental of noises in our offline environment—the click of the mouse, the taps of the keyboard. Unlike blind and visually impaired users (about whom more will later be said), sighted web users easily navigate the Web in a purely visual modality. It may seem, perhaps, as if Internet technology were, in fact, created for this very use. As Walter Murch puts it, there is "no natural and preexisting harmony between image and sound" (qtd. in Chion xvii). It would be a foregone conclusion that the Web would incorporate sound or that it could benefit from it at all. People have asked, "why should the Internet produce sound; what can sound do better than text?", but this betrays a stolidly occularcentrist position. Rarely does anyone consider why the Internet *should* contain *text*.

An argument usually cited to keep sound and other multimedia offline is the poor usability due to bandwidth, player modules, and file compatibility. Processors, servers, browsers and client computers are becoming much faster on a large scale, enabling greater use of audio. It is even possible that the data transfer technologies, known as 'the Grid,' developed for the Large Hadron Collider at CERN to provide access to the roughly 15 Petabytes of data which will be produced annually (CERN) to thousands of scientists worldwide, may "soon make the Internet obsolete" (Fox News). There can be few remaining reasons for text to exist online at the exclusion of sound. The progression of browser technology from the early nineteen-nineties to today bears out the fact that new ways of designing and thinking of the Web is an ongoing process[1]. Though the Web initially followed the model established by print, it has taken on so many characteristics not available to print (due to the limitations of print technology) that it has all but abandoned that model. Sound and text are equal in this respect, although out of sync

[1] Tim Berners-Lee introduced the first web browser, "World Wide Web," in the first quarter of 1991. It was quickly overtaken by Mosaic two years later, followed by Netscape in the late nineties, and since then by Internet Explorer, which holds the lion's share of users.

chronologically[2]. In fact, technologies once imagined capable of transmitting only auditory signals, such as the telephone system or radio, can now transmit visual information as well. The notion that the technology needed to implement sound is more complicated than that of text is likewise past its time. The extremely complicated webpages of the 21st century, many of them written in languages that treat the page as a movie, operate over 96% of all online computers, as they carry software capable of rendering auditory and visual data (W3Schools). As the visual aspect of the Web has been largely rethought "out of the box" of the print metaphor in the past few years, establishing particularly web conventions, so too must the aural aspect. The harmony between sound, the machine, and the interface must be made suitable for the Web.

To sidestep both the invalid arguments against sound and the valid ones, the notion of the 'experiential website,' has arisen. So far, it is the only (and apparently most convincing) argument for the implementation of sound online. It must be dispelled, as it brashly underestimates both Web genre and the capabilities and limitations of sound. This notion, what I refer to as 'the cult of experience,' has led to web audio content consisting of either intrusive banality or marketing hype. A full-length text on web audio published in 1996 asks the reader to imagine crowding our roommates around the television for a 30 minute show and turning off the sound, only to 'see what happens.' Naturally, the authors assert, distraction and confusion, diminished entertainment value and ineffectiveness ensue. Why then, the foregone conclusion bangs out, would anyone "sit for hours glazing glassy-eyed at a silent webpage" (Helmstetter and Simpson xvii-xviii). A second text, published five years later, seconds this supposed necessary relationship between sound and image, positing that, "if sound were not an integral part of the user experience, we'd still be sitting in silent-movie houses watching the likes of

[2] But just because the visual was first does not mean that it is somehow superior or that audio is 'tacked on.'

Gloria Swanson swoon and bat her eyelashes to convey a romantic interest" (1 Beggs and Thede). In their (silent) book, Beggs and Thede claim that 'because' the film soundtrack is the "element that makes us cry, sense fear, or feel the adrenaline of an exciting chase scene" (1), webpages should not only seek this same intensity of emotion, but can only do so with sound. Momentarily bracketing genre conventions and the finding that most users find answers or abandon the site after 30 seconds (Nielsen and Loranger 22), we take them at their word: the Web equals the movies. Proceeding down this line of thought, one makes the rather ironic discovery that "theories of the cinema until now [1994] have tended to elude the issue of sound, either by completely ignoring it or by relegating it to minor status" (Chion xxv). Over one hundred years of cinema and few film academics pay attention to sound, a move which has left film criticism and theory "resolutely image-bound" (Altman 44). But nearly every film made today, indeed for the past 70 years, has been made with sound. This begs the question, who *is* paying attention to film sound? Are we to conclude that all film sound *must be* completely natural and effortless, just as Helmstetter and Simpson, Beggs and Thede would like us to think? They invite the reader to ignore the fact that film sound is routinely ignored by critics and academics; that film was silent for the first 35 years of its existence; that the introduction of sound "wrought havoc on the visual of filmic composition…straightjacketing the camera" (Murch x); that the nature of timing was forced to change, that sound "temporalized the image" (Chion 17); that "there was great fear that sound would ruin the poetic art of cinema developed in the silent era" (Fischer 78); that sound plunged film into "dully photographed plays and vaudeville routines" (Knight 213); and that "the coming of sound during the late 1920s climaxed a decade of significant change within the American [film] industry" (Gomery 5). Even if audiated movies are produced pragmatically, it is not and never has been a haphazard or

unconscious endeavor. The web audio promoters above make a patently ridiculous argument. However, in turning to film, they unwittingly open an extremely ripe area of resemblance (of which more will later follow).

The cult of experience also requires that one ignore the simple and obvious fact that silence is often considered "golden," that current cultural practices arbitrarily enforce silence—polite concert performances demand silence from audiences, as do museum galleries and libraries; books have not been read aloud in the West for hundreds of years; silence is imposed on the congregations of funerals, weddings, Masses, and most church services; a certain kind of auditory discretion is expected on buses, trains, trams, and airplanes; some sounds are allowed at the dinner table while others certainly are not. If nothing else, no one uses the Web in an anechoic chamber. For those able to hear, the world is full of sound[3], including the spaces in which we use our desktops, laptops, and handhelds[4]. Phones ring, babies cry, Beethoven's Pathétique Sonata plays in the background—all occurring in a cohesive sound field that is difficult, if not impossible, to sterilize for a genuinely artificial 'experience.' Save for a few notable exceptions, the web community has shown no interest in 'experiences' that it hasn't asked for. Users typically ignore what does not help them to complete their task. Sound is subject to this as well. Sound must be explicitly requested (or reasonably expected) and must always be relevant. Otherwise, users turn off speakers, develop applications to eliminate the audio of any webpage, and discourage designers and clients from using sound on webpages. And all of these actions have been taken. Such pervasive reaction would be the end of the present examination of web audio if it were not for the wealth of approaches alternative to the cult of

[3] The world is full of acoustic energy even for those who cannot hear. The notion of sound need not be a purely internal, perceptual reality.

[4] It is, further, no requirement that gadgets make sound, especially those that make almost no sound on their own. Apart from that of annoying ring tones, the argument that the sounds of mobile gadgets are useful at all is severely limited (see McGookin and Brewster 2004).

experience. Much research in the field of Auditory Display has shown that sound is useful in the design of an interface; Rhetoric has established a set of conventions for sound that would be appropriate for online use; Internet technology has far surpassed many of the assumptions made in the early days of web audio. For the Web, designers can no longer simply rely on unsubstantiated notions of 'experience:' audience, genre and the technological situation—such as digitization, Internet protocols, browsers, and user machine capabilities—must be taken into account.

Though every sound necessarily intrudes into the physical environment of the listener, as well as her perceptual environment, it is not necessary to approach this reality with the language of 'experience.' Experience always imposes itself as a necessary interiorization, disregardful of the paths with which it may come to be interiorized. The cult of experience unsuccessfully attempts to supersede narrative and metaphor and nearly every technological and rhetorical implication. Through the mediation of the metaphor, the user is connected to the interface and the machine in a holistic fashion by her sheer will to be connected; the temporality of the narrative keeps her motivated. The physical properties of sound require three dimensions plus time, a reality which perfectly compliments narrative. As a visual *information source,*[5] the interface of the Web is composed of two-dimensional space of light and darkness. As a space, it may be the apotheosis of modern architecture, initially defined largely by the extent to which environmental sound was controlled or produced (Thompson 333). The Web may be a "silent world" but to enter it one does not "forsake both body and place [to become] a thing of words alone" (Barlow). Sound reinforces the metaphorical space of the Web, but it may just as easily be real, that is, embodied. The body is not forsaken; it is made transliterate. We have a constant:

a body consisting of the physical apparatus of the Internet, from the servers to the client computer. Audio is accordingly dependent on the particulars of the apparatus body and the prior bodies—the body (resonator) from which it originated, the sound of the room in which it was recorded, the microphones and processors in the signal chain, the sample rate and encoding schemes. This thesis presents a connection not only between user and machine, but the parts of the machine, and within the interface, content, and context of use. Online audio is almost never designed to exist within a symbiotic relationship with the webpage. There are many examples of sound that have little or nothing to do with the site from which they play. Attempting to actuate an internal experience, the real implications of the apparatus and content are disregarded by the cult of experience. This results in no new creation of environment but constitutes an aural invasion of the user's physical and cognitive environments. And, though web audio may always be plagued by the Internet's very own discouragement of time-based continuous content, the negative attitude of users and designers, and the fact that not every user has or uses their speakers, web audio is nonetheless long-overdue for its fair hearing.

THE SOUNDSCAPE OF THE AUDIO-ENABLED WEB

For a sizable portion of Internet users, the initial sounds of the Web must surely be bizarre. A dial tone, the rapid dialing of a phone number, a lull: beeps, whistles, pings, static suddenly washing over the speakers—then nothing, silence. A peppy "Welcome!" bursts out, most likely followed by "you've got mail"[6]. In the late nineties, when dial-up was the most popular route to the Web, a scene in the blockbuster science fiction movie "The Matrix" loosely followed this series of sonic events as the protagonist's mental projection of himself is called back into his body over a network connection[7]. In the movie, as in our daily lives, one may feel that upon "signing / logging -on" to the Web, one enters a strange world with strange and fleeting sounds. The very vehicle taking us there, the browser, often comes equipped with sounds all its own: a "sound pack" is built into Internet Explorer, which plays *click* sounds when a link is selected; the Firefox "find as you type" feature plays an alert when it cannot find the requested search string. But these are often swallowed up by the gargantuan amount of music, a commodity for which users have a ravenous appetite. Internet radio and music downloading applications of all kinds have flooded the Web, greatly effecting traditional production and distribution processes. There are tens of thousands of band websites, nearly all of which offer free samples of their music directly from the webpage. Add to that an overwhelming 2.2 million band pages on the social networking site MySpace, where bands feature full-length songs on their profiles, which can then be "added" to any of the other over 100 million user profiles (Sellers). A host of widgets[8] have made embedding a player with music common and a Firefox

---

[6] "You've Got Mail" is the title of a 1998 major motion picture and the iconic tag line for the over ten million America Online (AOL) users as of November 2007.

[7] In the scene entitled "Down the Rabbit Hole": a dial tone, a rotary phone being dialed, silence, Neo's scream digitally transposed, concluding with Morpheus' line, "Welcome to the real world."

[8] Google Audio, Odeo Music Player, WengoVisio, Xfacts Audio, dukaBUZZ, to name a few.

add-on application allows the user to control music on both the host machine and online[9]. The nearly 45,000 individual podcasts offering over three million individual episodes (PodcastAlley.com) and other webcasts numbering into the tens of thousand can be played directly from a webpage. Over three billion YouTube multimedia files are played on a monthly basis by over 79 million users (Yen), nearly all of which include sound. The social networking sites Voicethread and Gabcast[10] allow users to post and comment on blogs via voice. On Hipcast and Evoca users can record from any POTS[11] phone, Skype, or the in-browser recorder, where the audio is converted to MP3 format and is made available for immediate web publishing. Kompoz and eJamming enable musicians anywhere in the world to jam. Jamglue and Splice make it possible for users to remix each other's tracks.

From the bizarre to the banal to the innovative, there is a significant variety of sound online, making the application of any one theoretic apparatus developed for other mediums to the audio-enabled Web difficult. In order to both survey the variety of sound as it exists online in an overall scheme and to discuss particular elements of that scheme, two theories will be used: "Soundscape," repurposed from Acoustic Ecology and "Auditory Display," repurposed from Ecological Psychoacoustics. Coined by Canadian composer and aural theoretician Murray Schafer, the soundscape is defined as "any acoustic field of study…[consisting of] events *heard* not objects *seen*" (7-8; original emphasis)[12]. Used by Schafer to theorize all manner of sound from the natural world, the urban word and the sounds of "Creation," it is easily made suitable

---

[9] "Foxy Tunes," versions 0.5 (22 August, 2004) – 3.0.2 (16 June, 2008). Controls in the bottom of the browser control online players such as Pandora and Rhapsody Online, as well as Windows Media Player, iTunes, and Media Monkey.

[10] Audioblogger, terminated by its parent company Odeo on November 1st, 2006, functioned very similarly to Gabcast. However, like Gabcast, audio commentary was not possible.

[11] Plain Old Telephone System (POTS)

[12] To dispel an illusory snare: listening to the "acoustic field of study" of one's choosing is not to ignore the perception of the visual; it is, instead, to attend diligently to the world through our ears.

for online sound. The Schaferian soundscape includes not only sound, but is thoroughly involved with the relationship between societies and their sonic identities: the sonic changes brought on by the "dynamic hedonism…[of]…modern man," the "rich symbolism" sound contains for humanity and the reading of the "general acoustic environment of a society …as an indicator of social conditions which produce it and…the trending and evolution of that society" (5-7). To further make the Schaferian soundscape suitable for the web, Auditory Display[13], the use of "any type of [non-speech sound] sound to present information to a listener" (Walker and Kramer 149-150), is introduced. Particularly, the notions of "earcon" and "auditory icon" will be used to discuss particular elements of the soundscape.

As the space and time required would be prohibitive to analyze every instance (and the fact that one cannot hear the examples through this paper), I follow Schafer's recommendation that the analysis of any soundscape begin with the discovery of the "significant features of the soundscape, those sounds which are important either because of their individuality, their numerousness or their domination" (7). Those significant features, broadly categorized by "background sound," "signal," "introductory pages," and "interactive audio," are presented in the following section.

[13] From the field of Ecological Psychoacoustics.

The Peculiarities of Background Sound

Background sound is sound that plays as soon as the page has loaded, the duration of which is determined by the designer, often with no controls to stop it. This sound is often written into the code of the webpage or plays through a plug-in. The term 'background sound' is somewhat problematic. Two of the most immediate problems are the spatial and hierarchical implications. Spatially, this sound is not really 'in the background' in any listening environment, whether it be through headphones or in an open sound field. Both spatially and hierarchically, if the sound upsets the user's relation to the screen through sheer volume or novelty, it is not really 'in the background.' This also applies to instances in which sound commences when there is simply nothing on the screen. Although one might say that the current convention of its implementation postures it not so much 'in the background,' as simply not "to be listened to consciously" (9), users cannot help but 'listen to it consciously' as it defies the silent conventions of the Web. Most often, background sound is music, but is hardly ever designed or implemented as the so-called 'background music' of ambient music composers[14], who intend their work as "tapestries; large-scale, non-intrusive atmospheres which lend a consistent mood to the environments in which they are heard" (BOFOP). Schafer's term 'keynote,' meaning perpetual sound of an "archetypal significance" (10), would be appropriate if it were not for its referential qualification, what Schafer terms as a 'reference point' from which everything else takes on its special meaning" (9). Another problem is that the term 'background' may constitute an implicit visual expectation in two ways: 1. by relating sound to a figure and ground metaphor in which sound plays the ground to the ocular figure and 2. by connoting seeing one thing in relation to another. Background sound is difficult to describe because of the nature of the Web:

[14] Brian Eno, Cluster, Kraftwerk and Wendy Carlos, to name a few.

able to contain sound, interactive, but environmentally silent. The term[15] will continue to be used though it is not to be confused with the `bgsound` or `img` HTML coding schemes[16] that allow implementation of sound as an inline element.

Depending on the purpose of the site, background sound is used to serve either mimetic or extra-diegetic ends. It is often used without commentary or obvious connection to the textual intent of the site; just as often as it is a non-descript musical track, it is a sample of a particular sound commodity or service, such as songs on a band's website or a sample of a voice artist's work. Background sound used to promote the sound as a product is primarily mimetic. LedZeppelin.com, BonJovi.com, JohnnyCash.com, WhiteStripes.com, Sade.com, and thousands more feature samples of their material as background sound in the website. The intended scenario posits a user immediately hearing the band without searching for samples (a scenario in which unrequested sound may seem appropriate but that may in reality be a total intrusion). New material the user might not have heard through other media outlets can be promoted in this way, as can older material that the user will immediately recognize and with which she will immediately identify. Sites that do not sell music or other sound commodities or services use background sound in an extra-diegetic fashion. It is intended to 'say' something topical or to establish the institutional and/or the designer's own particular sense of character and competency. Some sites make this extra-diegetic approach to ethos obvious. The Prince.it plays an eight-second loop of the Prince funk song "Kiss," a combination of pun and the "coolness" of the song telling the user how light-hearted and fun this particular Italian spa and gym can be. Most, however take a more subliminal approach. In order to implicitly advertise that by "jazz" is

[15] The terms 'background sound,' 'incidental sound,' 'environmental sound,' and 'ostinato sound' are associated terms but carry their own undesirable connotations in this context.
[16] Incidentally, there is no CSS scheme for sound such as for images.

meant something other than big band, “smooth” jazz, or fusion, Atlanta Jazz Society.org uses a 16-second sample of a “straight-ahead” jazz quartet recording. Upon opening Disney Travel.com, the 18-second pulsing string-synthesizer-choir-harp-drum extravaganza sample reinforces the “where dreams come true” branding of Disney theme parks. On Digital Strife.com, one finds a full-screen shot of a natural beach under a blue sky accompanied by remixed Bob Marley tracks. Digital Strife.com, so we are subliminally told, is an alternative to the rule-governed and strife-inducing digital world, an unexpected juxtaposition in that one accesses such an alternative through the Internet. A similar juxtaposition is the confused message of Qssis.com, the website for a Toronto banquet hall. After the user selects the Flash page option from the landing page, a massive explosion, complete with a fire ball raging at the user, occurs. The background sound is a remix of Billy Idol’s “White Wedding,” with original vocals and additional overdubs, which plays for over five minutes. Oddly, Qssis bills their banquet hall as having “an elegant ambiance,” but seem to be unconcerned that not only is the phrase “white wedding” slang for cocaine, but that the song is centered on the themes of illegitimacy, infidelity and incest. One simply marvels at the rhetorical implications such an example poses.

Unclear use of background sound has nearly eradicated it from the Web. It is virtually a foregone conclusion in the minds of many users who believe that background sound inherently “doesn’t have a clear role” (Follett). Sound is contrary to the hush of the Web genre. Sound upsets the expectations of the user because it upsets the convention. But upsetting the convention does not assure the failure of anything. The success of this dissonance is largely dependent upon the relevancy of the sound and, to some degree, the prejudice the user has towards sound online. Sound captures attention, if only for a period of time long enough for the user to recognize it as useful, interesting, or to reach for the mute button of the speaker system. If the sound bears little

or no relation to the purpose of the website at all, this realignment may privilege the relevancy of the text over sound in the mind of the user, defeating the intentions of the sound designer. And the Web is always willing to demonstrate such examples of defeated intentions via auditory irrelevancy. Trois.com plays a 30-second 'ambient house' track, the sound of which is not readily associable with the upscale urban restaurant. Perhaps the track is played in the restaurant, but nothing about the track indicates that this is so or otherwise gives either a justification of its inclusion on the site or any information about the content of the site. Something similar can be written for Strictly Fashion.nl, and The GFD Group.com, the former a fashion site based in the Netherlands, the latter an American funding and development group, both of which feature rather pointless background sound. The rationale for irrelevant background sound requires that the user be taken by force, a maligned attempt to coerce relevancy where there is none. The text on "Vibrant" D W Fitness.com states that one "should already feel the energy" (as a result of the music), although the nearly five-minute background track proceeds at a rather slow tempo with occasional sound effects. Irrelevant background sound was ubiquitous for a time, most of it banking on the unfounded assertion that audio-enabled webpages "outperform silent pages" (Helmstetter and Simpson xviii). Not only is background sound a contra-conventional risk, scant adherence to usability conventions has greatly impacted the success of background sound and has contributed greatly to the distaste many web users have for online sound in general. The most flagrant of all offenses is that designers often did not provide users a way to turn the sound off. As a result, it is now the norm for tutorials to strongly caution designers against putting sound on a site[17] and nearly forbid sound without also providing controls[18]. Because the issue of

[17] "…be aware that many people find inline sound annoying" (W3Schools).
[18] "If you include sound on a page, be sure to provide a way for users to turn it off…the sound might start to irritate them" (Lemay and Colburn); "there always needs to be a user preference setting to turn [sound] off" (Nielsen Alertbox).

user control has been most forcibly commented upon in regards to audio, designers have largely abandoned the sound-*sans*-control model. Unfortunately, it still happens that background sound is created with poorly-visible, confusing, or non-existent controls.

The Variety of Signal

The signal is the basic unit of sound on a website, capable of "being organized into quite elaborate codes permitting messages of considerable complexity" (Schafer 10). Although background sound is related to signal in terms of its overall relation in the soundscape (where there is only background sound, the 'message' can take on greater significance through the very use of sound online), it usually has a subordinate relationship. Signal can be used as an explicit branding technique, such as the two-second abstract swoosh of Move With Velocity.com. It can also be an abstract sound with no clear relation to the site or institution at all, such as the 'spaceship control' sounds of GSU Recreational Services.edu and The GFD Group.com. Signal draws attention to itself, as it urges the user more forcefully than background sound to make sense of it; it communicates *something*, even if it turns out that the user subsequently ignores or terminates it. The basic unit of signal is the 'earcon,' which is a sound in an interface representing "a full range of 'messages and functions, as well as states and labels'" (Blattner et al. qtd. in Walker and Kramer 152). Earcons are designed to represent a simple hierarchical language of sounds, often musical in nature, in a metaphorical relationship with the action of the user (Walker and Kramer 152). Peroni Italy.com employs earcons in a near-textbook manner. Upon clicking the main navigation menu, a short percussive sound plays and a drop-down menu appears. When a sub-navigation link is activated from the menu, a second percussive sound plays (followed by an abstract 'swoosh' sound as the new page loads). The earcons thus give the user feedback that a certain action has occurred, as well as an indication of the hierarchy of the

site via pitch. The most prominent pitches of the first main navigation level are approximately 784 Hz and 988 Hz (the G and the B one octave above middle C), sounding nearly simultaneously. The sub-navigation pitch is 1047 Hz (two octaves above middle C). When the C pitch of the sub-level navigation sounds, Western ears tend to hear it as a resolution[19], a culmination of the now-destabilized G-B pitch structure. A non-musical, but related approach to signal is that of the Mediastorm.org website. A series of images take up most of the screen space. As these are rolled over, the rest of the images darken (in effect highlighting the rolled-over image) and a portion of audio sounds. Because the nature of the sound (a sample of the voice narrations that make up each page on the site) tells the user less about the site structure and more about the actual content, this type of signal may be more accurately described as an 'auditory icon.' Auditory icons represent "objects, functions, and actions to the user… intended to evoke the sound of an object or action" (Walker and Kramer 151-152). Though a branding device, the two-second abstract swoosh of Move With Velocity.com is an example of auditory icon as it represents a particular type of object (the 'swoosh' Flash movie that plays simultaneously). A door-chime sound on the homepage of Ciao Bella Hair Salon.com, which prominently features an image of the front door of the salon, denotes another example of auditory icon. Vip Homes 4 You.com sounds short auditory icons as the user mouses-over the man navigation links. These sounds are nondescript and do not relate to the site hierarchically and relate only loosely to the metaphor of the site-space in the capacity of matching a sound event with the movement of the cursor across the screen. Many sites use auditory icons in this rather hesitant fashion. George Carlin.com uses a 'futuristic' sine-wave sample when any of the main site navigation is clicked on, which does not relate to the page or to the late comedian, except, perhaps, in his penchant for

[19] A functional relationship is implied though not made explicit unless a third pitch sounds.

highly random gags. The musical personality site Unga-Unga.com uses a different sound when any of five out of the six main navigation areas of the site is clicked: an abstract with an ascending phaser effect (Info), an abstract with a sweeping phaser effect (Music), an emphysema-like cough (Zam), an explosion (Paul), and an obscure film quote (Links). Mildly humorous, bordering on ridiculous, this use of auditory icon, though indicative of action, is only marginally useful.

Voice-over narration represents the lion's share of signal on websites. It typically takes the form of a single voice (that of a professional voice-over artist) communicating a message about a particular product or service. Voice-over narration seems relatively resilient because of the belief created by advertisers that one can reach users "INSTANTLY, where it matters most... THEIR HEARTS!" (James). The empathetic nature of the human voice issuing forth from headphones or a set of computer speakers is promoted as one of the only remaining methods to "cut through all the meaningless, random white noise of life and penetrate the consciousness of [the] targeted audience" (Bader). There is some truth to this, not necessarily because our hearts are located in our ear canals or because we are distracted by meaningless noise, but because signal demands our attention. Like background sound on a website, narration is novel enough that we do not expect it. The human voice is the most striking example of sound used online. Ceding some credence to the marketers' position above, the voice is evocative, if nothing else, because there is almost no chance that a user would confuse that sound with a computer sound, not even the stilted 'computer voice' of screen readers. Biologically conditioned to recognize 'voice,' the brain immediately associates a plethora of "elaborate codes" (Schafer) with culturally-contextualized cues. "Manifest in every speech act" (Wennerstrom 60), these culturally-contextualized cues—the quality of the voice (raspy, breathy, lisping, clear, deep,

high), the sex, gender, sexual orientation, race, age of the speaker; the intonation, pronunciation, dialect—are bound together in a seemingly impenetrable conglomeration. Often, the brain's only immediately-available response is to recognize this signal as a particular brand of genre, an advertiser's genre. Such a genre implies fairly consistent standards of intonation, copy, and quality. Most recognizable is favorable mood of the speaker towards the copy. This is characterized by a rapid pace and excited intonation, or an 'extremely enthusiastic' quality of voice. Cues usually denoting new information to a discourse are used extensively: high pitch boundaries (usually indicating new information being added to the discourse) and high and low rising pitch boundaries (usually indicating subsequent discourse). Such devices are used subconsciously in everyday speech patterns to indicate where a speaker is going with any particular topic or how long the speaker demands to 'have the floor.' The ear is thus highly conditioned to hear and respond to these cues, waiting patiently for an indication that the speaker has 'made her point.' While some uniformity of intonation can be noted from the examples present, uniformity of pacing and duration of the signal is less apparent. The narration on Round Way Roofing lasts 15 seconds; the narration of Progressive Dyn lasts 92 seconds. One of the most succinct, Alan Dearing.com has a 14-second narration. The narration of Copy G.com runs 32 seconds. Furniture Wholesale Group.com and All Star Supply.com feature rather lengthy narrations at above 60 seconds but offer controls to stop the audio. The particular intonational approach of Progressive Dyn, All Star Supply.com and Furniture Wholesale Group.com becomes increasingly exhausting as the high-intensity of speaker intonation continues over time, while the more relaxed approach of Alan Dearing.com, and Copy G.com is less exhausting. But the perceived amount of relevant information can also change the perception of intonation, making us more sensitive to spurious information the longer it goes on and more excited the voice

sounds. Round Way Roofing and the 30-second narration-only adjective-gaunt description of Hotel North Wales.co.uk sound neither 'overly enthusiastic' nor redundant. However, the 49-second narration of Hair Paradise Salon.com advertising "the radiant beauty and confidence you've always dreamed of" with a smooth jazz background playing for another three minutes, then looping, in an overly-excited intonation and self-gratifying tone, encourages one to either surrender to it or categorize it as blatant hype.

Signal forces realignment of the otherwise dominant visual aspect of the Web. The page can be seen all in one glance; the cognitive process of the user's ear requires some length of time until the signal is identified and contextualized. Until this occurs, users may find it difficult to return to the visual text of the webpage. Branding signals can thus be powerful attention-grabbers. It is counterintuitive to note that narrations nearly always focus attention off-line. They serve as commercials for the product or service ultimately offered, thus revealing little about the actual content or organization of the website. One "scriptwriter and audio producer/director" actually encourages this non-diegetic approach, writing, "there's no point telling people what they already can see" (St. Mauer). While the narration of Hair Paradise Salon.com never makes any mention of the site in any way, Aerial Digital Media.com demonstrates an extreme by actually welcoming the user/visitor to "5700 Havana Drive in Northridgeland Hills," a property the real estate company for this website is promoting. The majority of sites with narration, such as Copy G.com, Furniture Wholesale Group.com, Camelot-inn.com, B L Crosses.com, give a cursory nod to the medium by stating "Welcome to (site name)." Narrations sometimes briefly reference some feature of the site, such as a featured area of the site to visit or a particular button to click on.

Not only is there a degree to which signal can comment on the hierarchy and space of the website, but also a degree to which the signal is relevant against background sound. As background sound and signal are frequently synchronized, they comment on each other in a reciprocal nature. In an intra-audio paradigm, the pairing of background (usually music, but sometimes ambient sound, such as on Whisper Audio.com) to signal may create any number of qualitative situations: ironic, zippy, ambivalent, excited, sad, etc. A highly-contextualized sonic hierarchy is created based on recorded musical conventions of tonality, era-specific medium characteristics, personality, tempo, encoding format, style, etc against the sex, gender, sexual orientation, race, age of the speaker; the intonation, pronunciation, dialect, encoding format, etc. In an inter-audio paradigm, the background-signal audio realigns the relation of audio to visual text, almost breaking apart the conventionally accepted medium for the communication of information on a website.

Introductory Pages

Once fashionable pre-home pages called "splash screens," introductory pages may serve any number of functional or rhetorical functions: present the user with options concerning bandwidth, language, local server, or simply establish the feel, intensity or sense of the style of the site. If audio is present on the main site, the introductory page usually serves to indicate and can give the user the option of proceeding to a non-audited version of the site (or allowing her to turn on/off, down/up her speakers). It may also be the only page on the site to use audio, such as Sweet Thang Unlimited.com, which plays a looped, MIDI version of Hot Chocolate's "You Sexy Thing." Mediastorm.org indicates the further use of sound on its introductory page by low-level 'pings' corresponding to visual blocks that load on the screen. Sunnydowns Hotel.co.uk

simply "Welcome[s you] to Sunnydowns Hotel" over a light background track. Digital Strife.com features a multimedia introductory page that attempts to establish the purpose of the site. Advertise Door To Door.com also features a multimedia introductory page followed by a video narration on the main site. The introductory page of Whisper Audio.com, a web audio design company, is silent with written and visual text denoting further use of sound: a large graphical rendition of a speaker cone outputting a stylized representation of sound is set between the written text phrases "audio logo," "sound effects," "music for websites," and "background music," among others. These phrases not only potentially indicate further sound but serve as an example of and advertisement for the company's services. Users are also given a choice of continuing to the main site or watching the multimedia introductory Flash movie. Strangely, there is no option of entering an audio-disabled site—and the Mute button on the main site does not work. The bands The Cure.com, IGCMusic.com, Steve Dancz.com, and Hörstreich.de use introductory pages to either provide samples of their music or other abstract audio to indicate further use of sound. While these introductory pages may somewhat further the concept of physical metaphor of entrance followed by a series of navigational options laid out in different positions on the page, they are not constructed in such a way so as to convey overall structure in terms of sound. In other words, the pages sound discreetly and do not relate to each other sonically. If the sound of the introductory page were to map to the main site, an audio continuum would be established between the introductory page and the rest of the site. The spatial metaphor would, in this case, present the user with an 'outside' that is nonetheless a part of site. Of the 120 sites surveyed for this paper[20], only Mediastorm.org is constructed in such a way as to integrate the 'outside' with the inside, showing that one leads to the other. Once the file for the main page

[20] See APPENDIX. A survey of 120 audio-enabled sites was assembled from June to September 2008. The criterion: site must prominently feature audio as an integrated feature of the page.

has loaded, the blocks, initially serving as visual reinforcement of the loading process, open out to form the sections of the main page. These sections are highly-evocative multi-media montages of people's lives: ex-junkies in Manhattan trying to get clean and raise a child, families of hurricane Katrina victims in Louisiana and Mississippi, a shell-shocked American soldier returned home from the Iraqi war, dying villagers in the Congo. The continuity between introductory page and main page implied by this particular design begins to establish that sense of space missing from every other audio-enabled introductory page and main site in the survey.

Interactive Audio

The metaphor of space is realized in its most profound sense in interactive audio applications. Usually artistic in scope and "entirely predicated on interaction" (Andrews), interactive audio may suggest potential directions for further development of web audio in general. Although interactive audio applications are predominantly musical in purpose, they firmly establish a figure to ground relationship through the use of background sound (which changes over time and in response to the movements of the mouse) and signal. While the designers of interactive audio scenarios obviously seek to construct a specifically interactive auditory environment, they do not neglect the visual space, often developing extremely interesting and unique active graphics, symbiotically wedded to the soundscape.

"Sticky Sound Elastic Structure" is a complex nine-pointed geometrical shape with a differently-pitched tone at each point inside a three-dimensional cube. The cube is rotated with the mouse at any speed and in any direction. As the cube moves, the geometrical shape inside is tossed around. As its points land on different surfaces of the three-dimensional cube, different sounds are produced. Although no auditory spatiality is established (through techniques of

panning, volume, or reverb), three-dimensional space and movement are established and explored. In any one of the eight works under the umbrella of "Altzero," one uses the arrow keys and mouse to navigate a simulated three-dimensional space. Where the user 'is' at any given moment determines the audio output. The reversal of the text to sound power dichotomy is apparent: "Altzero…envisages a gaming environment where everything is stripped out except the sound and the ability to move. The visuals then become a reference for the sounds, another means of deconstructing and understanding the structure of what is being listened to" (Squidsoup). Perhaps in the opposite direction is "Actionist Respoke," where the visuals are deliberately designed to be difficult. In fact, the entire interface "develops a chaotic/poetic dynamic of its own until it totally gets out of control" (Schlömer). Floating objects are 'captured' in a box attached to a string. Once captured, they play a low-end drum and bass loop and the screen functions as a two-dimensional map: as the mouse is moved towards the top of the screen, the volume of the sample increases; moving to the right or the left of the screen pans the sound in the respective direction, even as far as complete silence in the opposing speaker when the captured sample is completely to one side. New samples are introduced and are integrated by the user into an overall sound scheme. "The Square Root Of-1" features a simple visual interface in which the user can change the direction of the music, pitch, or speed by grabbing and dragging with the mouse. The individual sounds of the "Pianolina" are made up of taps on the strings of a piano with non-descript objects, strikes to the body of the piano and rakings of the strings, as well as the conventionally played keys. Sounds are represented by colored squares that are dragged and dropped into a sound space affected by metaphorical up-down gravity. There are controls that allow the user (or "player") to increase or decrease gravity as well as to highly randomize play. Individual tones (those representing the 'conventional' playing of a piano) may

be modified in terms of speed, direction, and spatial orientation in real time. There is, however, a polyphonic limit: as more elements enter the scenario fewer can sound their entire duration as the other ones coming into the scheme cut them off with their sound. In further support of the spatial metaphor, the sound of the tones follows the left-right movement of their visual equivalent in the visual interface. A multimedia site related to a band, In-Colour.net, is an interactive application which plays the band's music while allowing the user to activate certain sounds (and moving images) by mousing over them.

Interactive audio presents potential novel innovations for web sound. In spite of its innovative graphical design, the genre is overwhelmingly auditory and musical in scope. It presents a sonic environment in which the visual does not dominate the aural; in which the audio reacts and changes with the actions of the user; in which sounds are musically shaped: "a continuous field of possibilities lying *within the comprehensive dominion of music*" (Shafer 5, original emphasis). As exciting as this is for online audio, it may prove to be incomplete for the genre of the Web. The retrieval of discreet, textual and graphical information is the prevailing convention of the Web, a convention that need not and would not be rashly dismantled. Instead, the goal of audio-enabled designers should be to use the musically sensitive and spatially aware concepts of interactive audio applications in textually significant scenarios.

Discussion

The range of sound available on the Web is broad, from seemingly frivolous musical background tracks to interactive, 3-D sonic environments in which the user can create her own experience of the page. However, of the estimated 19.2 billion webpages in existence (Elert), it is unlikely that a statistically significant number incorporate sound as a functional component of

the page. Browsing the Web is most often navigating on a sea of silence. For all of the interesting and productive instances of the online soundscape, there are significant problems. If sound is present, it is more often than not poorly executed from a usability standpoint or simply banal. Users have learned this over the years and resist dealing with audio-enabled pages in any way possible. An add-on for Mozilla, Firefox, and Netscape called "Flashblock" simply blocks all instances of Flash, which often plays sound, from ever loading[21]. Unfortunately, too many users have been bogged down in sites designed with the unquestioned assertion that audio-enabled webpages make a website better, more compelling, thrilling even. In practice, however, background sound typically has not offered commentary on the content of a page, the site, and certainly not the site's place in/on the Web. However, due to the ease with which background sound is implemented and its (perhaps questionable) valid uses on band pages and the like, it is the most environmentally stable. We may conclude that certain personalities will always have a use for sound on their websites.

The apparent lack of understanding concerning even the most accessible of auditory display concepts is astounding. Researchers have repeated the finding that "displays using multiple parameters of sound generally outperformed unidimensional displays" (Pollack and Ficks in Walker and Kramer 164). Research has been conducted in order to verify the amount of annoyance auditory feedback conditions would cause when compared to a visual-only feedback condition. The results demonstrated that aesthetically pleasing (short interaction sound) auditory design was clearly preferred among users and could lead to performance benefits over not only a design with no auditory enhancements, but also a design with aesthetically less pleasing auditory enhancements (Ronkainen et al.) Other findings have shown that narration, auditory icons, and

[21] Other non-affiliated Java Script blockers (Annoyance Zapper and Flash Mute) have also been developed for these and other major browsers.

earcons may be used in conjunction to increase the effectiveness of an interface. Researchers Maria Vargas and Sven Anderson outline a scenario in which a musical hierarchy of earcons is used to convey additional location cues during menu navigation in a spoken automobile interface. These earcons are very brief percussive sounds tuned to represent the "depth" of an item in the hierarchy of an automobile console interface. The top level menu items are represented by a particular motif and simulated instrument (piano, chorus, bells, and horns), arpeggios and chords alternating to increase distinctiveness of the top level items, while sub-level menu items inherit the instrument and pitches of the top-level motif and are further differentiated on the basis of melody and rhythm. Vargas and Anderson conclude that after users become very familiar with the interface, earcons shorter than spoken words might make navigation quicker for spoken and visual menus alike. This design is additionally notable for the care with which was taken in regards to matching individual pieces of the sound scheme into a coherent whole, a design point which accounts for the cognitive problem that arises when incompatible sonifications are forced to compete in a single design (Walker and Kramer 163).

Outside of the work done by interactive audio designers, almost no progress has been made towards full audio-only interactivity. This reality has been noted by several in the field, and is stated most explicitly by Goose and Möller, who report that "little work has been conducted on interactive audio-only hypermedia systems" (363) and what has been done does not "exploit spatialization or simultaneity in…audio rendering or orientation aids" (364) and is not "concerned with the interactive browsing of HTML documents" (364). The focus of their work is to provide a conceptual framework in which a 3D audio-only browser may operate. Although that browser has not yet emerged, the ideas presented suggest approaches to the

implementation of sound on available browsers, as well as possible new features of open source browsers.

The Web remains in a very rudimentary stage of its auditory capabilities. Even if it were to remain a two-dimensional projection, the use of sound could allow us to experience its metaphor in a three-dimensional space. The metaphor can be exploited to establish a sense of spatial dimension and/or time to the interface, the two in tandem binding the webpage to the sound. Sound would then function over the entire site, following the user and reacting to the user's actions. But this is actively constructed by the designer based on how the user will construct it in her own mind. It could assist user orientation in the site overall, as well as permit the user, for example, in the case of handheld devices, to focus visual attention on other stimuli, interfacing with the machine's auditory interface (McGookin). With continuing non-standardization, it is unlikely that sounds used as functional elements of web pages available on the Web will ever become popular. Few industry standard sites use sound, most of the major sites not bothering with it all. One considers the attitudes towards web sound in relation to the 'talkies' upon their arrival: sound is plebian and unnecessary.

SITUATED MECHANICS

On the Star Trek television series[22], a device called the 'transporter' was capable of instantaneously relocating material up to 40,000 kilometers away by scanning atoms and "beaming up" the information. Mostly, this material was human. Although the 23rd century crew "frequently use[d] the transporter…," the original transporter of the 22nd century was somewhat of a notorious and avoided device, its occasional problems leading to the grotesque spectacle of a horribly misshapen body, existential conundrums of two personalities merged into a single body, the horror of people beamed into open space and a three-decade long disembodied suspension of a human in a buffer (Wikipedia). Given this usurpation of power in which the creator becomes more vulnerable than the machine, one is invited to ponder the transformative potential of any machine[23]. Though the closest the 21st century has come to the existential problems of the transporter is the cybernetic organism of social infrastructures and restorative bodily systems, we may also consider the effects of a transporter-type system on the artifacts of culture. The parallel is not difficult to draw in terms of online audio, which is especially subject to malfunction. The Internet is a *packet*-switched network. Unlike the *circuit*-switched telephone system, which uses a physical and electrical connection between two points, there is no constant physical connection between the Internet server and the user machine. Whereas *circuit*-switched networks are useful for very rapid transmission or where instantaneous interaction is required, *packet*-switched networks are more efficient when instantaneous reaction is not required, but "when very low

[22] Other filmic examples might be The Matrix franchise or the character "Vox 114" from the motion picture The Time Machine (2002)

[23] It is unsettling to note that such reflections are largely confined to science fiction; that, apart from a few highly politicized examples, technology and technological artifacts are considered innocuous or are not considered at all by the majority of society. By allowing technology to silently/transparently/weightlessly render our world, we subject ourselves to a fundamentalism of the perpetuity of advancement, the expectation that someone somewhere will drop something benevolent out of a laboratory to enable us to do the things we want to do. We later naively eulogize its "birth."

‘corruption’ of data is paramount” (Clark 105). Although individual data packets may not be corrupted, there is no stream of messages passing between the two same end-points. This is a crucial distinction as audio is inherently time-based, any interruption in the signal destroying the continuity of time. There are other challenges to real-time Internet delivery of time-based media as well. As the expense of bandwidth requires small file sizes, as much information as possible is gutted from audio files, causing audible differences between the compressed file and the non-compressed file. On the one hand, the Star Trek transporter would actually be somewhat of an ideal scenario for online audio: any amount of information transportable to vast distances with no loss of information giving the transported information the same qualities in the destination location as it had in the starting location. On the other hand, these challenges, though they have generated several alternative protocols and other work-around methods, are not necessarily ‘bad;’ they may actually constitute a component of online audio genre. Jonathan Sterne, documenting the early history of sound reproduction technology in The Audible Past, argues that similar issues flanked this emerging technology and that it was the job of advertisers to counteract the fact that listeners “knew very well that it was impossible to create a truly transparent sound-reproduction technology” (267). Sterne argues that listeners had to be convinced that, due to the supposed ‘amazing fidelity’ of early twentieth-century phonograph technology, live performance and recorded sound were, in fact, interchangeable. Fidelity, therefore, was marketed as the transparency of the medium, an assertion that has flanked every newly-introduced sound medium since. If one traces sound reproduction technology from the phonograph to the present technologies, one arrives at the Internet, which poses challenges to audio unprecedented by former media. The two most significant innovations have been ‘streaming’ and encoding / decoding technologies. These technologies have not only made real-

time online audio feasible, but constitute a particular kind of 'sound.' Just as obvious as it was to the late-nineteenth century audience that "different machines had sounds all their own" (Sterne 267), different compression and delivery methods carry 'sounds all their own.'

Compression / Decompression Schemes and Psychoacoustics

There is no native file format for sound online. Sound, therefore, is at a considerable disadvantage to other forms of data on the Internet. MP3, now "an Internet music standard," (About.com) has filled the gap somewhat by allowing designers to implement sound that is reasonably transmittable over lower connection speeds. Originally designed to transmit music over telephone connections, the MPEG Layer format quickly became popular in the late nineteen-nineties due to the convenience with which it could be used on portable audio devices. Since that time, few formats have been any serious competition in public opinion, which has embraced the MP3 and is investing massive sums of money in it (Thomson), even with its demonstrably lower resolution and clarity when compared to similar formats (Xiph.org). Although pop culture was swept off its feet years ago by the MP3, it is still possible to find cheerleading for its quality, convenience, and, obligatory in the current era of metaphysical advertising, transcendence: "mp3 is more than a technology. It is a sensational development that has reconnected musicians to music lovers, speakers to their listeners, creators to their audience" (Thomson). Such accolades do the work of inventing "the imagined correspondence between live and reproduced [sound]…along with the sound media" (Sterne 285). In the case of the MP3 format, fidelity and quality are invented with the help of the particular methodology used to create it. Of particular note is perceptual coding, a manipulation of tenets of the field of psychoacoustics. This encoding scheme states that certain information contained within an audio

file is outside of the spectrum of human hearing and can therefore be eliminated with no perceivable difference in the resultant audio. It is an interesting idea but one that does not live up to its promoters' pronouncements. Far from delivering audiophile sound, the format carries 'sounds all its own,' most crucially because the MPEG format is a 'lossy' compression scheme. There are two types of compression schemes: lossless and lossy. Lossless makes no impact on the resultant audio; it allows the exact original data to be reconstructed from the compressed data. The impact on the audio of lossy compression schemes, however, is highly variable. They attempt to evaluate the component of the output that is 'irrelevant' to human listeners because it falls outside the hearing threshold, (or because it will be masked by adjacent content) conveying "the essence of the sound rather than the waveform" (Stuart 17). MP3 is, therefore, not designed to reproduce high-fidelity audio. Against the claims of those promoting it as a high-fidelity, 'near-CD quality' format, there are many 'lossy' schemes in addition to the ubiquitous MPEG Layers 1-4: Ogg Vorbis, Advanced Audio Coding (AAC), and Musepack (MPC), to name a few. The culmination of internet infrastructure, bandwidth, user sound card and speaker set-up, and compression scheme, gives internet sound a place in sound production history.

Sound, Hearing, and Getting Sound to the Web

Any lossy compression scheme is not necessarily of poor quality. It is the result of simply one more step in the transmission of sound. Once sound captured in 'the real world' has been reproduced by technical means, it is something else altogether, at best only *similar* to the 'original.' Sound reproduction technology, though exterior to the body, is built to respond to natural stimuli in an anthropocentric paradigm (Sterne 11). In fact, the very nature of sound reproduction devices (microphone sensitivities and magnetic tape bandwidths, for example) are

constructed to respond to and contain only those sounds falling within the spectrum of human hearing, automatically excluding the infra- and suprasonic frequencies that stimulate our kinetic sense. Just the same, the microphone and the speaker alike do not hear sound, but feel it[24]. Sound is created in our brains by acoustical energy in our ears. Acoustical energy, a form of mechanical energy that propagates in waves in any particular medium, works by pushing molecules of air against themselves, other objects, bodies and eardrums. The molecules respond by swinging back toward their original state but usually overcompensate, swinging a bit too far in the opposite direction. This process of compression (the push) and rarefaction (the 'pull,' or the swinging back of molecules) is one 360° cycle. Acoustical energy is made up of two mutually-dependant elements: frequency and amplitude, what we generally refer to as pitch and loudness. Frequency is a function of time, telling us how many cycles occur during one second. A frequency of 1 is therefore one complete cycle every second[25]. Although acoustical frequency may be somewhat vague to most of us, we are familiar with its corollary, pitch, usually by way of comparison: a person singing 'off' or 'out of key' with music, the familiar change as a fire truck approaches and passes with sirens blaring[26], the difference of an engine turning at 1,000 rpm and at 5,000 rpm. These are the bare minimum acoustical concepts one should have when dealing with theoretical notions of sound. In order to hear sound from a webpage or across the Internet, it first has to go through a seemingly simple, though extremely complicated process beginning with a microphone and ending with a speaker. Every link in this chain, including the characteristics of the spaces in which the sound was recorded and is re-audiated, shapes the

---

[24] It is interesting to note that many of the commentaries of sound beginning with phenomenological explanations of cultural sound do so under the influence of the ephemeral nature of sound. While that is one position, it tends to deemphasize a more fundamental reality of the matter: sound is, after all, just as much mechanical as it is perceptual.

[25] Calculating wavelength as a quotient of the speed of sound (1,130 ft/s) over frequency returns an in-air wavelength of 1,130 feet!

[26] The Doppler Effect

sound. In order to show a small fraction of this complexity and the elements altering sound, I'll discuss the recording chain, digitization, and storage methods below.

In order to go online, sound must be digitized. The beginning of this process is the microphone, commonly referred to as the 'first link in the signal chain.' Microphones work on an electrical principle stating that mechanical energy can be converted to analogous electrical energy. As acoustical energy compresses the molecules surrounding the element within the microphone capsule, its electric field is disturbed, converting that mechanical energy into corresponding electrical signals (and vice versa), usually alternating current[27]. Acoustic energy becomes electrical current; sound has been converted to a possibility of containment and reproducibility. It then needs to be transferred (via cabling or radio waves—an additional step in the signal chain) to an amplifier, as the electric voltage[28] from a microphone is extremely low. Digitization is then a breaking-up of the mutual dependency between frequency and amplitude so as to render them reunited back into the world. The machine accomplishes this first step by imposing its own frequency on the original signal and making measurements of the original amplitudes resulting in "a string of pulses" (Gibilisco 487). Each pulse of the string is called a *sample*, each sample "having [an amplitude] equal to the instantaneous value of the signal [amplitude]" (Everest 126). Each sample is then expressed in binary form as a numeric value. The quantity of samples made per second is referred to as the *sample rate*. Though sampling rates can range from 8,000 to 2,822,400 samples per second, 44,100, often written as "44.1k," is the most widely-used for "listening-quality" audio. The samples are then enclosed in a container, along with other information, for storage. The most common types of containers are Audio Interchange File Format (AIFF), AU, Waveform Audio Format (WAV), and XMF (Extensible

[27] Alternating current is an electric current whose direction reverses cyclically
[28] Voltage is a measure of potential energy between two points of an electrical circuit

Music Format). These containers often are too large for the bandwidth available and to stream uninterrupted over the Internet. In order to improve data transmission, "any of three forms [is used]: reducing the bitrate, reducing the sample rate, or employing compression techniques" (Huber and Runstein 283). This means that the accuracy of the sample can be trimmed, the sample rate can be scaled down, or that 'non essential' information from the resulting file can be removed. Any of these methods creates discernable differences in the resulting audio.

Online and Back Out Into the World

The term 'sound reproduction technology' is perhaps a misleading term. It assumes the existence of absolute fidelity and dissuades us from thinking of the actual situations in which sound plays. Instead of *re*producing sound, sound systems produce sound unique to a given context. "*Re*production technology" privileges the original signal over the signal at any point in the chain. Sound holds on to the bodies that create it; resonant frequencies of chests, mouths, rooms, microphones[29], signal processing equipment, all leaving their aural fingerprint on the sounds coming from and affected by them. The effects of the actual transducer compressing air (a set of speakers, usually, hooked up to some kind of player) also have an effect. There is, in fact, no way for a body not to be communicated. The user system, consisting of processor, sound card, speakers, headphones, contrary to many assertions otherwise, has its own impact on sound. This system makes its presence heard through its relative ability to render the digital information continuously and accurately. As 'fidelity' is somewhat of a loaded term, one may only

[29] Although there are three types of microphones—ribbon, dynamic, and condenser—they all work on the same (transducer) principle. Because microphone elements are constructed from different materials, which react differently to the pressure of the molecules acting on them, a different output is rendered. The lighter the material of the element, the more low energy frequencies it can capture, which are typically the upper partials. Alos, there will be sounds present in the eventual audio signal that someone standing in the room at the time the recording was made never heard.

responsibly comment from a relativist position, which nearly always contains assumptions of value: how ‘warm,’ ‘full,’ ‘clear,’ ‘clean,’ ‘loud,’ ‘soft’ is the sound? How does it sound in one environment as opposed to another—a question that is made nearly unanswerable with any breadth with the use of handheld devices that can access the Web from nearly anywhere on the planet. The concerns of the effects of playback equipment, though they may seem esoteric in this setting, have been the concerns of audio engineers for decades. It is standard procedure for music engineers, for example, to mix and master music based on all the scenarios in which they anticipate the music to be played. This is often done simply by making several mixes and testing them out on small computer speakers, MP3 players, in automobiles, and in strip clubs. Web audio should be designed with the same sensitivity to listening environment and mobility. Sound carries with it the characteristics of every thing the signal comes into contact with through the recording process and takes on the characteristics of the space in which it is audiated. Unfortunately, web audio designers have been so preoccupied with the difficulties of getting sound to the Web that they have not adequately understood the complexity of getting sound back out into the world.

Discussion

Apart from certain artifacts, such as sea shells and hollowed-out tree trunks, sound reproduction technology is a thoroughly modern and thoroughly documentable endeavor. And it has never gone silent; our hearing, engaged in the modern world, engages a technological sound, whether we realize it or not; indeed what is natural and what is not is increasingly less apparent for humans as well as other creatures (Pescovitz). However, it may go unnoticed for long periods of time, largely because we take little notice of the major impact we have on our sonic

environment. To differ with Sterne, who asks one to "consider how rare it is for histories of newspapers or literature to begin with naturalistic descriptions of light and phenomenologies of reading" (14), I hold that for those charged with the task of technologizing and theorizing sound, such descriptions are necessary. Perhaps a detailed account of the "million-year adventure" of a photon from the center of the Sun to our eyeballs (deGrasse Tyson 68) is overkill, but an understanding of ink and the properties of paper is. For film to do what it is supposed to do, concepts of light and time had to be understood, from the chemicals on the film to the camera and lighting, aperture settings and film speeds, to the darkroom and editing room. Getting sound to stick to the medium is an endeavor that requires us to know something about how sound works and how the medium will respond. Without this simple knowledge, phenomenologies and philosophies threaten to subsume all other discourse. Online audio is at a particular disadvantage because 1. it requires an extra layer knowledge on top of sound engineering and because 2. there seems to be little motivation to develop audio tools, to bring them "up to the level of the tools create[d] for the visual side" (Burke 9). However, progress made in other areas is leading us into a potentially exciting moment for online digital audio. Just as technologies arose to make data files smaller and, thus, acceptable for transmission, bandwidths and processor speeds have steadily increased. It is in this cultural moment that online audio is less and less confined to the expediency of the lesser of two evils, making the techno-rhetorical undercurrents more accessible. Understanding of both these undercurrents and the nature of the phenomena is "the technological nature of our mediated world…that we have not yet sufficiently attended to" (Rickert and Salvo 298). In the mediation of the Web operating over the Internet is contained the same principles of embodiment that are contained in acetates and CDs. That makes us more vulnerable to the machine than we ever thought we would be.

TRANSMOGRIFIED DELIVERY

Delivery, a term denoting both vocal and physical aspects of a speech, is necessarily embodied—necessarily aural and necessarily visual. It is the study and practice of whatever actual and meta content and persuasion any speaker generates though the use of his voice and the spectacle of his physicality. For much of the history of rhetoric, transmission of the actual content of a speech was considered so akin to the canons of arrangement and style that it was eventually disregarded through annexation. Meta-content, however, was so highly subjective and situational that only a small number of authors ever committed its principles to writing, and even then left it shot through with over-simplifications and gaping holes. As writing and introspection replaced public performance and discourse as methods to "Truth," the delivery of actual content was transformed into the page, the book, the printed word; meta-content was entirely abandoned. However, electronic technologies are continually providing new ways of allowing us to communication apart from the written word. Radio, VoIP, telephone and spoken dialogue systems allow us to use our most basic instinct: spoken language. Human communication is conditioned to respond to and transmit all manner of physical and vocal cues, alongside what is 'actually' being communicated. Out of the nearly 15,000 words heard and spoken each day by the average person (Rayson et al.), we hear what we think is intended not by listening alone, but by listening and watching (Werner and Leibold 201-203). The auditory and visual modalities are locked together by the principle of synchresis, "the forging of an immediate and necessary relationship between something ones sees and something one hears" (Chion 5). This formation helps us hear better, both in the actual words spoken (Neuhoff 233-238) and the intentions 'behind' those words (Tannen) and cues us for our own speech. City dwellers are exposed to as many as 5,000 advertisements per day (Story), plastered on every conceivable space, from

subway turnstiles to the average 4 hours and 35 minutes of television watched each day (Holmes), to the 52 minutes of radio, audio books, or CDs listened to during the daily commute (Langer). By this audio-visual modality we perceive embodied phenomena. We see an advertisement and we see the physical component supporting it/making it possible: a billboard, a sign, a banner trailing a single-prop plane in the sky, an image projected unto a building at night, bumper stickers, T-shirts, vending machines. Just as much as we see what is 'on' the television, the movie screen, the overhead projector, the telescope and the computer, we see the television, movie screen, overhead projector, telescope, and computer. The physical properties do not disappear and never go silent. It is the very physicality of technologies that contextualizes our relation to the sounds and images produced (Rosenblum 233). The medium places conditions on the content and our relation to the content. But it is also true that the content places conditions on the medium. Contained in the study of delivery, this was an extremely important concern for ancient orators and it should concern us today: ethos is also derived from the perceived quality, timeliness, and appropriateness of presentation. There are various levels in this vein: how much differently do we hear a 45 rpm record than an eight-track when we know to which medium we are listening? How different might Ansel Adams' *Mount Williamson* look printed on cotton paper or projected in an IMAX theater or on the Macintosh 128K computer screen? Each of these artifacts not only changes the way the image appears to us and contextualizes our relation to that image via our relationship with the artifact, but the image itself changes the way we see the artifact. In other words, by contrasting the immediate presence of the image with our prior experiences to it, our relation to the artifact is changed. It is a hypermediacy in the sense that we are made intensely aware of the medium (Bolter 25). This does not go so far as to proclaim, as others have, that the medium subsumes or dominates all other aspects of a performance or

artifact. It posits, instead, a relationship between the two, neither of which is dominant. It is enough to consider technologies, the cultures that produced them and the perception of them holistically. One may discover that delivery encompasses this holistic notion, as it is inherently forced to reckon with the implications of rhetoric that have been side-stepped by many of its practitioners.

The Tacit Subordination of Delivery

Except for a few isolated examples, ancients as well as moderns have routinely ignored, even dismissed, a substantive approach to delivery. Not only is its description notoriously difficult, what little time it has been given, has been taken up by either praise or vilification. Perhaps no one was more acrimonious than Aristotle, who referred to delivery as necessary but "unworthy," its power due to the corruption of the audience (Rhetoric III, v, 1404). But his predecessors, the ancient Greek orators Demosthenes and Gorgias, considered it differently, the former calling it "the only virtue of oratory" (qtd. in Quintilian XI 3, 5), the latter writing that "the effect of speech upon the condition of the soul is comparable to the power of drugs over the nature of bodies" (Gorgias 14). Ancient Greeks Theophrastus, Athanasius, and Athenaeus likewise considered delivery the "most important thing for persuasion in Rhetoric" (qtd. in Caplan 190). Plato almost completely avoids the subject, making only passing remarks in Phaedrus and Gorgias. Ancient Roman writers seemed to have been equally polarized concerning delivery. While the author of Rhetorica ad Herennium judiciously explained that "skilful invention, elegant style…artistic arrangement…careful memory…will be of no more value without delivery, than delivery alone and independent of these" (III 19-20; 189-191),

Lucian viciously lampooned it[30]. Though Quintilian did warn against the mode of reading made "effeminate with unnatural softness" (I viii, 1-2), he provided very little instruction on delivery and, most peculiarly, no explanation as to how one goes about not sounding 'effeminate.' Augustine was content to caution against "the immoderate pleasure of the ears" (IV, 19). Contemporary rhetoricians, textbook writers, and theorists have inherited these attitudes and only gussied up the short shrift in modern garb. One of the great proponents of aurality / orality, Walter Ong, suggests the inferiority of this mode of communication with the assertions that analysis was made possible with the "interiorization of writing" (9), that sound is "essentially evanescent" (32) and, through what may be called the 'audio-visual litany,' that, in oral / aural cultures, thought and expression "tend to be" (37) additive, "aggregative rather than analytic," redundant, traditionalist, agonistically toned, homeostatic, and situational (37-49). In light of the relative lack of anthropological evidence with which Ong makes such claims (however philosophically persuasive they may be), one need not accept that such a dichotomy exists or has ever existed. Kathleen Welch believes, as did Cicero and Quintilian, that "there is no speaking without writing" (8), but, unlike Cicero and Quintilian, takes this as license to completely ignore the fact that the oral / aural modality (referred to by Welsh as "noise" and a "mere act of speaking") actually involves sound[31]. Although Welch acknowledges a principle of the transmogrification of delivery, she resolutely confines delivery to written communication (Welch in Reynolds 1993). Edward P. J. Corbett, in the self-acclaimed "most widely used textbook of its kind in advanced composition and writing," manages to describe in three short paragraphs

---

[30] "Bring then above all ignorance, to which add confidence, audacity….the loudest voice you can come by, please, a ready falsetto, and a gait modeled on my own…" (p. 224; 12). "Appeal constantly to the pathetic instinct, smite your thigh, mouth your words well, punctuate with loud sighs, and let your very back be eloquent as you pace to and fro" (pg.227; 19).

[31] From a position of aurality, all the ink of a supposed "unconscious deprivileging of writing" (16) is both laughable and blatantly offensive: if one's "multiply constructed" identity is inherently gendered, classed, and raced (24), so too are the sounds produced by that identity.

(22-23) the great importance delivery instruction held in ancient Greek and Roman culture and its descent following the invention of printing. But that proves only to be the wind-up: delivery training, according to Corbett, amounts to acting lessons and, besides, he remarks, "all the great orators in history have been great 'hams'" (22-23). Perhaps this quip is intended to be taken tongue-in-cheek, but in light of the fact that hardly another word is written concerning delivery proper in the remainder of the book, one may confidently come to other conclusions. In addition to Corbett, other textbook authors have "chronically, perhaps unintentionally…promoted a truncated version of the five classical canons . . . ignoring, misrepresenting, or failing to engage…delivery issues"(Reynolds 2). To follow Reynolds' understated phrasing here, writers are out of their league: so little has been written concerning delivery because of the difficulty of describing and proscribing delivery. A valid concern though not very persuasive. Composers have developed a system for notating sound in a fraction of the time. It seems strange that the most often restated advice concerning delivery would be 'do not deliver it badly.' The second most popular statement from orators laments the difficulty of writing the rules of delivery. The writer of Rhetorica ad Herennium says it outright: "no one has written carefully on [delivery]—all have thought it scarcely possible for voice, mien, and gesture to be lucidly described" (III xi 19)[32]. However, concepts of style, indeed, what we like to call "voice" in writing, are just as difficult, just as pragmatically sensitive, and our proscriptions can be every bit as sterile as those used to describe delivery.

[32] At III xv 27, the writer states, "I was not confident that it was possible to treat these matters [delivery] adequately in writing."

Embodiment and Gesture

Delivery has become an afterthought because scholars have played along with the idea of idealized truth[33] and the religiously-informed principle of the renunciation of the body. Because delivery is necessarily embodied, by purging delivery the soul could be purged of the body. Alongside the already-perceived difficulty of rules of delivery, delivery was left an impoverished intellectual pursuit. Whatever delivery has lost in tradition can now be refashioned to fit modern technology. The once public and visual spaces of where, in previous years, the audience was confined, are now subsumed by technological hardware. But this requires neither disembodiment nor denial of the senses: the speaker may still be seen and heard through the use of audiovisual technology. Where one or both of the audio-visual lock is not present, that aspect is contained in the artifact of the medium; the medium contains whatever sounds issues from it and the design of the interface contains the visual. There are strong precedents to this modern embodiment in ancient rhetorical treatises. The most comprehensive of these treatises is the Rhetorica ad Herennium. Much emphasis was placed on matching the physical movements, such as posture, hand/arm positions, facial gestures, motion, to the voice (III xv. 26-28)[34] and matching the voice to the intent of the message. An orator with no access to public address equipment could account for the difficulties of space, noise, and the relative strength of the voice, in part, with his gestures, directly reinforcing his words and imparting clarity to his audience. Physical movements included gestural techniques that produced aural affects of their own, such as stomping the foot, clapping the hands, and slapping the thigh. The motions and positions of the hands and arms, though they did not produce any sound loud enough to have any effect, were nonetheless important and proscribed.

---

[33] This process, at full steam in the late Roman empire, has been documented a number of times by Peter Brown.
[34] Also in Quintilian. On the Teaching of Speaking and Writing. I, xi, 8-11, 15-19

A computer interface-incarnation of these ideas is real-time graphic equalization display and the type and function of controls provided to engage with any sound. Real-time graphic equalization is a direct link between sound and seeing, visually reinforcing the audio at any given time. These visualizations, modeled on the graphic equalizers of conventional digital players (which were modeled on the Volume Unit (VU) meter of analog units), display a set of vertically-aligned bars that light up (from top to bottom) as the amplitude of the signal increases. Other types of visualizations, though they react in real-time to the audio being played, do not necessarily reinforce either the content carried as audio or the status of the audio itself. Windows Media Player, for example, synchronizes various abstract graphics with amplitude variations of the particular track it is playing. Like the proscriptions concerning gesture, which arose from convention and shifted with taste, there are proscriptions concerning the physical appearance of the interface. Conventional audio players, such as CD and tape players, have employed buttons with typography standardized since the early days of consumer cassette players. The filled-in right-pointing wedge has denoted “play,” the square has denoted “stop,” a set of parallel lines has denoted “pause,” an outlined right- or left-pointing single wedge has denoted “progress / regress” by a single track respectively, and outlined right or left-pointing double wedges have denoted “fast-forward” and “fast-backward” respectively. Major multimedia players such as iTunes, Windows Media Player, QuickTime, and RealPlayer have also adopted these conventions, as well as many web-based players. Apart from a handful of conventional sites with alternative audio interface schemes, the most widespread online deviations from this scheme are to be found in interactive audio, as discussed previously.

The Voice

Delivery is necessarily embodied not only to the degree to which it is visual, but also because of the relation of the size of the space and the relative power of the voice. Mastering the voice has always been a challenge to the orator, unaided as he was until the mass-production of public address sound reinforcement (PA) equipment, modern architectural acoustics, and electronic networks in the twentieth century. Outdoor spaces presented much the same difficulties as they do today. Though much advancement improved the transmission of the voice, it was severely limited in scope. Ancient theatres were ingeniously designed to focus the vocal performances of actors and move the audience closer to the performers, but could only do so for a relatively elite group of people and were always subject to the destructive forces of the wind (Blesser and Salter 97). Architecture prior to the twentieth century, though it created spaces in which sound had energy enough to carry through the entire building, wrecked intelligibility on account of the high signal to noise (in this case, reverberation) ratio. Sometimes, however, even this quality was used as an asset, such as Echo Hall in Olympia, where matters of state could be discussed 'privately' due to the extremely strong echoes created (Blesser and Salter 95). In every other discourse requiring intelligibility, the voice needed to "carry" over a large distance and over the noise of the crowd, all the while remaining distinct and controlled. Rhetors were thus constantly concerned with improving the voice to meet these demands. Indeed, two of the three aspects of voice quality in the section on delivery in Rhetorica ad Herennium address volume and stability. It is said that Isokrates, because he believed that he did not have a good voice, refrained from oratory altogether (Pernot 28). Demosthenes famously practiced his speeches by the roar of the sea, in an underground room, spoke with pebbles in his mouth...and recited prose and poetry while running all in an effort to improve his vocal delivery (Plutarch qtd. in Pernot

30-31). Quintilian believed simply that "a weak voice is incompatible with first-rate excellence in delivery" (XI, 13). Not only was the orator's voice required to be a "good, strong voice," but one that is not "ugly" or undependable (Quintilian XI, 13). It was to be clear and flexible, with no hoarseness, what is considered 'noise' or distortion in a communication chain. Students of rhetoric were advised to maintain a clear, even tone of voice and to stop practicing or speaking when that tone was no longer sustainable (Cicero III 224-227; Rhetorica ad Herennium III xi 21). Though orators were not in direct control of the environment, they did have some control over the mechanisms that produced sound, even if the general consensus tended to be somewhat cynical towards improvement due to a belief in the innate physical abilities of the orator and taste of the audience (Quintilian XI, 177).

A Negative Implication of Hypermediacy: The Barbarism

Beauty and clarity of tone could become more than simple aesthetic preference. They were indications of the state of a speaker's mind and soul. Especially in the later Empire, any tone that intimated anger or any emotion other than the idealized detached calm of the orator, was extremely distasteful, potentially leading to the loss of stature (Brown 48). An orator was expected to diligently and without fail demonstrate the excellence of his education, his paedia. Though such controlled conduct most likely contributed to a safer environment for all involved, it also demonstrated the speaker's willingness to submit to the values of the true wielders of power. In short, he was to demonstrate his trustworthiness for the ear of the ruler. Paedia was a notion particularly prevalent in the later Roman empire. But that code of conduct had always been present in relation to delivery. And it always translated to the demonstration of trust and to the state of the orator's soul. Any audience, from emperor to peasant, would have recognized

certain characteristics. This leads us to a notion related to delivery which many today find particularly insidious, the barbarism. Although we tend to recognize it as motivated by xenophobia, ancient cultures may have found it a necessity of cultural survival. From the mouth of the orator, the barbarism was a flagrant offensive to the "correct" sound of the language. Quintilian gives perhaps the most thorough explanation of the barbarism as, "an offense committed in regard to a single word…caused by…separation or confusion of syllables, aspiration, or other faults of sound" (Quintilian V, 6-7)…[and] that which we regard as proceeding from the natural disposition, when he, by whom anything has been uttered insolently, or threateningly, or cruelly, is said to have spoken like a barbarian" (Quintilian I, v, 9). From this final piece, it becomes apparent that the term was ethnocentrically motivated[35], which, put another way, means it is an acknowledgement of the speaker's non-conformity, confused allegiance, poor education, and weak soul. By pronouncing a barbarism, he demonstrates that his values are not those of the community. He is therefore to be distrusted. Where the "goal of…transparent presentation" (Bolter 25) is realized, an audience does not question the values of the speaker, but the content of his speech. However, wherever prosodic or gestural incongruence surfaced would detract from the actual content. An accent and prosody used by the people in the centers of power, i.e. Athens and Rome (though one was never to be substituted for the other) was a prerequisite, as demonstrated in Cicero's De Oratore. Speaking of "correct Latin," Crassus lauds the "accent characteristic of the Romans who are from the city itself," other dialects, he

[35] Judgments predicated on non-standard pronunciation in conversational and oratorical speech are considerably less acceptable than in eras prior to the last century. However impolite, extra-linguistic signals continue to access the symbolic power of discourse in which utterances are signs of wealth and signs of authority (Bourdieu 502); in that every manner of speaker from Standard American English natives, non-natives, females, speakers of vernaculars, homosexuals, etc have disparate footings; that "discourse is not only used for exchanging information but is also equally functional in shaping interpersonal and inter-group relations" (Jaworski and Coupland 291); that "we define ourselves by what we say, but more notably by how we say it" and that "our pronunciation allies or isolates us from a community of speakers" (Parrino 171).

suggests, "give offense," are "unpleasant," "provoke criticism," and in short, "sound or smell of foreignness" (III 44). Crassus even goes so far as to say, "let us...learn to avoid...peculiar foreign pronunciation" (44). To moderns, it was a nearly unbridled ethnocentrism; to ancients, it resolutely pointed out those who were most likely to betray Rome due to their non-Roman origins.

Although one cannot make a direct connection between technological media and the barbarism, there is a degree to which they are both representatives of the values of any particular community. Sound reproduction technology is the pursuit of fidelity, which is defined as a "set of social and sonic relations in which participants [can] have faith" (Sterne 274). If technological media betray their belonging to other value systems, for example, to prior eras, "social relations," or to otherwise differing aesthetic tastes, they may suffer the indignity of being directly stigmatized as inferior, inadequate, or anachronistic. Although hypermediacy often imparts a kind of innocence to 'inferior' technologies, its theoretical construct will be tolerated for only so long by an audience. The penalty against authority for sub-standard sound reproduction is as subject to the whims of taste and patience as is the requirement of 'correct Latin.' For moderns, as well as the Ancient Romans, this penalty seems quite high[36] (52). The barbarism, as the term most often used to signify both an affront to delivery and social and intellectual norms, bears some relation to those devices and technologies that fail the test of fidelity, the sound-reproduction equivalent of incorrect Latin. It is often the case, especially in online environments, that audio need be constructed in such a way as to diminish this very "virtue" of fidelity.

[36] Although of those present in De Oratore, no one is willing to construct an exhaustive list of how a speaker may fail.

Discussion: The Limitations of Preservation and Transmittal

In a technological sense, the limitations of preservation and transmittal have been and are the challenges of "properly and accurately" recording sound and making it available to times and/or locations existing apart from the original source of the sound. It is perhaps ironic to note that the modern quality of 'voice' (or any auditory information) is subject to the power of the transmitting medium, a problem managed, to at least some degree, by speakers prior to the harnessing of electricity. Though orators were not in direct control of the environment, they did have some control over the mechanisms that produced sound even if the general consensus tended to be somewhat cynical towards improvement due to a belief in the innate physical abilities of the orator and taste of the audience (Quintilian XI, 177). Orators could improve their voices and their gestures. They could choose locations or at least practice in locations suitable for speaking. The theatre of ancient Greece itself was not "merely…a place for actors to perform and the audience to listen, [but] an extension of the actors' mouths, producing, shaping, and propagating their voices" (Blesser and Salter 98). Yet the modern orator receives no assistance in the online environment, the end of which is completely of the user's own choosing: a living room, office, airplane, or park. He also has no control or influence over the quality, power, or design of the speakers and soundcards of the auditor. Whereas there was a small number of traditional spaces to manage for the ancient orator, modern electronic speakers can run the gamut, from the small, built-in speakers of most laptops to sophisticated, studio-grade systems routed to a computer. Such mobility and range of listening environments presents unprecedented challenges to the modern era. An additional challenge is that compression ratios are inversely proportionate to bandwidths, hosting plans, and user systems and configurations: the lower the average quality of the transmission apparatus, the higher the compression ratios used. These

limitations of connection speeds and processor rates has forced audio into a small number of file formats, none of them native to the Web or of especially high fidelity. This creates additional difficulties for the seamless delivery of any continuous content.

In eras prior to digitization, concealing the medium of either the vinyl record or magnetic tape, even FM and AM radio waves and other analogue transmission technologies, was a constant concern. Other types of technology throughout the recording chain, such as reverberation methods, transducer types, transistor electronics, and production methods of all kinds, carried with them a certain aural signature that were as often unwanted as they were characteristic. But these characteristics as seldom recognized. ‘Faults of sound,’ whether those produced by vinyl disc and record player, consumer magnetic cassette tapes, temperaments (musical tuning systems) of previous eras, or encoding schemes prior to those currently popular, are instantly heard. If they do not charm us for nostalgic or artistic reasons, they offend us: one need only consider the chagrin on the faces of the “dearly beloved” as the vocalist sings the lighting of the Unity Candle song in Renaissance-era Meantone temperament to the accompaniment of an Edison cylinder phonograph. Both were perfectly acceptable in their day, made acceptable to our ears only as museum pieces. We would certainly never allow *our own* music to be played through them.

The cues imbedded in not only the sounds of the voices in our digital transmissions but also the meta-meta content of the particular transmission medium chosen, indicate underlying assumptions of value. The value most in demand in nearly any sound reproduction and transmission technology is fidelity, which is a particular kind belonging that the auditor deems as ‘correct.’ Ancient Romans expected 'correct Latin'; Ancient Greeks expected ‘The Truth.’ All three rely on their own understanding of the Ultimate, and even lend a certain amount of faith

and assistance users to those technologies, speakers, and philosophies, enabling them to function flawlessly in the appropriate paradigm. Perhaps as modern-day physiognomists, we rarely resist judging the quality of any technology or artifact when the spectacle of its physicality is revealed. By doing so, we necessarily alter the artifact and our relation to it. We, indeed, have a strange way of thinking about the technologies current at any given time: Bob Dylan in/famously angering thousands at the Newport Folk Festival in 1965, not necessarily by having 'gone electric,' but by using amplification (and a soundman) ill-suited to 'electric' music. His performance had similar effects on subsequent audiences aside from those making the equally in/famous cat-calls 'traitor' and 'Judas,' of the 1966 Manchester Free Trade Hall. Bob Dylan, folk purist, was allowed to draw attention to social injustice in a puritanical American folk tradition, but only at a moderately sound-reinforced volume.

CONCLUSION

Although the user machine and the Internet have been defined as disembodied technologies, they are the opposite: the construction of the Internet, the mechanics of sound and digitization, the history of sound reproduction technology and the history of online audio produce a sufficiently embodied technology. Sound is 'captured' and produced by a physical apparatus—this is a necessity. The apparatus, like the persona of any orator, communicates visually, aurally, and kinetically. The apparatus, like the sound intended to issue from it, is a culmination of not only technological necessity, but cultural ideas of technology: it looks, feels, and sounds the way it does due to a variety of reasons, not all of them necessarily functional. Instead of placing the potential for persuasion on either 'ends' of the network, we should discover what the apparatus itself is capable of. It is in and through the apparatus that the physical presence of the orator can be seen, touched, and heard. The information, already traveling over the Internet to some 1.5 billion people worldwide (Anderson 14), already serves to inform, persuade, entertain, to obfuscate, and spur to action. However, delivery is often assumed out of the equation, as if only the arrangement, style, and 'genius' of the text anywhere online speaks. The morphic orator is always present and his delivery is transmogrified into the apparatus itself, the screen it presents and the sounds it makes. But the possibilities implied for Rhetoric have been routinely ignored; the possibilities for sound online have been so obscured by the frivolity with which sound has been used for nearly 15 years now that it is unlikely that the current structure of the Internet will change in any significant way regarding delivery.

Developers of the late nineties were persuaded that the silent Web was akin to a dead screen. To fill the enormous gap left by the "glassy-eyed" group crowded around the television set, "hungry authors" claimed that the use of sound online

> adds dimension…[gives a] competitive edge…can dramatically increase visitor interest, return visits and information retention…stimulate[s] power emotional synaptic responses that cannot be activated through the written word…[and that sound] is novel.
>
> Helmstetter and Simpson 45-46

Other authors seconded this assertion, arguing that the "silent experience" of the Web should yield, "simply put, to enhance the user experience" (Beggs and Thede 1). This carried with it numerous implications: the Internet is for entertainment, users are seeking an 'experience,' and that technology determines experience and cultural need. They leaned on the 'negatives' of "a silent experience…no more real or gripping than a postcard" (1) instead of properly developing theory and practice that would have resulted in more useful applications of web sound. The claims of the less-ethical web audio developers soared to incredible heights: "it has been proven that audio can increase your sales by as much as 300% and some report up to 426% increase" (Weavers). However the 'proof' is not accessible. The claims of salespeople and the assertions of the web audio authorities are misguided. Unfortunately, the results of such rampant case-making are the privileging of the visual over the auditory, of locking web metaphors into print metaphors, and in the promotion of visual literacy to the detriment of sonic literacy.

And this is not a dichotomy that has gone unnoticed. Sherwani et al. sought to understand why no "purely voice-based access to the unstructured information on websites" (1) exists, then

developed a web search application for users in the developing world where web search is not only an unfamiliar metaphor, but where "there may be no advantage in mimicking a GUI interface" (4). While users in this study tended to prefer a graphic user interface, other research, such as that of Prasad et al., has demonstrated the potential of mixed modal presentations for illiterate users. Some of the most interesting work seeks to further the metaphor of web space and increase the interactivity a user can have with a website. "Media space" projects, such as those of Bly and Irwin, Buxton, Finn et al., and Isaacs and Tang, bring people together "in a mixed physical and virtual environment…blending…online and physical form" (Karahalios and Dobson 1958) with the use of physical "telepresence sculptures." These forms are human scale physical avatars that are controlled by a user in a remote location but that interact with live people in an offline environment. Pop culture, never to be outdone by those on the forefront of the sensational, has invented its own physical correspondents to carry out online commands and vice versa: we may soon be able to scent our homes by sending an e-mail (Williams and Castaňeda) or "spice up" our online sex lives with motion-capture suits (Lynn).

Although teachers of rhetoric and practicing orators have demonstrated their apprehension not only of developing a theory of delivery, but of even making any concerted and substantive arguments. One realizes the high value many of the ancient Greek and Roman orators placed on delivery, but must make what would seem like a theoretical leap from their notions of correct presentation to our own. One could consider "media bias" as a viable point of contact. The experiment of Riegelsberger et al revealed that users have implicit feelings of trust with automated video and audio interfaces. They conclude that, "if the design goal is engagement rather than trust, our data suggests that [video, audio, avatar, and photo] can be effective" (Riegelsberger et al. 1748). Online, as well as any other kind of rhetorical interaction,

sound production is subject to many of the same concerns that ancient authors so reluctantly wrote about. The first requirement is the acknowledgment that tone, texture, volume, pitch, in other words, what would otherwise be aesthetic concerns, be components available for persuasion or dissuasion. Above all, one should acknowledge that whatever technological means accomplishes sound production is a valid concern. Online audio is in need of cultivation because of its vastly underutilized ability to sound and behave uniquely. The quality of audio currently runs the gamut from MIDI to the most popular encoding format, the MP3. For the time being, the MP3 format seems to have found some favor with consumer audiences who do not seem to be aware of or care for the many other competing encoding schemes that, if one cannot argue sound better, have a unique sound. But, if the trends thus far established for sound technologies online, the MP3 will fall out of favor. Transitions between technologies are so swift that it is difficult for the user audience to develop much attachment to them. One may draw an analogy to prior music containers, such as the rather lengthy transition times between vinyl, analogue information on magnetic tape, digital information on magnetic tape, compact disc, and encoding schemes. Until the advent of the MPEG system, the WAV format was secure offline for nearly 15 years. Now it has begun to lose ground to encoding schemes initially developed not for their high "quality," but for their quality relative to their ability to be transmitted over networks. Streaming technology, which became available in 1996, has almost completely replaced modem pinging sounds and MIDI files. Ever since then, a series of improvements to sound quality has come and gone across the Web, each one with its own 'sound.' The current moment of the MP3 sound is a fleeting one. Bandwidth will continue to increase, which will allow greater file sizes with far less compression. Compression schemes may be less-necessary than they once were. What remains to be seen is whether the tastes of users will follow suit. Like the generation of

people that came of age with vinyl recording technologies and continue to hear its sound production as the most pleasing and accurate, what will the current generation hear when 'improved' technologies become available? And what will that mean in regards to notions of 'quality' and 'fidelity' as the 'real' world becomes more and more simulated?

As is often the case with multi-disciplinary studies, a consensus of purpose and vocabulary must be reached if the field is to remain fertile. From psychoacoustics, semantics, general studies of aurality and soundscape studies, to the auditory computer-human interaction, there must be a concerted effort to identify what each discipline can contribute to the field as a whole. Societies such as The Acoustic Ecology Institute, The World Forum for Acoustic Ecology, The New Media Consortium, and several others, exist at least in part for this purpose. The trendsetting books and papers by individual authors are obviously greatly needed, but it is collaboration on a large scale that is best suited to deal with the many aspects of sound studies. Anthologies, bringing together the ideas of many different authors under the same jacket, if not directly collaborative in nature, do serve to get us all 'on the same page,' as it were. Notable examples of these are Hearing Cultures: Essays on Sound, Listening, and Modernity, The Auditory Culture Reader, and Hearing History: A Reader. Work needs to continue in the way of standardizing the vocabulary of the field. This work is being done, not only inventorying sound vocabulary discrepancies between different groups of people (McGregor et al.), but in developing "robust principles for a semantically effective design of [a] common audio-vocabulary…[that are] based of the application of rhetorical schemes" (Polotti and Benzi). Lastly, on the most accessible level, professors should not only attempt to substantively address delivery when teaching rhetoric, but should encourage their students to examine it further. This

thesis has come about because the professors around me welcomed discussion and assignments on this topic. It has been, therefore, my first contribution of what I hope to be many to the field of sound studies.

BIBLIOGRAPHY

Introduction

Altman, Rick. "The Evolution of Sound Technology." Film Sound: Theory and Practice. Ed. Elisabeth Weis and John Belton. New York: Columbia UP, 1985. 44-53.

Barlow, John Perry. "Crime and Puzzlement." June 8, 1990. 17 September 2008 <http://www.sjgames.com/SS/crimpuzz.html>.

Beggs, Josh and Dylan Thede. Designing Web Audio. Cambridge: O'Reilly, 2001.

CERN. "LCG Project: Overview." Worldwide LHC Computing Grid. 17 September 2008 <http://lcg.web.cern.ch/LCG/overview.htm>.

Chion, Michel. Audio-Vision: Sound On Screen. Ed. and Trans. Claudia Gorbman. New York: Columbia UP, 1994.

Fischer, Lucy. "Film Editing." A Companion to Film Theory. Ed. Toby Miller and Robert Stam. Hoboken, New Jersey: Blackwell, 2004. 64-83.

FoxNews. "'The Grid' Could Soon Make the Internet Obsolete," 7 April 2008. 17 September 2008 < http://www.foxnews.com/story/0,2933,347212,00.html>.

Gomery, Douglas. "The Coming of Sound: Technological Change in the American Film Industry." Film Sound: Theory and Practice. Ed. Elisabeth Weis and John Belton. New York: Columbia UP, 1985. 24-36.

Helmstetter, Anthony and Ron Simpson. Web Developers Guide to Sound and Music. Scottsdale, Arizona: Coriolis Group Books, 1996.

Knight, Arthur. "The Movies Learn to Talk: Ernst Lubitsch, René Clair, and Rouben Mamoulian." Film Sound: Theory and Practice. Ed. Elisabeth Weis and John Belton. New York: Columbia UP, 1985. 213-220.

McGookin, David A. and Stephen A Brewster. "Understanding Concurrent Earcons: Applying Auditory Scene Analysis Principles to Concurrent Earcon Recognition." ACM Transactions on Applied Perceptions 1 (2004): 130-155.

Ratzan, Lee. "Making sense of the Web: A Metaphorical Approach." Information Research, 6.1 (2000). 17 September 2008 <http://InformationR.net/ir/6-1/paper85.html>.

Murch, Walter. Forward. Audio-Vision: Sound On Screen. By Michel Chion. Ed. and Trans. Claudia Gorbman. New York: Columbia UP, 1994. vii-xxiv.

Nielsen, Jakob and Hoa Loranger. Prioritizing Web Usability. Berkeley, California: New Riders, 2006.

Wesch, Michael. The Machine is Us/ing Us, 31 January 2007. 17 September 2008 <http://www.youtube.com/watch?v=NLlGopyXT_g>.

White, Frederick A. Our Acoustic Environment. New York: John Wiley and Sons, 1975.

"Flash How To," W3Schools. 17 September 2008 <http://www.w3schools.com/Flash/flash_howto.asp>.

Yost, William A. "Overview: Psychoacoustics." Human Psychophysics. Ed. William A. Yost, Arthur N. Popper, and Richard R. Fay. New York: Springer-Verlag, 1993. 1-12.

The Soundscape of the Audio-Enabled Web

Advertise Door To Door, Nationwide Direct Marketing, Inc. 2006. 17 September 2008 <http://advertisedoortodoor.com/>.

Aerial Digital Media, Montclair Homes. 2007. 17 September 2008 <http://www.aerialdigitalmedia.com/vttest/index.html>.

Alan Dearing, Alan Dearing Photography. 17 September 2008 <http://www.alandearing.com>.

Allstar Supply, Allstar Supply, Inc. 17 September 2008 <http://www.allstarsupply.com/>.

Andrews, Jim. "Interactive Audio on the Web 1" Trace Online Writing Centre. 22 September 2003. 17 September 2008 <http://tracearchive.ntu.ac.uk/Review/index.cfm?article=80>.

Atlanta Jazz Society, The Atlanta Jazz Society, Inc. 17 September 2008 <http://www.atlantajazzsociety.org/>.

Bader, Jerry. "Enhancing Web Effectiveness With Audio Design." 2008. Jayde Online. <http://www.sitepronews.com/archives/2008/may/7.html>. July 24, 2008.

Bahktin, M.M. "The Problem of Speech Genres." The Discourse Reader. Ed. Adam Jaworski and Nikolas Coupland. London: Routledge, 1999. 121-132.

Blattner, Meera M., Denise A. Sumikawa, and Robert M. Greenberg. "Earcons and Icons: Their Structure and Common Design Principles." Human-Computer Interaction. 4 (1989): 11-44.

B L Crosses, 17 September 2008 <http://blcrosses.com>.

"BOFOP, C.S.J." In *Thursday Afternoon* [CD liner notes]. Virgin. 1985, 2005.

Camelot-inn, The Camelot Inn. 17 September 2008 <http://www.camelot-inn.com>.

Copy G, 10 August 2008 <http://www.copyg.com>.

Crispien, Kai, Würz, Woldgang, Weber, Gerhard. "Using Spatial Audio for the Enhanced Presentation of Synthesized Speech within Screen Readers for Blind Computer Users." Proceedings of International Conference on Computers Helping People with Special Needs, 1994. Vienna: Springer Verlag, 1994.

The Cure, The Cure. 17 September 2008 <http://thecure.com>.

Digital Strife, 17 September 2008 <http://www.digitalstrife.com/>.

Disney Travel, Disney. 17 September 2008 <http://home.disney.go.com/travel/>.

D W Fitness, Deandera "Dee" Washington. 17 September 2008 <http://www.dwfitness.com/index.html>.

Elert, Glenn. "Number of Web Pages." The Physics Factbook. Glenn Elert. 29 May 2007. 17 September 2008 <http://hypertextbook.com/facts/2007/LorantLee.shtml>.

Firefox Add-ons. 17 September 2008 <https://addons.mozilla.org/en-US/firefox/addon/433>.

Follet, Jonathan. "Audio and the User Experience." UX matters. 18 June, 2007. 17 September 2008 <http://uxmatters.com/MT/archives/000200.php>.

Furniture Wholesale Group, Furniture Wholesale Group. 17 September 2008 <http://www.furniturewholesalegroup.com/>.

George Carlin, 17 September 2008 <http://www.georgecarlin.com/home/home.html>.

The GFD Group, Genesis Funding and Development. 17 September 2008 <http://www.thegfdgroup.com/>.

Goose, Stuart and Carsten Möller. "A 3D Audio Only Interactive Web Browser: Using Spatialization to Convey Hypermedia Document Structure." International Multimedia Conference: Proceedings of the Seventh ACM International Conference on Multimedia, Part 1. 1999. 363 – 371.

GSU Recreational Services, Georgia State University Recreational Services. 17 September 2008 <http://www2.gsu.edu/~wwwrec/clubs/index.html>.

Hair Paradise Salon, Hair Paradise Salon. 17 September 2008 <http://www.hairparadisesalon.com>.

Hörstreich, Hörstreich 2008. 17 September 2008 <http://www.hoerstreich.de>.

IGC Music, International Groove Conspiracy 2006. 17 September 2008 <http://www.igcmusic.com>.

In-Colour, The Concretes. 17 September 2008 <http://www.in-colour.net>.

James, Nick. "Here's Exactly What Makes Web Audio So Powerful." 4HB. 2004. 17 September 2008 <http://www.4hb.com/20041117201601.html>.

Lemay, Laura and Rafe Colburn. Web Publishing with HTML and CSS. Indianapolis, Indiana: Sams, 2006.

Mediastorm, Washington Post. 17 September 2008 <http://mediastorm.org/>.

Move With Velocity, 10 August 2008 <http://www.movewithvelocity.com/homepage.html>.

"Guidelines for Multimedia on the Web," Jakob Nielsen's Alertbox, ed. Jakob Nielsen, December 1995, 17 September 2008 <http://www.useit.com/alertbox/9512.html>.

Peroni Italy, Birra Peroni Internazionale. 17 September 2008. <http://www.peroniitaly.com/us/>.

Podcast Alley, "Site Statistics." Chris McIntyre and PodcastAlley.com. 17 September 2008. <http://www.podcastalley.com/index.php>.

Pollack, Irwin and Lawrence Ficks. "Information of Elementary Multidimensional Auditory Displays." Journal of the Acoustical Society of America 26 (1954): 155-158.

The Prince, The Prince: La Palestra Di Montelupo. 17 September 2008 <http://www.theprince.it/home.htm>.

Progressive Dyn, Progressive Dynamics, Inc. 17 September 2008 <http://www.progressivedyn.com/>.

Qssis, 2004. Qssis Banquet Halls. 17 September 2008 <http://www.qssis.com/about.html>.

Ronkainen, Sami, Jonna Häkkiliä, and Leena Pasanen. "Effect of Aesthetics on Audio-Enhanced Graphical Buttons." Proceedings of ICAD 05-Eleventh Meeting of the International Conference on Auditory Display. 2005. 121-126.

Roundway Roofing.co.uk. Roundway Roofing Contractors. 17 September 2008 <http://www.roundwayroofing.co.uk>.

Schafer, R. Murray. The Soundscape: Our Sonic Environment and the Tuning of the World. Rochester, Vermont: Destiny Books, 1993.

Schlömer, Rüdiger. 17 September 2008 < http://www.stromgasse.de/projects/actionist/>.

Sellers, Patricia. "Myspace Cowboys." Fortune 26 August 2006. 17 September 2008 <http://money.cnn.com/magazines/fortune/fortune_archive/2006/09/04/8384727/index.htm>.

Sherwani, J., Dong Yu, Tim Paek, Mary Czerwinski, Y. C. Ju, and Alex Acero. "VoicePedia: Towards Speech-based Access to Unstructured Information." Interspeech (2007): 146-149.

Squidsoup. 2006. 17 September 2008 <http://www.squidsoup.org/altzero/>.

Steve Dancz, 17 September 2008 <http://www.stevedancz.com>.

St Mauer, Susan. "Words that Work for Audio: A Little Sound Advice." 17 September 2008 <http://209.85.215.104/search?q=cache:_XttEMV5XUJ:www.marketingprofs.com/3/maur5.asp+how+the+web+works+audio&hl=en&ct=clnk&cd=44&gl=us>.

Strictly Fashion, Strictly Fashion. 17 September 2008 <http://www.strictlyfashion.nl/>.

Sunnydowns, Mike Willington. 17 September 2008 <http://www.hotelnorthwales.co.uk>.

Sweet Thang Unlimited, 2006. 17 September 2008 <http://www.sweethangunlimited.com>.

Trois, Trois Atlanta. 17 September 2008 <http://www.trois3.com/home.html>.

Unga-Unga, Zam Johnson and Paul Griesbach. 17 September 2008 <http://www.unga-unga.com/frame.html>.

Vargas, Maria L.M. and Sven Anderson. "Combining Speech and Earcons to Assist Menu Navigation." Proceedings of the 2003 International Conference on Auditory Display, July 6-9, 2003. Boston.

Vip Homes 4 You, Cash Homebuyers, LP. 17 September 2008 <http://www.wesellusahomes.com/>.

W3Schools. 17 September 2008 <http://www.w3schools.com >.

Walker, Bruce and Gregory Kramer. "Ecological Psychoacoustics and Auditory Displays: Hearing, Grouping, and Meaning Making." Ecological Psychoacoustics. Ed. John G. Neuhoff. Boston: Elsevier Academic Press, 2004. 150-170.

Wennerstrom, Ann. The Music of Everyday Speech: Prosody and Discourse Analysis. New York: Oxford UP, 2001.

Whisper Audio, 17 September 2008 <http://www.whisperaudio.com>.

Yen , Yi-Wyn. "YouTube Looks For The Money Clip." Fortune 25 March 2008. 17 September 2008 <http://techland.blogs.fortune.cnn.com/2008/03/25/ youtube-looks-for-the-money-clip/>.

Situated Mechanics

About.com, "Karlheinz Brandenburg and The Secret History of MP3." BBC.com 17 August 2000. 17 September 2008 <http://inventors.about.com/gi/dynamic/offsite.htm?zi=1/XJ&sdn=inventors&cdn=money&tm=15&f=00&tt=2&bt=1&bts=0&zu=http%3A// www.bbc.co.uk/h2g2/guide/A406973>.

Burke, Chris. "Interactive Audio on the Web: Review and Recommendations." Los Angeles: MIDI Manufacturers Association, 2002.

Carlson, A. Bruce, Praul B. Crilly and Janet C. Rutledge. Communication Systems: An Introduction to Signals and Noise in Electrical Communication. 4th ed. Boston: McGraw Hill, 2002.

Clark, Martin P. Data Networks, IP and the Internet: Protocols, Design and Operation. West Sussex, England: Wiley, 2003.

Everest, F. Alton. Master Handbook of Acoustics. 4th ed. New York: McGraw-Hill, 2001.

Gibilisco, Stan. Teach Yourself Electricity and Electronics. 3rd ed. New York: McGraw-Hill, 2001.

deGrasse Tyson, Neil. Death by Black Hole and Other Cosmic Quandaries. New York: W.W. Norton, 2007.

Thomson, "The History of mp3." Thomson 2005. 17 September 2008 <http://www.mp3licensing.com/mp3/history.html>.

Huber, David Miles and Robert E. Runstein. Modern Recording Techniques. 5th ed. Boston: Focal Press, 2001.

Maguire, James and Jim Loudermilk. TechTV's Secrets of the Digital Studio: Insider's Guide to Desktop Recording. Indianapolis: Que, 2002.

Pescovitz, David. "Birds Mimic Ringtones." Happy Mutants LLC 18 July 2005. 17 September 2008 <http://www.boingboing.net/2005/07/18/birds-mimic-ringtone.html>.

Rickert, Thomas and Michael Salvo. "The Distributed *Gesamptkunstwerk*: Sound, Worlding, and New Media Culture." Computers and Composition 23. (2006) 296-316.

Sterne, Jonathan. The Audible Past: Cultural Origins of Sound Reproduction. Durham, North Carolina: Duke UP, 2002.

Stuart, J. Robert. "Coding High Quality Digital Audio." Huntingdon, United Kingdom: Meridian Audio Ltd, 1998.

Wikipedia, "Transporter (Star Trek)." Wikimedia Foundation. 6 September 2008. 17 September 2008 <http://en.wikipedia.org/wiki/Transporter_%28Star_Trek%29>.

Xiph, "Dare to Compare." Xiph.org. 17 September 2008 <http://www.xiph.org/vorbis/listen.html>.

Delivery

Anon; [Cicero]. Rhetorica ad Herennium. Harry Caplan, Ed. Cambridge: Harvard UP, 2004.

Aristotle. Rhetoric. The Complete Works of Aristotle: The Revised Oxford Translation. Ed. Jonathan Barnes. Princeton: Princeton UP, 1995.

Bourdieu, Pierre. "Language and Symbolic Power." Trans. Gino Raymond and Matthew Adamson. The Discourse Reader. Ed. Adam Jaworski and Nikolas Coupland. London: Routledge, 1999. 98-106.

Bull, Michael. "Soundscapes of the Car: A Critical Study of Automobile Habitation." The Auditory Culture Reader. Ed. Michael Bull and Les Back. Oxford: Berg Publishers, 2004. 357-374.

Cicero. On the Ideal Orator. Trans James M. May and Jakob Wisse. New York: Oxford UP, 2001.

Corbett, Edward P. J. and Robert J. Connors. Classical Rhetoric for the Modern Student. 4th ed. New York: Oxford UP, 1999.

Cusack, Rhodri and Robert Carlyon. "Auditory Perceptual Organization Inside and Outside the Laboratory." Ecological Psychoacoustics. Ed. John G. Neuhoff. Boston: Elsevier Academic Press, 2004. 16-44.

Erlmann, Veit. "But What of the Ethnographic Ear? Anthrophology, Sound, and the Senses." Hearing Cultures: Essays on Sound, Listening and Modernity. Ed. Veit Erlmann. Oxford: Berg Publishers, 2004. 1-20.

Gaudio, Rudolf P. "Sounding Gay: Pitch Properties in the Speech of Gay and Straight Men." American Speech, 69. 1. (Spring, 1994), pp. 30-57.

Gorgias. "Encomium of Helen." The Older Sophists: A Complete Translation by Several Hands of the Fragments in Die Fragmente Der Vorsokraticker Edited by Diels-Kranz with a New Edition of Antiphon and of Euthydemus. Rosamond Kent Sprague, ed. Columbia, South Carolina: U of South Carolina P, 1972. 50-54.

Holmes, Gary. "Nielsen Media Research Reports Television's Popularity Is Still Growing." Nielsen Company 21 September 2006. 17 September 2008 <http://www.nielsenmedia.com/nc/portal/site/Public/menuitem.55dc65b4a7d5adff3f65936147a062a0/?vgnextoid=4156527aacccd010VgnVCM100000ac0a260aRCRD>.

Jaworski, Adam and Nikolas Coupland. The Discourse Reader. London: Routledge, 1999.

Langer, Gary. "Poll: Traffic in the United States: A Look Under the Hood of a Nation on Wheels." ABC News 13 February 2005. 17 September 2008 <http://abcnews.go.com/Technology/Traffic/story?id=485098>.

Neuhoff, John G. "Interacting Perceptual Dimensions." Ecological Psychoacoustics. Ed. John G. Neuhoff. Boston: Elsevier Academic Press, 2004. 250-263.

Ong, Walter J. Orality and Literacy: The Technologizing of the Word. London: Mehuen & Co., 1982.

Parrino, Angela. "The Politics of Pronunciation and the Adult Learner." Adult ESL: Politics, Pedagogy, and Participation in Classroom and Community Programs. Ed. Trudy Smoke. Mahwah, New Jersey: Lawrence Erlbaum Associates, 1998.

Pernot, Laurent. Rhetoric in Antiquity. Trans. W.E. Higgins. Washington, D.C.: Catholic U of America P, 2005.

Plato. "Phaedrus." Alexander Nehamas and Paul Woodruff, trans. Plato: Complete Works. John M. Cooper, ed. Indianapolis, Indiana: Hackett, 1997. 506-556.

Quintilian. On the Teaching of Speaking and Writing. Ed. James J. Murphy. Carbondale: Southern Illinois UP, 1987.

Rayson, P., Leech, G., and Hodges, M. "Social Differentiation in the use of English Vocabulary: Some Analyses of the Conversational Component of the British National Corpus." International Journal of Corpus Linguistics. 2.1. pp 133 - 152. John Benjamins: Amsterdam, 1997.

Riegelsberger, Jens M., M. Angela Sasse, and John D. McCarthy. "Do People Trust Their Eyes More Than Their Ears: Media Bias Wile Seeking Expert Advice." CHI April 2-7, 2005.

Rosenblum, Lawrence D. "Perceiving Articulatory Events: Lessons for an Ecological Psychoacoustics." Ecological Psychoacoustics. Ed. John G. Neuhoff. Boston: Elsevier Academic Press, 2004. 220-241.

Sterne, Jonathan. The Audible Past: Cultural Origins of Sound Reproduction. Durham, North Carolina: Duke UP, 2002.

Story, Louise. "Anywhere the Eye Can See, It's Likely to See an Ad." New York Times. 15 January 2007. 17 September 2008<http://www.nytimes.com/2007/01 /15/business/media/15everywhere.html>.

Tannen, Deborah. That's Not What I Meant! How Conversational Style Makes of Breaks Relationships. New York: Ballantine Books, 1986.

Welch, Kathleen E. Electric Rhetoric: Classical Rhetoric, Oralism, and a New Literacy. Cambridge: MIT Press, 1999.

Wennerstrom, Ann. The Music of Everyday Speech: Prosody and Discourse Analysis. New York: Oxford UP, 2001.

Conclusion

Anderson, Mark. "Rebuilding the Internet." Discover. Dec. 2008: 14.

Evoca, Evoca, LLC. 2008. 17 September 2008 <http://www.evoca.com/customers>.

Prasad, Archana, Indrani Medhi, Kentaro Toyama, Ravin Balakristnan. "Exploring the Feasibility of Video Mail for Illiterate Users." Conference AVI 2008. May 28-30 2008.

Karahalios, Karrie G. and Kelly Dobson. "Chit Chat Club: Bridging Virtual and Physical Space for Social Interaction." CHI 2005 April 2-7, 2005.

Lynn, Regina. "Motion-Capture Suits Will Spice Up Virtual Sex." Wired Online. 9 May 2008. 17 September 2008 <http://www.wired.com/culture/lifestyle/commentary/sexdrive/2008/05/sexdrive_0509>.

McGregor, Iain, Grégory Leplâtre, Alson Crerar and David Benyon. Sound and Soundscape Classification: Establish Key Auditory Dimensions and Their Relative Importance. Proceedings of the 12th International Conference on Auditory Display. 2006. 105-112.

Polotti, Pietro and Carlo Benzi. "Rhetorical Schemes for Audio Communication." Proceedings of the 14th International Conference on Auditory Display. 2008.

Weavers, Roger. "'Using The Right Software You Can Add Audio To Your Website In Just A Few Minutes!'" Web Site Audio Reviews. 17 September 2008 <http://www.websiteaudioreviews.com>.

Williams, Martyn and Chiara Castaňeda. "When Roses Won't Do, E-Mail a Fragrance." The New York Times 8 April 2008. 17 September 2008 <http://www.nytimes.com/idg/IDG_002570DE00740E1800257425001B1F2B.html>.

APPENDIX: SURVEY OF 120 AUDIO-ENABLED WEBSITES

Background Sound

America Web Works. <http://www.americawebworks.com/>.

Atlanta Jazz Society. <http://www.atlantajazzsociety.org/>.

Big Snow. <http://www.bigsnowent.com/site.html>.

Bobby Mcferrin. <http://www.bobbymcferrin.com>.

Bon Jovi. <http://www.bonjovi.com/bonjovi/>.

Chris Nicholson Salon. <http://www.chrisnicholsonsalon.com/>.

Ciao Bella Hair Salon. <http://ciaobellahairsalon.com/>.

The Citizens Band. <http://www.thecitizensband.net/>.

Daily Wav. <http://www.dailywav.com/>.

Digitalstrife. <http://www.digitalstrife.com/>.

Disney Travel. <http://home.disney.go.com/travel/index>.

D W Fitness. <http://www.dwfitness.com/index.html>.

Farm Rio. <http://www.farmrio.com.br/>.

George Strait. <http://www.georgestrait.com/>.

The GFD Group. <http://www.thegfdgroup.com/>.

The Hair Color Salon. <http://www.thehaircolorsalon.com/>.

Jessica Simpson. <http://www.jessicasimpson.com/>.

Johnny Cash. <http://www.johnnycash.com/>.

Led Zeppelin. <http://www.ledzeppelin.com>.

Lenny Kravitz. <http://www.lennykravitz.com/>.

Move With Velocity. <http://www.movewithvelocity.com/homepage.html>.

Muse. <http://www.muse.mu/index.php>.

Myspace Example. <http://profile.myspace.com/
index.cfm?fuseaction=user.viewprofile&friendid=50439974>.

GSU Clubs. <http://www2.gsu.edu/~wwwrec/clubs/index.html>.

Qssis. <http://www.qssis.com/>.

The Prince. <http://www.theprince.it/home.htm>.

Sade. <http://www.sade.com/>.

Smoken Dudes BBQ. <http://www.smokendudesbbq.com/index2.html>.

Sound of the Web. <http://www.soundoftheweb.net/>.

Steve Dancz. <http://www.stevedancz.com/>.

Strictly Fashion. <http://www.strictlyfashion.nl/>.

Sweethang Unlimited. <http://www.sweethangunlimited.com/>.

Tony Cao Salon Spa. <http://www.tonycaosalonspa.com/>.

Trendz Salon. <http://www.trendzsalon.com/>.

Trois 3. <http://www.trois3.com/home.html>.

White Stripes. <http://www.whitestripes.com/home/home.html>.

Background Sound and Narration

Aerial Digital Media. <http://www.aerialdigitalmedia.com/vttest/index.html>.

Alan Dearing. <http://www.alandearing.com/>.

All Star Supply. <http://www.allstarsupply.com/>.

BellSouth Real Web Audio. <http://www.realwebaudio.com/>.

Bewerley Hall Farm. <http://www.bewerleyhallfarm.co.uk/>

B L Crosses. <http://blcrosses.com/store/>.

Camelot-Inn. <http://www.camelot-inn.com/>.

Copy G. <http://www.copyg.com/>.

Crystal Mountain. <http://talkinsites.com/demos/crystalmountain/crystalmountain.htm>.

Furniture Wholesale Group. <http://www.furniturewholesalegroup.com/>.

Hair Paradise Salon. <http://www.hairparadisesalon.com/>.

Hotel North Wales. <http://www.hotelnorthwales.co.uk/>

Manchester Plastics. <http://www.manchesterplastics.com/>.

My Instant Web Audio. <http://www.myinstantwebaudio.com/>.

Progressive Dyn. <http://www.progressivedyn.com/>.

Roundway Roofing. <http://www.roundwayroofing.co.uk/>.

Synergy Soup. <http://www.synergysoup.com/>.

Talkin Sites. <http://www.talkinsites.com/index.html>.

Web Audio Marketing. <http://www.webaudiomarketing.com/>.

Westford Legacy. <http://www.westfordlegacy.com/home.htm>.

Westwood One. <http://www.westwoodone.com/
offtherecord;jsessionid=AF19B5F0DE850C3AC0A20627D03021B9>.

Signal and Auditory Display

Essential Freight. <http://www.essentialfreight.com/default.asp>.

George Carlin. <http://www.georgecarlin.com/>.

National Response Center. <http://www.nrc.uscg.mil/nrchp.html>.

Sad Trombone. <http://www.sadtrombone.com/>.

Unga-unga. <http://www.unga-unga.com/frame.html>.

VIP Homes 4 You. <http://www.viphomes4you.com/Home.cfm>.

Multimedia

Amy Winehouse. <http://www.amywinehouse.co.uk/>.

Buehnekunst Theatre. <http://www.buehnenkunsttheatre.org/home.html>.

Democracy Now. <http://www.democracynow.org/>.

Denise Richards. <http://gallery.deniserichards.com/>.

Dtours360. <http://www.dtours360.com/>.

IKEA Dream Kitchen. <http://demo.fb.se/e/ikea/dreamkitchen/site/default.html>.

In-Colour. <http://www.in-colour.net/>.

Info-chip. <http://www.info-chip.com/>.

Mediastorm. <http://mediastorm.org/>.

Mevio. <http://www.mevio.com/>.

Michael Jackson. <http://www.michaeljackson.com/>.

Peroni Italy. <http://www.peroniitaly.com/us/>.

Sonar. <http://www.sonar.es/2007/eng/s2007.cfm>

Spam. <http://www.spam.com/>.

Sun Studio. <http://www.sunstudio.com/index.aspx?bhcp=1>.

Voki. <http://www.voki.com/>.

Your Brain on Music. <http://www.yourbrainonmusic.com/>.

Sitepal Avatar. <http://www.sitepal.com/>.

Virtual Presenter. <http://www.virtualpresenter.com/>.

Introductory Pages

ACA Music. <http://www.acamusic.co.uk/>.

Advertise Door to Door .<http://advertisedoortodoor.com/>.

Aimeenez. <http://aimeenez.net/>.

Arborealinn. <http://www.arborealinn.com/>.

BSC. <http://www.bsc.com/>

The Cure. <http://www.thecure.com/>

Denise Richards. <http://www.deniserichards.com/Pages/intro.htm>.

Digitalstrife. <http://www.digitalstrife.com/>.

Geocities Example. <http://www.geocities.com/Heartland/Plains/5514/index.html>.

Höstreich. <http://www.hoerstreich.de/>

IGC Music. <http://www.igcmusic.com/>

Jerry Spinelli. <http://www.jerryspinelli.com/newbery_001.htm>.

National Response Center. <http://www.nrc.uscg.mil/nrchp.html>.

SC Blues. <http://www.scblues.com/>.

Taknbystorm. <http://www.taknbystorm.net/>.

Westford Legacy. <http://www.westfordlegacy.com/>.

UK Web Audio. <http://www.ukwebaudio.com/>

Interactive Audio

Actionist Respoke. <http://www.stromgasse.de/projects/actionist/>.

Altzero. <http://www.squidsoup.org/altzero/>

Neonlight. <http://www2u.biglobe.ne.jp/~macopism/noct/data/index.html>.

Pixel by Pixel. <http://www.lecielestbleu.com/media/pixelbypixelframe.htm>.

Pianographique. <http://www.pianographique.net/>.

Pianolina. <http://www.grotrian.de/spiel/e/info.html>.

Sound Garden. <http://www.cddc.vt.edu/host/404/interactive/soundgarden/>.

The Square Root of -1. <http://www.atomless.com/sqrt-1/>.

Sticky Sound Elastic Structure. <http://moebio.com/santiago/pegajoso/>.

Community

Audioblogger. <http://forum.odeo.com/index.php?pg=kb.page&id=67>.

eJamming. <http://www.ejamming.com/>.

Evoca. <http://www.evoca.com/index.php>.

Gabcast. <http://www.gabcast.com/>.

Hearingvoices. <http://hearingvoices.com/webworks.php>.

Hipcast. <http://www.hipcast.com/>.

Jamglue. <http://www.jamglue.com/>.

Kompoz. <http://www.kompoz.com/compose-collaborate/home.music>.

Splice. <http://www.splicemusic.com/>.

Voicealizer <http://www.voicealizer.com/>.

Voicethread <http://voicethread.com/#home>.