An Aesthetic of the Irreducible

Ryan Patrick Maguire
Milwaukee, WI

Master of Arts, University of Virginia, 2015
Master of Arts, Dartmouth College, 2013
Undergraduate Diploma, New England Conservatory, 2010
Bachelor of Arts, Beloit College, 2008

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Music

University of Virginia
May, 2019

Edward Coffey (chair)
A.D. Carson
Luke Dahl
Mona Kasra
Noel Lobley

Contents

List of Figures

Abstract

This practice-focused dissertation is about my artistic efforts to develop an aesthetic of the irreducible. In the portfolio of almost thirty individual works created over a period of five years, I locate the irreducible in sounds and images discarded by perceptual coding algorithms such as the MP3. Perceptual coding economizes signals by excising information deemed perceptually irrelevant to formalized models of human audiovisual sensation. When digital signals are encoded using such algorithms, subtle details are irreversibly dissolved. These abandoned particulars, which could not be compressed into perceptually coded signal, exemplify the irreducible. Attending to them is to affirm their value.

In addition to the irreducible, this dissertation explores adjacent concepts through varied acts of composition. First, I develop the breath as a vital metaphor for its ephemerality and dynamism. Following this, I relate ghosts to the information deleted by the compression process. I value noise (that which resists reduction) as a wellspring of the atopic (that without place). In practice, I employ sonic allusions to allow multiple meanings to emerge. I accumulate textures through parallel and serial processes of embodiment and materialization. Throughout, I demonstrate how the techniques I've developed for working with perceptual coding detritus affirm the irreducible.

A brief introduction establishes key concepts supporting an aesthetic of the irreducible. Chapter one describes my work with Suzanne Vega's "Tom's Diner," famously used as a test track in the development of the MP3. A detailed account of practical techniques for working with compression detritus is interspersed with poetic, historic, technical, and theoretical excursions. A meditation at the midpoint of the document develops the idea of the *atopic* as that which exceeds a given conceptual system. I identify my compositional materials as atopic and consider their varied potentialities. The second chapter describes additional creative experiments with compression remainders, reinforcing key

themes throughout. A final reflection is followed by detailed computer code used in the creation of the portfolio.

Acknowledgements

I would like to express my deepest appreciation to my committee—Ted Coffey, A.D. Carson, Luke Dahl, Mona Kasra, and Noel Lobley. It is a privilege to know each of you. Thank you for your feedback, encouragement, and patience. Ted—you were already a musical and intellectual inspiration to me before we personally met. It has been my great fortune that, in addition to your brilliance, you are also good humored, deeply compassionate, and a supportive mentor. Beyond my committee, I would like to thank Tara Rodgers, Larry Polansky, Matthew Burtner, Jon Appleton, Judith Shatin, and Michael Casey for their guidance and tutelage. Your support means everything to me. To everyone whose work has inspired me whether cited in the text of this dissertation or not—thank you.

I owe a debt of gratitude to Dean Barnaby for making this dissertation possible—I am honored to have received the Dean's Dissertation Completion Fellowship which supported me while writing. This work was also generously supported by the Scholars Lab at the University of Virginia and, in particular, by Ammon Shepard. Thank you! I would also like to thank Michael Puri and Bonnie Gordon (as Directors of Graduate Studies) and the staffs of the Department of Music and Music Library for their assistance.

Completing this dissertation would have been miserable, at best, without the camaraderie of my colleagues, collaborators, friends, and family. To my colleagues, collaborators, and friends—thank you for inspiring me on a daily basis. To my Mom—completing this dissertation would have been impossible without your care. You gave me hope when I thought I couldn't do it. Thank you. I owe a special thanks to Pearl the cat for being literally by my side while I wrote this dissertation. You are cute. Finally, to Paige—your love, support, and encouragement renewed me every day. Thank you for making me smile.

Inhalation

This dissertation is comprised of twenty-eight original art works that adapt and express an aesthetic of the irreducible. First, the breath, in its combination of transience and vitality, becomes a powerful metaphor for the irreducible. Comparably immaterial, I relate ghosts to the sounds deleted by the compression process, inviting the attendant undertones. Noise becomes a primary resource for developing a practical aesthetics and I employ it in various media throughout the portfolio. In reference to established cultural forms and sound percepts, I find allusion (as opposed to direct statement) to be a valuable tool for inviting multiple interpretations. Throughout the dissertation, I transform materials compositionally, through technological means, by corporeal means, and in collaboration with others. In these processes, I value the distinct accumulations of noise that emerge, texturing the resultant artifacts.

Chapter one describes a constellation of works created from audiovisual material deleted during lossy compression of the recording "Tom's Diner" and its accompanying music video by Suzanne Vega. Following a brief description of the original song, a detailed account of the technical implementation of *moDernisT*, the central work in this dissertation, ensues. This recounting begins with extreme granularity—describing the first 1500 milliseconds over six pages—interspersed with thoughts about breath, texture, and virtual space. Over the ensuing twenty-nine pages, the techniques deployed in creating the entire two minute composition are thoroughly presented. This reportage is relieved by asides on the nature of sound, on historic threads connecting Suzanne Vega's song with the MP3, and on the MP3's particular instantiation of perceptual technics. Throughout, the technical description is accompanied by aesthetic motivations with philosophical and poetic references woven through as asides or footnotes.

The chapter concludes with an expansion of this work into visual and media art. First, two videos from the portfolio are described. I explain how to translate my core technique for recovering compression detritus to the visual domain using OpenCV and Python. The videos are evoked both aesthetically and technically, with a supplementary still frame from the first video concluding the section. Following this, I describe a handmade, handheld device built using a microcomputer. The device gives an experimental physical form to the sounds and images curated in this chapter. A largely technical description of its implementation is motivated by concerns for materiality and embodiment.

A meditation on *atopia* bisects the written portion of this dissertation. I generalize my interest in the sounds deleted by perceptual coding algorithms to an interest in that which exceeds any given system. I term this the atopic, borrowing from the Greek for *without place* or *unusual*. Following an account of the atopic as manifested in photography, I return to recorded sound in order to situate the works comprising this dissertation within an atopic frame.

The second half of this essay describes the remaining art works in less detail. Across twenty-six pages, the creations are sorted into thirteen projects. I describe their technical implementations and return to the dissertation's major themes throughout. *Ghost/Ambient* is a seven track album of minimalist works created using convolution for spectral ends. *$†@®* develops texture through embodied performance while regarding the material experience of low-frequency sound. Built on an ambient backdrop, feelings of longing and regret are drawn out by subtle allusions.

freeLanguage considers how atopia is manifested in language, with a conversation between Ornette Coleman and Jacques Derrida footnoting the technical account. *Harvest Light, Rest,* and *white, pink, brown* are live performances exploring the tension between fidelity to notation and the creative agency of individual performers. *Ghost in the Codec* is a gallery installation in the form of a photo booth allowing visitors to print out their own compression detritus. *Pure Beauty* and *Resolution Disputes* are two collaborations arising from my work with MP3 compression detritus. Recent work developing and performing with a sample library of audio compression ghosts is followed by efforts to transfer visual compression ghosts to clothing and other material forms. Finally, a work in progress commissioned by Laura Thompson and Émilie Fortin is described. Expanding the notion of the atopic to notational systems, I create a score from the non-essential notes in Haydn's *Trumpet Concerto*, a recording of which was used in the MP3 listening tests.

Following a brief, summarizing conclusion, computer code capable of creating audio and visual compression detritus is provided. Algorithms used to segment and rearrange sound files and to trace semi-random binaural trajectories in *moDernisT* are given. The dissertation concludes with the approximately five hundred lines of code used to create the handheld, interactive version of *moDernisT* and an accompanying bibliography.

> "The mp3 encoding process puts the body on a sonic austerity program. It decides for its listeners what they need to hear and gives them only that."[1]

Every MP3 contains within it codified assumptions regarding supposedly universal characteristics of the human ear and the human listening body. While the results of auditory perceptual research are considered statistically significant in the signal processing literature, this dissertation questions

1 Sterne, Jonathan. "The mp3 as cultural artifact." *New media & society* 8, no. 5 (2006): 825-842.

presumed universal aspects of human sound experience and perception.[2] My aesthetic goal is to create immersive sensory environments through sound. To reduce music to only that which can be perceived according to a discrete model of auditory perception is counter to the aesthetics of the irreducible.[3] In this dissertation, I develop art works which immerse the listener in complex soundscapes possibly exceeding their perceptual facilities. In doing so, my hope is that the individual may come into contact with the auditory sublime. Reducing listening environments to a minimum necessary stimulus is tangential to that effort.

I work with sonic materials routinely discarded by compression technologies, finding a creative spark at the intersection of perpendicular logics: those of the aesthetic and those of perceptual technics. Ironically, I distribute this material online, where it is subjected to compression and network logics anew. This is a compromise I am willing to make—by doing so, the work has reached a far greater public than it would have otherwise. Fortunately, the ideas seem to survive the compression, even if the sonic details don't.[4] Remarkably, the project went viral, with a digital audience eager to hear sounds from beyond the compressed confines of the Internet. Some listeners even requested uncompressed downloads after hearing the project streamed. In all of this, I've observed the aesthetic of the irreducible and the logic of perceptual technics coexisting and intersecting in unexpected ways.

Michael Bull's study of iPod users suggests that music listening, for many people, is not an end in itself. It folds itself in to the rhythms of daily life.[5] As Sterne notes, portable music "aids its listeners in various

2 Regarding perceptual coding, Sterne calls for an ontology of hearing that accounts for plurality, noting "psychoacoustics and information theory contain within them pregiven answers to the normative questions that animate communication—why communicate, to what end." Sterne, Jonathan. *MP3: The meaning of a format*. Duke University Press, 2012.

3 Consider, for example, how the real numbers are uncountably infinite while the integers are only countably so.

4 "Perhaps the best attitude, then, is a certain ambivalence towards mp3s." Sterne, Jonathan. "The mp3 as cultural artifact." *New media & society* 8, no. 5 (2006): 825-842.

5 Bull, Michael. *Sound Moves: iPod culture and urban experience.* Routledge, 2015.

subjective acts, but it is not itself an object of direct contemplation. Listening is not an end in itself. It does not aspire to a single aesthetic standard or political project."[6] As such, I understand that the works created for this dissertation may take on many (perhaps compressed) forms and serve many (perhaps unanticipated) purposes.

Though this work may present as a critique of digitality, the usage of digital technologies in production and presentation implies that it is not an outright rejection of such logics. The universalizing tendencies of rationalist epistemology (i.e., mistrust of sensory experience and of the body in favor of immaterial abstraction) are aided by digital technologies which have discretization and abstraction at their core. As Aden Evens notes, "The digital makes an excellent companion to rationalist epistemology, instrumental reason, positivist notions of truth, and so colludes with ways of thinking on the rise for a number of centuries."[7] Describing how such reductive strategies can be traced to the universalizing strategies of Western epistemology, the philosopher Adriana Cavarero, in her plea for an ontology that takes the singularity of the individual, embodied voice as its foundation, charges that "the unrepeatable singularity of each human being, the embodied uniqueness that distinguishes each one from every other is, for the universalizing tastes of philosophy, a superfluity. Uniqueness is epistemologically inappropriate"[8]

The digital is ascendant because it instantiates such already influential epistemologies. By way of the bit, digital technologies make abstraction concrete. This abstraction is both the power and limit of the digital. As with the digital in general, perceptual coding should be placed in proper perspective as a technique limited in scope.[9] These limits do not necessitate rejection, but they do compel a degree of

6 Sterne, Jonathan. *MP3: The meaning of a format*. Duke University Press, 2012.
7 Evens, Aden. *Logic of the Digitial.* Bloomsbury Publishing, 2015.
8 Cavarero, Adriana. *For more than one voice: Toward a philosophy of vocal expression.* Stanford University Press, 2005.
9 Infrasound and ultrasound, both defined as lying beyond the bounds of human auditory perception, have been shown in

skepticism and sensitivity. By listening to sounds deleted by perceptual coding, I affirm the actual which lies beyond the digital, the material which exceeds the formal, and the irreducible which defies prediction and explanation.

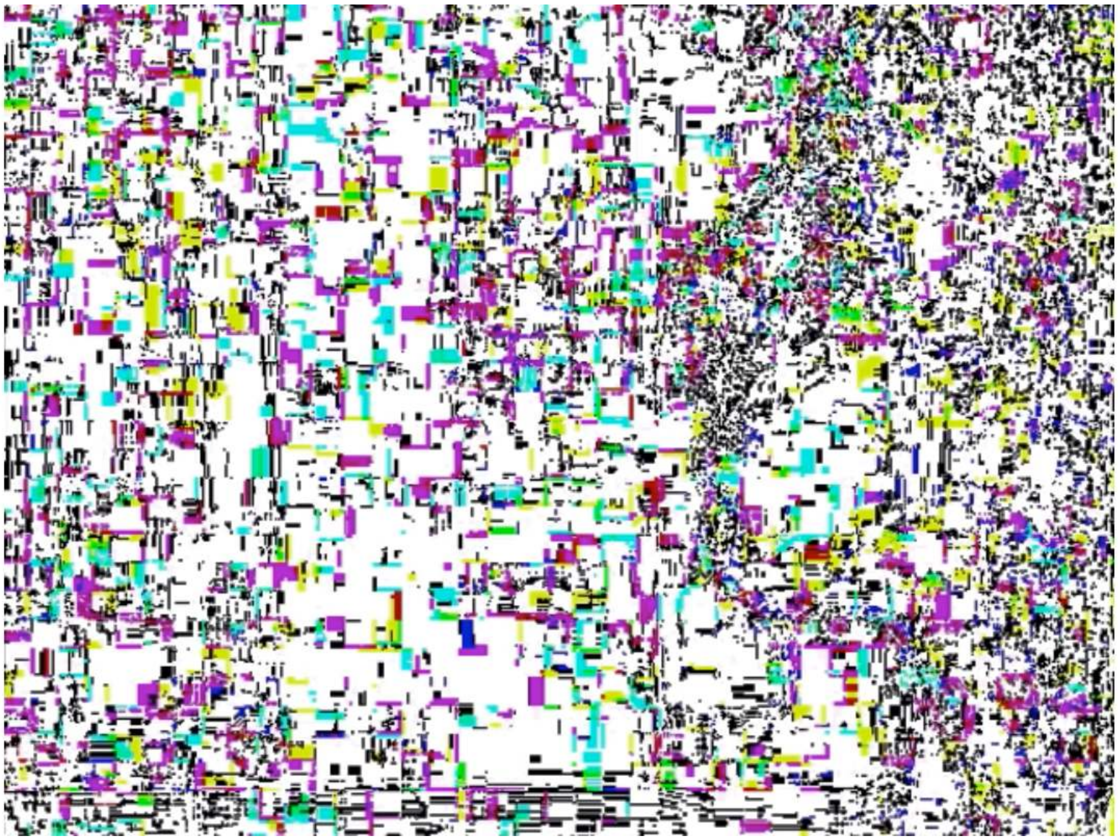

Figure 1. MPEG-4 detritus from the video "Tom's Diner"

It would be impossible to exhaustively list the artistic practices and practitioners concomitant with the creative works in this dissertation owing both to the the limited capacity of my own perspective and to the infinitely expanding network of ideas and techniques that my central concerns are a part of. Nevertheless, in the following paragraphs, I will briefly describe several aesthetic movements, individual

recent years to activate the auditory cortex. See: Dommes, E., H.C. Bauknecht, G. Scholz, Y. Rothemund, J. Hensel, and R. Klingebiel. "Auditory cortex stimulation by low-frequency tones—an fMRI study." *Brain Research* 1304 (2009): 129-137. Also see: Oohashi, Tsutomu, Emi Nishina, Manabu Honda, Yoshiharu Yonekura, Yoshitaka Fuwamoto, Norie Kawai, Tadao Maekawa, Satoshi Nakamura, Hidenao Fukuyama, and Hiroshi Shibasaki. "Inaudible high-frequency sounds affect brain activity: hypersonic effect." *Journal of neurophysiology* 83, no. 6 (2000): 3548-3558.

artists, and particular works which relate to the creative portfolio described in this dissertation.

Subtraction is a central technique in this dissertation—my materials are all obtained by time-domain subtraction of compressed audio from uncompressed audio. In visual art, subtractive works evoke Robert Rauschenberg's *Erased de Kooning Drawing* (1953), in which the artist erased a drawing by Willem de Kooning to create a new work. The new media artwork *Super Mario Clouds* by Corey Arcangel is similarly revealed through a process of digital subtraction. To create the work, Arcangel hacked the classic video game Super Mario Brothers and erased everything in the original game except the clouds. The resulting animation is exhibited as a large-scale digital projection with accompanying silk-screen prints for sale. In its appropriation of 1980s popular culture, open source hacker ethics, stripped down aesthetics, and foregrounding of something hitherto considered background, I find resonances with the works presented in this dissertation.[10]

In music, subtraction evokes minimalism, where linear subtractive processes might, for example, subtract a note from a looping melody on each iteration, as in Rzewski's "Les Moutons de Panurge."[11] Taken to its extreme, subtraction results in silence, calling to mind John Cage's famous work *4'33"*. In his book, *Silence*, Cage traces subtractive processes at least back to Debussy, whose compositional process was anecdotally described as: "I take all the tones there are, leave out the ones I don't want, and use all the others."[12] More recently, and involving recording technology, William Basinski's *Disintegration Loops*, was created by recording the sound of aged magnetic tape loops materially decaying as they played, increasingly disintegrating with each cycle before the listener's ear. This is a material manifestation of the linear subtractive process found Rzewski's work. Finally, in electronic music,

10 Tribe, Mark, Reena Jana, and Uta Grosenick. *New media art*. Los Angeles: Taschen, 2006.
11 Additive and linear subtractive processes are described, among other key concepts in minimal music, in Warburton, Dan. "A working terminology for minimal music." *Intégral* 2 (1988): 135-159.
12 Cage, John. *Silence: lectures and writings*. Wesleyan University Press, 2011.

subtractive synthesis is a foundational sound production technique typified by the use of filters to pare noisy input signals into band-limited outputs.[13]

The material uniting all of the works in this dissertation is compression detritus. Compression detritus —that which is deleted by compression—is closely related to compression artifact—distortion introduced by the compression process. The piece *I Am Sitting in a Video Room* by artist Patrick Lidell is a Youtube-based homage to Alvin Lucier's 1969 piece *I Am Sitting in a Room*, which substitutes the accumulation of audio compression artifacts caused by repeated Youtube uploads for the accumulation of architectural-acoustic resonances in Lucier's original.[14] Compression artifacts are a foundational material in glitch arts across media. Carolyn Kane, in *Compression Aesthetics: Glitch From the Avant-Garde to Kanye West*, goes even further, noting that "the entire history of modern art could be construed as a glitch and compression of Enlightenment epistemology."[15] The aesthetic developed in this dissertation has as precursors Cubism, in its denial of the sole validity of linear perspective, Impressionism and Post-Impressionism, in their emphasis on abstract color over classical realism, and Futurism, in its poeticization of the speed and compression of modern life in the mid-twentieth century. Russolo, in *The Art of Noises*, advocated using modern technologies to create sound—especially noise—in place of old world orchestras.[16] The "orchestra" used to create the works in this dissertation is populated primarily by the MP3 compression "instrument family"—variable, constant, and average bit rate encodings of all resolutions, from 8kbps to 320kbps.

13 Smith, Julius O. "Viewpoints on the history of digital synthesis." In *Proceedings of the International Computer Music Conference*, pp. 1-1. INTERNATIONAL COMPUTER MUSIC ACCOCIATION, 1991.
14 Lucier, Alvin, Michael Roth, Nicolas Collins, and Ronald Kuivila. *Alvin Lucier: A Celebration*. Wesleyan University Press, 2012.
15 Kane, Carolyn. "Compression Aesthetics: Glitch From the Avant-Garde to Kanye West." Unpublished *article. http://ivc.lib.rochester.edu/wp-content/uploads/2014/10/Kane-Compression-Aesthetics.pdf* (2014).
16 Russolo, Luigi. "The Art of Noises (1913)." *Futurist Manifestos* 74 (1973).

A musical interest in previously nonmusical sounds could be considered a defining feature of experimental music. In *musique concrète,* an early example, recordings of everyday sounds are aestheticized, as in Schaeffer's "Etude aux chemins de fer." Sound recording and production technologies are commonly centered in such musics—a narrow focus unfortunately typical of conventional masculinity.[17] For example, circuit bending musicians modify the electronic components of commercial music hardware, instruments, and toys to create new, unpredictable sounds. Christian Marclay's *Record Without A Cover* aestheticizes the inadvertent skips and pops caused by the physical degradation of a vinyl record. Oval's *94diskont* uses the sounds of skipping CD's to create ambient, minimal techno. Yasunao Tone creates abstract compositions by dissecting and corrupting MP3 files for his *MP3 Deviations* series. Musical encounters with previously nonmusical sounds also typify noise and punk musics, as well as genres featuring record scratching, extended performance techniques, and other creative subversions of musical instruments and technologies.

Media artifacts feature prominently in many hauntological and afrofuturist musics, for example in the crackle of vinyl records foregrounded in recordings by the artists Burial and Tricky (and Kendrick Lamar, Shabazz Palaces, Ariel Pink, Low, and many others). Mark Fisher distinguishes hauntology from postmodernism by noting that, whereas postmodernism glosses over the temporal disjunctions of twenty-first century life, hauntological and afrofuturist artists highlight them—as in the foregrounded crackle of a vinyl record which opens *To Pimp a Butterfly*.[18] Simon Reynolds describes the broken sense of time common to hauntology and afrofuturism as *dyschronic*. Dyschronia is a situation in which the

17 Hannah Bosma, in her critique of the masculinity of glitch aesthetics, writes, "With respect to gender, it is crucial to differentiate between a narrow notion of technology and music, which fetishises and isolates its material, and a broad, hybrid conception that takes the situatedness of its practices into account. A masculine aesthetics of digital failure, focused on technology for technology's sake, turns out to be a deadend. Alternative conceptions of glitch, as intertwined with the practices and social relations in which it is embedded, may yet revitalise the notion of glitch and sensitise it to gender." Bosma, Hannah. "Gender and technological failures in Glitch music." *Contemporary Music Review* 35, no. 1 (2016): 102-114.

18 Fisher, Mark. "The metaphysics of crackle: Afrofuturism and hauntology." *Dancecult: Journal of Electronic Dance Music Culture* 5, no. 2 (2013): 42-55.

present and can not be sealed off from the past and into which traces of unrealized futures unpredictably bubble up.[19]

There are additional resonances between the aesthetic developed in this dissertation and certain concepts from Japanese aesthetics. In *Essays in Idleness*, Yoshida Kenkō gives an example of the understated beauty expressed by the idea of *wabi*: "Are we to look at cherry blossoms only in full bloom, at the moon only when it is cloudless? … Branches about to blossom or gardens strewn with faded flowers are worthier of our admiration."[20] The idea of *yūgen* is nearly ineffable, roughly translating to mysterious grace or dark, and placing importance on the imagination. Kamo no Chōmei characterizes it by example: "When looking at autumn mountains through mist, the view may be indistinct yet have great depth. Although few autumn leaves may be visible through the mist, the view is alluring. The limitless vista created in imagination far surpasses anything one can see more clearly."[21] This passage indicates a general predilection for the allusive rather than the explicit.

Throughout this dissertation, I reference various discourses beyond the disciplinary boundaries of music composition. To begin, the last word of my title, irreducibility, indicates a relationship to the basic question of reductionism: whether phenomena can be sufficiently described in terms of their constituent parts.[22] Reductionism is inverted by various antireductionist arguments—holism, emergentism—made manifest in such disciplines as cybernetics, systems theory, and ecology.[23] Antireductionist views maintain that a complete description in terms of constituent

19 Reynolds, Simon. "Haunted audio." *The Wire* 273 (2006): 26-33.
20 Kenko, Yoshida. *Essays in idleness*. Cosimo, Inc., 2005.
21 no Chomei, Kamo. "An Account of My Hut." *Anthology of Japanese Literature* (1955): 197-212.
22 Ney, Alyssa. "Reductionism." (2008).
23 Block, Ned. "Anti-reductionism slaps back." (1996).

parts is epistemologically or ontologically impossible for a given phenomena. My project, being concerned with irreducibility, implies that human audiovisual perception (and aesthetic experience) is insufficiently described by discrete perceptual coding models. Beyond philosophy, irreducibility arises in more mathematical subjects than should be enumerated here. In these contexts, the concept generally refers to an entity that can not be broken down into smaller parts.

The idea of atopia also appears throughout this dissertation. This term, borrowed from the Greek for *without place* or *unusual*, is used to describe a society without territorial borders or national identity by the German sociologist Helmut Willke.[24] In relation to the aesthetic, I am interested in Barthes' conception of atopia which suggests a poetic register of language where the need to constantly make rhetorical sense is contested.[25] This register allows creativity to emerge from an indeterminate state before convention reduces non-specified phenomena to concrete predicates. The works in this portfolio strive to express just such an indeterminate state.

Skepticism towards reductionism and an embrace of the atopic in this dissertation suggests a parallel rejection of the ways in which people and groups are othered by society through exclusion and displacement socially, politically, economically, and geographically. The concept of the other in western philosophy is developed variously through the writings of Hegel, Husserl, Sartre, de Beauvoir, and related philosophers. Hegel, in *Philosophy of Mind*, describes how knowledge of self can only arise through examination of one's relationship with others.[26] This struggle for self-consciousness is undermined in unequal relationships by the inadequate recognition of the self through the other.

24 Willke, Helmut. "Atopia." *Studien zur atopischen Gesellschaft. Frankfurt a. M.: Suhrkamp* 58 (2001).
25 Barthes, Roland. *Roland Barthes by Roland Barthes*. Macmillan, 2010.
26 Hegel, Georg Wilhelm Friedrich, and Michael Inwood. *Hegel: Philosophy of Mind: Translated with Introduction and Commentary*. Oxford University Press, 2007.

Nearly one hundred and fifty years later, Simone de Beauvoir describes the unequal treatment of women in society in terms of this so-called "Lord and Bondsman" dialectic.[27] Othering is also critiqued by postcolonial theory in the work of Gayatri Chakravorty Spivak, for example, who develops (from Gramsci) the idea of the *subaltern*—the marginalized other in a given socio-economic context.[28] [29] In *The Souls of Black Folk*, W.E.B. DuBois conceives of the black experience in America as “a world which leaves him no true self-consciousness, but only lets him see himself through the revelation of the other world. It is a peculiar sensation, this double-consciousness, this sense of always looking at one’s self through the eyes of others, of measuring one’s soul by the tape of a world that looks on in amused contempt and pity.”[30] In general, reductionistic ideologies homogenize, essentialize, and subordinate *othered* social identities, undergirding harmful discriminatory practices.[31] My work translates ethical critiques of othering into a creative practice which centers and celebrates previously discarded sounds.

This dissertation develops an aesthetic of the irreducible through creative engagement with perceptual audio coding. This engagement situates the project in close relation to technology studies, media studies, and sound studies. In particular, I am indebted to the work of Jonathan Sterne who contributes to all three of these discourses. His book, *MP3: The Meaning of a Format*, provided the initial inspiration for this project.[32] What began as a simple exploration into the particularities of the MP3 format and its affordances for music composition has blossomed into a much broader inquiry regarding the ineffable, the non-quantifiable, and the nonpredicable. It is in relation to these disparate discourses that I develop an aesthetic of the irreducible.

27 Beauvoir, Simone. *The Second Sex.* Alfred A. Knopf, 1971.
28 Spivak, Gayatri Chakravorty. "Can the subaltern speak?." *Can the subaltern speak? Reflections on the history of an idea* (1988): 21-78.
29 Gramsci, Antonio. *Prison notebooks*. Vol. 2. Columbia University Press, 1992.
30 Du Bois, W. E. B., Henry Louis Gates, and Terri Hume Oliver. "The souls of black folk: A norton critical edition." *Edited by Henry Louis Gates, Jr. and Terri Hume Oliver. New York and London: WW Norton & Company* (1999).
31 See, for example, Said’s discussion of the West’s polarizing and patronizing representations of “The East” in Said, Edward W. *Orientalism*. Vintage, 1979.
32 With thanks to Tara Rodgers, who introduced me to his work during my studies with her.

Chapter 1

A Capella

Suzanne Vega's "Tom's Diner" (1981) is an a capella meditation on a single melodic idea, five verses in strophic form.[33] It begins with her description of a morning scene in a coffee shop, observing the comings and goings of various patrons and employees. The prototypical folk form becomes unsettled at the beginning of the fifth verse: here, the lyrics shift suddenly "from perception to memory, from exteriority to interiority," set in relief by marked silences.[34] This arresting move occurs on the dominant and subdominant, rather than at the end of the phrase.

To the bells
Of the cathedral . . .

I am thinking
Of your voice . . .

Vega proceeds with a corkscrewed reworking of the original melody, working her way down to the mediant and a third fermata, continuing the *Urlinie* 5 – 4 – 3, and leaving the listener in a state of suspension: "And of the midnight picnic, once upon a time before the rain began . . ."

Finally, the narrator snaps out of her reverie, emerging from the sphere of memory to reengage with the world of perception. Vega resumes the playful affect encountered earlier in the song, her detailed account of the day's events picking up balletically where it left off: "And I finish up my coffee and it's

33 The song has inspired dozens of remixes and, owing to its usage in the MP3 listening tests, has become a standard test item in digital audio, similar to the "Lena" photograph in digital imaging. Suzanne Vega's account of the song's genesis and the cultural life it has since lived was published in the New York Times. Vega, Suzanne. "Tom's Essay." Measure for Measure. http://measureformeasure.blogs.nytimes.com/2008/09/23/toms-essay/index.html

34 Richardson, John. "Vega Meets Weill: Alienation Effects and Archaeological Strategies in 'Tom's Diner' and 'Luka'." Unpublished draft provided to author.

time to catch the train." The song concludes with the now iconic wordless melody humming along, the scene fades to silence, the narrator striding off to catch her train.[35]

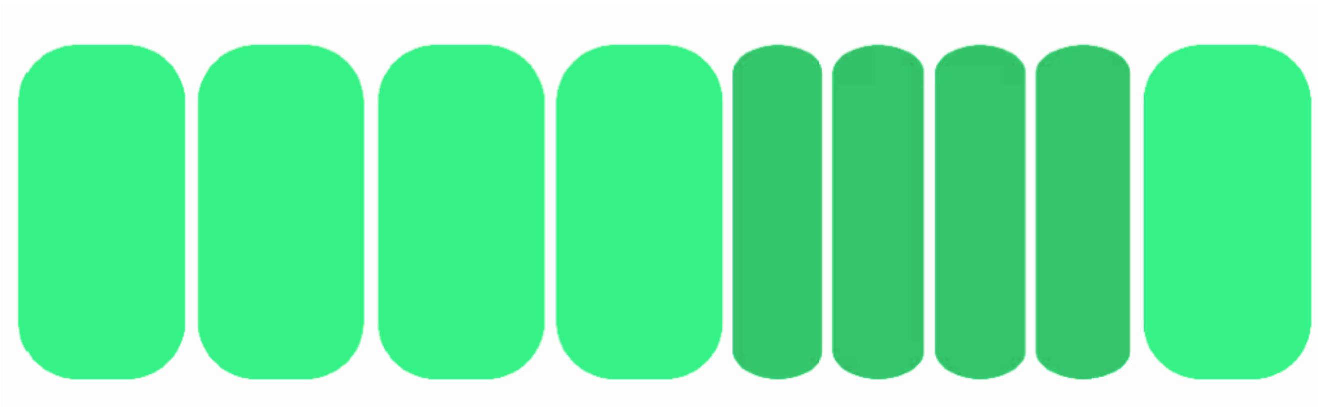

Figure 2. Sections of the song "Tom's Diner"

moDernisT

First, there is *breath*.[36] An inhalation. It's short, lasting just two tenths of a second. The abbreviated duration makes it read like a gasp. One might ask: who is gasping? Are they surprised by something? What information does this inhalation carry? Breathing is among the most fundamental human phenomena, alongside eating, sleeping, and the heart beating.[37] Certainly, the breath is a vital

35 Ibid.

36 Breath is a principle of life. In respiration, it suggests a binary rhythm—"the opening and closing of the Gates of Heaven" in Taoism. The Persian poet Sa'dī writes: "Each respiration holds two blessings. Life is inhaled, and stale, foul air exhaled. Therefore thank God twice every breath you take." In some yoga texts, a third state, *Kumbhaka,* is emphasized: the pause between inhalation and exhalation.

37 Medically, breathing is an automatic, essential bodily function. It's symbolic importance across cultures is described in Chevalier, Jean, and Alain Gheerbrant. *A dictionary of symbols*. Oxford: Blackwell, 1994.

force. I'm not sure if this particular breath indexes gender or age, but it is relatively high pitched, with a mostly flat spectral distribution from 1k to 10k. It's not white noise, but it is close. It emanates from the center of the stereo field and, with headphones on, it sounds like it's coming from inside of your own head. Are you surprised? An inhalation like this, at the beginning of a recording raises questions of intention. Was the breath left here on purpose? If so, what purpose does it serve? At the very least, this breath implies life beyond the recording, it suggests that something lies beyond the frame of the song. In this case, it's the only element of the original recording that I've left in, uncompressed.[38] In two hundred short milliseconds, the breath announces that a singer is here, she's about to begin, and you might want to listen.

A click emanates from both channels and at the same time, you hear something reminiscent of a voice, but digitally obfuscated. The voice that inhaled recedes into the distance as a new logic takes precedence. It's still only been thirty-five hundredths of a second, but information is dense here. It's clearly in F-sharp minor, and the first two scale degrees are audible in a rising stepwise succession, but the melodic line is submerged in a digital wash. The entire F-sharp minor scale, save for E natural, enters simultaneously, spread across two octaves and with a timbre reminiscent of digitally generated breath. This new texture contains four layers, each layer consisting of the MP3 detritus of a different quality MP3 file. Three of the layers derive from constant bit rate MP3s, two encoded at 320kbps, the maximum quality encoding, and one encoded at 64kbps, considered moderate quality. The fourth layer is encoded using a variable bit rate (VBR) encoding that automatically applies a low pass filter at approximately 19.5kHz and fluctuates between 190 and 250 kbps.[39] [40]

38 Full disclosure—I am amplifying it.

39 All MP3's for this project were encoded using LAME (LAME Ain't an MP3 Encoder), recommended as the highest quality MP3 encoder for most purposes by Hydrogenaudio. LAME [Computer Software]. Version 3.1. http://lame.sourceforge.net

40 The technical underpinnings of the MP3 are detailed in a variety of sources. A good starting point is: Brandenburg, Karlheinz. "MP3 and AAC explained." In *Audio Engineering Society Conference: 17th International Conference: High-*

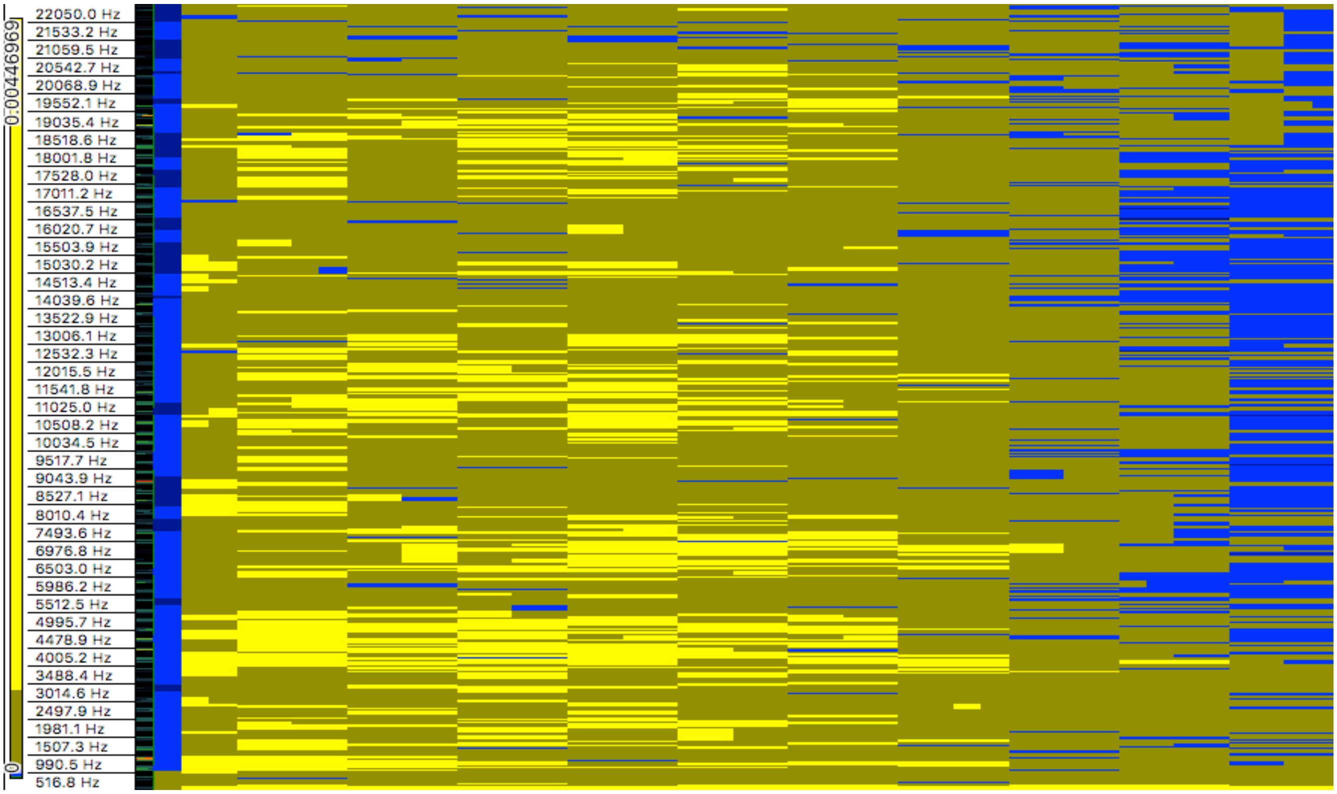

Figure 3. The opening breath in *moDernisT* as adaptive spectrogram[41 42]

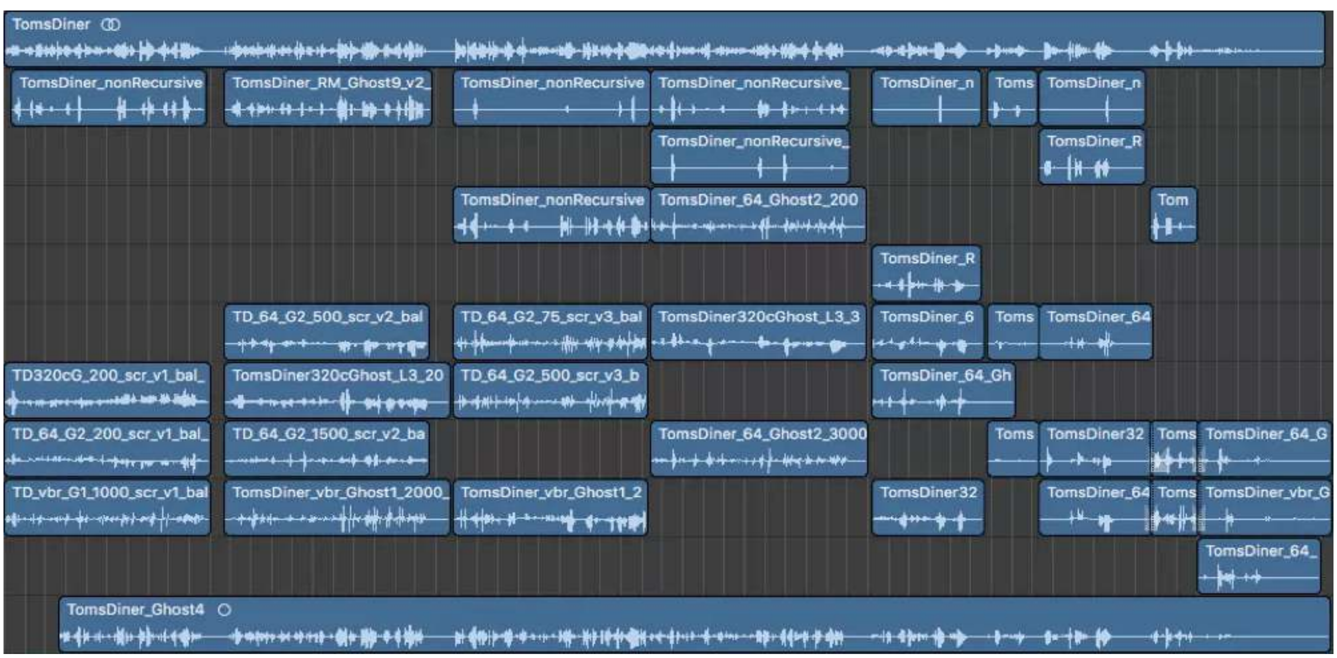

Figure 4. moDernisT in the digital audio workstation

Quality Audio Coding. Audio Engineering Society, 1999.

41 The time-frequency resolution of an adaptive spectrogram varies in direct proportion to the complexity of the signal.

42 Sonic Visualiser [Computer Software]. Version 3.0.3, 2017. Retrieved from http://www.sonicvisualiser.org

The first layer is extracted from the first verse of "Tom's Diner" using a simple frequency domain masking script that I write in Python.[43] I compress the verse and convert it to the frequency domain using a linear frequency spectrum short-time Fourier transform with an FFT filter bank length of 16,384 samples, an FFT signal window length of 8,192 samples with a Hamming window, and an FFT hop size of 4,410 samples, one tenth of the sampling rate. I perform the same transform to the original uncompressed audio and then verify that the two files are aligned in time before doing comparative processing.

Both of these files are now stored as 'numpy' arrays in Python and are thus available to further processing using linear algebra operations.[44] I use a simple entrywise matrix subtraction to find the difference between the compressed and uncompressed audio files. I then scan this difference matrix for entries below an adjustable threshold. I mark every entry that lies below the threshold. These entries correspond with time-frequency points where the compressed and uncompressed audio files are similar. Next, in the uncompressed matrix, I erase every entry that I've marked as being similar to the MP3. The remainder represents those time-frequencies present in the uncompressed audio that are different in the compressed audio. By adjusting the threshold, I can selectively isolate time-frequencies by their degree of difference from the original.

I now use an inverse short-time Fourier transform to render the audio as a waveform in the time domain, estimating the phases via phase vocoder resynthesis. The resulting audio contains clear undulations at prominent frequencies, almost sine-like, giving this layer a distinctive character in comparison to the diffuse, textural, breath-like quality of the previous layer. I write this waveform to a

43 Python Software Foundation. Python Language Reference, version 2.7. Available at http://www.python.org
44 NumPy (colloquially known as 'numpy') is the primary scientific computing package in Python. NumPy [Computer Software]. Version 1.15.4. http://www.numpy.org/

file and pass it through an HRTF script (detailed below) using semi-random time, azimuth, and elevation settings. The chosen settings situate the implied sound source behind the listener—an impossible perceptual effect in traditional two-channel stereo imaging. I import this processed audio into a digital audio workstation for mixing, dropping it in at the one second mark.

I apply reverb, using a large hall impulse response with a 6.8 second decay time, and set the reverb output at about 8 decibels below the dry output. I pan the reverberated signal back and forth slowly during the verse several times, moving plus or minus twenty percent from the center every few seconds. I situate all of this quietly in the mix again, nearly 20 db below unity gain for the stereo track as a whole.

The second layer is created by encoding the source song as a 320kbps constant bit rate, joint stereo MP3. This MP3 is compared with the original uncompressed audio file using a null test. The resultant difference between the two versions forms the basis for the layer. I process this further by removing the original semantic and melodic information. This is achieved by repeatedly indexing into the verse at pseudorandom time-points, extracting segments of bounded indeterminate duration, and compiling these in linear succession (with cross-fades) until the total duration of the new assemblage is greater than or equal to the duration of the original verse. The resultant audio maintains the timbre of vocal MP3 detritus and a general sense of key but has a new rhythmic character. Specific syllables are no longer present. The original melody is obscured. I generate several candidates using this procedure and select the most appealing composite by ear.

Next, I attend to spatialization by converting the stereo audio file to mono and processing it using a digital signal processing HRTF library called Headspace, written in Python for performing head-related transfer functions based on KEMAR dummy head microphone measurements by the MIT Machine Listening Group.[45] [46] I write a Python script to interpolate between user-defined time, azimuth, and elevation points. The script generates automated 3d panning which can be applied to any sound file. I pass this layer through the algorithm specifying its starting location front and center with a trajectory that travels around the head to the left, circling three quarters of the way around to the right ear, before traveling over the listeners head and ending in the back, behind the head, at the end of the twenty-second-long first verse. In the final mix, I place this layer in the background, nearly 30 db below unity gain. It is just audible at this level, since the mix overall is quiet, and it adds a subtle texture.[47]

The third layer is constructed in a similar fashion to the second layer. It first diverges by having an encoding bit rate of 64 kbps. This bit rate leaves behind more detailed sonic information than does 320 kbps compression. I create composite melodies from randomly selected segments two hundred milliseconds or less in Python. I generate several candidates with different random orderings and select the composite that catches my ear. I next determine a binaural trajectory for the chosen composite to take. Beginning on the right side, the melody flips rapidly to the left, hovering there for several seconds before rebounding to the right, oscillating vertically, and tumbling backwards to settle behind the listener.[48] I mix this layer somewhat louder than the previous two, around 10 db below unity.

The fourth layer derives from an MP3 encoded at a variable bit rate between 200 and 250 kbps. This encoding is phase inverted and time-aligned with the original uncompressed audio. As similar to the

45 Rabl, Christopher and Curtis Litchfield. Headspace [Computer Software]. Https://github.com/crabl/HeadSpace
46 Gardner, Bill and Keith Martin. “HRTF Measurements of a KEMAR Dummy-head Microphone.” (1994).
47 Materially and aesthetically, this work presents an accumulation of subtle textures. In layering MP3 detritus, I hope for an effect similar to the “hum of London” that Brian Eno describes. “If you sit in Hyde Park just far enough away from the traffic so that you don’t perceive any of its details, you just hear the average of the whole thing. And it’s such a beautiful sound. For me that’s as good as going to a concert hall at night.” Eno, Brian and Anthony Korner. “Aurora Musicalis.” In *Artforum*, vol. 24, no. 10, pp. 26-36. 1986.
48 Cheng, Corey I., and Gregory H. Wakefield. “Moving sound source synthesis for binaural electroacoustic music using interpolated head-related transfer functions (HRTFs).” *Computer Music Journal* 25, no. 4 (2001): 57-80.

second layer, I use a Python script to reassemble the track from randomly selected segments with a maximum duration of one second each. These segments are up to five times longer than the segments used in the first layer. At this length, they retain some of their original semantic content (with recognizable syllables and words) but longer phrases are rendered incoherent. I curate the final composite from a pool of several candidates generated by the aforementioned script.[49] I am interested in how this compositional process results in a loss of information, paralleling the information loss that MP3 compression enacts on the source recording, in the following sense: the transformation turns discernible language into something only language-referring. The loss creates a semantic vacuum which affords multiple interpretive readings by encouraging less literal, more intuitive encounters with the material.[50] Additionally, I believe such materials both afford and reward close timbral listening—a mode of listening which I hope to encourage for people encountering these timbrally distinctive perceptual coding artifacts.[51] Ultimately, I also mix this layer quietly at almost 20 decibels below unity gain.[52]

Verse One

The first 1500 milliseconds contain a great deal of new information which I've now described in detail,

---

49 "Somebody asked Debussy how he wrote music. He said: I take all the tones there are, leave out the ones I don't want, and use all the others." Cage, John. *Cage.* Cunningham Dance Foundation, 1970.

50 The approach is similar to Lansky's. "*Idle Chatter* uses a kind of stochastic distribution, random selection without replacement, of LPC-synthesized voice fragments in which words are edited so that they are unintelligible..." Lansky, Paul. "Reflections on spent time." In ICMC. 2009.

51 Lansky again: "What was noticeable, however, was that my listeners had to do some work while they listened. The combination of this and the random textures seemed to be a step in the right direction with respect to the problem of decay." Ibid.

52 "We have all had the experience, on a visit to one of the great temples of Kyoto or Nara, of being shown a scroll, one of the temple's treasures, hanging in a large, deeply recessed alcove. So dark are these alcoves, even in bright daylight, that we can hardly discern the outlines of the work; all we can do is listen to the explanation of the guide, follow as best we can the all-but-invisible brush strokes, and tell ourselves how magnificent a painting it must be." Tanizaki, Jun'ichirō. *In praise of shadows.* Vol. 24. Random House, 2001.

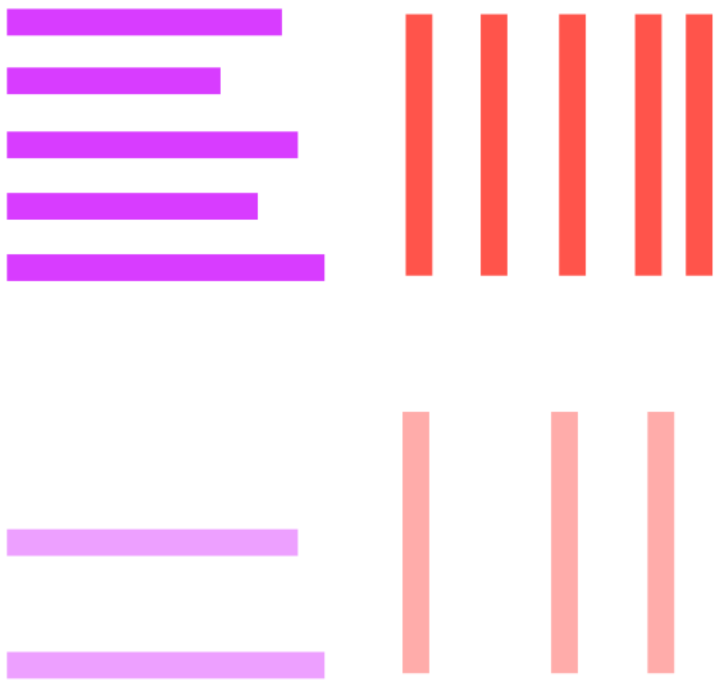

Figure 5.

but the listener does not (usually) have the means to slow time to this level of granularity when encountering the composition as a work of fixed media. A discussion of the first verse taken in aggregate, a structural unit which lasts roughly 20 seconds, tracks more closely to the experience of listening to the recording. Broadly, this section of music reads as having two primary textures—the first comprised of clearly pitched, reverberant waveforms in a range typical of the human voice, and the second comprised of three sublayers of shifting, breath-like, noisy material made to flock or wander restlessly around the listener.[53] Continuing as in the first 1500 milliseconds, the first texture presents a faint outline of the original melody, informationally compressed by jettisoning the finer details while retaining the overall structure. This is set in relief against the second texture which, as described above, presents rich, shifting, grain-like sounds swirling around the listener. Over the full twenty second verse, this second texture's center of acoustic energy rolls from the right side of the listener to the left side. The verse, which opens with a sharp breath and a click sending spectral energy up to 20k, is largely band-limited to frequencies 10k and below. The timbre gradually brightness as the verse proceeds, with frequencies up to 15k re-emerging by the end of the section. The gesture truncates into near silence, as the reverb tail from the first layer recedes in the wake of a fermata.

> "Here, what is normally seen as a disturbance—such as the slightest breath or the most distant echo—instead becomes a contribution and, thus, a material."[54]

53 Flocking is a celebrated natural phenomenon in spatialized electroacoustic music. For example, see Stine, Eli and Kevin Davis. *The Murmurator: A Flocking Simulation-Driven MultiChannel Software Instrument for Collaborative Improvisation*. In *International Computer Music Conference (ICMC).* 2018.

54 Bailly, Jean-Christophe, and Matthew H. Anderson. "The Slightest Breath (On Living)." *CR: The New Centennial Review* 10, no. 3 (2010): 1-11.

Verse Two

An inhalation primes the second verse, same as the first, which shifts quickly into a subtly transformed texture. The reverberant sine tones are now nearly inaudible, sitting 30-plus db below unity gain in the mix. However, a new, fifth layer has been added to the shifting, grain-like texture, resulting in more overall volume and greater timbral density. The new layer is derived from a null test using a 64 kbps constant bit rate encoding of the song's second verse. This is algorithmically spliced with a maximum splice size of 500 ms—large enough that syllables are audible but too short for most words. This layer is rendered in binaural space using the Headspace library in Python, animated by a custom script with pseudorandom numbers chosen for the time, azimuth, and elevation, with the final path chosen by the composer from a pool of candidates output by the script.

Additionally, three of the layers that were already present in the first verse, specifically those derived from a null test (layers two, three, and four above) rather than a frequency domain masking script (layer one above) have had their maximum segmentation size increased significantly. The layer deriving from a 320kbps MP3, corresponding to layer two above, has seen its maximum segmentation size increased from 200 milliseconds to 2000 milliseconds. Layer three, derived from a 64kbps MP3, its maximum segmentation size increased from 200 milliseconds to 1500 milliseconds. Finally, layer four, derived from a variable bit rate MP3, its maximum segmentation size increased from 1000 to 2000 milliseconds. Perceptually, this technical change increases the relative intelligibility of the songs lyrics, adding a layer of semantic complexity to the final artifact, and in the heightened presence of Suzanne Vega's voice, a change that lends a greater warmth and human presence to an otherwise cold soundscape filled with digital compression detritus.

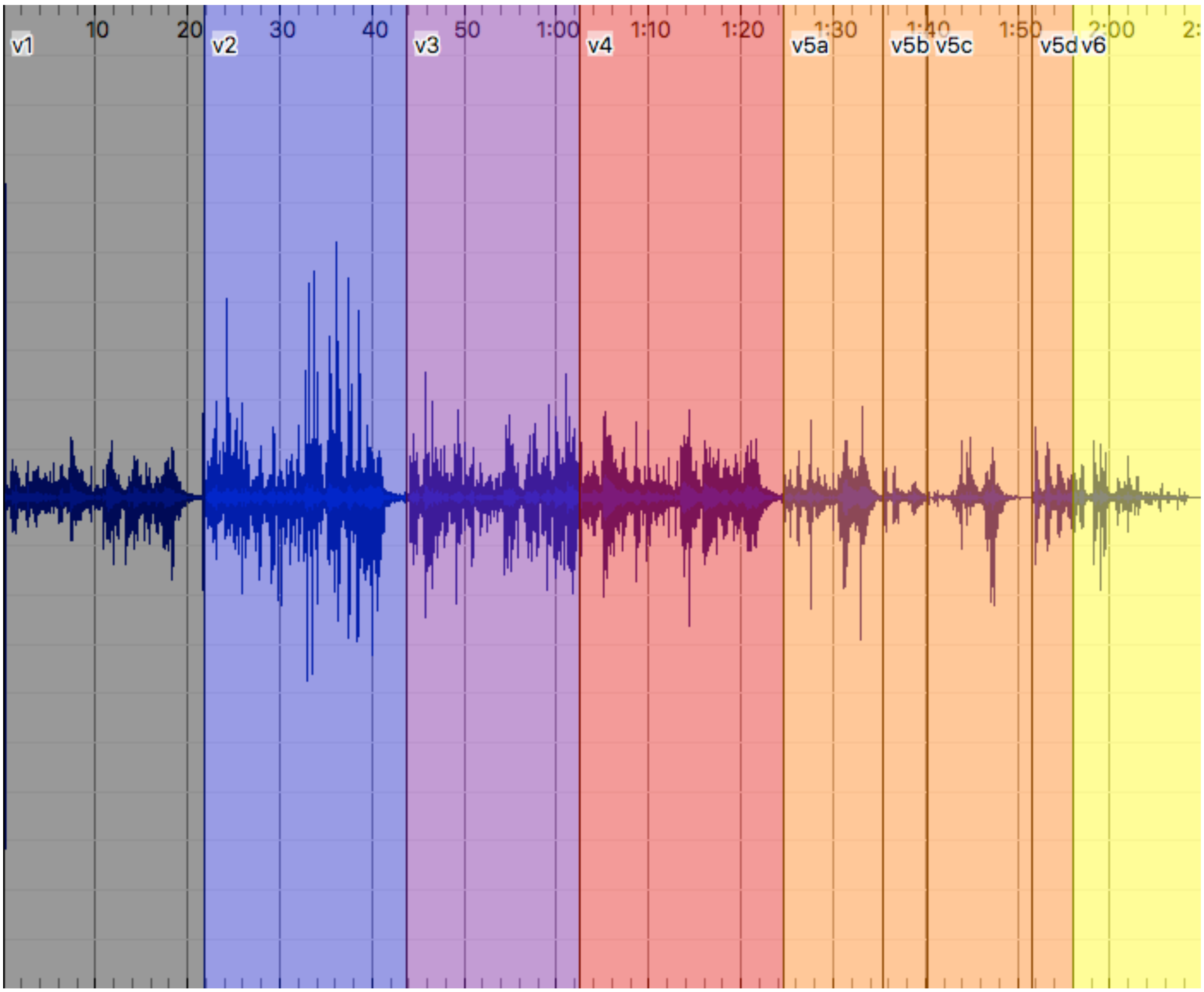

Figure 6. moDernisT as waveform

The lyrics, somewhat more discernible in this verse than the first, manifest as various short phrases occasionally emerging from a granular mix. The opening and closing clauses are particularly audible: the ascending melodic phrase walks up scale degrees 1, 2, 3, and 5 for "It is always...," and the final coiled up phrase passes through scale degrees 5, 6, 5, 5, 4, and 5 for "...before I pour the milk." The melodic content sampled from the original song is clearly displayed (as just described) at the start and finish of the verse, with the opening ascending four note melodic phrase quickly giving way to a scrambled stochastic melody.[55] The verse ends with the final six note cadential phrase clearly voiced. Thus, the overall melodic design is as if the original melody has been sliced into puzzle pieces, with the first and last pieces left in their proper place, and the intervening connective pieces randomly scrambled, re-oriented, and obscured.

Spectrally, this verse, the second of six, contains the greatest spectral flux in the entire composition. As codified by the BBC Research and Development VAMP plugins, the spectral flux measures the change in magnitude in each frequency bin between successive frames, restricted to positive changes and summed across all frequency bins.[56] Measuring this flux using the L1 norm (as opposed to the Euclidean distance) and taking the mean of both left and right audio channels, with a Hann window of size and increment 1024, I detect a peak spectral flux of nearly 470 at 33.7 seconds.[57] In comparison, the preceding verse peaks at 302, and no subsequent verse exceeds a spectral flux of 209, with most frames measuring in the single digits. Taking a wider view, the first verse has low to moderate spectral flux values, the second verse has moderate to high values, and the spectral flux gradually diminishes to low values over the remainder of the composition. This spectral design is not shared by the original

55 For a discussion of the aesthetics of indeterminacy, see Hoogerwerf, Frank W. "Cage Contra Stravinsky, or Delineating the Aleatory Aesthetic." *International Review of the Aesthetics and Sociology of Music* (1976): 235-247.
56 Baume, Chris and Yves Raimond. BBC Vamp plugin collection. https://github.com/bbc/bbc-vamp-plugins
57 Following empirical testing by Dixon in Dixon, Simon. "Simple spectrum-based onset detection." *MIREX 2006* (2006): 62.

song, which has its peak flux near the golden section, roughly two thirds of the way through the composition.[58]

Space is also treated differently in the second verse, particularly through reverberation and panning. A virtual space with multiple planes is created, corresponding with distinct fictional spaces suggested by the lyrics. The new layer, layer five, runs through a space simulation reverb based on a 6.759 second sampled impulse response suggesting the narrator's withdrawal into her own personal space. The dry signal is at unity gain and the reverb output sits at 8db below unity. The reverb output has a 0.68 second attack time and a 6.08 second exponential decay time.

And I'm pretending not to see them...

In addition, layer three now runs through a compressor and into its own space simulation reverb, whereas during verse one the signal was completely dry. This correlates with the activity the narrator observes in the diner. The compressor is a digital emulation of the Focusrite Red 3. The compression ratio is 8 to 1, with the attack time set fast at 9.5ms and the release moderate at 97ms. The compressor threshold is set for a gain reduction of approximately -3db. Coming out of the compressor, the signal passes through another space simulation reverb with a 0.546 second sampled impulse response taken from a restaurant. The reverb output is set at -10db below the dry output which is at unity gain. The reverb output has a volume envelope with a very fast 0.05 second attack time and a 0.49 second exponential decay to 0.

58 Krimphoff finds spectral flux to be one of three dimensions (along with spectral centroid and rise time) correlating with the multidimensional perception of timbre. As cited in Wu, Bin, Andrew Horner, and Chung Lee. "Musical timbre and emotion: The identification of salient timbral features in sustained musical instrument tones equalized in attack time and spectral centroid." In *ICMC*. 2014.

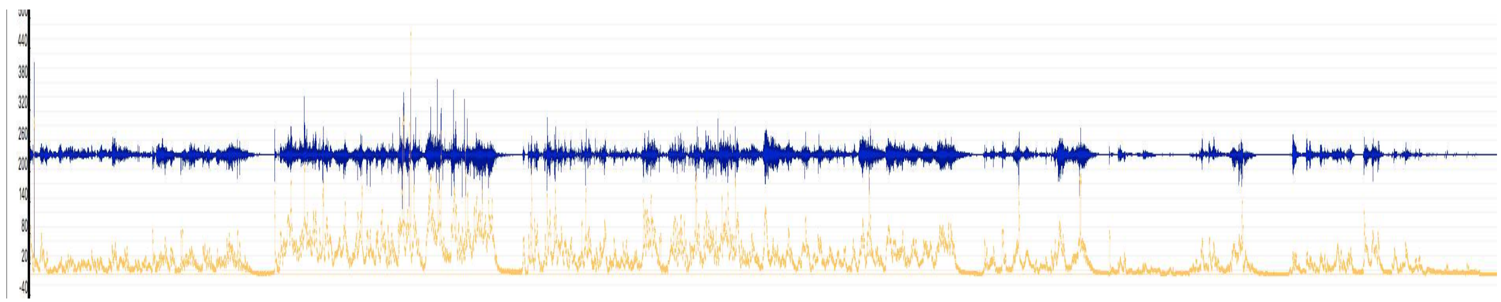

Figure 7. moDernisT spectral flux

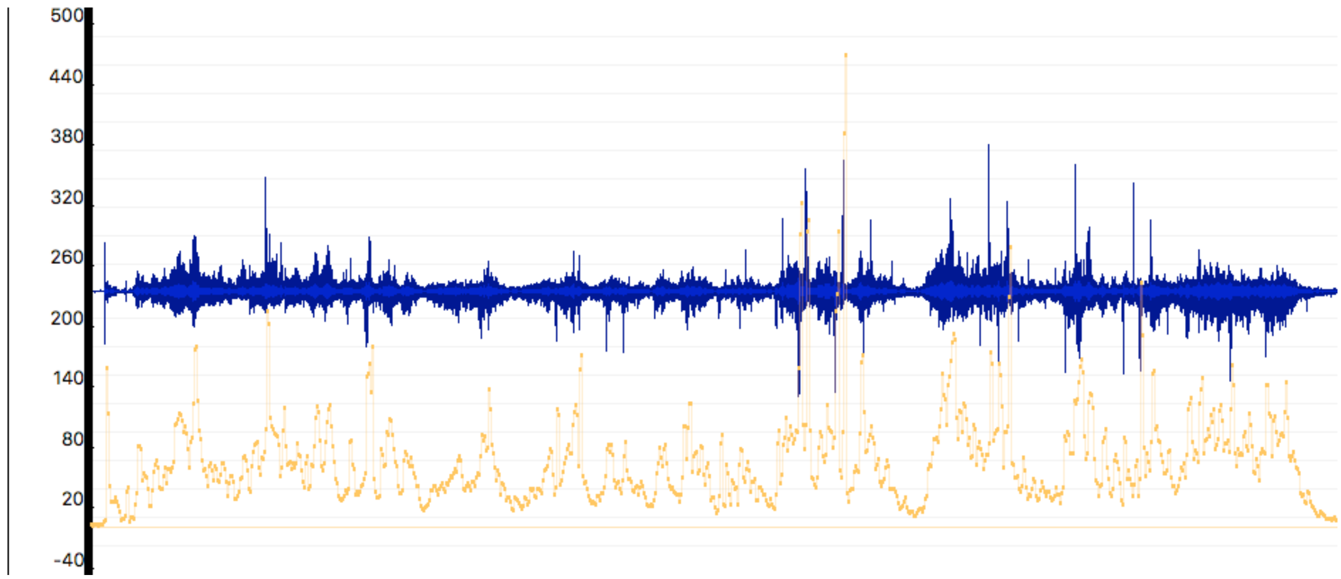

Figure 8. Spectral flux during verse two

"It is always nice to see you," says the man behind the counter...

Layers one and four, rather than featuring reverberation, contribute to the new spatial character of the verse through panning. Layer one, which had been panning subtle back and forth plus or minus 15% left and right during the first verse returns to the center and remains static. Layer four takes up the spatial movement instead, representing the uncomfortable reaction the narrator has to the bustle of activity in the diner. Whereas during verse one, layer four had remained firmly in the center of the stereo field, it now shifts slowly to the left ultimately achieving a 42% displacement, before moving back to the center and ultimately 44% to the right. It then slowly moves back to the left, turning around at 29% left, returning to the right before turning around at 30% right and ultimately settling in the center of the stereo field as the reverb tail fades out.

She is shaking her umbrella and I look the other way...

All of these compositional effects combine to create a surrealistic simulated space which is particularly lucid when the composition is listened to privately via headphones.[59]

> "With the sound system bodies are placed inside sound, whereas with earphone listening it's the opposite, sound is placed inside bodies."[60]

BRK

Sound is, by its very nature, a dynamic and vibratory process, one that can never be fully expressed outside of its embodiment in a physical medium. Owing to this embodiment, sound itself is never simply a medium for encoding meaning. It has "direct sensorial affects and effects, as with smells, tastes, and gestures." Answering Cavarero's call for an ontology of uniqueness, Julian Henriques notes that sonic events are "entirely distinctive, unique and unrepeatable." While text draws the mind to certain conventional objects of inquiry, sound is not so convincingly bound up with language, symbol, and notation. Sound is not properly an object at all, but a process or event. The apparent thing "sound" is a figment, "an abstraction of the action, whose reality vanishes as soon as we examine it at all closely."[61] In the digital, we find all phenomena reduced to symbol by way of the bit. This abstraction makes

59 In addition to subtly transforming the timbre of recorded music, the MP3 has made private music listening increasingly mobile. Personal listening experiences increasingly occur in public spaces, i.e., earbuds on the train, rather than in the confines of the home. For an elucidation of some possibilities for musicians interested in composing specifically for headphone listening, see Begault, Durand R., and Larry Polansky. "The composition of auditory space: recent developments in headphone music." *Leonardo* 23, no. 1 (1990): 45-52.

60 Henriques, Julian. *Sonic bodies: Reggae sound systems, performance techniques, and ways of knowing.* Bloomsbury Publishing USA, 2011.

61 Small, Christopher. Musicking: The meanings of performing and listening. Wesleyan University Press, 2011.

sound appear as if it were a contingent, infinitely reconfigurable thing and in doing so produces a logic unfamiliar to empirical sensibility.[62]

Verse Three

As the reverb tail fades out, gentle tape noise slowly fades in, culminating in the sharp inhalation which launches the third verse. The texture again mutates, the spacious reverb collapsing inward and the granular texture becoming increasingly particulate to the point of clutter. This is a sonic translation of the original lyrics in which the narrator, after describing the morning scene in a diner during the first two verses, buries her head in a newspaper and becomes absorbed in a more personal, private experience.

I open up the paper
There's a story of an actor
Who had died while he was drinking
It was no one I had heard of...

The text painting is carried out first and foremost by the fifth layer. This layer emerges from the background fray with a seven decibel gain increase overall. It is split into two parts, sublayer 5a and sublayer 5b. Both are relatively dry, with the reverb output 48 decibels below the dry output for sublayer 5a, and sublayer 5b bypassing reverb entirely. With this layer given new prominence in amplitude, virtual proximity, and textural thickness, the maximum granular segment size is decreased dramatically for sublayer 5a, from 500 milliseconds to only 75 milliseconds. Sublayer 5b holds steady at 500 milliseconds. The 75 millisecond segment size limit for sublayer 5a is the smallest yet encountered,

62 Evens, Aden. *Logic of the Digital.* Bloomsbury Publishing, 2015.

37.5% of the previous smallest segment size limit which occurred during verse one. The lyrical melody is thus transformed into an indeterminate cloud of vocal timbres.[63]

> WORDS   Discern too concretely to leave room for the mind.
> Forget the true measures of expression: suggestions.
> Let infrarealities disappear.
> Sift without restoring.[64]

Over the final eight seconds of this nineteen second verse, the reverb output for sublayer 5a increases linearly by 24 decibels, culminating in a clearly audible reverb tail sounding out the virtual space. This gradual expansion of virtual space aligns with the text of the lyrics. Over the final eight seconds, the narrator gradually becomes aware of someone watching her as she pages through the newspaper, ultimately raising her head at the final moment: "*And I'm turning to the horoscope / and looking for the funnies / when I'm feeling someone watching me / and so I raise my head*." Here, the lyrics have been translated into a reverberant sonic gesture: just as the narrator raises her head, turning her attention to her surroundings, the entire virtual space is made clearly audible.[65]

> "The walls of an electronically created virtual acoustic space can expand or contract and assume new angles or virtual surfaces. The resulting resonances and reflections, which change continuously during the course of a performance, can be used to create spatial progressions, much as one creates chord progressions or timbre transformations by changing the tone quality of an instrument while performing a single pitch."

63 This technique, *dissolution of semantic meaning through deconstruction*, is discussed in Lane, Cathy. "Voices from the Past: compositional approaches to using recorded speech." *Organised Sound* 11, no. 1 (2006): 3-11. The idea also calls to mind asemic writing, a wordless, open semantic form of writing. Gaze, Tim, and Michael Jacobson. *An Anthology of Asemic Handwriting*. Punctum Books, 2013.

64 Isou, Isidore. "Manifesto of Letterist Poetry." *Manifesto: A Century of Isms* (1983): 545-6.

65 Oliveros, Pauline. "Acoustic and virtual space as a dynamic element of music." *Leonardo Music Journal* 5, no. 1 (1995): 19-22.

Layers one and four support the text as well; layers two and three have been removed. The present layers shift to a smaller simulated space referencing the narrator's diminished awareness of the room, wherein she concentrates her attention on a constricted physical space while reading the newspaper, only peripherally aware of the activity happening in the surrounding diner.[66] Some description: layer one, derived from frequency domain masking, is split into two sublayers (similarly to layer five). Sublayer 1a retains the same heavy reverb as in previous verses, with a 6.8 second impulse response and the reverb output set at just 8 decibels below the dry output. Unlike previous verses though, the masking threshold is set just below 100%, resulting in a sparse texture with only four distinct sounds clearing the threshold during the verse. This sublayer is complemented by sublayer 1b. Sublayer 1b sits back in the mix at 31 decibels below unity, but is situated closer in virtual space and retains presence by having greater rhythmic density. The masking threshold is set much lower resulting in almost ten times as many rhythmic onsets as sublayer 1a. The reverb for sublayer 1b is constructed from a half second long impulse response recorded at a restaurant rather than the cavernous 6.8 second impulse response used for sublayer 1a.

Layer four drops thirteen decibels from verse two to verse three, placing it in a more supportive role in the overall texture. The maximum segment size remains at 2000 milliseconds but the layer is run through an HRTF Python script which creates an illusion of three dimensional motion around the listener's head. This also has a noticeable effect on the timbre—midrange frequencies are relatively well preserved, but low frequencies below 200 hertz and high frequencies near 20k are attenuated.[67] The sense of three-dimensional space created by the head related transfer function is distorted by linear

66 Studies demonstrating attention-dependent changes in the spectro-temporal receptive fields of the auditory cortex are reviewed in Fritz, J.B., Mounya Elhilali, Stephen V. David, and Shihab A. Shamma. "Auditory attention—focusing the searchlight on sound." *Current opinion in neurobiology* 17, no. 4 (2007): 437-455.

67 Recent research has sought to diminish the unintended timbral coloration produced by HRTFs in binaural synthesis. For example, see Merimaa, Juha. "Modification of HRTF filters to reduce timbral effects in binaural synthesis." In *Audio Engineering Society Convention* 127. Audio Engineering Society, 2009.

shifts in the balance between the left and right stereo channels. The verse begins with both channels even, before shifting approximately 20% to the right, then left, then right, then left again, before returning to an equal balance at the end of the verse. This subtle spatial distortion creates a somewhat disorienting effect, making it difficult to track the three-dimensional trajectory created by the HRTF. Cumulatively, and in relation to the previous verse, these effects subdue this layer relative to the overall texture.

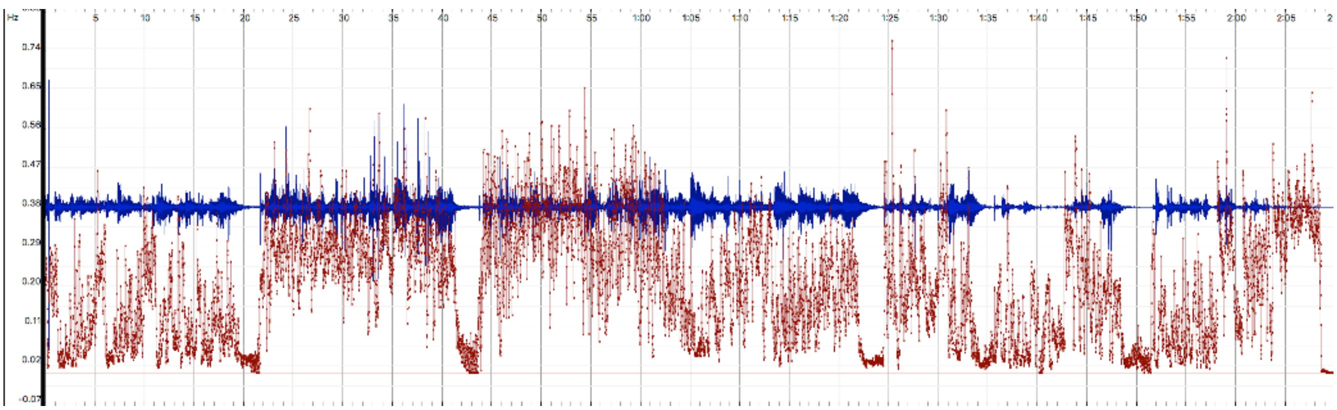

Figure 9. moDernisT zero-crossing rate

Now taken as a whole, verse three is quieter than the previous verse and yet features noisy spectra more prominently, resulting in a surface texture with greater temporal detail. This is reflected in the zero-crossing rate, which peaks during this section. For comparison, the average zero-crossing rate (i.e., the average number of zero crossings per second) for verse one is 0.11 Hz, for verse two it is 0.26 Hz, and for verse three it is 0.32 Hz, the highest average ZCR of any section of the composition. The zero-crossing rate is effective at discriminating whether a given sonic texture is dominated by noise-like spectral elements, like fricatives, or steady tonal elements, like vowels. A higher zero-crossing rate indicates a noisier sonic texture. Speech signals tend to exhibit high ZCR variance as speakers alternate between frication and voiced speech, while music tends to exhibit a relatively stable ZCR. Under this interpretation of zero-crossing rate, *moDernisT* is spectrally closer to speech than music, and, as the ZCR increases from verse one to verse three, increasingly fricatival.[68]

"Siri Suri sei []!"[69]

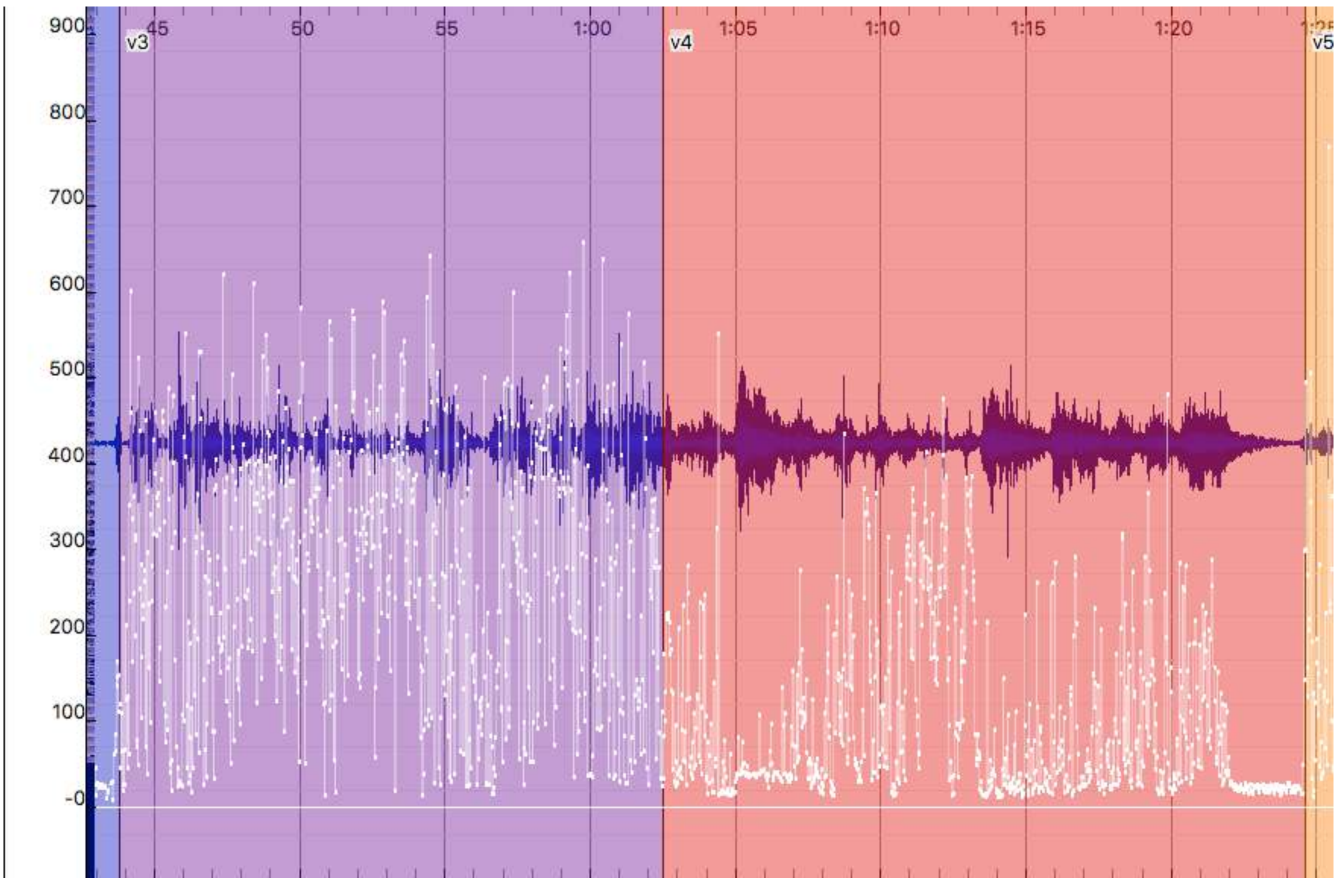

Figure 10. Inharmonicity (white) in verses three and four

68 Consonants have been shown to carry symbolic properties. In English, fricatives are often judged as relatively harsh, rough, active, sharp, difficult, and angular. Greenberg, Joseph H., and James J. Jenkins. "Studies in the psychological correlates of the sound system of American English." *Word* 20, no. 2 (1964): 157-177.

69 Fricative-laden excerpt from an early sound poem (1905!). Morgenstern, Christian, Anne Gabrisch, and Horst Hussel. *Galgenlieder.* B. Cassirer, 1921.

Verse Four

After verse three, verse four begins immediately, without pause or fermata, just as in the original song, creating a compound structure. Reverberation predominates and the micro-segmentation of the preceding verse disappears. The overall spectrum is more harmonic here, returning to a similar level of harmonicity as the first verse. These changes help to differentiate between verse three and verse four even as the rhythmic structure of the song elides them.

There's a woman
on the outside
Looking inside
does she see me?

Layer one is more reverberant than at any previous point in the composition, corresponding with the shift from interior to exterior as the narrator turns her attention to a woman outside of the cafe. It is again split into two sublayers, as in verse three. The first sublayer, 1a, has a slightly lower masking threshold than it had in verse three, but still results in a sparse texture compared to the first two verses. The reverb uses the same 6.8 second impulse response for both sublayers. In sublayer 1a, the reverb output is set at -2 dB while the dry output is at -7.8 dB. This is the first time in the composition that a layer has had a reverb output exceeding the dry output. This creates a distancing effect, making the sublayer sound as if it is located some moderate distance away from the listener.

Sublayer 1b charts a dynamic course through virtual space. Mixed 3.8 dB quieter than sublayer 1a at -22 dB, it has a fixed reverb output at -8 dB and a variable dry output. The dry output begins the verse at -10.5 dB and ends the verse at -7 dB, having interpolated between these two values linearly over the time span of the verse. Thus, sublayer 1b begins the verse at a moderate virtual distance from the listener, with the reverb output exceeding the dry output, and ends the verse having moved steadily nearer to the listener, with the dry output now exceeding the reverb output by the end.

Partially inverting the layers present in verse three, verse four features layers 2 and 3, while layers 4 and 5 are absent. Notably, words and even short phrases are now intelligible. Layer 2 is split into two sublayers, 2a and 2b. As before, layer two is derived from a 64 kbps MP3 inverted using a null test.

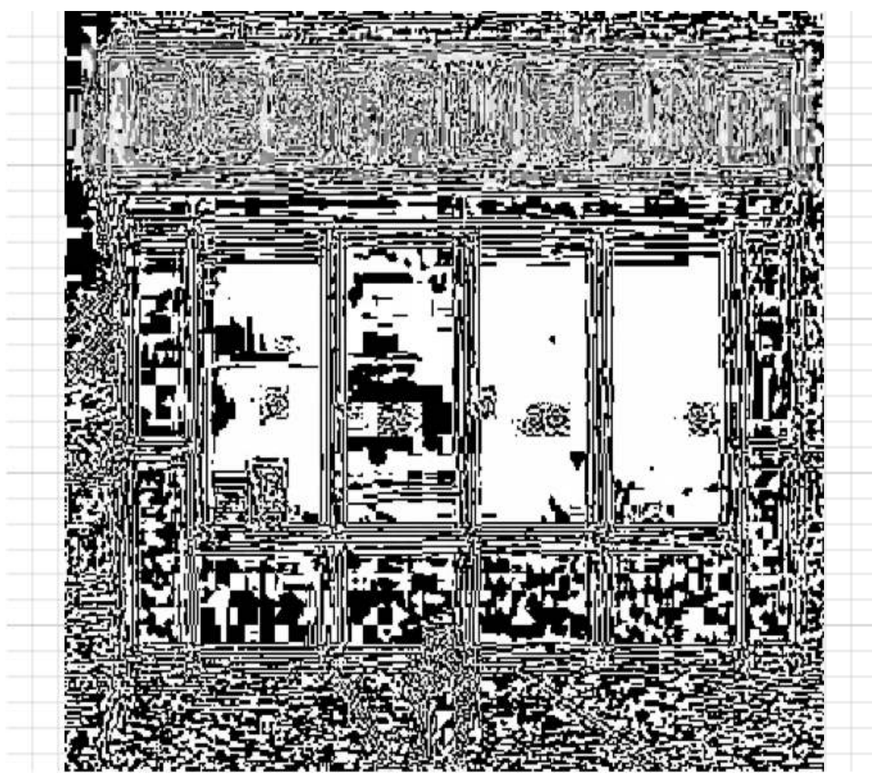

Figure 11. The windows at Tom's Restaurant (JPEG detritus)

Sublayer 2a is scrambled with a 2000 millisecond maximum segmentation size, rendering full words audible. It is mixed at -12.4 dB, roughly even with layer 1. Creating an overlay of virtual spaces, the sublayer is run through a reverb using the same 0.5 second restaurant impulse response as before, the reverb output level set at -10 dB and the dry output level set at unity gain. Sublayer 2b is set at -12.4 dB, but once the output is run through a compressor peaks just below unity gain, placing it in the foreground, ahead of both layer 1 and sublayer 2a. The maximum segment size for sublayer 2b is the longest yet encountered—3000 milliseconds. This results in full phrases being audible during the verse.

…she's straightening…
…does she see me?
her hair is…
her hair is…

The compressor is the same as used in layer three previously—a digital simulation of the Focusrite Red 3. This noticeably sharpens the sound, creating a rhythmic, "punchy" effect overall and a heightened presence for the decaying tail of each syllable. The input threshold is set at -20.5 dB, resulting in gain reductions hovering around 3 dB; however, an automatic make-up gain is applied to the output, normalizing the signal to just below unity. As before, the compression ratio is 8 to 1, the attack is set fast at 9.5ms and the release is moderate at 97 milliseconds. The output of the compressor is fed into the restaurant simulation reverb, the reverb output 10 decibels below the dry output, set to unity gain.

Layer three features the identical 3000 millisecond maximum segment size as sublayer 2b. In contrast however, it is derived from a 320 kbps MP3—the highest quality encoding possible. At this quality level, less information is deleted during the compression stage; as such, the recovered detritus isn't as dense as in layer two. Additionally, layer three is spatialized using the Python HRTF script to create 3D motion, inadvertently softening the timbre as well. The layer is set at -16 dB, somewhat in the background, and also passes through a large reverb with a 6.8 second impulse response. The reverb output is set at -24 dB while the dry output is at 0 dB, creating the perception that the sound source is close to the listener, but situated in a large reverberant space. The verse concludes with a significant fermata, during which the reverb tails come into the foreground with clearly resonant pitch, driving the inharmonicity to near zero.

Her hair has gotten wet.

BCKSTRY

In 1990, an audio engineer with the Fraunhofer Institute named Karlheinz Brandenburg was fine-tuning the MP3 compression algorithm in the Bavarian city Erlangen when he heard a voice singing somewhere down the corridor. He was "electrified."[70] The voice was of American musician Suzanne Vega, singing the song "Tom's Diner," airing on German radio that day. Brandenburg was jolted by the intuition that "it would be nearly impossible to compress this warm a capella voice." Over the ensuing months he encoded the song thousands of times, listening and then modifying the algorithm over and over again. Hilmar Schmundt, writing in *Business 2.0*, suggests, "When an MP3 player compresses music by anyone from Courtney Love to Kenny G, it is replicating the way that Brandenburg heard Suzanne Vega."

As Brandenburg's development continued, an English computer scientist named Tim Berners-Lee was

70 Vega, Suzanne. "Tom's Essay." Measure for Measure. New York Times, 23 September 2008.

working on a seemingly unrelated project at CERN. On August 2nd, 1991, a message went out on the alt.hypertext newsgroup asking, "Is anyone reading this newsgroup aware of research or development efforts in ... Hypertext links enabling retrieval from multiple heterogenous sources of information?"[71] No one responded for three full days. Finally, on August 6th, Tim Berners-Lee replied from his office in France, "The WorldWideWeb (WWW) project aims to allow links to be made to any information anywhere," an audacious claim marking the first public announcement of the fledgling Internet technology. He continues, "The WWW project was started to allow high energy physicists to share data, news, and documentation. We are very interested in spreading the web to other areas, and having gateway servers for other data. Collaborators welcome!" A typical silence fell over the newsgroup again until ten days later, when Berners-Lee announced bug fixes and included version 0.12 of a NeXTStep hypertext browser and editor. He signed off with an apology for "clogging up the news."

A month later, on September 24, 1991, A&M Records released *Tom's Album*, a full length album of remixes of the song "Tom's Diner."[72] This was the unlikely outcome of a situation that had seemed destined for litigation just months earlier, when an unauthorized version of Suzanne Vega's moderately popular alt-folk tune had been reworked by British remixers DNA, prompting skepticism from A&M's legal team. But, at Suzanne Vega's insistence, the label not only decided not to sue, but instead promoted DNA's track along with 10 other remixes. Whereas the initial, Kurt Weill-inspired a capella had gained passing commercial attention, the DNA remake, with its hip hop-inspired rhythm track punctuated by gritty brass samples, became the highest charting song of Vega's career to date—reaching the top 20 in over a dozen countries, number 1 in Germany, Austria, and Switzerland, number 2 in the UK, and number 5 in the US.[73]

71 Google Groups. Accessed March 18, 2018. https://groups.google.com/forum/#!msg/alt.hypertext/eCTkkOoWTAY/bJGhZyooXzkJ.
72 In 1989, A&M Records was acquired by Polygram, a subsidiary of Phillips, co-creators of the compact disc.
73 Richardson, John. "Vega Meets Weill: Alienation Effects and Archaeological Strategies in 'Tom's Diner' and 'Luka'."

That November, Brandenburg's working group, an international communications industry consortium, would announce the finalization of the MPEG-1 layer I, II, and III codecs. Layer III, the MP3, was supported by Fraunhofer, Thomson, and AT&T. The competing layer II, a higher-fidelity standard, was backed by Panasonic and Philips, the parent company of A&M Records. In a press release, the MPEG Working Group wrote, "The storage media targeted by MPEG include CD-ROM, DAT, and computer disks and it is expected that MPEG-based technologies will eventually be used in a variety of communication channels such as ISDN and local area networks and even in broadcasting applications."[74] Shortly thereafter, Layer II was chosen over Layer III as the standard for two formidable emerging digital audio applications: audio on video compact discs and stereo digital audio broadcast, including satellite radio, terrestrial digital radio, and satellite TV.[75] It seemed as though the MP3 might end up being nothing more than a footnote.

With the Web still in its formative stages, it was difficult to foresee the coming sea change. While Philips was securing the traditional broadcast markets, Fraunhofer looked for new avenues to promote their standard. In July 1994, they released two command-line programs, L3Enc and L3Dec, that would allow users to encode and decode stereo audio files. In keeping with Tim Berners-Lee's initial credo for the WWW that "information should be freely available to anyone,"[76] an Australian hacker promptly stole and reverse-engineered the programs, re-wrote the GUIs, and released them for free under the name "thank you Fraunhofer."[77] Amazingly, this violation turned out to be Fraunhofer's redemption.

Unpublished draft provided to author.

74 Adler, Mark. "MPEG Press Release." Email. 1991. Retrieved from http://web.mit.edu/graphics/doc/ISO-11172-summary

75 Sterne, Jonathan. *MP3: The meaning of a format*. Duke University Press, 2012.

76 Chiariglione, Leonardo. "MPEG Press Release." 1991. Retrieved from https://web.archive.org/web/20110503174827/http://mpeg.chiariglione.org/meetings/kurihama91/kurihama_press.htm. Accessed March 18, 2018.

77 Denegri-Knott, Janice, and Mark Tadajewski. "The emergence of MP3 technology." Journal of Historical Research in Marketing 2, no. 4 (2010): 397-425.

The format, now freely available to tech-savvy early adopters, would spread quickly from hacker circles to college campuses in the mid-90's.[78] Optimized for fast download speeds, the MP3 rapidly blossomed into the primary attraction on the Internet—in 1999, "MP3" was the most popular search term on the Internet, surpassing pornography terms. That August, the file-sharing server Napster was released. Over 28 million people downloaded the software, making the MP3s on their hard drives accessible to millions of other users around the globe. The music industry was shellshocked and in 2001 A&M Records, the former Philips subsidiary and Suzanne Vega's record label, sued Napster, ultimately bringing them down in a landmark case in intellectual property history.[79]

Verse Five

The next song section, seemingly verse five, breaks up the strophic song structure. It is the crux of the piece. This fifth section is constructed from four subsections with three fermatas in the inter-section breaks. In other words, the first three subsections each end with a fermata. These subsections settle on the dominant, subdominant, and mediant respectively, creating a broad 5-4-3 melodic line across the phrases, an expansion of the motivic walkdown paired with the lyrics "*the diner*" in the opening verse. My treatment maintains this structure while altering various inner details.

78 Perhaps seeing the troubled waters ahead, Philips sold off A&M Records with PolyGram in 1998.

79 Witt, Stephen. *How music got free: The end of an industry, the turn of the century, and the patient zero of piracy*. Penguin, 2015.

Subsection 5a features the head related transfer function prominently. It is comprised of layers one, two, and three, continuing a texture similar to that produced in verse four. From this noise-like texture a single word emerges, drenched in reverb and suggesting the narrator's experience listening to distant bells. Layer one is derived from frequency domain masking, while layer two is derived from a 64kbps MP3 null test and layer three from a 320kbps MP3 null test. I will now detail each layer in turn.

Layer one establishes continuity between verses four and five. It is a connective thread split into two sublayers, 1a and 1b. Sublayer 1a features a single bell-like tone, a slightly flat E, corresponding to the subtonic. It is mixed at -18.2 dB, with a reverb output of -2dB and a dry output of -7.8dB. Sublayer 1b features two melodic skips between the mediant and the dominant interrupted by a silence (filled by the E in sublayer 1a, completing a III chord) which is followed by a final skip from the mediant to a sustained C-sharp—the dominant. The sublayer is initially mixed at -12.4dB before fading to silence in the middle of the subsection, and then returning to -12.4dB at the end, via linear fades. The signal here is run through a subtle frequency shifter—creating a light quivering effect—before passing into the space simulation reverb. The frequency shift is set at 3.706 Hz with the LFO at 3.33 Hz and the envelope follower on. Frequency shifted output is set at 79% and the dry signal at 21%. This signal is sent out through a space simulation reverb using an impulse response of 3.37 seconds recorded in a cathedral, supporting the corresponding lyrics: "*Oh, this rain it will continue, through the morning as I'm listening, to the bells of the cathedral...*" The dry output is set at unity gain with the reverb output set at -10dB. Both sublayers 1a and 1b are spatialized using the HRTF Python script previously described.

> "Consider the echoes attached to a liturgy in a cathedral ... Space … provides a great deal of contextual and allusive meaning even though these meanings change over time, with context, and according to the people listening."[80]

Layer two is also constructed from two sublayers. Sublayer 2a has a maximum segmentation size of 800 milliseconds, allowing at most for syllables and words to come through. The word "morning" is discernible twice. The signal is spatialized using the HRTF, and seems to swirl around the listener's head. It is run through the same 6.8 second reverb used previously, the reverb output set at -24dB and the dry output at unity gain. Sublayer 2b has a maximum segmentation size of 4000 milliseconds, the largest segmentation size in the composition. The phrase "to the bells of the cathedral..." is clearly audible. This sublayer is, in another instance of sound/text painting, passed through a space simulation reverb with an impulse response taken from a cathedral. The reverb signal is automated carefully so that only the word "cathedral" is sent into the reverb. The dry output remains constant at unity gain. This is the only sublayer of the subsection not processed using a head related transfer function. Owing to this, it has the broadest frequency spectrum in the subsection. Finally, layer three is scrambled using a 1000 millisecond maximum segmentation size. It is processed using the HRTF Python script and mixed at 20dB below unity gain, placing it in the background.

5b

I am thinking of your voice...

Subsection 5b is the shortest segment in the entire composition, a single melodic phrase. It also exhibits the lowest peak RMS energy in the song, calculated as a rolling average over three second intervals,

80 Sterne, Jonathan. "Space within space: Artificial reverb and the detachable echo." *Grey Room* (2015): 110-131.

with intermittent values measured using a window size of 2048 samples and a 1024 sample increment. In other words, it is both quiet and fleeting. The subsection is comprised of three layers: layers one, two, and three. In contrast to subsection 5a, it has no sublayers.

I imagine this section as a distant memory, almost a daydream. Layer one was created by masking the uncompressed WAV with the MP3 in the frequency domain, then converting this difference to MP3 and masking it recursively nine times. This process returns fewer artifacts with each pass, like an aging recollection, increasingly remote. This signal is run through a 6.8 second reverb, the reverb output set 26dB below the dry signal, which is at unity gain. Layer two derives from a null test of a 64kbps MP3, scrambled with a maximum segment size of 1000 milliseconds, enough for short words to emerge (e.g., *voice*). The output is run through a semi-randomized HRTF script and into a 6.8 second reverb. The reverb output is set at 24 decibels below unity, dry output at unity gain. The overall track is mixed 16 decibels below unity. Layer three derives from a null test of a 320kbps constant bit rate MP3, scrambled with a maximum segmentation size of 1000 milliseconds. The signal is processed using a digital compressor at 8:1, with a fast attack and moderate release, and then processed using a 6.8 second reverb, with the reverb output 10 decibels below the dry output. The whole is mixed quiet at -25.6 decibels. The final sibilant in the word *voice* is released at the end of the phrase.[81]

5c

...and of the midnight picnic
once upon a time
before the rain began...

81 A voice*less* alveolar sibilant, ironically.

Subsection 5c again contains three layers, continuing the memory from 5b. Layers one and two both have two sublayers. Sublayer 1a contains one note—scale degree three—on which the phrase ends. It is derived from a frequency domain mask, resulting in relatively clear tones, and located in binaural space using an HRTF. The signal is mostly dry, with the 6.8 second reverb set 26 decibels below the dry signal at unity gain. The sublayer is mixed low at -18dB. Sublayer 1b derives from a frequency mask performed recursively nine times. The track begins dal niente, fading up to -22dB. The 6.8 second reverb output begins at -18dB and interpolates linearly up to -8dB for the final note. The dry level begins at -3.4dB and interpolates linearly downwards to -10.5dB for the final note. Thus, the dry and wet levels cross just before the final note, creating the illusion of a sound source moving away from the listener in a large space. Correspondingly, the narrator trails off in recollection of "the time before the rain began."

midnight picnic once before

Sublayers 2a, 2b, and layer 3 are scrambled with a maximum segmentation size of 1000 milliseconds. For sublayer 2a, the 6.8 second reverb is set at -15.5dB. Sublayer 2b is completely dry, and is given a spatial trajectory using the HRTF script. Sublayer 2a is mixed at -16dB and sublayer 2b, the loudest sublayer here, is mixed at -10dB. Layer 3 is processed using the HRTF script, passed through a digital compressor with an 8:1 ratio, and run through a half-second reverb with an impulse response taken from a restaurant, the output set at 10 decibels below unity. This layer is mixed very quiet at -37.6dB, filling in the gaps between words with faint verbal fragments.[82]

82 "The pleasure of vaporwave is therefore understood as a pleasure of remembering for the sake of the act of remembering itself." Glitsos, Laura. "Vaporwave, or music optimised for abandoned malls." *Popular Music* 37, no. 1 (2018): 100-118.

5d

I finish up my coffee
and it's time to catch the train

Subsection 5d concludes section 5 with the narrator abruptly ending her reverie. There are three layers. Layer one, derived from frequency masking, is mixed at -12.4dB. The signal is dry except for the last note, the dominant, which is passed through the 0.5 second restaurant impulse response reverb with a reverb output at -5.8 dB. The timbre is reminiscent of a vocoder, suggesting the digital substratum, and the word "coffee" is almost discernible. Layer two has a maximum segmentation size of 400 milliseconds and is mixed at -25.1dB to start the subsection, linearly increasing to -17.5dB at the end of the section. The *m* in "my", is audible repeatedly, as if the narrator is stuttering, with the full word finally uttered at the end. Finally, layer three has a maximum segmentation size of 500 milliseconds. It also has a rushed, stuttering quality. Beginning at -37.6dB, it increases linearly to -25.6dB at the end of the section. The compression and reverb settings are the same as they were during subsection 5c, placing us in the diner.

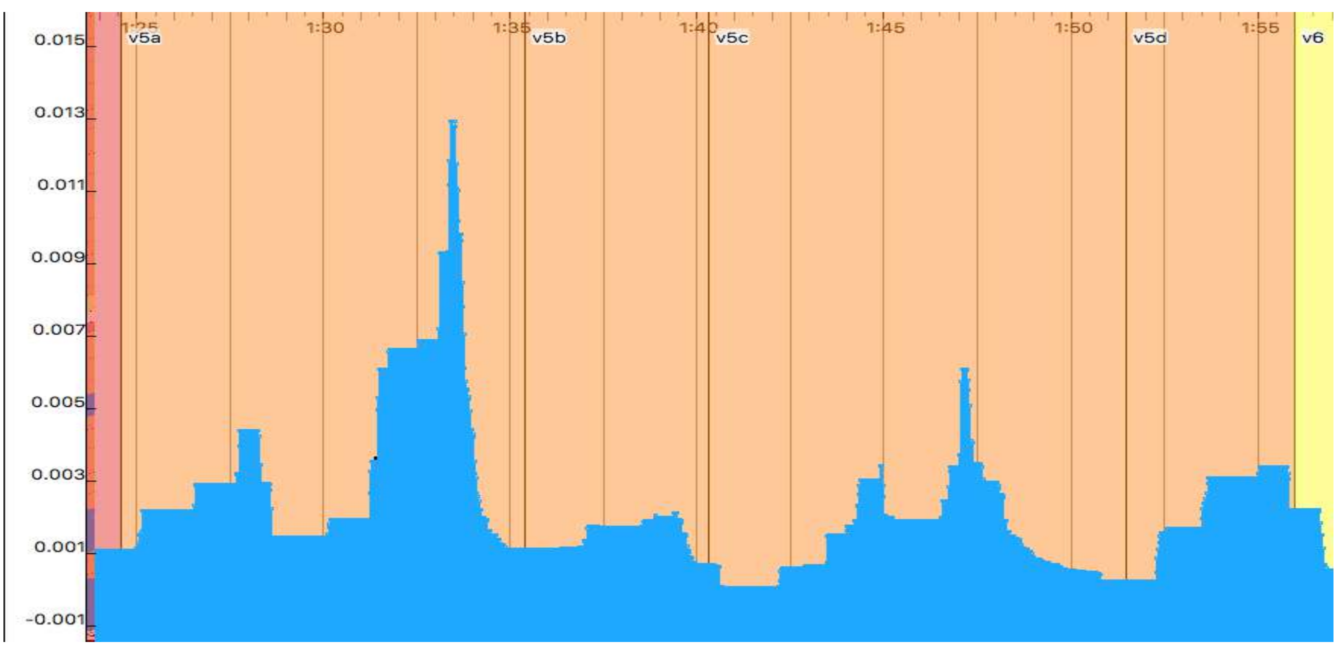

Figure 12. Average RMS Energy (bright blue) for verse five

TCHNCS

The modified discrete cosine transform is an example of a class of transforms called Time Domain Aliasing Cancellation utilized in perceptual audio coders such as the MP3. After mapping a time series audio signal into the frequency domain using the MDCT, a perceptual coder can examine the allocation of the bit pool and delete selected information to achieve the data compression goals common to communications technologies designed for network transmission. Economizing signals by applying perceptual research is a common story in the telecommunications industry, a process referred to by Jonathan Sterne as *perceptual technics*. The creative works in this dissertation invert the technologies developed from such research.

Section Six

Section six, the outro, begins with a breath and is split by another halfway through. Bits of the melody stall and burst forth sporadically, as the narrator recedes into the distance, headed off to catch her train. The section features layers two, four, and a new layer. All layers contribute to the sense of a sound source fading into the distance, last breaths sounding out sporadically, obscured by an intervening digital mist.

dʊ dʊ dʊ dʊ
dʊ dʊ dʊ dʊ
dʊ dʊ dʊ dʊ
dʊ dʊ dʊ dʊ

Layer two is divided into two sublayers, both noisy. Sublayer 2a has a maximum segment size of 1000 milliseconds. It's short, 0.5 second reverb output increases linearly from -10dB to 0dB at the end of the song, supporting the illusion of the sound source receding. The sublayer is mixed at -25.6dB. Sublayer 2b has a maximum segment size of 500 milliseconds. At this length, we mostly hear the *d* consonant at the beginning of each syllable, the sustained *ʊ* abruptly truncated. Its 4.3 second reverb is derived from the impulse response of a concrete room, and the reverb output sits at -10dB. The sublayer is mixed at -13dB.

Layer four, raspy and distorted, based on a variable bit rate encoding, is mixed at -10.1dB. It has a maximum segment size of 1 second and is presented without reverberation. In contrast, the new layer is bell-like, its sine tones ringing out in a reverberant space. Derived from frequency domain masking set with a fairly low threshold, the signal is passed through a 4.4 second reverb simulating a natural outdoor space. This text painting is meant to suggest the narrator heading outdoors at the end of the song, as she journeys off to catch her train. The main melodic phrase, which walks up from the tonic, supertonic, and mediant, to the dominant is foregrounded via volume automation. Subsequent melodic notes are attenuated. Later, a brief descent from the mediant to the supertonic is brought forward before receding again. Finally, the last two notes of the wordless melodic phrase, the mediant and the dominant, are emphasized as the song fades to silence

Figure 13. Tom’s Restaurant (ghost)

Visions

> "I have kept only the images which enthrall me, without my knowing why (such ignorance is the very nature of fascination, and what I shall say about each image will never be anything but . . . imaginary)."[83]

Perceptual coding is not limited to the auditory domain. Algorithms in this class exist for image and video data as well. In the following section, I will detail some of my work with visual compression detritus. These experiments have largely been carried out using the OpenCV computer vision library in Python. I create and use scripts that take the frame-by-frame, pixel-by-pixel differences between uncompressed and lossy compressed versions of identical digital images and video clips. I experiment with different ways of normalizing the data from these scripts until arriving at interesting results. Converting RGB pixel values (red, green, blue) to HSV pixel values (hue, saturation, value) has proven particularly useful.[84]

I have created two different videos to accompany *moDernisT*. Both are hosted for free online viewing on Vimeo.[85] The first is a collaboration with video artist Takahiro Suzuki, with whom I had previously collaborated. Takahiro is an experimental filmmaker who works with found archival footage. After listening to the composition, he created a video sourced from archival analog film footage. The aesthetic impression of his video was of something mysterious, grainy, and somehow tactile, including several shots of insects crawling on a t-shirt seemingly suspended in mid-air. I compressed the video using MPEG-4 part 10, Advanced Video Coding compression and took the frame-by-frame, pixel-by-

---

83 Barthes, Roland. *Roland Barthes by Roland Barthes*. Macmillan, 2010.

84 HSV and HSL (hue, saturation, lightness) are alternative representations of RGB color space that align more closely to the measured human perception of color values, making them useful for extracting perceptual coding detritus.

85 And thus, they are compressed but mobile, available for streaming. When presenting the work live or in a gallery setting, I show the full uncompressed renderings.

Figure 14. Uncompressed frame from "Tom's Diner" music video

Figure 15. Compressed frame from "Tom's Diner" music video

Figure 16. Zoomed in view of uncompressed frame (bottom right corner)

Figure 17. Zoomed in view of compressed frame (bottom right corner)

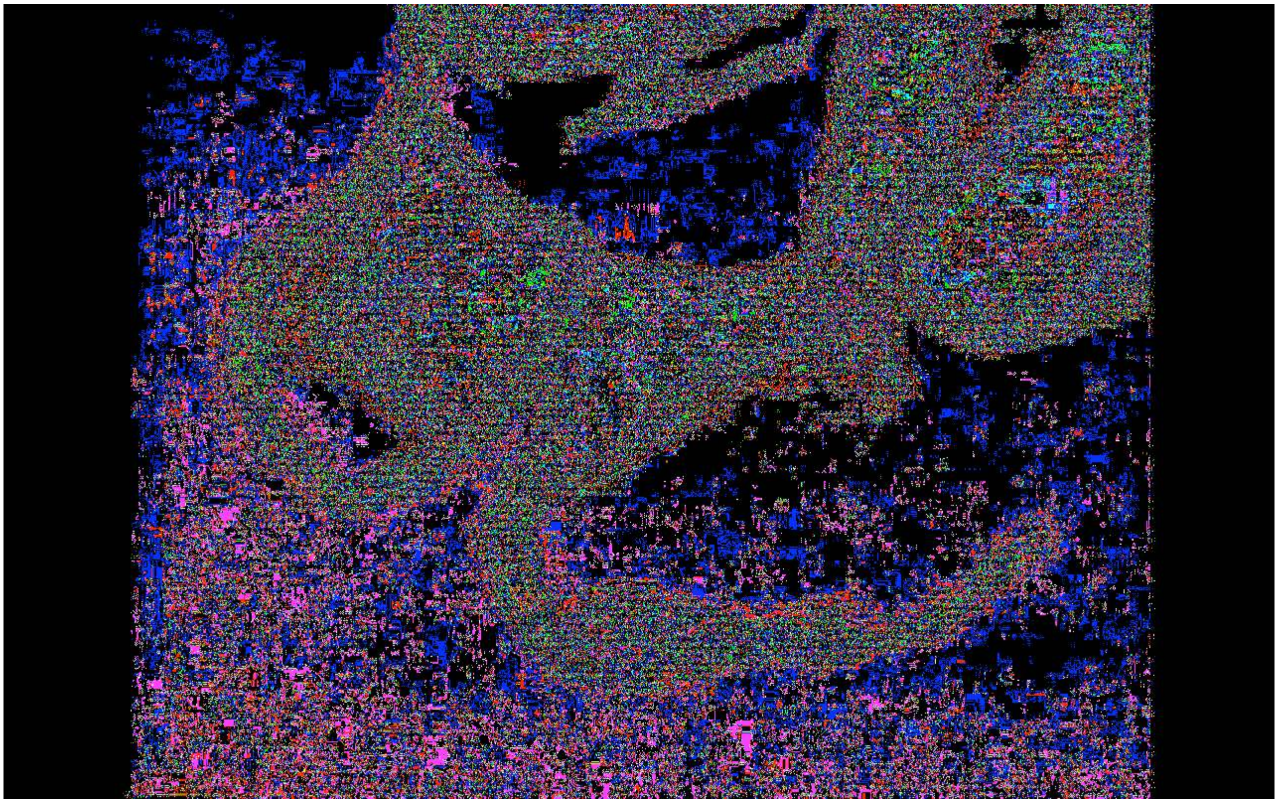

Figure 18. Pixel-by-pixel difference in RGB space

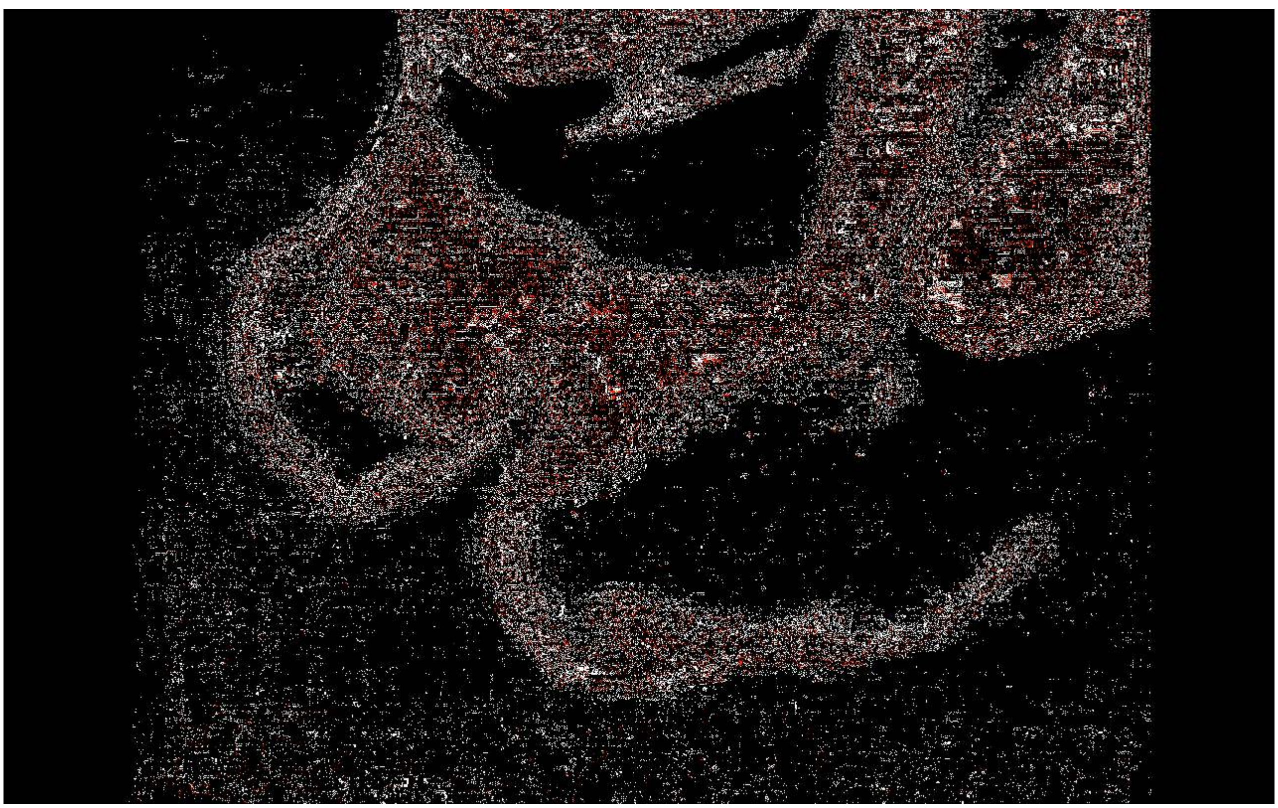

Figure 19. Pixel-by-pixel difference in HSV space

pixel difference between the uncompressed and compressed video using my OpenCV Python scripts. The result is visually compelling. The original grainy film footage is largely reduced to black and white with a few splashes of magenta, textural geometric patterns emerging. The source of the footage and the recognizable images and symbols were even further obscured. I use the resultant video as a visual counterpoint to the recording, *moDernisT*, produced using an analogous algorithm in the audio domain.

After creating this version, I compressed the original "Tom's Diner" music video using H.264 compression and isolated the deleted information using my Python script. The original a capella version of the music video was hard to find—the file I tracked down was a digitized conversion of a VHS recording from an airing of the video on MTV. Thus, the artifact I created contains an accumulation of media-specific noise from analog film and television transmission. The uncompressed video was already low fidelity to begin with—bearing the usual hallmarks of VHS video. MP4 compression adds additional (perceptually discrete) quantization noise. The deleted information is surprisingly colorful compared to the video from Takahiro Suzuki. This process maintains the primary visual structure—boundaries between visual objects, for example—but gives hard lines a restless, noise-like texture. The effect is a kind of shimmering—light playing on the surface of digital water.

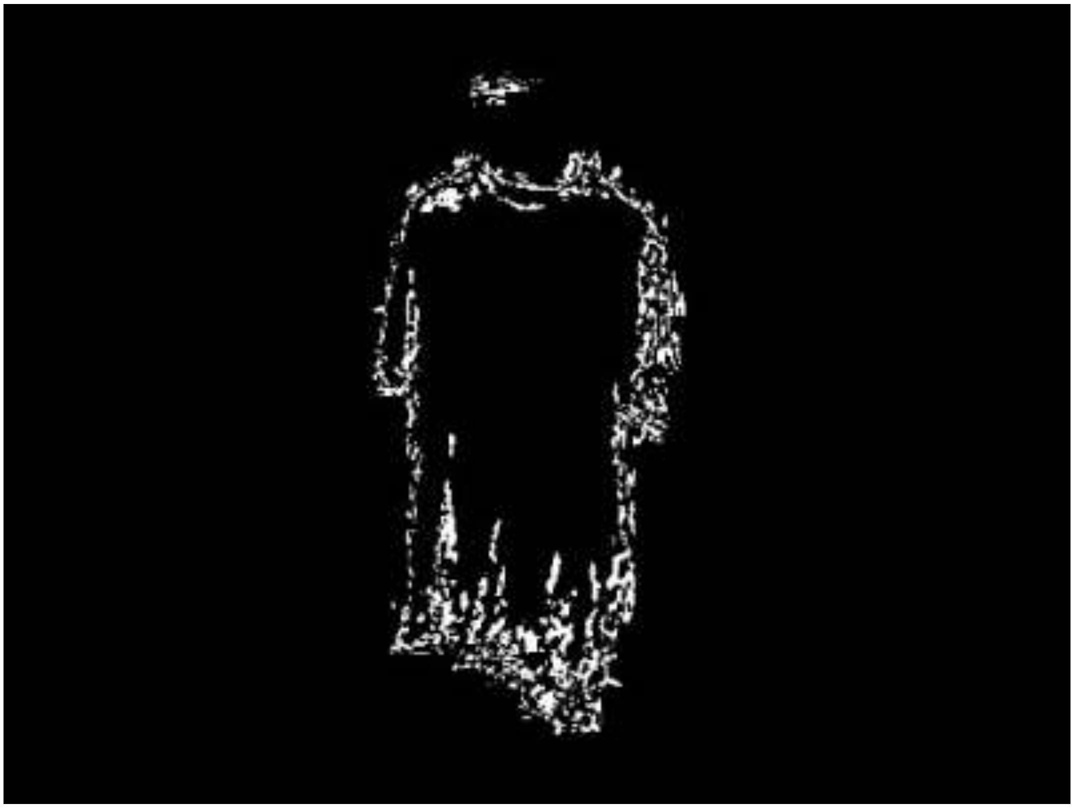

Figure 20. Screenshot from *moDernisT_v1*

Materiality

I will next describe an interactive, handheld object created as an experimental physical format for this composition. I will highlight aesthetic and philosophical themes along the way. The object described proposes a particular technological configuration of music designed to explore the (im)materiality of lossy compression detritus. In doing so, it cracks open the fixed form of the original composition, centers the listening experience in a semi-private space, blurs the boundary between listener and composer, between composer and interaction designer, between composition and improvisation, and between music and media art.[86]

86 These boundaries are all vibrant sites of creative activity in early twenty-first century music and sound practices, as can be discerned by paging through the most recent issue of *The Wire* at any given time. The twenty-four interviews with women in electronic music and sound featured in *Pink Noises* present a rich cross-section of twenty-first century

I have provided a detailed tutorial for constructing this object online using the open source code sharing platform Github.[87] The material object and its technical implementation are thus relatively transparent, while the aesthetics presentation is somewhat opaque. The object is constructed from the following hardware: Raspberry Pi 3, 2.8" PiTFT capacitive touchscreen, ADS1115 16-bit ADC, Adafruit PowerBoost charger, 5V 2.4A Switching Power Supply, Lithium Ion Polymer Battery (3.7v 2500mAh), White LED Backlight Module (23mm x 75mm), Mini Fan, 10k logarithmic rotary potentiometer, volume knob, 32GB SDHC card, USB A to MicroUSB B converter, 1/8th inch headphone extender, LED Arcade Buttons, Metal LED On/Off Switches, breakout breadboards for wire routing. These components are hand-assembled and mounted in a customized 3d printed enclosure.

The enclosure was designed in code using OpenSCAD.[88] OpenSCAD encourages modular design, wherein small parts can be dynamically resized, moved, and assembled into complex designs. I created four modules 1) constructing the main body of the enclosure, 2) arranging the holes for electronic components in the enclosure, 3) adding a surface texture based on extruded lossy image compression detritus, and 4) assembling all of the pieces together. After running these four modules I split the model into two halves, rendered these as STL files, and 3D printed the results. I used a translucent filament in the final design which, along with the backlit white LED and bright white arcade-style buttons, is meant to suggest the transient, ephemeral nature of compression detritus.

practice. Rodgers, Tara. *Pink Noises.* Duke University Press, 2010.

87 https://github.com/magwhyr/lossystuff

88 I have published a tutorial and all of the code online through the University of Virginia Scholars Lab. Maguire, Ryan. "3D Printed Enclosures with OpenSCAD". Published November 14, 2017. https://scholarslab.lib.virginia.edu/blog/3d-printed-enclosures-with-openscad/.

Just as m*oDernisT* is given a new, particular materiality through this object, its musical structure is pried open, given freely to the listener-performer-player. The original fixed media composition is embedded in an interactive, handheld, portable game-like system containing a granular synthesis buffer and a flexible signal chain with several audio effects coded in Processing. The object does not have a speaker—instead users are provided headphones which connect to an 1/8th inch stereo output jack. Through headphone listening and close interaction with the physical object, held in both hands like a book, musical reception and interaction become private experiences. Here, design decisions are considered compositional decisions. These decisions shape the final form of the musical work by presenting the listener/performer with a field of possibilities. Inspired by the form factor of handheld video games from the 1990s, the design encourages interaction with the visual, sonic, and material presence of the art work. It makes MP3 and MP4 detritus accessible to tactile engagement.

On startup, the computer automatically launch a Java sketch containing the audiovisual material. The audio is the carefully composed MP3 detritus detailed earlier in this chapter. In its default state (e.g., sitting on a table without activating the touchscreen or indenting the arcade buttons) the recording plays at normal speed using a granular sample player. The sketch features both a waveform visualization and a spectrogram visualization. The waveform is displayed by default and the spectrogram is displayed when a user presses any of the four arcade buttons. The X and Y position of any touch screen interaction controls various audio effects, depending on which (if any) of the arcade buttons are being simultaneously depressed.

Without any buttons depressed, the y-position controls the rate and direction of playback, visualized as a vertical playhead superimposed over a waveform representation of the current audio frame. If any button is pressed, the visual representation changes to a spectrogram instead of a waveform, the x- and y-position of any touch screen interactions varying the coloration of the visualization. If button one is pressed, the x-position controls the location of the playhead in the audio file. If button two, the x-position controls grain size and y-position controls grain interval. If button three, x- and y-position combine to determine pitch. If button four, mouse position determines the center frequency and Q of a bandpass filter. These various functions can be combined at any time. Two button combinations trigger special visual effects—one replaces the spectrogram view with frames from the original a capella music video for "Tom's Diner," while the other shows the pixel information deleted during MP4 compression from those frames. Users (a.k.a., co-composers, performers, and audients) are not presented with any user guide or instruction manual—they are instead encouraged to explore and uncover the various audiovisual possibilities through play. The aesthetic experience is meant to be mysterious and allusive, even as the underlying technics are made clearly available.

The handheld form factor affords and encourages active exploration of curated video and composed sounds. In contrast to many participatory art works, this is done in a semi-private space, typically of mobile listening experiences, with headphones on, holding a bespoke physical object with both hands. It engages the senses of sight, hearing, and touch. The lone audience member is empowered to direct the performance. They can filter, sort, and replay these sounds, lingering, departing, and returning as desired. This interactivity exemplifies the flexibility gained by digitization, and yet, by using these abilities to meditate on audiovisual compression detritus, its power is tempered by an appreciation for the sonic detail and material presence sacrificed in its service.

I encourage physical attention to the material aspects of musical experience through handheld, interactive physical media. This is an effort to de-center musical experience from the sense of hearing (via the ear) and orient it to the sense of touch (via the body).[89]

Figure 21. Interactive *moDernisT*

89 Maus, Fred Everett. "Somaesthetics of Music." *Action, criticism, and theory for music education* 9, no. 1 (2010): 9-25.

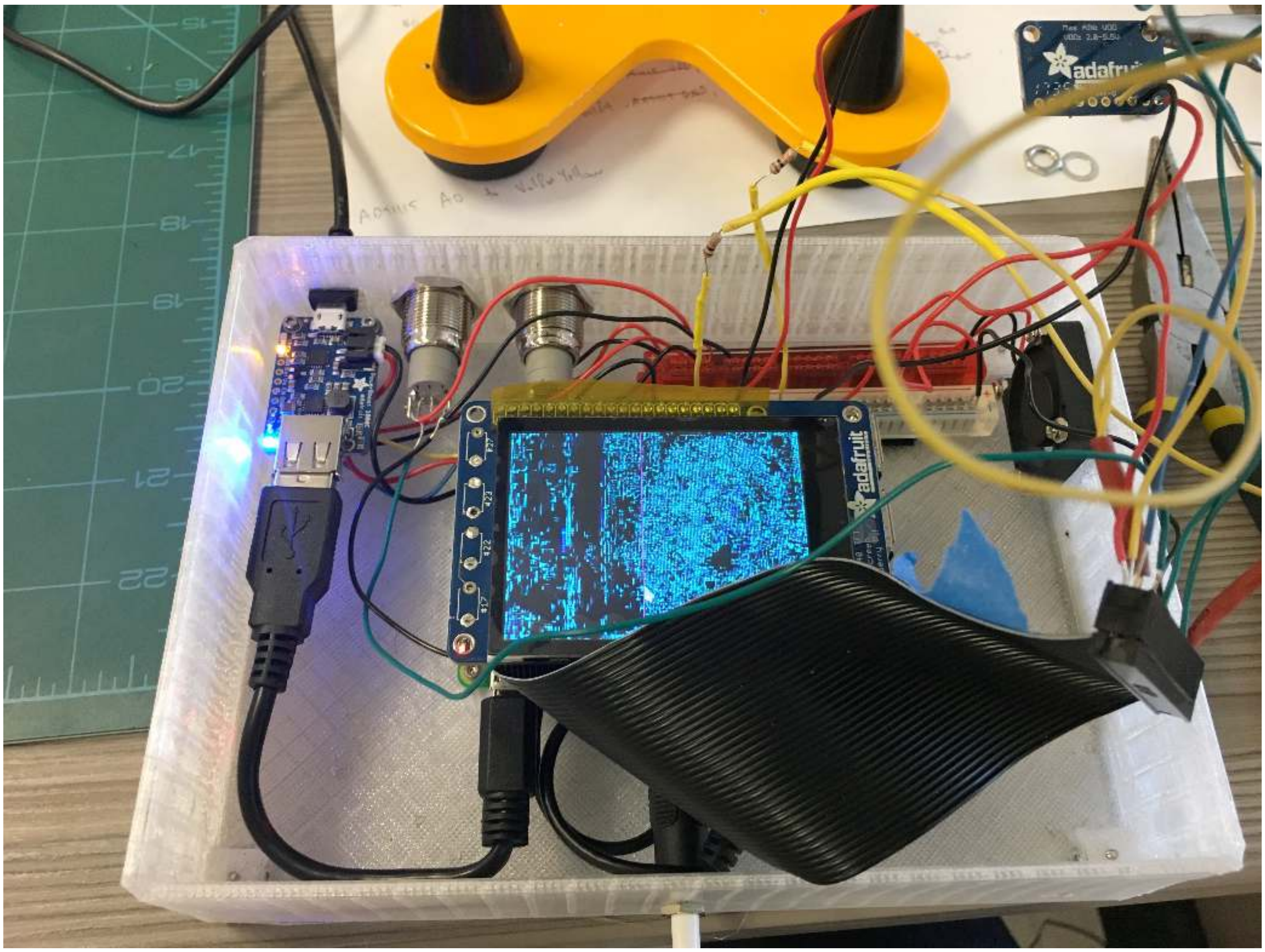

Figure 22. Testing the screen and power supply during assembly

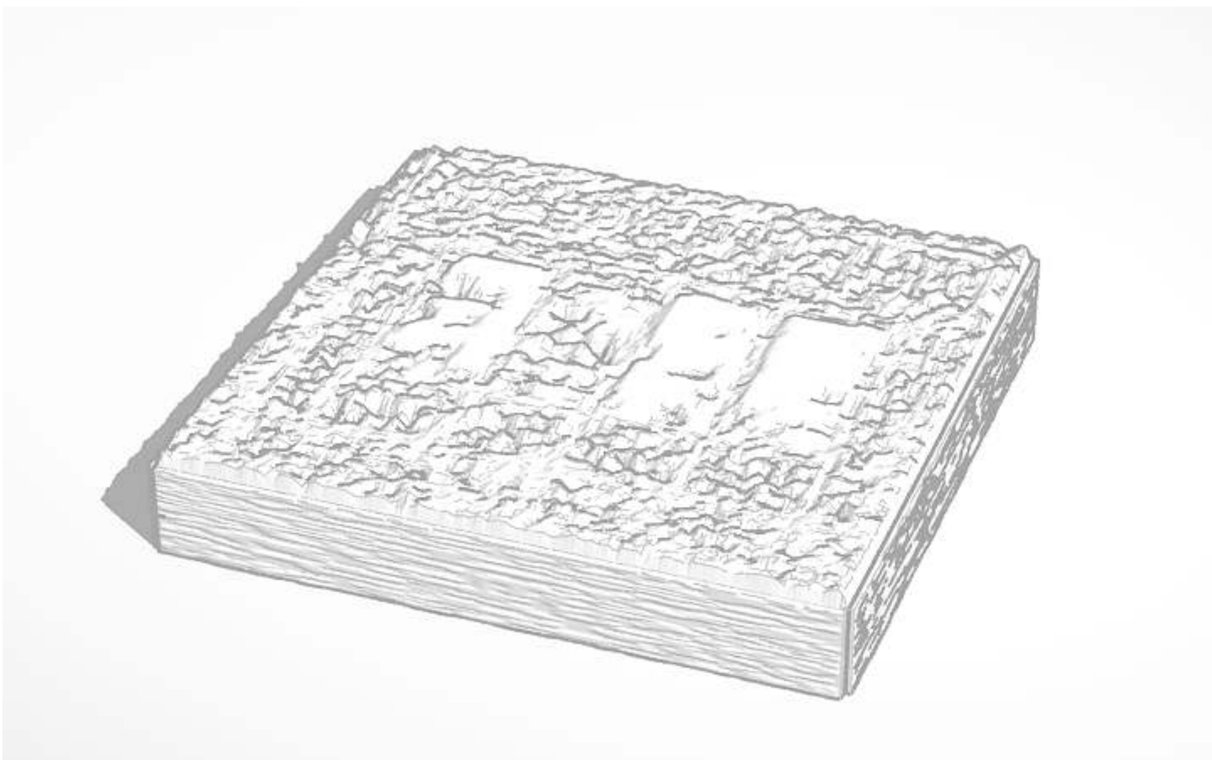

Figure 23. Original SCAD concept for surface texture from JPEG detritus

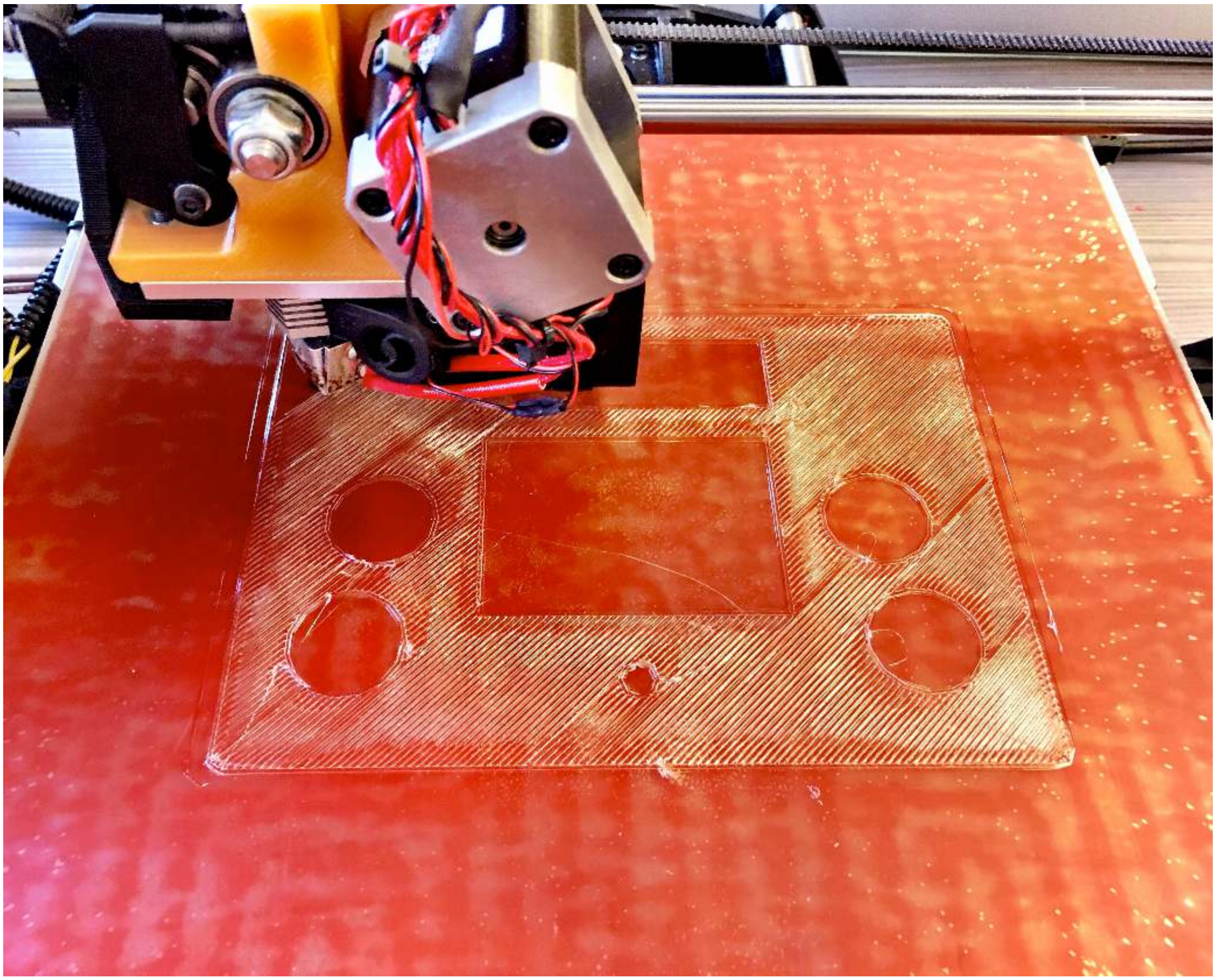

Figure 24. 3D printing the top plate

Meditation

> "This is not to apply a mythic embellishment to the past, or to express regrets for a lost youth … This is to say that the art of living has no history: it does not evolve: the pleasure which vanishes for good, there is no substitute for it. Other pleasures come, which replace nothing. *No progress in pleasures*, nothing but mutations."[90]

Direct physical indexicality is lost when converting a recording from analog tape, vinyl record, or compact disc to MP3. As the MP3 format came into wide use on file-sharing networks in the 1990s and 2000s, most MP3s were created as DIY bootlegs "ripped" from commercial recordings distributed via physical media.[91] Twenty years later, most MP3s are encountered via online streaming services (such as Spotify) and were never engaged with physical media (beyond the equipment used to record in the studio) to begin with. In this case, it seems less appropriate to speak of a loss of physical indexicality (from CD, vinyl, or analog tape), and more of a gain from nothing. MP3s have expanded access to millions of recordings for billions of people, recordings that most music listeners would not be able to access readily via physical media. They have also revealed that the aspiration towards high fidelity media is not shared by everyone—quite the opposite in many cases.[92] We might thus ask what sounds are created by MP3 compression in addition to asking what sounds are deleted. What are the distinctive characteristics of sounds transformed by MP3 compression?[93] My project answers these questions by presenting compression detritus directly to the listener. These sounds are both created and deleted by MP3 compression. They exist in counterpose to the MP3. The MP3 and its detritus are co-constituted. Through acts of composition, I engage in direct experience of such sounds.

---

90 Barthes, Roland. *Roland Barthes by Roland Barthes*. Macmillan, 2010/

91 Witt, Stephen. *How music got free: The end of an industry, the turn of the century, and the patient zero of piracy.* Penguin, 2015.

92 An informal survey by a Stanford professor showed that students increasingly prefer the sound of MP3 compressed audio to uncompressed. Plambeck, Joseph. "In mobile age, sound quality steps back." *New York Times* 9 (2010): B1.

93 A beautiful album composed with MP3 compressed sounds by Alberto Ricca, also known as Bienoise, was recently released on 3.5" Floppy Disk (or digital download) by Mille Plateaux. *Most Beautiful Design*. Mille Plateaux, 2018.

No sound begins as an MP3—there is always an uncompressed source. The sounds in MP3 files are akin to spectral summaries of these sources. A given source recording is analyzed, pruned, and encoded during MP3 compression. The decoded MP3 file thus presents to the listener sounds filtered through its own critical listening framework. This sonic-technologic configuration places the engineers and designers of the MP3 format in a sonic-curatorial role between the source and the listener.[94]

The sounds that are deleted by the MP3 algorithm are sounds that, in many cases, escape notice. They are high frequencies, inharmonic spectra, textures perceived as noise, quiet pitches overshadowed by louder ones in their proximity, and more. As I consider the differences between compressed and uncompressed digital audio files, I abut ongoing conversations regarding the differences between analog and digital audio, and between live and recorded sound. These pairings all involve a translation from one medium to another, e.g., from analog to digital, or, more generally, a reduction from some multidimensional space, say a concert hall, to a space of fewer dimensions, say a compact disc.

Though I am interested in the specific byproducts of the reduction from uncompressed digital audio to perceptually coded audio, I am also broadly interested in detritus as it occurs in any translation between media, between representational systems, or between languages. In one word, this could be described as an abiding interest in the irreducible, i.e., emergent properties that cannot be sufficiently described by constituent parts and which are, thus, lost in translation. I associate the sounds deleted during MP3 compression, that is, the irreducible sounds which comprise my *materials* or even my *medium* for this project, with the *atopic:* that which exceeds a given perceptual or conceptual frame.[95] In other words, the atopic is that which exceeds a given method of representation. It is that which is imperceptible via a

94 Such is the nature of all media technologies.

95 Atopic is derived from the Greek *atopia*, meaning *without place*. My definition differs from but is related to Barthes' usage of the term in *A Lover's Discourse: Fragments*: "unclassifiable, of a ceaselessly unforeseen originality." Barthes, Roland. *A lover's discourse: Fragments*. Macmillan, 1978.

given mode of perception, or that which is incomprehensible by a given conceptual framework.

A photograph compresses a rich, multidimensional scene in the real world to a 2-dimensional plane with variations in color. The real world scene is now constricted within a finite visual frame size and limited to the optical, geometric, and chromatic resolution of the photographic medium. An object or body in the real world has no such two-dimensional bound—it can be viewed at increasingly fine levels of detail down to the subatomic, beneath which we have not yet found a limit.[96] In a photograph, transcribed light and/or color information goes no deeper than the resolution of the medium. We lose the ability to interact in real time using our full sensory apparatuses with the light and color bearing properties of the photographed object or body, inclusive of technologies which extend our sensory abilities (ultraviolet and infrared sensors, for example). Further, we can no longer see from any perspective beyond the one specified by the photograph. However, a photograph also has its own material depth tangential to that of the original visual scene, and any encounter with the photograph will make its own unique material configuration available to aesthetic perception. Our own sensory capacities are not usually degraded by the change in media. As in this example, particular methods of representation, even reductive ones, are idiosyncratic and, thus, potentially valuable when viewed in proper perspective. Nonetheless, by highlighting that which exceeds the resolution of such transformational processes, I prioritize the irreducible properties of that which is so transformed.

> "Atopia lets us foresee a moment of respiration where mobile and discontinuous qualities provide sufficient reason to say that a *thing*, a *being*, or an *event* first exists as a free fact of consciousness before it eventually becomes the object of judgement or of intellectual determination. This respiration is none other than the moment of poetry that allows what Barthes labels *nuance*, or variation, to emerge. It is the source of intensive differences."[97]

---

96 Of course, human perception is also finite. I nonetheless value immersive aesthetic environments with the potential to exceed those capacities.

97 Millet, Yves. "Atopia & Aesthetics. A Modal Perspective." *Contemporary Aesthetics* 11, no. 1 (2013): 17.

Chapter 2

Accumulation

In this chapter, I will describe additional creative experiments with compression detritus. I will take a broader view than in the previous chapter, describing the technical implementations more briefly.

Ghost / Ambient

The transformation at the core of my technique is subtraction. When the breath leaves the body, folklore suggests that a ghost remains behind. Similarly, when compression removes the breath from a signal, its difference assumes a spectral presence. Staying with folklore, the corpse, reanimated, becomes a zombie. And so it is with the compressed file, which by some strange magic, dances almost like the real thing.

I access the ghosts left behind by mp3 compression through subtraction in the time domain. I create the following works through multiplication in the frequency domain.[98] Uncompressed digital audio signals are stored as discrete time series data, typically using Pulse-code modulation. In contrast, compressed audio is encoded in the frequency domain. In order to bring compressed and uncompressed audio into contact then, one or the other needs to be transformed. Thus, I work in both time and frequency domains to obtain my spectral material.

Ghost/Ambient is a seven track album featuring MP3 detritus from recordings used during various MP3

98 Thus, both of the fundamental arithmetic operations—the additive, in the form of its inverse (subtraction), and the multiplicative—and both poles of the time-frequency view of a signal are utilized.

listening tests.[99] These source recordings are of Chapman's "Fast Car", Vega's "Tom's Diner", and Haydn's *Trumpet Concerto* from the Fraunhofer tests, and Stravinsky's *Rite of Spring*, Ellington's "Take the A Train", Chopin's Polonaise Opus 71, No. 1, and La India's "Sobre el Fuego" from other publicly available listening tests.[100] [101] [102] [103] [104] [105] [106] Each of the seven tracks for the album was composed in a similar fashion. These can be considered algorithmic, 'formalized' works.[107] The procedure itself is simple (described below) and based on a single algorithm: convolution.[108] Therefore, I also consider these to be in some sense 'minimalist' works.[109]

After creating an MP3 from a 16-bit, 44.1kHz digital encoding of a source recording, I align the two files (WAV and MP3) in an audio editing software application.[110] I work at the sample level to compensate for the bit padding done by the MP3 encoder.[111] Once time-aligned, I invert the phase of one of the files. This cancels out time points at which the two files are identical, leaving audio only at the time points where the two files differ. This procedure is commonly referred to as a null test.[112] I then normalize the results of the null test and render it as a 16-bit 44.1kHz audio file.

99 Track lists for the Fraunhofer listening tests are provided in Sterne, Jonathan. "MP3: The meaning of a format." (2012). Various public listening tests have also been performed to test MP3 quality. Hydrogenaudio is a community audio website that has hosted many such tests. https://hydrogenaud.io/

100 Chapman, Tracy. *Tracy Chapman*. Elektra, 1988, CD.

101 Vega, Suzanne. *Solitude Standing.* A&M Records, 1987, CD.

102 Haydn, Joseph. Håkan Hardenberger, trumpet ; Academy of St. Martin-in-the-Fields ; Sir Neville Marriner, conductor. *Trompetenkonzerte = Trumpet Concertos*. Schwann Musica Mundi, 1987, CD.

103 Stravinsky, Igor. *Le Sacre du Printemps: Symphony in Three Movements*. Sony Classical, 1990, CD.

104 Ellington, Duke. *Take the "A" Train*. Vintage Jazz Classics, 1990, CD.

105 Chopin, Frédéric. Michel Dussault, piano.*Chopin Méconnu: The Unknown Chopin.*Toronto, Ontario, Canada : Les disques SRC, p1992, CD.

106 India. *Sobre el Fuego.* RMM Records: RMD 82157, 1997, CD.

107 My interest in formalized music was spurred along by an early encounter with: Xenakis, Iannis. *Formalized music: thought and mathematics in composition*. No. 6. Pendragon Press, 1992.

108 Roads, Curtis. "Sound transformation by convolution." *Musical signal processing* (1997): 411-438.

109 A range of contemporary perspectives on minimalism and postminimalism can be found in Potter, Keith, and Kyle Gann. *The Ashgate research companion to minimalist and postminimalist music*. Routledge, 2016.

110 Audacity [Computer Software]. Version 2.1.0, 2015. Retrieved from http://audacity.sourceforge.net/

111 For detailed technical information on implementing perceptual audio coders, including MP3, see Bosi, Marina, and Richard E. Goldberg. *Introduction to digital audio coding and standards*. Vol. 721. Springer Science & Business Media, 2012.

112 On phase in audio recording: https://www.soundonsound.com/techniques/phase-demystified

These ghosts, unprocessed, are sonically interesting, if harsh—in the sense that new frequency components not heard in the original recordings now predominate.[113] If I did not feel so compelled to explicate their compositional affordances, I might simply present these files directly to the public. On my website, I have made a limited number of such files available.[114] I am additionally planning a more comprehensive web interface for such artifacts in collaboration with a web programmer and data visualization specialist. Nonetheless, my primary objective in this dissertation is to understand these artifacts through acts of composition.[115] In the body of work developed for the creative portfolio, *Ghost/Ambient* serves as a baseline—it is the work where I have applied the lightest compositional touch, so to speak.

I take each *ghost* and pair it with its original 16-bit 44.1kHz source recording once more. I convolve these two files together. Conceptually, the algorithm expands common frequencies and suppresses those which are not shared between the two files. Mathematically, convolution in the time domain is equivalent to multiplication in the frequency domain, and vice versa. Aesthetically, the resultant audio sounds as if one file has been played into a room whose resonances were created from the other, creating a slowly evolving reverberant soundscape.[116] Metaphorically, the ghost echoes inside the body of its former digital being.

For each of the seven tracks, I used the full track length of the original recording for the window size.

---

113 Marui, Atsushi, and William L. Martens. "Timbre of nonlinear distortion effects: Perceptual attributes beyond sharpness." In *Proceedings of the Conference on Interdisciplinary Musicology*. 2005.

114 http://theghostinthemp3.com

115 On composition as research: Pace, Ian. "Composition and Performance can be, and often have been, Research." *Tempo* 70, no. 275 (2016): 60-70.

116 Such effects are only possible because of artificial reverb technology. Jonathan Sterne describes the creation of the "detachable echo" while developing a theory of space and its construction and representation by artificial reverberation in Sterne, Jonathan. "Space within space: Artificial reverb and the detachable echo." *Grey Room* (2015): 110-131.

This creates resonances in proportion to the amplitude and sounding time for each frequency bin in the spectrum. I complete the process by normalizing the resultant sound file. These minimally processed recordings are then presented as the finished works.

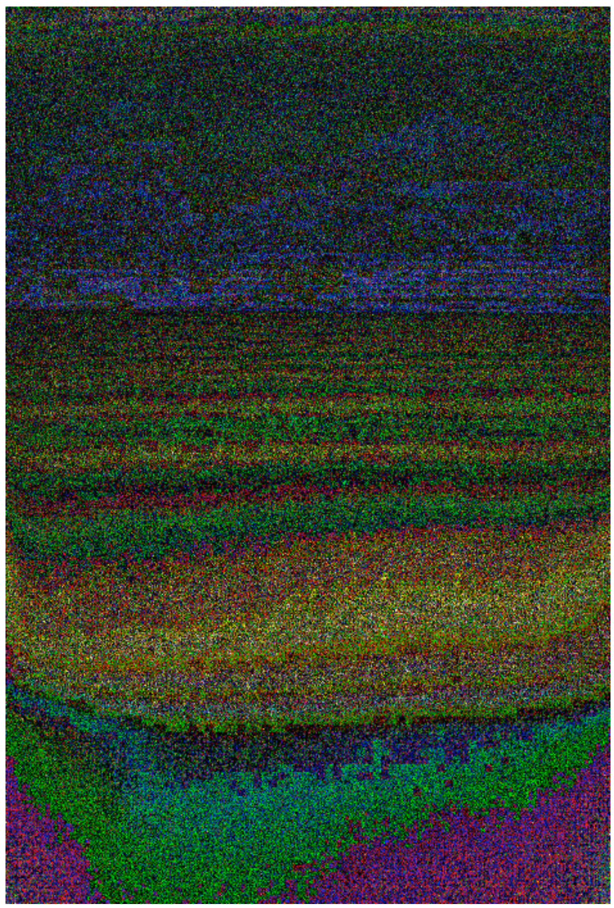

Figure 25. Ghost/Ambient

$†@®

Like *moDernisT*, *$†@®* is sourced from a recording that was used as listening test track during the development of the MP3 format at the Fraunhofer Institute in the early 1990s. *$†@®* was created from the audiovisual information deleted during MP3 and MP4 compression of the music video for Tracy Chapman's 1988 Grammy-nominated song "Fast Car". The original song's reflective lyrics offer a first person narrative from the perspective of a working poor woman living in a homeless shelter while trying to escape the cycle of poverty. After noting the similarities between her alcoholic father (whom her mother left) and her current partner (to whom the refrain is addressed), the song ends with the narrator presenting an ultimatum: "You gotta make a decision—leave tonight or live and die this way." In *$†@®*, these lyrics are obscured—only traces of the original recording remain, as for the most part it has been deleted by the digital compression process. To someone familiar with the song, the detritus is particularly engaging. In draws the listener in by leaving space for the memory to complete the sonic picture that the faded artifacts present.[117]

The compositional process begins by establishing an ambient backdrop, similar to those created for *Ghost/Ambient*. In this case, rather than setting a window length equal to the length of the song, I experiment with a shorter moving window of under 10 seconds. Moving next to the foreground, I analyze the song structure, detecting eleven sections. In addition to the ghost file created in the time domain via null test, I prepare material in the frequency domain using my Python masking script to create files with a variety of difference thresholds. I then scramble these eleven subsections in a similar

117 “Whose voice, no one's, there is no one, there's a voice without a mouth, and somewhere a kind of hearing, something compelled to hear, and somewhere a hand, it calls that a hand, it wants to make a hand, or if not a hand something somewhere that can leave a trace, of what is made, of what is said, you can't do with less, no, that's romancing, more romancing, there is nothing but a voice murmuring a trace.” Beckett, Samuel. “Texts for Nothing," in *Collected Shorter Prose: 1945-1980*. John Calder, 1984.

manner to *moDernisT*, disrupting the linear flow of time at the level of phrases, oscillating between levels of relative un/intelligibility.[118]

I next turn my attention to low frequencies.[119] Using a low pass filter, I isolate the bass guitar.[120] I then splice several notes from the bass line and load them into a sampler instrument. I perform these phase inverted samples live in the studio, resulting in a new bass groove spun out from the old bass line. I then run this through a complex signal chain passing from a denoiser, into a noise gate, into a compressor, into an exciter, into a reverb, into an equalizer, into a binaural filter, and finally through a high pass filter. Following this, I isolate a bass drum from the ghost recording using a low pass filter and careful time splicing. I process the bass drum using a room simulation reverb, an exciter, and a sub-bass enhancer. I place this processed bass drum at two key climactic moments in the composition—marking the beginnings of the ninth and tenth sections.

My final additions thicken the musical texture. I make two copies of the phase inverted ghost and pan them hard left and hard right, respectively. Each is filtered through a bandpass filter and has its volume carefully hand automated. This allows me to shift these parts between foreground and background dynamically. The left channel is gated using a noise gate with a 30ms attack and a very slow 2000ms release.

In the final mixing stage, I equalize the fifth section differently than the rest of the composition, with a

118 "'Twas brillig, and the slithy toves, Did gyre and gimble in the wabe; All mimsy were the borogoves, And the mome raths outgrabe." Carroll, Lewis. *Through the looking glass: And what Alice found there*. Rand, McNally, 1917.

119 For a theorization of low frequency sound's vital materiality, see Jasen, Paul C. *Low end theory: Bass, bodies and the materiality of sonic experience*. Bloomsbury Publishing USA, 2016.

120 Filtering can be used as a rudimentary source separation technique. I take a machine listening approach to source separation in "Creating musical structure from the temporal dynamics of soundscapes." In *2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA)*, pp. 1432-1433. IEEE, 2012.

broad boost centered at 1600Hz and a more focused boost at 400Hz, marking this section as sonically and structurally significant. The entire mix is then run through a multiband compressor and a final equalization which boosts both the very high end and the very low end slightly, creating more high frequency articulation and adding weight to the bass and sub-bass.

The video is created using the same algorithm as for *moDernisT*—in Python, I utilize the OpenCV library to calculate the frame-by-frame and pixel-by-pixel difference between compressed and uncompressed renderings of the music video. This is saved as a matrix using a numerical computing package (numpy) and finally rendered as a high-definition video file using H.264/x264 lossless compression.[121] I work in RGB space for this video rather than converting to HSV before calculating the difference. This keeps calculations to a minimum. The streamlined result is less perceptually interpretable as the "difference" between the uncompressed and compressed videos than it would be if converted to HSV space before calculation, but the colors are brighter this way, and the resulting video is dynamic.[122] [123]

121 I attempted two transformations of the data before deciding on which rendering to use for the final art work. First, I took the absolute value of each difference rather than allowing negative pixel values (which wrap around to 255 from 0). Second, I subtracted each frame of the absolute value difference from a frame with all pixel values set to zero (white pixels). As an alternate, I subtracted each frame of the absolute value difference from a frame with all pixel values set to 255 (black pixels). Neither of these variations were used for the final video.

122 This is an interesting variation on a mapping problem. The key variable here is not how the information is mapped to a visualization but rather how the data is formatted before processing. I call this a formatting problem.

123 An alternative algorithm for finding the so-called lost information is this: each pixel in each frame that is identical between uncompressed and compressed renderings of the video is set to either 0 or 255, depending on whether you prefer to represent zero change as the color white or the color black. Each pixel that has changed from the uncompressed video in the compressed video is set as the original uncompressed pixel value. From an information standpoint, this amounts to a presentation of the pixels that are altered by compression. In this arrangement, no information regarding the amount by which each pixel is changed is preserved, but the information which is ultimately most vulnerable to compression is more clearly represented.

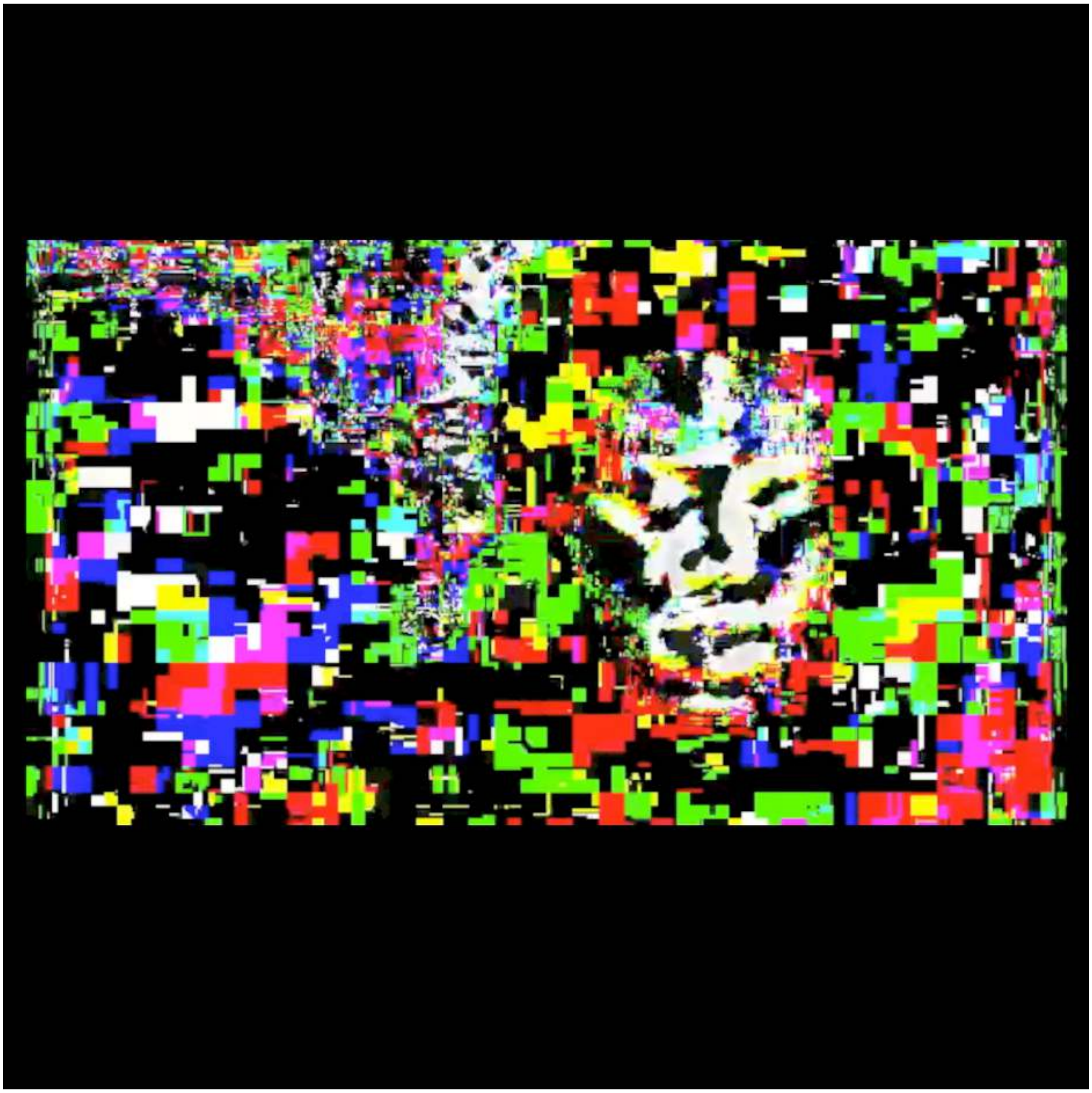

Figure 26. $†@®

I employ additional techniques while creating *freeLanguage.* The material is sourced from "In All Languages," a recording by Ornette Coleman used in the Fraunhofer listening tests. I improvise on digital hardware to perform this material, listening and reacting intuitively during the recording process. In this improvisational act, committed to tape, I articulate the composition-improvisation dichotomy. Connecting light and sound, I collaborate with artist Heather Mease to create a visual counterpoint.[124]

I prepare the sonic material via phase inversion, masking, and simulated real-time MP3 compression.[125] I bandpass filter several copies of the phase inversion detritus, varying center frequency and Q. I record these changes in real-time, navigating by ear. I then isolate prominent frequencies in the detritus using a Python masking script. I perform a spectral analysis and notch out both high frequencies and short transients. I resynthesize the remaining material and re-amplify it through a Peavey electronic drum amplifier. Working at high gain in the studio, I adjust the equalization to taste, tracking the output via condenser microphone. This conversion, from digital to analog, through one transducer and back into another, is a means of accumulating texture. It is a transformational process undertaken for the sake of the artifacts it precipitates. I sculpt one final layer by performing MP3 detritus simulation in real-time

124 In the summer of 1997, saxophonist/composer Ornette Coleman and philosopher Jacques Derrida met in Paris before and between Coleman's three concerts at La Villette. During these meetings, Derrida interviewed Coleman about his views on composition, improvisation, language, and racism. Their respective ideas about "languages of origin" and their experiences of racial prejudice converge remarkably, with Coleman musing, "Do you ever ask yourself if the language that you speak now interferes with your actual thoughts?" Derrida responds, "It is an enigma for me … I know that something speaks through me, a language that I don't understand, that I sometimes translate more or less easily into my language." In this sense, freeLanguage is an attempt at recording "actual thoughts" through studio improvisation with sonic material that was lost in the translational act of compression.

125 In response to my project, the company Goodhertz created a *Ghost in the MP3* mode for their lossy compression simulation plug-in *Lossy.* I was invited to serve as a beta tester for this plug-in and requested a free copy of the plug-in in exchange for their using the title of my project. The plug-in has a pleasant timbral signature which differs from my algorithms.

using a hardware MIDI controller linked to an MP3 detritus plug-in. This engages the body in the creative process, centering physical gesture rather than the resultant reified object. The precise trajectory of a physical gesture is noisy and slightly unpredictable—qualities which I relish.[126]

I then focus on the head of the original composition. I time stretch its phase inversion detritus, expanding the introductory head to the length of the original composition. I then blend this material with the other layers, automating volume levels and panning for each track, and equalizing and compressing the track as a whole. I load this software into an open-source DJ'ing software application which I interact with via a Korg Kaoss DJ Mixer.[127] This allows me to perform the piece one last time —translated through the perceptual system and altered by embodied reactions—adding rubato or accelerating, filtering dynamically, and adding real-time signal processing effects.

The video is MP4 detritus from the biographical Ornette Coleman documentary *Made in America.*[128] I use a customized script in Natron to calculate and render the difference between uncompressed and compressed versions of the video.[129] I process a twenty minute excerpt from the film in this manner. I deliver this material to my collaborator, Heather Mease. She segments, arranges, and layers subsections from this material. Her edit creates a visual counterpoint to the music. Of particular interest is the rhythmic language she develops by cutting between contrasting clips. The effect is of a (silent)

126 For a perspective on human-computer musical interfaces which centers the body and physical movement, see Bellona, Jon. "Physical Composition: The Musicality of Body Movement on Digital Musical Instruments." PhD dissertation, University of Virginia, 2018.

127 Mixxx [Computer Software]. Version 2.1.5, 2018. Retrieved from https://www.mixxx.org

128 Clarke, Shirley, and Ornette Coleman. DVD. *Ornette: Made in America*. Directed by Shirley Clark.New Jersey: Milestone Film & Video, 2014..

129 Natron [Computer Software]. Version 2.0.3, 2016. Retrieved from http://natron.fr

percussion part communicated visually.[130] [131]

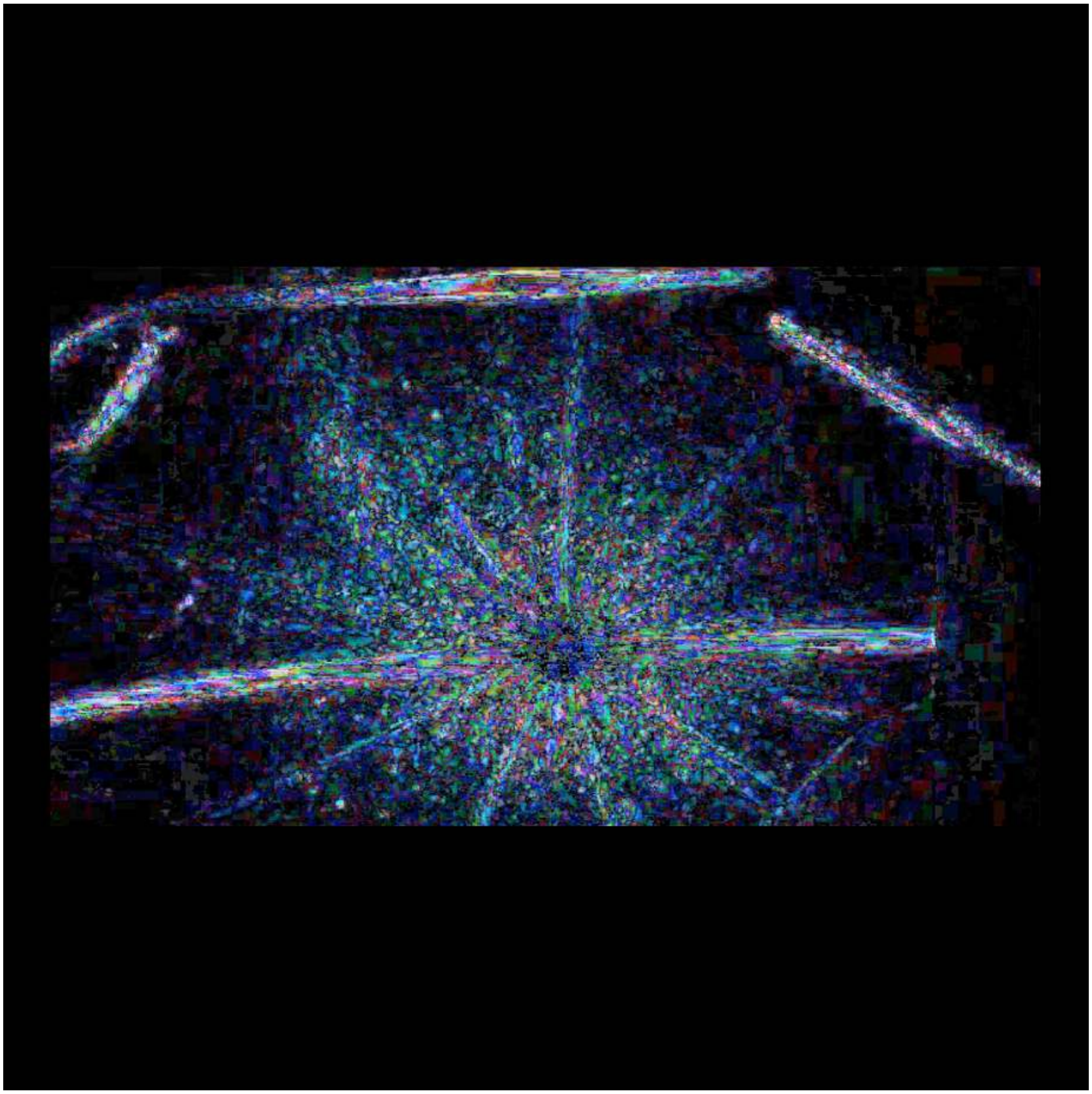

Figure 27. freeLanguage

130 My work beyond this dissertation is often collaborative. I value collaboration as a means of human connection and for its transformative potential, as ideas move across boundaries between individuals, accumulating artifacts and resonances. There is a growing field of research highlighting the benefits and challenges of collaborative composition, especially in educational settings. For example: Rusinek, Gabriel. "Action-research on collaborative composition: an analysis of research questions and designs." *Musical creativity: Insights from music education research* (2012): 185-200.

131 Non-cochlear sonic art is discussed at length in Kim-Cohen, Seth. *In the blink of an ear: toward a non-cochlear sonic art*. A&C Black, 2009.

Harvest Light

Neil Young has been a public advocate for high resolution digital, going so far as to create a portable high resolution audio player through his company PonoMusic, and subsequently releasing his entire back catalogue in high resolution digital audio for free download.[132] [133] *Harvest Light* is a first attempt at considering how to perform this material live. I develop the difference between compressed and uncompressed versions of his album *Harvest* as compositional material.[134] I convolve the compression detritus with the uncompressed audio—the same technique used for *Ghost/Ambient.* Live, I perform a deconstruction of the lyrical, harmonic, and melodic content of the song singing and playing electric guitar. I process my voice and guitar in real-time using the *Ghost in the MP3* simulation plug-in. I project a heavily compressed digital rendering of Neil Young's film *Journey Through the Past* on stage, transforming my composition into a soundtrack to the (now silent) film.[135]

Figure 28. Harvest Light, in performance

132 Edmund, Mark. "Music to Whose Ears?." *Quality Progress* 48, no. 2 (2015): 14.
133 https://neilyoungarchives.com/
134 Young, Neil, Ben Keith, Kenneth Buttrey, Tim Drummond, Jack Nitzsche, David Crosby, Stephen Stills, James Taylor, and Linda Ronstadt. *Harvest*. Reprise, 1972.
135 The IMDB record for *Journey Through the Past* can be found at https://www.imdb.com/title/tt0068776/. Retrieved April 10, 2019.

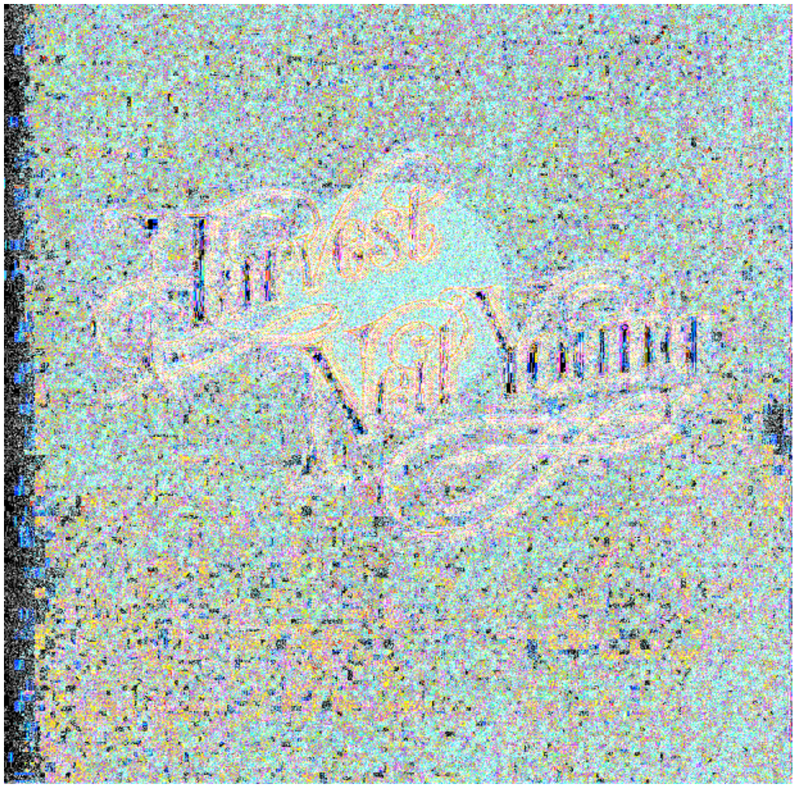

Figure 29. JPEG detritus from the *Harvest* album cover

white, pink, brown

white, pink, brown is an acoustic composition for an indeterminate number of instrumentalists playing pitched instruments. It is the first work which centers notation in this dissertation. The musical material is derived from computer-assisted transcriptions of maximally MP3 compressed white, pink, and brown noise.[136] [137] The performance structure is game-like.[138] Performers select three transcriptions from a shuffled deck of notational fragments just before performance. Their task is to sight read these

136 This builds off of preliminary experiments described in my masters thesis. Maguire, Ryan. 2013. *Distort them as you please: the sonic artifacts of compositional transformation*.

137 Unlike other compositions in the portfolio, this piece uses maximally compressed MP3 artifacts. In other words, rather than working with the information that is deleted by MP3 compression, I am utilizing the information that is preserved after MP3 compression. I used the lowest possible quality MP3 compression settings (8kbps).

138 The game piece is only superficially competitive. Performers self-score according to a partially subjective rubric that they self-determine individually. For a survey on cooperation and competition in music composition see Hermans, Phillip. 2013. *Cooperation and competition in music composition*.

fragments as accurately as possible while satisfying different musical criteria for each of the three sections of the composition. The player with the highest self-assigned score at the end is awarded the choice of either performing an improvised cadenza or electing another member of the ensemble to perform one.[139]

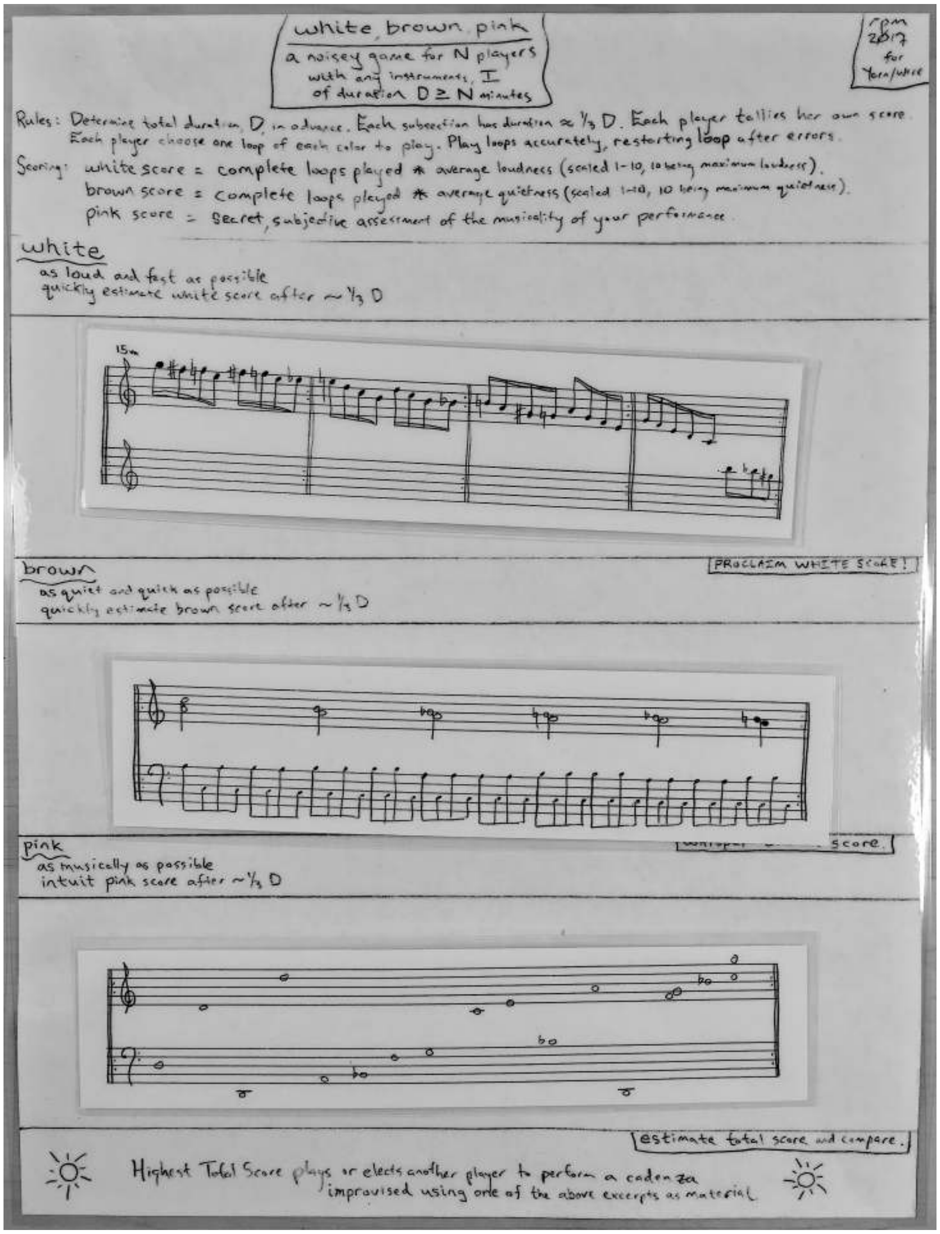

Figure 30. white, pink, brown

139 My intention is to foreground the unique presence of the performers through this game piece. Their contributions transform notated fragments of noise into a vital aesthetic experience.

Rest, 1-3

During the summer of 2014, I wrote a collection of nine songs in memory of my late Aunt Kathy. Over the intervening years, I reworked the songs into different forms in collaboration with numerous musicians as part of a personal grieving process. Their lyrics meditate on familial love, lost childhood, the persistence of memory, the passage of time, and the irreversible loss of death.[140] *Rest, 1-3* is created from the compression detritus of assembled recordings of the first three songs of the collection.

I've performed the piece publicly twice, once at Duke University and once in Charlottesville, at the Bridge Progressive Arts Initiative. Live, I create a 6-channel diffusion from the sonic remainders of these songs, allowing their presence to completely fill the room and surround its inhabitants. Two channels are sourced from recordings of the songs with a trio (drums, electric bass, and electric guitar) and the remaining four channels are sourced from original digitally processed vocal performances. I insert myself into the detritus by singing and playing electric guitar, digitally compressing my voice live, working around the song structures and melodies in an impressionistic semi-improvisation. The guitar signal is split between two channels, arranged in stereo, and the vocals are routed through a stereo PA-system. The 6 channel ambient soundscape surrounds the audience, with each layer given its own speaker for output. The result is ten channels of audio surrounding the audience. I project light onto the entire stage, created from the erased remainder of original film footage shot in collaboration with cinematographer Andy Patch: a sunrise over the ocean, fog rolling in to a moonlit valley, an aimless drive through empty city streets.

140 I relate my use of compression detritus to grieve the death of a loved one to William Basinski's use of disintegrating tape loops to both mourn and accept loss. As he notes in the liner notes to *Disintegration Loops,* "tied up in these melodies were my youth, my paradise lost, the American pastoral landscape, all dying gently, gracefully, beautifully." William Basinski, Liner Notes in The Disintegration Loops. For a discussion of Basinski's work and theorization about the poetic impermanence of sound, see Guimond, David. "The Sounds of Disappearance." *Intermédialités: Histoire et théorie des arts, des lettres et des techniques/Intermediality: History and Theory of the Arts, Literature and Technologies* 10 (2007): 115-130.

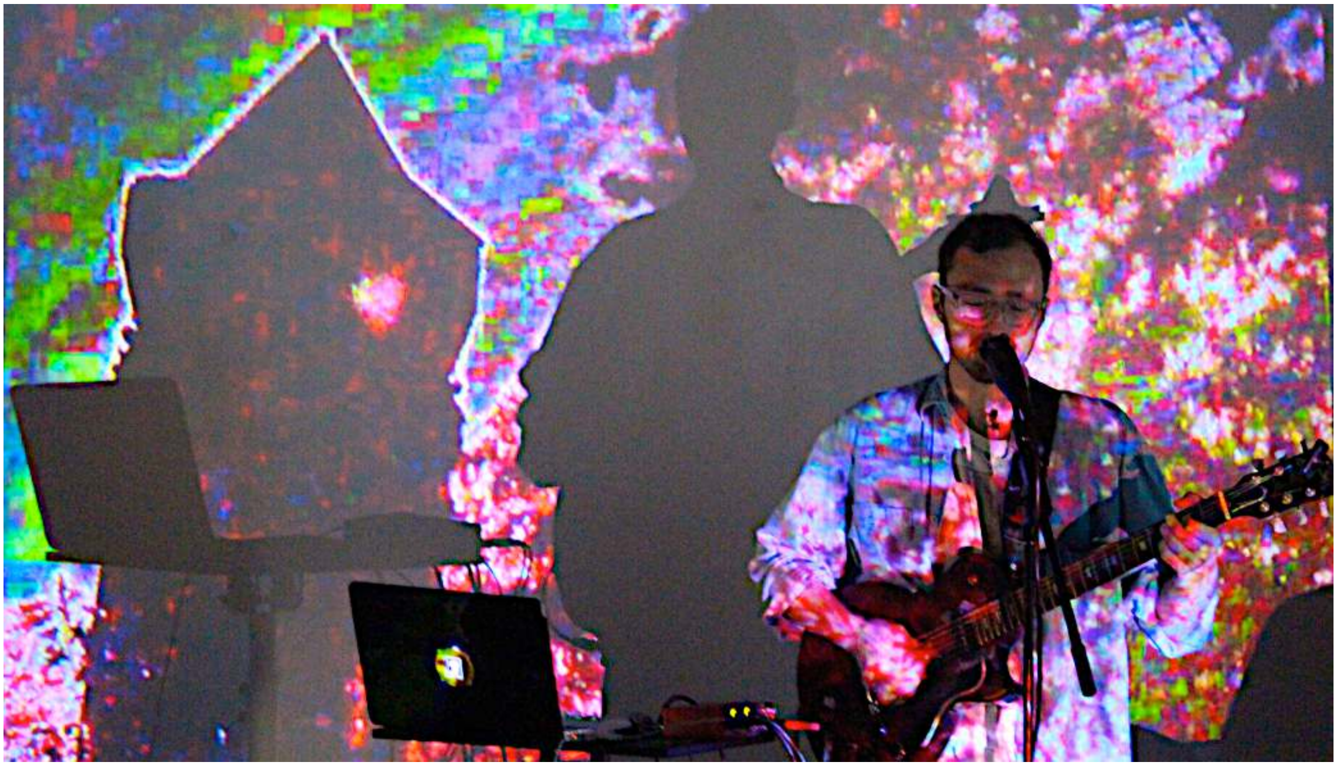

Figure 31. Rest, 1-3

Ghost in the Codec

Ghost in the Codec extends my creative research in lossy compression detritus to image file formats. Specifically, I work with JPEG compression to create an interactive audiovisual installation.[141] The installation is comprised of a computer terminal running Processing with a mouse, camera, and printer attached.[142] Users sit at the terminal and take a self-portrait using the camera. This digital image is first encoded as a lossless PNG file. If users are satisfied with the photograph, they may then compress the file using the JPEG compression format. The information which is deleted during compression is then

141 Wallace, Gregory K. "The JPEG still picture compression standard." *IEEE transactions on consumer electronics* 38, no. 1 (1992): xviii-xxxiv. For an object-oriented theorization of JPEG compression, together with practice-research works, see Caplan, Paul Lomax. "JPEG: the quadruple object." (2013).

142 Processing [Computer Software]. Version 2.2.1, May 19 2014. Retrieved from http://processing.org

automatically rendered to a new image file and printed via the attached color printer. Users are provided with adhesive and are encouraged to affix their compression ghosts to the walls of the installation room. Over the course of the installation, the walls are slowly filled with these ghost portraits.[143]

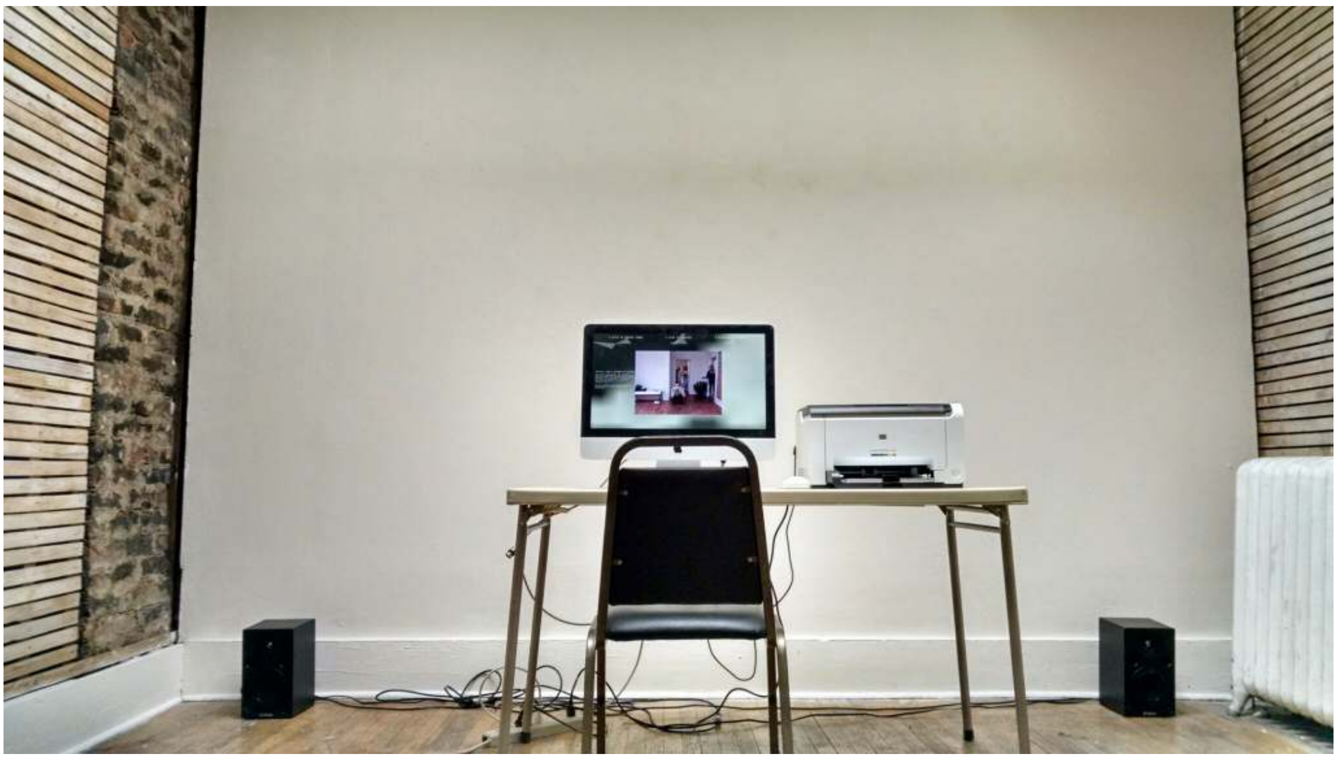

Figure 32. Before *Ghost in the Codec*

143 An international collection of contemporary art dealing with ghostly images, absence, loss, ephemera, and immateriality is Roca, José. *Phantasmagoria: specters of absence*. Independent Curators, 2007. Going forward, I hope to situate my work more regularly in gallery settings and art spaces.

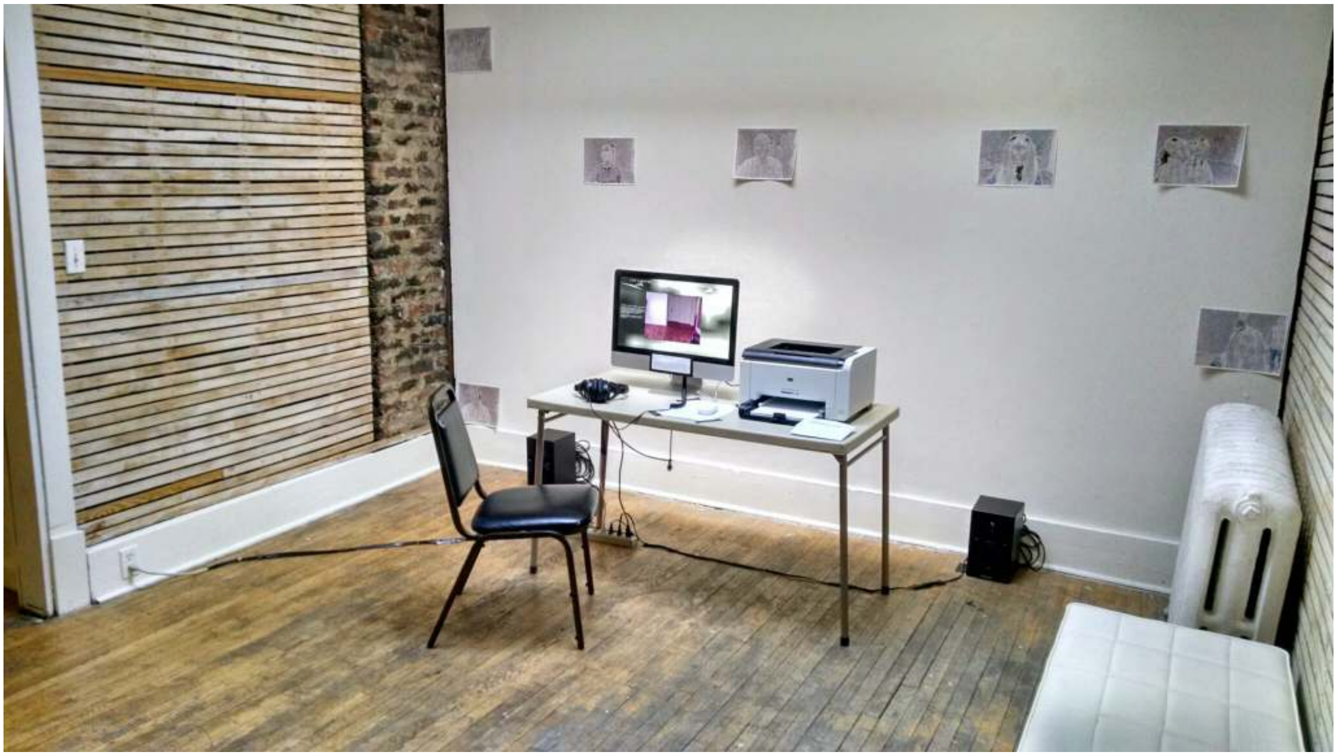

Figure 33. Ghost in the Codec, pt. 1

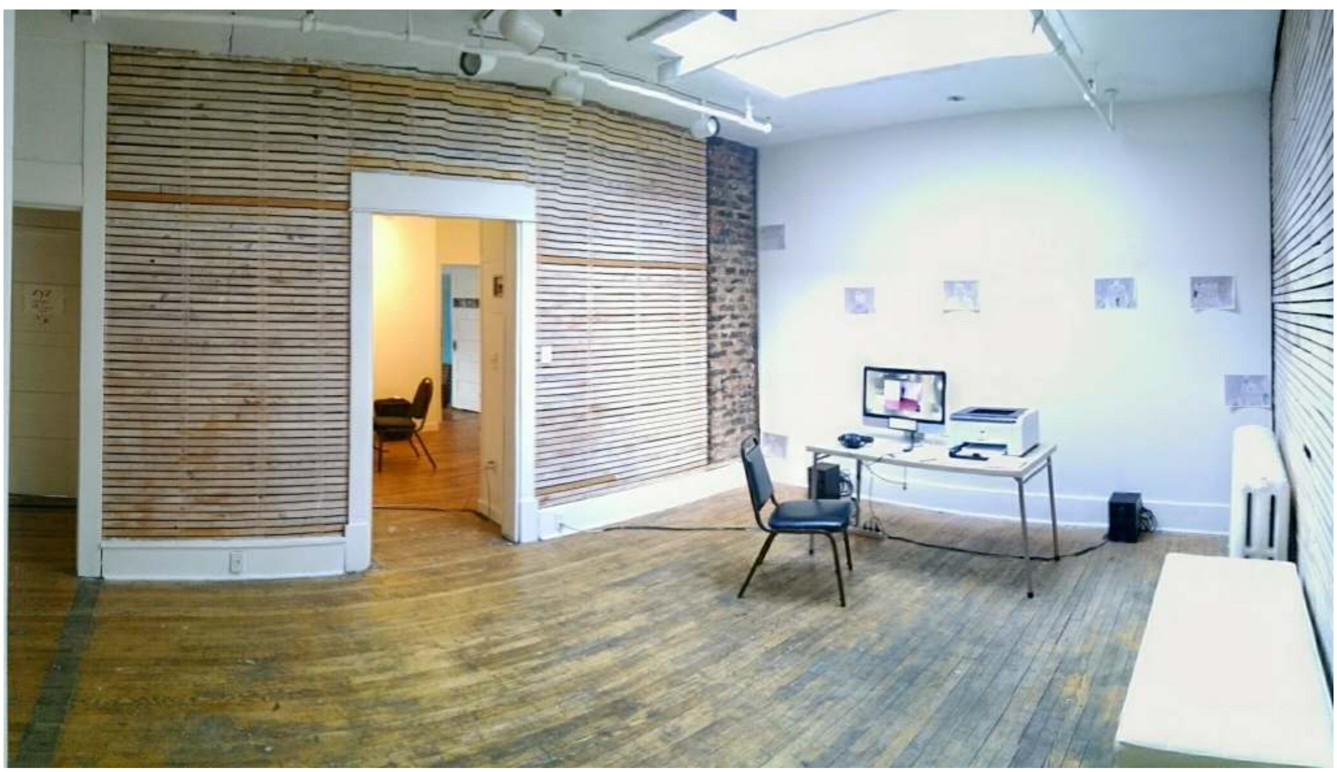

Figure 34. Ghost in the Codec, pt. 2

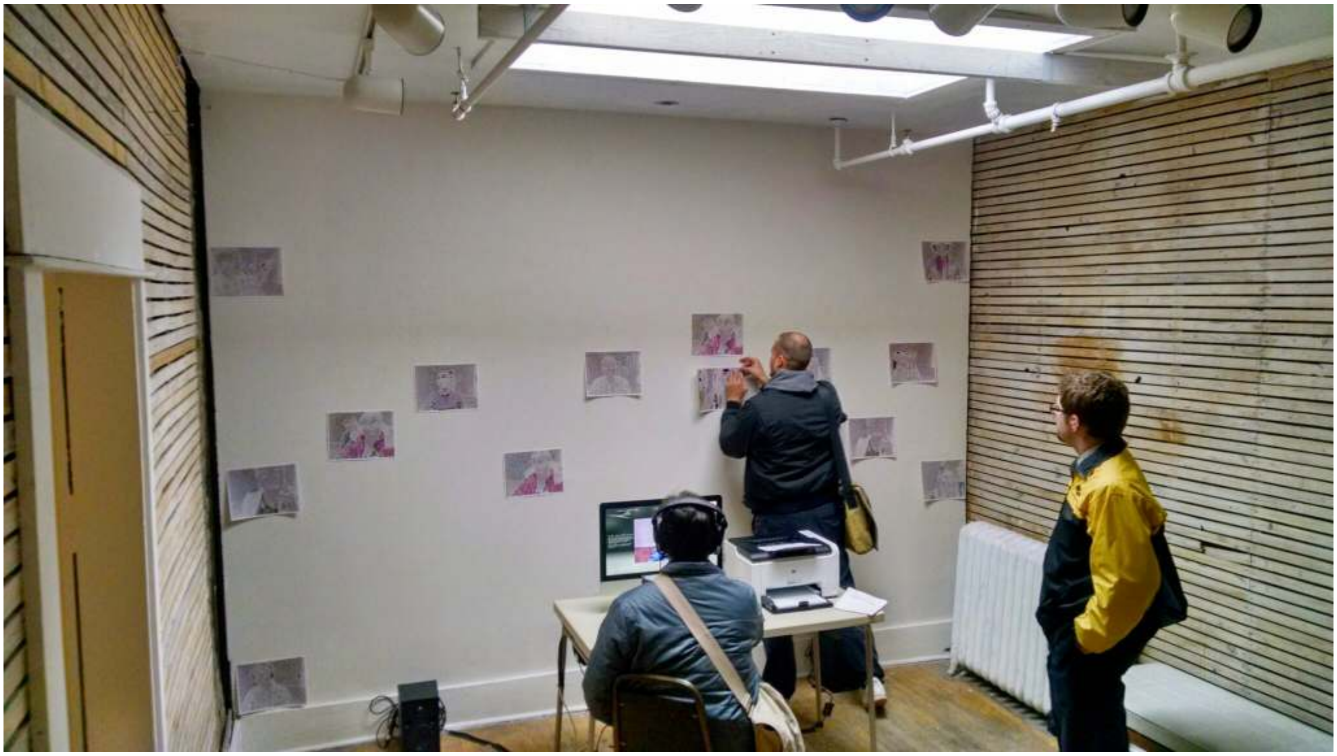

Figure 35. Ghost in the Codec, pt. 3

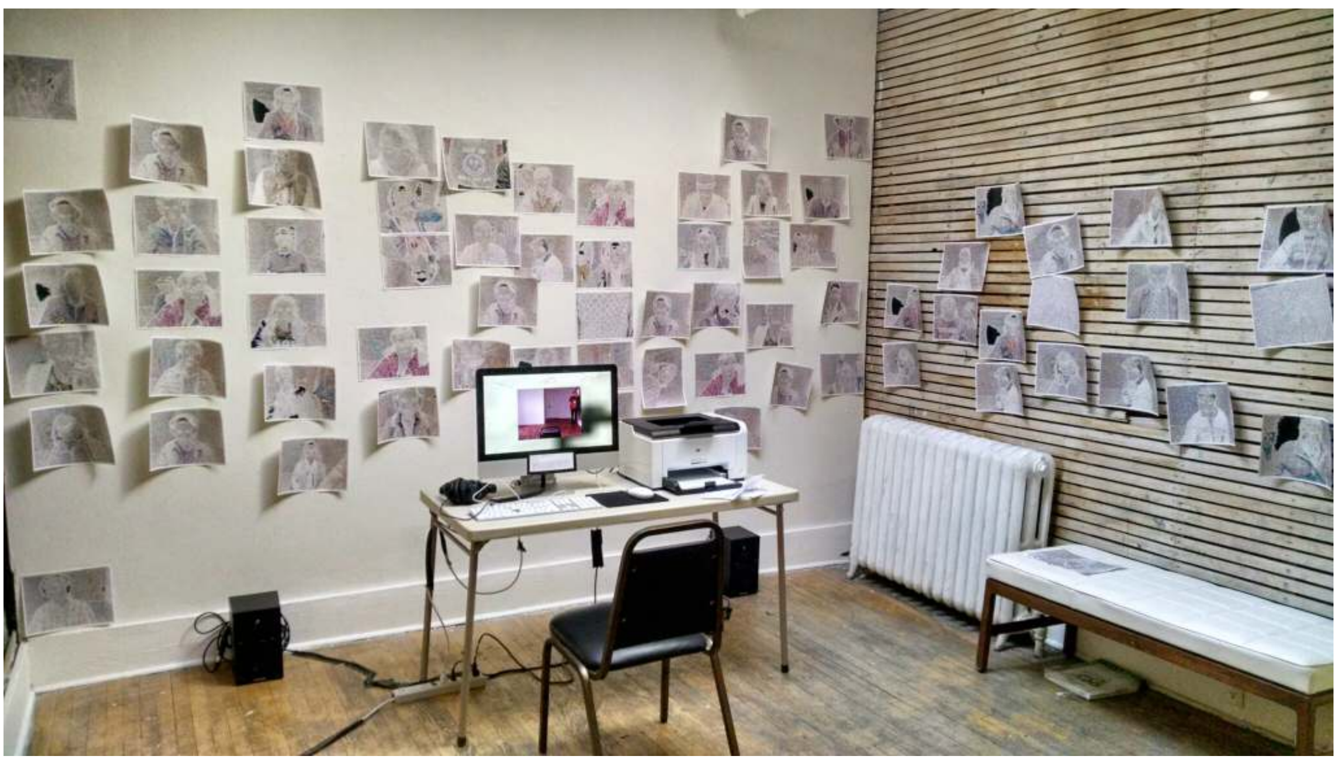

Figure 36. After *Ghost in the Codec*

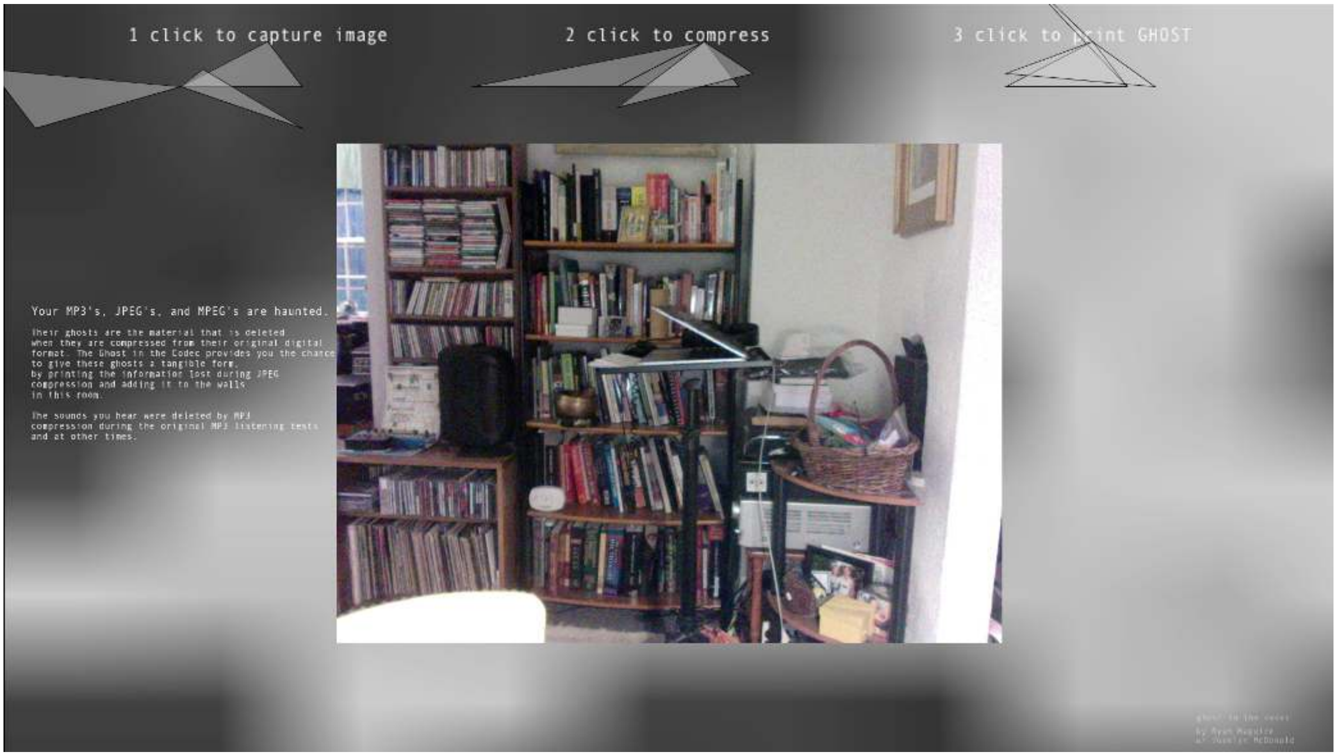

Figure 37. Ghost in the Codec computer terminal view

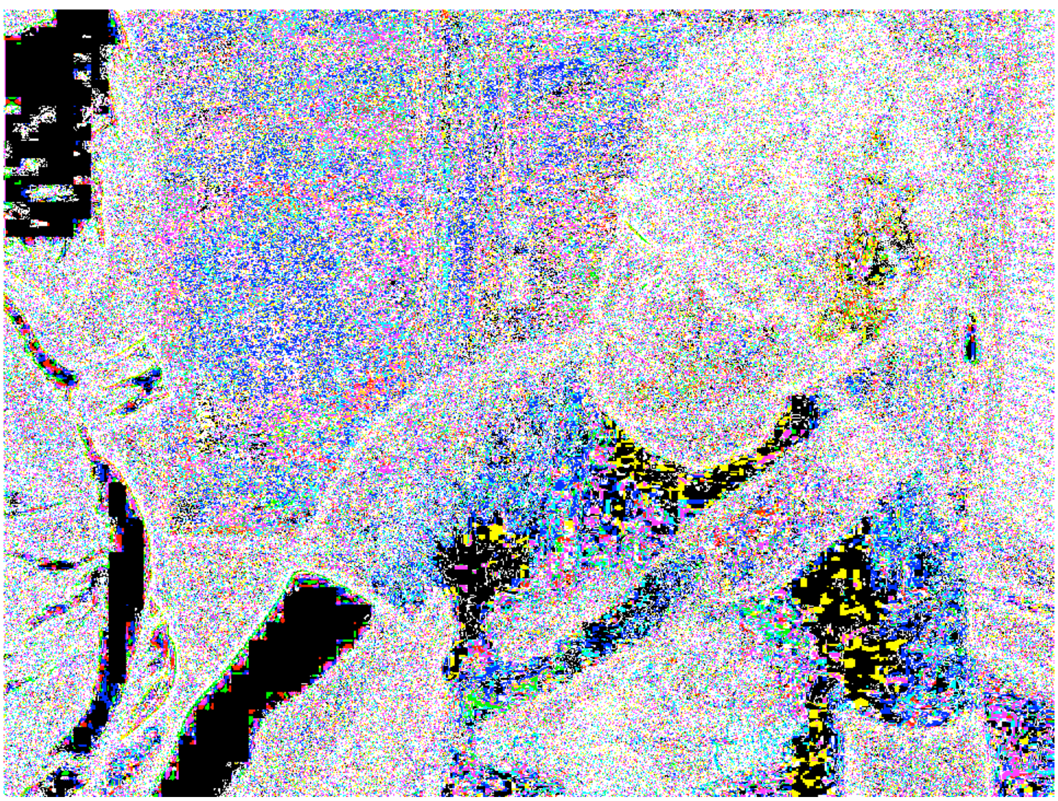

Figure 38. Ghost in the Codec printed detritus

Pure Beauty

The next two works are collaborations that I was invited to participate in. Sebastian Zimmerhackl, from Berlin-based design studio Selam-X, contacted me about collaborating for a multimedia exhibition titled *Pure Beauty*.[144] Specifically, they were working on an installation using generative adaptive networks to synthesize and morph between computer-generated make-up advertisement photographs. The resulting images are somewhat ghastly, calling to mind the uncanny valley.[145]

Following discussion, I created several drafts using MP3 artifacts. The first draft was a variation on the material from *Harvest Light*. In search of something harsher, I worked with some original recordings of extended vocal techniques. I finally ended up working with samples of mechanical sounds and drum machines. I created an algorithmic texture using Sonic Pi to create textures of varying density from the MP3 encoding detritus of my sample set.[146] For the installation, I recorded 20-minutes of the output from Sonic Pi which I then live-mixed using the Kaoss Mixer, as in *freeLanguage*. The resultant 20-minute audio was looped throughout the installation as an accompaniment to the video Selam-X created.

144 https://selam-x.com/

145 Mori, Masahiro, Kari F. MacDorman, and Norri Kageki. "The uncanny valley: The original essay by Masahiro Mori." *IEEE Spectrum* (2012): 98-100.

146 Sonic Pi [Computer Software]. Version 3.1.0, 2018. Retrieved from https://sonic-pi.net/

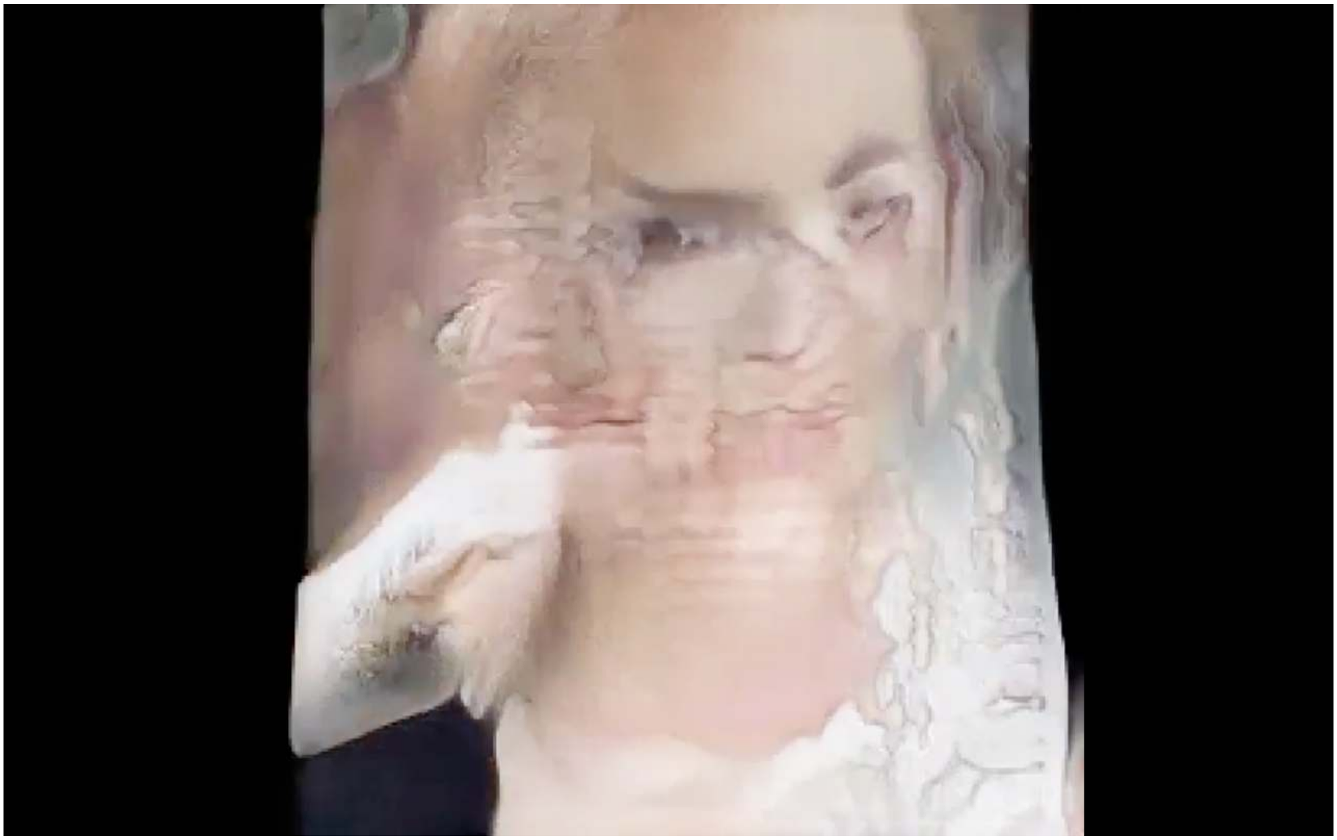

Figure 39. Still frame from *Pure Beauty* (video by Selam-X)

Resolution Disputes

I was invited by glitch artist and theorist Rosa Menkman to collaborate on the sound for a gallery installation titled *Beyond Resolution* at Transfer Gallery in Brooklyn, NY.[147] Rosa provided me with the beginnings of a sound design file in WAV format. I converted this to MP3 and found the difference using phase inversion. After mastering, I delivered the ghost to Rosa, who used it as the soundscape for the installation. Building from this work, I am currently in conversation with both Rosa and Sebastian about future collaborations.

147 Menkman, Rosa. "Glitch studies manifesto." *Video vortex reader II: Moving images beyond YouTube* (2011): 336-347.

Live/Improvisation

I am currently developing a live performance practice with hardware samplers. The first performance of this material was a headphones only concert at a gallery space. I work with a Bastl Instruments Microgranny granular synthesizer, a Teenage Engineering P.O. 33 sampler, a Critter and Guitari Organelle, and mix these signals using the Korg Kaoss Mixer. I fill the samplers with segments from my sample library of MP3 detritus, discussed below. Building from my work as a free improviser, I set flexible structures for myself through which I explore the range of sonic transformations that the hardware enables my material to undertake.

Figure 40. Headphone concert: MP3 detritus from *The Rite of Spring*

Photography, Clothing, Stickers, etc.

Extending my work with visual compression detritus, I practice photography and work with the resultant compression detritus to create a range of objects. These objects include stickers, mugs, shirts, and art prints. My goal is not only to develop a distinctive visual language but also to explore and define my own identity.

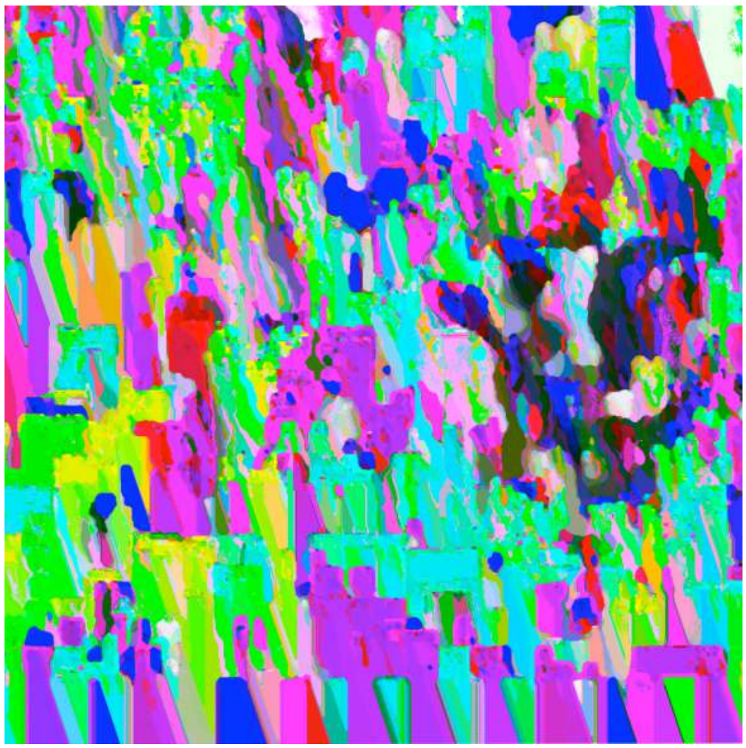

Figure 41. Processed JPEG detritus

Figure 42. JPEG detritus on shirt and background

Sample Library

I am currently developing a compression detritus sample library for use in live performance. The library is currently comprised of bass, drum, and synthesizer sounds, and fragments of recorded music. I plan to extend this to acoustic instrument samples in the future. The musical recordings are drawn freely from a diverse range of recordings. These include a compilation of 1960's British folk revival music titled *Gather in the Mushrooms*, George Crumb's *Ancient Voices of Children,* MC5's "Kick Out the Jams," and Public Enemy's "Bring the Noise," for example. Drums are sampled from a range of classic drum machines including the TR-808, TR-909, and Linn Drum LM1. Bass sounds are sourced from the TB-303 and synth sounds sampled via the SP-12, with additional synthesizer samples planned.

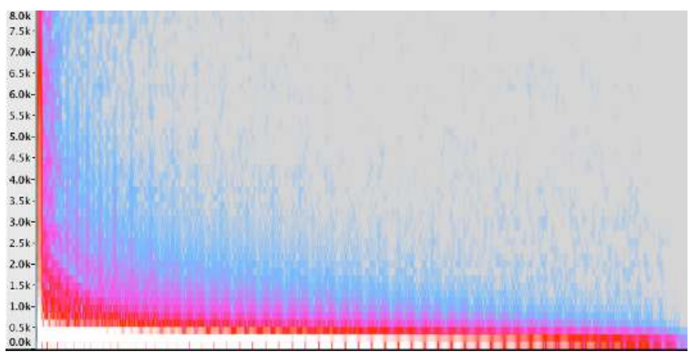

Figure 43. TR-909 Bass Drum as 16-bit PCM audio

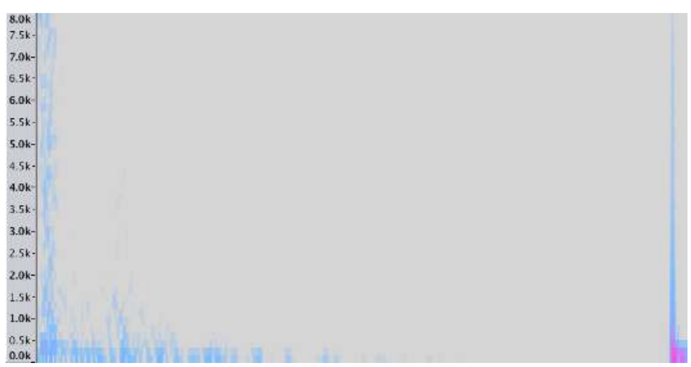

Figure 44. MP3 detritus of TR-909 Bass Drum

To consider what is lost when digital audio is compressed is to raise longstanding questions about fidelity, authenticity, and change. These concepts are debated at inflection points between old and new media: for example, when digital audio was first introduced, in comparison to the analog, when magnetic tape was adopted, in comparison to records, and when sound recording was invented, in comparison to live performance. When notation first developed, the primacy of the text came into conflict with the primacy of the aural. The field of historic performance practice is centrally concerned with the tension between maintaining fidelity to notation and allowing individual performers interpretive agency.

Of the recordings used during the original MP3 listening tests at Fraunhofer, Haydn's *Trumpet Concerto* is the only one that has a score.[148] In my work, I consider how systems (whether they be mathematical, notational, linguistic, etc.) are necessarily exclusionary in order to set a boundary between elements within and outside of the system.[149] [150] For example, in music, theoretical systems describe how individual components of a composition create various substructures through their interrelationships. A simple example of the kind of distinction made with these systems is between chord tones and non-chord tones in the analysis of tonal music. Under such an analysis, certain notes are considered to be foundational in expressing the harmony and others are considered decorative. A compressed representation of a given composition might, using such a system, describe its primary harmonic development (and thus *how it works* musically) through noting relationships between simultaneous and

148 Sterne, Jonathan. "MP3: The meaning of a format." (2012).

149 For a naturalistic account of the inside-outside dichotomy, see Moreno, Alvaro, and Xabier Barandiaran. "A Naturalized Account of the Inside-Outside Dichotomy." (2004).

150 "There was neither non-existence nor existence then; there was neither the realm of space nor the sky which is beyond. What stirred? Where?" O'Flaherty, Wendy Doniger. "trans. 1981. The Rig Veda." (1988): 343.

successive chord tones. Such a description is valuable analytically. My project, however, is to work with the material that such an account will have categorized as superfluous or decorative. In other words, I am interested, in this example, in non-chord tones.[151]

I begin with a reduction of the orchestral score arranged for trumpet and piano.[152] This reduction is, by analogy, a kind of lossy compression. I take this already compressed score and use a high resolution KIC Bookeye scanner to digitize each page of the score. I then perform a harmonic analysis of the composition, systematically noting each chord tone. I open each page individually in GIMP, a raster graphics editing software application, and erase the note heads for every chord tone.[153] Non-chord tones are left unedited, resulting in a notation comprised mostly of note stems without definite note heads. Only the excess remains, the non-chord tones, unedited. The score also contains text, written in English. In an analogous process to the deletion of chord tones, I delete the primary information bearing letters, the consonants, leaving only the inessential remainder, the vowels.[154]

---

151 There is an ongoing search in Physics for an all-encompassing *theory of everything*, a pursuit which many (Stephen Hawking, for example) consider impossible owing to Gödel's incompleteness theorem. An interesting spin on this kind of totalizing project emerged in the social sciences in the mid-twentieth century, when a group of social scientists collected massive amounts of data on the seemingly ephemeral aspects of life such as daydreams and free associations. Their database could be viewed as a precursor to the large scale data collections which corporations like Google now monetize in the early twenty-first century. For a detailed account of the so-called *Database of Dreams*, see Lemov, Rebecca Maura. *Database of dreams: The lost quest to catalog humanity*. Yale University Press, 2015. For more on theories of everything, see Weinberg, Steven. *Dreams of a final theory*. Vintage, 1994. For Hawking's refutation of an ultimate theory, see "Gödel and the End of Physics." *http://www.hawking.org.uk/index.php/lectures/91*

152 Thus my *original* has already been subjected to a kind of lossy compression, which I will further exacerbate by compression in another dimension.

153 GNU Image Manipulation Program [Computer Software]. GIMP 2.8.22, 2017. Retrieved from http://www.gimp.org/

154 Consonants carry more information than vowels in written English. As evidence, compare the following text reduced to only vowels or consonants. Only vowels: **e *ui** **o** *o* *u*** o*e* **e *a** *o*. Only consonants: th* q**ck br*wn f*x j*mps *v*r th* l*zy d*g. Which is easier to decipher? The case of speech is more complicated. See Lewicki, Michael S. "Information theory: A signal take on speech." *Nature* 466, no. 7308 (2010): 821.

Figure 45. OEO score excerpt

Exhalation

In information theory, an incompressible string is defined as a string whose Kolmogorov complexity is equal to its length, so that it has no shorter encodings. The Kolmogorov complexity, which determines the boundary between compressibility and incompressibility, defines the number of bits into which a string can be compressed without losing information. This measure is used in the formulation of Chaitin's incompleteness theorem—an impossibility result similar to Gödel's incompleteness theorem, Turing's halting problem, Richard's paradox, and Heisenberg's uncertainty principle. Impossibility results set boundaries on the knowledge available through dominant rationalist epistemologies. Often described as paradoxical, these results provide an important counterbalance to the seemingly infinite force of rationalism.

In this dissertation, I have created art from the incompressible remainders of compressive systems. I begin by imaging these remainders as breath. m*oDernisT* begins with a literal breath, Suzanne Vega's sharp inhalation, and ends with the metaphorical breath of digital compression detritus fading out in exhalation. The dissertation concludes with *OEO*, the remainders of an 18th century composition compressed by an 18th century analytical technique. This connects the reductive strategies of lossy compression to the categorical distinctions of Kant, Haydn's contemporary, and Enlightenment thought.[155]

There is value in spending time outside of the compressed sites of twenty-first century culture. Digital technology continues to asymptotically approach its limit, offering infinite progress but never

155 This connection is made more robustly, and in a context with greater ethical stakes than experimental music composition, by Ashon Crawley in *Blackpentacostal Breath*. Crawley, Ashon T. *Blackpentecostal Breath: The Aesthetics of Possibility*. Oxford University Press, 2017.

transcending its discrete nature. In recognition of the margin, we can look beyond currently constituted technologies and systems. We can imagine new configurations. We can shift between alternatives. We can prioritize the irreducible.

Code

Audio Detritus

```
# MASKING PROCEDURE

# Import Libraries
import librosa
import librosa.display
import IPython.display as ipd
import numpy as np
import matplotlib.pyplot as plt

# File path
cd /path/to/audio/file/folder/

# Import Files
a1 = 'file1.wav'
a2 = 'file2.wav'

# Load into arrays
y1, sr1 = librosa.load(a1, sr=44100, mono=False)
y2, sr2 = librosa.load(a2, sr=44100, mono=False)

# Verify same dimensions
y1 = y1[:,:y2.shape[1]]

# STFT of both channels for both files
F1l = librosa.stft(y1[0])
F1r = librosa.stft(y1[1])
F2l = librosa.stft(y2[0])
F2r = librosa.stft(y2[1])

# Difference for each channel
diffl = np.subtract(np.abs(F1l), np.abs(F2l))
diffr = np.subtract(np.abs(F1r), np.abs(F2r))

# Apply mask to uncompressed, setting threshold as desired
F1l[np.where(np.abs(diffl)<0.07)]=0
F1r[np.where(np.abs(diffr)<0.07)]=0

# Inverse STFT
ygl = librosa.istft(F1l)
ygr = librosa.istft(F1r)

# Create ghost array
```

```
ghost = np.vstack((ygl, ygr))

# Write to new file
librosa.output.write_wav('librosaTest3.wav', ghost, sr1, norm=False)


# PHASE INVERSION (using sox)

# Create MP3 from WAV
sox Example.wav -C -4 ExampleMP3VBR4.mp3

# Trim 2257 samples to align, use -1 to invert phase
sox ExampleMP3VBR4.mp3 ExampleMP3vbr4.wav trim 2257s vol -1

# Mix the original and the inverted MP3 to perform null test
sox -m Example.wav ExampleMP3vbr4.wav HopFLiMP3vbr4PhGh.wav




# HRTF Batch Processing

# Import Libraries
from pyo import *
from headspace import *
import os
from random import randint

# Sstart Server
s = Server(audio="offline")

# Define HRTF Path
def pat():
     voice.setAzimuth(pathAz.get())
     voice.setElevation(pathEl.get())

# Set file path
folder_path = "/Path/to/files"
output_folder = os.path.join(folder_path, "processed")
if not os.path.isdir(output_folder):
     os.mkdir(output_folder)

# get the list of files to process
sounds = [file for file in os.listdir(folder_path) if
sndinfo(os.path.join(folder_path, file)) != None]

# enter the batch processing loop
```

```
for sound in sounds:
     # retrieve info about the sound
     path = os.path.join(folder_path, sound)
     info = sndinfo(path)
     dur = info[1]
     sr = info[2]
     chnls = info[3]
     fformat = {'WAVE': 0, 'AIFF': 1}[info[4]]
     samptype = {'16 bit int': 0, '24 bit int': 1, '32 bit int': 2,
          '32 bit float': 3}[info[5]]

     # set server parameters
     s.setSamplingRate(sr)
     s.setNchnls(chnls)
     s.boot()
     s.recordOptions(dur=dur, filename=os.path.join(output_folder,
          os.path.splitext(sound)[0]+"_hrtf"+os.path.splitext(sound)
          [1]), fileformat=fformat, sampletype=samptype)

     # example of processing
     sf = SfPlayer(path)
     b = Mix(sf, voices=1)
     voice = HeadSpace(b, 0, 0).out()
     m = Metro(dur).play()
     m2 = Metro(dur).play()
     pathAz = TrigLinseg(m, [(0, randint(-40,40)),((dur/6),randint(-
          1,1)*(randint(0,70))),((dur/5),randint(-
          1,1)*(randint(35,70))),((dur/4),randint(60,90)),
          ((dur/3),randint(-100,30)),((dur/2),randint(int(-
          dur),int(dur))),((2*dur/3),randint(60,100)),
          ((3*dur/4),randint(30,140)),(dur,randint(-20,50))])
     pathEl = TrigLinseg(m2, [(0, randint(0,60)),((dur/4),randint(-
          30,30)),((dur/2),randint(-40,40)),((3*dur/4),randint(-
          20,80)),(dur,randint(-25,25))])
     p = Pattern(pat, .01)
     p.play()

     # start the render
     s.start()

     # cleanup
     s.shutdown()
```

```
# Verse Scrambling Procedure

# Import Libraries
from random import randint

# Define Verses
verse1 = ["v1", 666, 20333, 20, 1000]
verse2 = ["v2", 21934, 42146, 100, 2000]
verse3 = ["v3", 44146, 63394, 20, 200]
verse4 = ["v4", 63394, 82768, 200, 3000]
verse5a = ["v5a", 84996, 95671, 200, 1000]
verse5b = ["v5b", 96265, 101198, 200, 1000]
verse5c = ["v5c", 101198, 111515, 100, 1000]
verse5d = ["v5d", 112000, 116732, 50, 500]
verse6 = ["v6", 116745, 129000, 20, 1000]


verses = [verse1, verse2, verse3, verse4, verse5a, verse5b, verse5c,
verse5d, verse6]

# Scramble Verse
def verseScram(songName, verse, min, max, originName):
     seg = prepSeg(songName, verse[1], verse[2])
     scramble(seg, min, max, verse[0])
     s.recordOptions(dur=((verse[2]-verse[1]+max)/1000),
          filename="%s_%d_scr_%s_bal.wav" % (songName, max,
          verse[0]), fileformat=0)
     sf = SfPlayer('%s_%d_scr_%s.wav' % (songName, max, verse[0]),
          speed=[.99,1], loop=False, mul=.3)
     comp = SfPlayer('%s_%s.wav' % (originName, verse[0]),
          speed=[.99,1], loop=False, mul=.3)
     out = Balance(sf, comp, freq=10).out()
     s.start()

# Alternate Scrambling
def verseScram2(songName, verse, originName):
     seg = prepSeg(songName, verse[1], verse[2])
     scramble(seg, verse[3], verse[4], verse[0])
     s.recordOptions(dur=((verse[2]-verse[1]+verse[4])/1000),
          filename="%s_%d_scr_%s_bal.wav" % (songName, verse[4],
          verse[0]), fileformat=0)
     sf = SfPlayer('%s_%d_scr_%s.wav' % (songName, verse[4],
          verse[0]), speed=[.99,1], loop=False, mul=.3)
     comp = SfPlayer('%s_%s.wav' % (originName, verse[0]),
          speed=[.99,1], loop=False, mul=.3)
     out = Balance(sf, comp, freq=10).out()
     s.start()
```

```
# Process all verses
def allVerses(songName, verses, originName):
     for verse in verses:
          verseScram(songName, verse, verse[3], verse[4], originName)

# Alt Verse Process
def allVerses2(songName, verses, originName):
     i = 0
     while i < len(verses):
          verseScram2(songName, verses[i], originName)
          i += 1

# Prepare Song Segments for Concatenation
def prepSeg(songAlias, start_msc, end_msc):
     songName = "%s" % songAlias
     song = AudioSegment.from_wav("%s.wav" % (songName))
     audioSegment = song[start_msc:end_msc]
     return audioSegment

# Make verses
songName = "TomsDiner_nonRecursive_Ghost6"
seg = prepSeg(songName, 666, 20333)
seg.export("%s_v1.wav" % (songName), format="wav")
seg = prepSeg(songName, 21934, 42146)
seg.export("%s_v2.wav" % (songName), format="wav")
seg = prepSeg(songName, 44164, 63394)
seg.export("%s_v3.wav" % (songName), format="wav")
seg = prepSeg(songName, 63394, 82768)
seg.export("%s_v4.wav" % (songName), format="wav")
seg = prepSeg(songName, 84996, 95671)
seg.export("%s_v5a.wav" % (songName), format="wav")
seg = prepSeg(songName, 96265, 101198)
seg.export("%s_v5b.wav" % (songName), format="wav")
seg = prepSeg(songName, 101198, 111515)
seg.export("%s_v5c.wav" % (songName), format="wav")
seg = prepSeg(songName, 112000, 116732)
seg.export("%s_v5d.wav" % (songName), format="wav")
seg = prepSeg(songName, 116745, 129100)
seg.export("%s_v6.wav" % (songName), format="wav")
```

Image Detritus

```
# Import Libraries
import imageio
import numpy as np

# Load Images
png = imageio.imread('image.png')
jpg = imageio.imread('image.jpg')

# Crop to identical shapes
png = png[:, :, :jpg.shape[2]]

# Find max and min value pixel-by-pixel
max = np.maximum(png,jpg)
min = np.minimum(png,jpg)

# Absolute value of difference
absDiff = max - min

# Subtract absolute value of difference from black background
black = np.tile(0, png.shape)
bGhost = black + absDiff
imageio.imsave('bGhost.png', bGhost)

# Subtract Abs Val of Diff from white background
white = np.tile(255, (png.shape)
wGhost = white - absDiff
imageio.imsave('wGhost.png', wGhost)

# Simple difference mask
diff = png - jpg
png[np.where(diff==0)]=255
misc.imsave('ghost.png', png)
```

Video Detritus

```
# Import Libraries
from numpy import *
import cv2
from cv2 import VideoWriter_fourcc

# Create blank arrays
hi = []
```

```
lo = []

# Import Videos
hiCap = cv2.VideoCapture('hiClip0.m4v')
loCap = cv2.VideoCapture('loClip0.m4v')

# Load videos into blank arrays
ret, frame = hiCap.read()
while ret:
     hi.append(frame)
     ret, frame = hiCap.read()
hiCap.release()

ret, frame = loCap.read()
while ret:
     lo.append(frame)
     ret, frame = loCap.read()
loCap.release()

# Reformat the data
hiArray = asarray(hi)
loArray = asarray(lo)

# Find the difference
ghost = hiArray - loArray

# Specify encoding codec
fourcc = VideoWriter_fourcc(*'HFYU')

# Write ghost to file at appropriate resolution
out = cv2.VideoWriter('clipGhost0.avi',fourcc, 24.0, (720,478))
for fr in ghost:
     out.write(fr)
```

Interactive Detritus

```
// import libraries
import beads.*;
import org.jaudiolibs.beads.*;
import java.util.Arrays;
import gohai.glvideo.*;
import processing.io.*;

AudioContext ac; // declare our parent AudioContext
PowerSpectrum ps; // declare our future FFT display

// what file will be granulated?
String sourceFile = "data/moDernisT.wav";
String sourceFile2 = "data/tomsDinerQ.wav";
String movieFileLocation = "moDernisT_v2smaller2.mp4";
String movieFile2Location = "tomsDiner.mp4";
Gain masterGain; // our master gain
GranularSamplePlayer gsp; // our GranularSamplePlayer object
SamplePlayer sp; // Sample Player
SamplePlayer sp2; // Sample Player 2

// Movie
GLMovie mov;
GLMovie mov2;

// these unit generators will be connected to various granulation
parameters
Glide gainValue;
Glide grainSizeValue;
Glide intervalValue;
Glide positionValue;
Glide pitchValue;
Glide randomnessValue;
Glide rateValue;
Glide frValue;
Glide queValue;

// this object will hold the audio data that will be granulated
Sample sourceSample = null;
Sample tomSample = null;
Sample modSample = null;

// this float will hold the length of the audio data, so that
// we don't go out of bounds when setting the granulation
// position
float sampleLength = 0;
```

```
// this UGen will be our Band Pass Filter
BiquadFilter bpfilter;
BiquadFilter bpfilter2;

Compressor c; // Compressor unit generator

// this boolean tells whether or not the screen is being touched /
mouse is pressed
boolean locked = false;

// keytracker
boolean[] keys;

// video mode toggle
boolean videoToggle = false;
boolean ghost = true;

// floats to store values for params
float posVal;
float grVal;
float intVal;
float freqVal;
float frVal;
float quVal;
float playhead;

// movie frames
int newFrame = 0;

// waveform colors
color cursor;
color fore;
color back;



/*-----------------------------------------
---------------------------------*/

// set up the applet
void setup() {
  //fullScreen(P2D);
  size(320,240, P2D);
  colorMode(HSB, 360, 100, 100);
  cursor = color(300,100,100);
  fore = color(275, 100, 100);
```

```
back = color(0,0,0);
background(back);
stroke(fore);

ac = new AudioContext(); // initialize our AudioContext

// Create a new compressor with a single output channel.
c = new Compressor(ac, 1);
c.setAttack(2);
c.setDecay(200);
c.setRatio(100.0);
c.setThreshold(0.5);
c.setKnee(0);

// key tracking stuff
keys=new boolean[4];
keys[0]=false;
keys[1]=false;
keys[2]=false;
keys[3]=false;

// we encapsulate the file-loading in a try-catch
// block, just in case there is an error with file access
try {
   // load the audio file which will be used in granulation
   sourceSample = new Sample(sketchPath("") + sourceFile);
   tomSample = new Sample(sketchPath("") + sourceFile2);
   modSample = new Sample(sketchPath("") + sourceFile);
}
// catch any errors that occur in file loading
catch(Exception e) {
   println("Exception while attempting to load sample!");
   e.printStackTrace();
   exit();
}

// store the sample length - this will be used when
// determining where in the file we want to position our
// granulation pointer
sampleLength = (float)sourceSample.getLength();

// set up our master gain
gainValue = new Glide(ac, 100.0, 20);
masterGain = new Gain(ac, 1, gainValue);

// initialize our GranularSamplePlayer
gsp = new GranularSamplePlayer(ac, sourceSample);
```

```
sp = new SamplePlayer(ac, modSample);
sp2 = new SamplePlayer(ac, tomSample);

// these ugens will control aspects of the granular sample
// player
// remember the arguments on the Glide constructor
// (AudioContext, Initial Value, Glide Time)
grainSizeValue = new Glide(ac, 125, 50);
intervalValue = new Glide(ac, 100, 100);
positionValue = new Glide(ac, random(sampleLength), 30);
pitchValue = new Glide(ac, 1, 20);
randomnessValue = new Glide(ac, 0.5, 10);
rateValue = new Glide(ac, 1, 100);
frValue = new Glide(ac, 640.0, 100);
queValue = new Glide(ac, 0.0156, 100);

// initial param float values
posVal = 0;
freqVal = 1;
grVal = 80;
intVal = 100;
frVal = 640.0;
quVal = 0.0156;

// connect all of our Glide objects to the previously
// created GranularSamplePlayer
gsp.setGrainSize(grainSizeValue);
gsp.setGrainInterval(intervalValue);
gsp.setPitch(pitchValue);
gsp.setRandomness(randomnessValue);
gsp.setRate(rateValue);
gsp.setLoopType(SamplePlayer.LoopType.LOOP_FORWARDS);
gsp.getLoopStartUGen().setValue(0);
gsp.getLoopEndUGen().setValue((float)sourceSample.getLength());

// connect our GranularSamplePlayer to the Band Pass Filter
bpfilter = new BiquadFilter(ac, BiquadFilter.BP_SKIRT, 640.0,
   0.0156);
bpfilter2 = new BiquadFilter(ac, BiquadFilter.BP_SKIRT, 640.0,
   0.0156);
bpfilter.addInput(gsp);
bpfilter.setGain(1.0);
bpfilter2.addInput(sp);
bpfilter2.addInput(sp2);
bpfilter2.setGain(1.0);

// add bp filter to compressor
```

```
  c.addInput(bpfilter);
  c.addInput(bpfilter2);

  // connect our compressor to the master gain
  masterGain.addInput(c);
  //

  gsp.start(); // start the granular sample player
  sp.start();
  sp.pause(true);
  sp2.start();
  sp2.pause(true);

  // connect the master gain to the AudioContext's master output
  ac.out.addInput(masterGain);

  background(0); // set the background to black

  // FFT Prep
  ShortFrameSegmenter sfs = new ShortFrameSegmenter(ac);
  sfs.addInput(ac.out);
  FFT fft = new FFT();
  sfs.addListener(fft);
  ps = new PowerSpectrum();
  fft.addListener(ps);
  ac.out.addDependent(sfs);

  // start processing audio
  ac.start(); // begin audio processing

  // load movie file for later
  mov = new GLMovie(this, movieFileLocation);
  mov2 = new GLMovie(this, movieFile2Location);

  // the push buttons
  GPIO.pinMode(16, GPIO.INPUT);
  GPIO.pinMode(4, GPIO.INPUT);
  GPIO.pinMode(5, GPIO.INPUT);
  GPIO.pinMode(6, GPIO.INPUT);
}


//------------
// The Draw Loop

void draw()
{
```

```
if (GPIO.digitalRead(16) == GPIO.HIGH) {
  if (keys[0]) {
    buttonReleased(0);
  }
}
else {
  if (!keys[0]) {
    buttonPressed(0);
  }
}
if (GPIO.digitalRead(4) == GPIO.HIGH) {
  if (keys[1]) {
    buttonReleased(1);
  }
}
else {
  if (!keys[1]) {
    buttonPressed(1);
  }
}
if (GPIO.digitalRead(5) == GPIO.HIGH) {
  if (keys[2]) {
    buttonReleased(2);
  }
}
else {
  if (!keys[2]) {
    buttonPressed(2);
  }
}
if (GPIO.digitalRead(6) == GPIO.HIGH) {
  if (keys[3]) {
    buttonReleased(3);
  }
}
else {
  if (!keys[3]) {
    buttonPressed(3);
  }
}
if (keys[0] && keys[1]) {
  if (mov.available()) {
    float location = map((float)gsp.getPosition(), 0.0,
        sampleLength, 0.0, mov.duration());
    mov.jump(location + random(-randomnessValue.getValue(),
        randomnessValue.getValue()));
    mov.read();
```

```
    }
    background(0);
    if(mousePressed) {
      tint(map(mouseY, 0, height, 0, 360), map(mouseX, 0, width, 0,
          100), 256-pow(2, map(mouseY, 0, height, 0, 8)));
    }
    else {
      tint(360/2, 100, 100);
    }
    image(mov, 0, 0, width, height);
    playhead = width*((float)gsp.getPosition()/
     (float)sourceSample.getLength());
    stroke(cursor);
    line(playhead, 0, playhead, height);
  }
  else if (keys[2] && keys[3]) {
    if (mov2.available()) {
      float location = map((float)gsp.getPosition(), 0.0,
          sampleLength, 0.0, mov2.duration());
      mov2.jump(location + random(-randomnessValue.getValue(),
          randomnessValue.getValue()));
      mov2.read();
    }
    background(0);
    if(mousePressed) {
      tint(map(mouseY, 0, height, 0, 360), map(mouseX, 0, width, 0,
          100), 256-pow(2, map(mouseY, 0, height, 0, 8)));
    }
    else {
      tint(360/2, 100, 100);
    }
    image(mov2, 0, 0, width, height);
    playhead = width*((float)gsp.getPosition()/
     (float)sourceSample.getLength());
    stroke(cursor);
    line(playhead, 0, playhead, height);
  }
  else if (videoToggle && ghost) {
    float location = map((float)gsp.getPosition(), 0.0, sampleLength,
     0.0, mov.duration());
    mov.jump(location);
    mov.read();
    background(0);
    if(mousePressed) {
      tint(map(mouseY, 0, height, 0, 360), map(mouseX, 0, width, 0,
          100), 256-pow(2, map(mouseY, 0, height, 0, 8)));
    }
```

```
    else {
      tint(360/2, 100, 100);
    }
    image(mov, 0, 0, width, height);
    playhead = width*((float)gsp.getPosition()/
     (float)sourceSample.getLength());
    stroke(cursor);
    line(playhead, 0, playhead, height);
    stroke(fore);
    loadPixels();
    for(int i = 0; i < width; i++) {
      int buffIndex = i * ac.getBufferSize() / width;
      int vOffset = abs((int)(((1 + ac.out.getValue(0, buffIndex)) *
          width / 2)-1));
      vOffset = min(vOffset, height);
      pixels[abs(vOffset * height + i)] = fore;
    }
    updatePixels();
  }
  else if (videoToggle) {
    float location = map((float)gsp.getPosition(), 0.0, sampleLength,
     0.0, mov2.duration());
    mov2.jump(location);
    mov2.read();
    background(0);
    if(mousePressed) {
      tint(map(mouseY, 0, height, 0, 360), map(mouseX, 0, width, 0,
          100), 256-pow(2, map(mouseY, 0, height, 0, 8)));
    }
    else {
      tint(360/2, 100, 100);
    }
    image(mov2, 0, 0, width, height);
    playhead = width*((float)gsp.getPosition()/
     (float)sourceSample.getLength());
    stroke(cursor);
    line(playhead, 0, playhead, height);
    stroke(fore);
    loadPixels();
    for(int i = 0; i < width; i++) {
      int buffIndex = i * ac.getBufferSize() / width;
      int vOffset = abs((int)(((1 + ac.out.getValue(0, buffIndex)) *
          width / 2)-1));
      vOffset = min(vOffset, height);
      pixels[abs(vOffset * height + i)] = fore;
    }
    updatePixels();
```

```
}
else if (keys[0]) {
  background(back);
  stroke(fore);
  playhead = width*((float)gsp.getPosition()/
   (float)sourceSample.getLength());
  stroke(cursor);
  line(playhead, 0, playhead, height);
  stroke(fore);
  float[] features = ps.getFeatures();
  if(features != null) {
    for(int x = 0; x < width; x++) {
      int featureIndex = (x * features.length) / width;
      int barHeight = Math.min((int)(features[featureIndex] *
        height), height - 1);
      line(x, height, x, height - barHeight);
    }
  }
}
else if (keys[1] || keys[2] || keys[3]) {
  background(back);
  stroke(fore);
  float[] features = ps.getFeatures();
  if(features != null) {
    for(int x = 0; x < width; x++) {
      int featureIndex = (x * features.length) / width;
      int barHeight = Math.min((int)(features[featureIndex] *
        height), height - 1);
      line(x, height, x, height - barHeight);
    }
  }
  playhead = width*((float)gsp.getPosition()/
   (float)sourceSample.getLength());
  stroke(cursor);
  line(playhead, 0, playhead, height);
}
else {
  // waveform display
  loadPixels();
  Arrays.fill(pixels, back);
  stroke(fore);
  for(int i = 0; i < width; i++) {
    int buffIndex = i * ac.getBufferSize() / width;
    int vOffset = abs((int)(((1 + ac.out.getValue(0, buffIndex)) *
        width / 2)-1));
    vOffset = min(vOffset, height);
    pixels[abs(vOffset * height + i)] = fore;
```

```
    }
    updatePixels();
    playhead = width*((float)gsp.getPosition()/
     (float)sourceSample.getLength());
    stroke(cursor);
    line(playhead, 0, playhead, height);
  }
}

void mousePressed() {
  // track mouse movements for rate control
  locked = true;
  // print(gsp.getPosition()/sourceSample.getLength());
  if(!keys[0] && !keys[1] && !keys[2] && !keys[3]) {
    rateValue.setGlideTime(100);
  }
  if(keys[0] && keys[1]) {
    videoToggle = true;
    ghost = true;
    sp.setPosition((float)gsp.getPosition());
    sp.pause(false);
  }
  else if(keys[2] && keys[3]) {
    videoToggle = true;
    ghost = false;
    sp2.setPosition((float)gsp.getPosition());
    sp2.pause(false);
  }
}


void mouseDragged() {
  fore = color(hue(fore)+map(mouseY, 0, height, -5, 5),
saturation(fore), brightness(fore));
  if(keys[0]) {
    // The X-axis is used to control the position in the wave file
          that is being granulated.
    posVal = (float)((float)mouseX / (float)width) * (sampleLength -
     400);
    positionValue.setValue(posVal);
    gsp.setPosition(positionValue);
  }
  if(keys[1]) {
    // dragging mouse while holding key 1 changes grain size and
          interval
    grVal = map(mouseX, 0, width, 20, 1000);
    intVal = map(mouseY, 0, height, 20, 1000);
```

```
    grainSizeValue.setValue(grVal);
    intervalValue.setValue(intVal);
  }
  if(keys[2]) {
    // change pitch with mouse
    freqVal = log(pow(map(mouseX, 0, width, 1, 10), map(height-
     mouseY, 0, height, 1, 10)));
    pitchValue.setValue(freqVal);
  }
  if(keys[3]) {
    frValue.setValue(frVal);
    queValue.setValue(quVal);
    bpfilter.setFrequency(frValue);
    bpfilter.setQ(quVal);
  }
  if(!keys[0] && !keys[1] && !keys[2] && !keys[3]) {
    //moving the mouse changes the playback rate and direction
    rateValue.setValue(map(mouseY, 0, height, 2, -2));
  }
}

void mouseReleased() {
  fore = color(275, 100, 100);
  locked = false;
  videoToggle = false;
  sp.pause(true);
  sp2.pause(true);
  rateValue.setGlideTime(60000);
  rateValue.setValue(1);
}

void buttonPressed(int button) {
  if(button==0) {
    keys[0] = true;
    gsp.setPosition(positionValue);
    fore = color(285, 100, 100);
  }
  if(button==1) {
    keys[1]=true;
    grainSizeValue.setValue(grVal);
    intervalValue.setValue(intVal);
    fore = color(280, 100, 100);
  }
  if(button==2) {
    keys[2]=true;
    pitchValue.setValue(freqVal);
    fore = color(270, 100, 100);
```

```
  }
  if(button==3) {
    keys[3]=true;
    frValue.setValue(frVal);
    queValue.setValue(quVal);
    bpfilter.setFrequency(frValue);
    bpfilter.setQ(queValue);
    fore = color(265, 100, 100);
  }
}

void buttonReleased(int button) {
    fore = color(275, 100, 100);
    if(button==0) {
    keys[0]=false;
    gsp.setPosition(null);
  }
  if(button==1) {
    keys[1]=false;
    grainSizeValue.setValue(125);
    intervalValue.setValue(100);
  }
  if(button==2) {
    keys[2]=false;
    pitchValue.setValue(1);
  }
  if(button==3) {
    keys[3]=false;
    //reset filter
    frValue.setValue(640.0);
    queValue.setValue(0.0156);
    bpfilter.setFrequency(frValue);
    bpfilter.setQ(queValue);
  }
}
```

Bibliography

Adler, Mark. “MPEG Press Release.” Email. 1991. Retrieved from http://web.mit.edu/graphics/doc/ISO-11172-summary

Audacity [Computer Software]. Version 2.1.0, 2015. Retrieved from http://audacity.sourceforge.net/

Bailly, Jean-Christophe, and Matthew H. Anderson. “The Slightest Breath (On Living).” *CR: The New Centennial Review* 10, no. 3 (2010): 1-11.

Barthes, Roland. *A lover's discourse: Fragments*. Macmillan, 1978.

———. *Roland Barthes by Roland Barthes*. Macmillan, 2010.

Basinski, William. “The Disintegration Loops (2002-2003).” *Temporary Residence 5* (2012).

Baume, Chris and Yves Raimond. BBC Vamp plugin collection. Retrieved from https://github.com/bbc/bbc-vamp-plugins

Beauvoir, Simone. *The Second Sex.* Alfred A. Knopf, 1971.

Beckett, Samuel. “Texts for Nothing," in *Collected Shorter Prose: 1945-1980*. John Calder, 1984.

Begault, Durand R., and Larry Polansky. “The composition of auditory space: recent developments in headphone music.” *Leonardo* 23, no. 1 (1990): 45-52.

Bellona, Jon. “Physical Composition: The Musicality of Body Movement on Digital Musical Instruments.” PhD dissertation, University of Virginia, 2018.

Block, Ned. "Anti-reductionism slaps back." (1996).

Bosi, Marina, and Richard E. Goldberg. *Introduction to digital audio coding and standards*. Vol. 721. Springer Science & Business Media, 2012.

Bosma, Hannah. "Gender and technological failures in Glitch music." *Contemporary Music Review* 35, no. 1 (2016): 102-114.

Brandenburg, Karlheinz. "MP3 and AAC explained." In *Audio Engineering Society Conference: 17th International Conference: High-Quality Audio Coding*. Audio Engineering Society, 1999.

Bull, Michael. *Sound Moves: iPod culture and urban experience.* Routledge, 2015.

Cage, John. *Cage.* Cunningham Dance Foundation, 1970.

———. *Silence: lectures and writings*. Wesleyan University Press, 2011.

Carroll, Lewis. *Through the looking glass: And what Alice found there*. Rand, McNally, 1917.

Caplan, Paul Lomax. "JPEG: the quadruple object." (2013).

Cavarero, Adriana. *For more than one voice: Toward a philosophy of vocal expression.* Stanford University Press, 2005.

Chapman, Tracy. *Tracy Chapman*. Elektra, 1988, CD.

Cheng, Corey I., and Gregory H. Wakefield. "Moving sound source synthesis for binaural electroacoustic music using interpolated head-related transfer functions (HRTFs)." *Computer Music Journal* 25, no. 4 (2001): 57-80.

Chevalier, Jean, and Alain Gheerbrant. *A dictionary of symbols*. Oxford: Blackwell, 1994.

Chiariglione, Leonardo. "MPEG Press Release." 1991. Retrieved from https://web.archive.org/web/20110503174827/http://mpeg.chiariglione.org/meetings/kurihama91/kurihama_press.htm. Accessed March 18, 2018.

Chopin, Frédéric. Michel Dussault, piano.Chopin Méconnu: The Unknown Chopin.Toronto, Ontario, Canada : Les disques SRC, p1992, CD.

Clarke, Shirley, and Ornette Coleman. DVD. *Ornette: Made in America*. Directed by Shirley Clark.New Jersey: Milestone Film & Video, 2014..

Crawley, Ashon T. *Blackpentecostal Breath: The Aesthetics of Possibility.* Oxford University Press, 2017.

Denegri-Knott, Janice, and Mark Tadajewski. "The emergence of MP3 technology." Journal of Historical Research in Marketing 2, no. 4 (2010): 397-425.

Dixon, Simon. "Simple spectrum-based onset detection." *MIREX 2006* (2006): 62.

Dommes, E., H.C. Bauknecht, G. Scholz, Y. Rothemund, J. Hensel, and R. Klingebiel. "Auditory cortex stimulation by low-frequency tones—an fMRI study." *Brain Research* 1304 (2009): 129-137.

Du Bois, W. E. B., Henry Louis Gates, and Terri Hume Oliver. "The souls of black folk: A norton critical edition." *Edited by Henry Louis Gates, Jr. and Terri Hume Oliver. New York and London: WW Norton & Company* (1999).

Edmund, Mark. "Music to Whose Ears?." *Quality Progress* 48, no. 2 (2015): 14.

Ellington, Duke. *Take the "A" Train*. Vintage Jazz Classics, 1990, CD.

Eno, Brian and Anthony Korner. "Aurora Musicalis." In *Artforum*, vol. 24, no. 10, pp. 26-36. 1986.

Evens, Aden. *Logic of the Digital.* Bloomsbury Publishing, 2015.

Fisher, Mark. "The metaphysics of crackle: Afrofuturism and hauntology." *Dancecult: Journal of Electronic Dance Music Culture* 5, no. 2 (2013): 42-55.

Fritz, J.B., Mounya Elhilali, Stephen V. David, and Shihab A. Shamma. "Auditory attention—focusing the searchlight on sound." *Current opinion in neurobiology* 17, no. 4 (2007): 437-455.

Gardner, Bill and Keith Martin. "HRTF Measurements of a KEMAR Dummy-head Microphone." (1994).

Gaze, Tim, and Michael Jacobson. *An Anthology of Asemic Handwriting*. Punctum Books, 2013.

Glitsos, Laura. "Vaporwave, or music optimised for abandoned malls." *Popular Music* 37, no. 1 (2018): 100-118.

GNU Image Manipulation Program [Computer Software]. GIMP 2.8.22, 2017. Retrieved from http://www.gimp.org/

Google Groups. Accessed March 18, 2018. https://groups.google.com/forum/#!msg/alt.hypertext/eCTkkOoWTAY/bJGhZyooXzkJ.

Gramsci, Antonio. *Prison notebooks*. Vol. 2. Columbia University Press, 1992.

Greenberg, Joseph H., and James J. Jenkins. "Studies in the psychological correlates of the sound system of American English." *Word* 20, no. 2 (1964): 157-177.

Guimond, David. "The Sounds of Disappearance." *Intermédialités: Histoire et théorie des arts, des lettres et des techniques/Intermediality: History and Theory of the Arts, Literature and Technologies* 10 (2007): 115-130.

Hawking, Stephen. "Gödel and the End of Physics." Retrieved from *http://www.hawking.org.uk/index.php/lectures/91*

Haydn, Joseph. Håkan Hardenberger, trumpet ; Academy of St. Martin-in-the-Fields ; Sir Neville Marriner, conductor. *Trompetenkonzerte = Trumpet Concertos*. Schwann Musica Mundi, 1987, CD.

Hegel, Georg Wilhelm Friedrich, and Michael Inwood. *Hegel: Philosophy of Mind: Translated with Introduction and Commentary*. Oxford University Press, 2007.

Henriques, Julian. *Sonic bodies: Reggae sound systems, performance techniques, and ways of knowing.* Bloomsbury Publishing USA, 2011.

Hermans, Phillip. 2013. *Cooperation and competition in music composition*.

Hoogerwerf, Frank W. "Cage Contra Stravinsky, or Delineating the Aleatory Aesthetic." *International Review of the Aesthetics and Sociology of Music* (1976): 235-247.

India. *Sobre el Fuego.* RMM Records: RMD 82157, 1997, CD.

Isou, Isidore. "Manifesto of Letterist Poetry." *Manifesto: A Century of Isms* (1983): 545-6.

Jasen, Paul C. *Low end theory: Bass, bodies and the materiality of sonic experience*. Bloomsbury Publishing USA, 2016.

Kane, Carolyn. "Compression Aesthetics: Glitch From the Avant-Garde to Kanye West." Unpublished *article. http://ivc.lib.rochester.edu/wp-content/uploads/2014/10/Kane-Compression-Aesthetics.pdf* (2014).

Kenko, Yoshida. *Essays in idleness*. Cosimo, Inc., 2005.

Kim-Cohen, Seth. *In the blink of an ear: toward a non-cochlear sonic art*. A&C Black, 2009.

LAME [Computer Software]. Version 3.1. http://lame.sourceforge.net

Lane, Cathy. "Voices from the Past: compositional approaches to using recorded speech." *Organised Sound* 11, no. 1 (2006): 3-11.

Lansky, Paul. "Reflections on spent time." In ICMC. 2009.

Lemov, Rebecca Maura. *Database of dreams: The lost quest to catalog humanity*. Yale University Press, 2015.

Lewicki, Michael S. "Information theory: A signal take on speech." *Nature* 466, no. 7308 (2010): 821.

Lucier, Alvin, Michael Roth, Nicolas Collins, and Ronald Kuivila. *Alvin Lucier: A Celebration*. Wesleyan University Press, 2012.

Maguire, Ryan. "3D Printed Enclosures with OpenSCAD". Published November 14, 2017. https://scholarslab.lib.virginia.edu/blog/3d-printed-enclosures-with-openscad/.

———. "Creating musical structure from the temporal dynamics of soundscapes." In *2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA)*, pp. 1432-1433. IEEE, 2012.

———. 2013. *Distort them as you please: the sonic artifacts of compositional transformation*.

Marui, Atsushi, and William L. Martens. "Timbre of nonlinear distortion effects: Perceptual attributes beyond sharpness." In *Proceedings of the Conference on Interdisciplinary Musicology*. 2005.

Maus, Fred Everett. "Somaesthetics of Music." *Action, criticism, and theory for music education* 9, no. 1 (2010): 9-25.

Merimaa, Juha. "Modification of HRTF filters to reduce timbral effects in binaural synthesis." In *Audio Engineering Society Convention* 127. Audio Engineering Society, 2009.

Menkman, Rosa. "Glitch studies manifesto." *Video vortex reader II: Moving images beyond YouTube* (2011): 336-347.

Millet, Yves. “Atopia & Aesthetics. A Modal Perspective.” *Contemporary Aesthetics* 11, no. 1 (2013): 17.

Mixxx [Computer Software]. Version 2.1.5, 2018. Retrieved from https://www.mixxx.org

Moreno, Alvaro, and Xabier Barandiaran. "A Naturalized Account of the Inside-Outside Dichotomy." (2004).

Morgenstern, Christian, Anne Gabrisch, and Horst Hussel. *Galgenlieder.* B. Cassirer, 1921.

Mori, Masahiro, Kari F. MacDorman, and Norri Kageki. "The uncanny valley: The original essay by Masahiro Mori." *IEEE Spectrum* (2012): 98-100.

Murphy, Timothy S. “The Other’s Language: Jacques Derrida Interviews Ornette Coleman, 23 June 1997.” *Trans. Timothy S. Murphy. Genre: Forms of Discourse and Culture* 37, no. 2 (2004): 319-328.

Natron [Computer Software]. Version 2.0.3, 2016. Retrieved from http://natron.fr

Navas, Eduardo. "Remix theory: The aesthetics of sampling." (2014).

Ney, Alyssa. "Reductionism." (2008).

no Chomei, Kamo. "An Account of My Hut." *Anthology of Japanese Literature* (1955): 197-212.

NumPy [Computer Software]. Version 1.15.4. http://www.numpy.org/

O'Flaherty, Wendy Doniger. "trans. 1981. The Rig Veda." (1988): 343.

Oliveros, Pauline. “Acoustic and virtual space as a dynamic element of music.” *Leonardo Music Journal* 5, no. 1 (1995): 19-22.

Oohashi, Tsutomu, Emi Nishina, Manabu Honda, Yoshiharu Yonekura, Yoshitaka Fuwamoto, Norie Kawai, Tadao Maekawa, Satoshi Nakamura, Hidenao Fukuyama, and Hiroshi Shibasaki. “Inaudible high-frequency sounds affect brain activity: hypersonic effect.” *Journal of neurophysiology* 83, no. 6 (2000): 3548-3558.

Pace, Ian. "Composition and Performance can be, and often have been, Research." *Tempo* 70, no. 275 (2016): 60-70.

Plambeck, Joseph. "In mobile age, sound quality steps back." *New York Times* 9 (2010): B1.

Potter, Keith, and Kyle Gann. *The Ashgate research companion to minimalist and postminimalist music*. Routledge, 2016.

Processing [Computer Software]. Version 2.2.1, May 19 2014. Retrieved from http://processing.org

Python Software Foundation. Python Language Reference, version 2.7. Retrieved from http://www.python.org

Rabl, Christopher and Curtis Litchfield. Headspace [Computer Software]. Retrieved from https://github.com/crabl/HeadSpace

Reynolds, Simon. "Haunted audio." *The Wire* 273 (2006): 26-33.

Ricca, Alberto. Bienoise. *Most Beautiful Design*. Mille Plateaux, 2018.

Richardson, John. "Vega Meets Weill: Alienation Effects and Archaeological Strategies in 'Tom's Diner' and 'Luka'." Unpublished draft provided to author.

Roads, Curtis. "Sound transformation by convolution." *Musical signal processing* (1997): 411-438.

Roca, José. *Phantasmagoria: specters of absence*. Independent Curators, 2007.

Rodgers, Tara. *Pink Noises.* Duke University Press, 2010.

Rusinek, Gabriel. "Action-research on collaborative composition: an analysis of research questions and designs." *Musical creativity: Insights from music education research* (2012): 185-200.

Russolo, Luigi. "The Art of Noises (1913)." *Futurist Manifestos* 74 (1973).

Said, Edward W. *Orientalism*. Vintage, 1979.

Small, Christopher. Musicking: The meanings of performing and listening. Wesleyan University Press, 2011.

Smith, Julius O. "Viewpoints on the history of digital synthesis." In *Proceedings of the International Computer Music Conference*, pp. 1-1. International Computer Music Association, 1991.

Sonic Pi [Computer Software]. Version 3.1.0, 2018. Retrieved from https://sonic-pi.net/

Sonic Visualiser [Computer Software]. Version 3.0.3, 2017. Retrieved from http://www.sonicvisualiser.org

Spivak, Gayatri Chakravorty. "Can the subaltern speak?." *Can the subaltern speak? Reflections on the history of an idea* (1988): 21-78.

Sterne, Jonathan. *MP3: The meaning of a format*. Duke University Press, 2012.

———. "The mp3 as cultural artifact." *New media & society* 8, no. 5 (2006): 825-842.

———. "Space within space: Artificial reverb and the detachable echo." *Grey Room* (2015): 110-131.

Stine, Eli and Kevin Davis. *The Murmurator: A Flocking Simulation-Driven MultiChannel Software Instrument for*

Collaborative Improvisation. In *International Computer Music Conference (ICMC).* 2018.

Stravinsky, Igor. *Le Sacre du Printemps: Symphony in Three Movements*. Sony Classical, 1990, CD.

Tanizaki, Jun'ichirō. *In praise of shadows.* Vol. 24. Random House, 2001.

Tribe, Mark, Reena Jana, and Uta Grosenick. *New media art*. Los Angeles: Taschen, 2006.

Vega, Suzanne. *Solitude Standing.* A&M Records, 1987, CD.

————. "Tom's Essay." Measure for Measure. New York Times, 23 September 2008. http://measureformeasure.blogs.nytimes.com/2008/09/23/toms-essay/index.html

Wallace, Gregory K. "The JPEG still picture compression standard." *IEEE transactions on consumer electronics* 38, no. 1 (1992): xviii-xxxiv.

Warburton, Dan. "A working terminology for minimal music." *Intégral* 2 (1988): 135-159.

Weinberg, Steven. *Dreams of a final theory*. Vintage, 1994.

Willke, Helmut. "Atopia." *Studien zur atopischen Gesellschaft. Frankfurt a. M.: Suhrkamp* 58 (2001).

Witt, Stephen. *How music got free: The end of an industry, the turn of the century, and the patient zero of piracy.* Penguin, 2015.

Wu, Bin, Andrew Horner, and Chung Lee. "Musical timbre and emotion: The identification of salient timbral features in sustained musical instrument tones equalized in attack time and spectral centroid." In *ICMC*. 2014.

Xenakis, Iannis. *Formalized music: thought and mathematics in composition*. No. 6. Pendragon Press, 1992.

Young, Neil, Ben Keith, Kenneth Buttrey, Tim Drummond, Jack Nitzsche, David Crosby, Stephen Stills, James Taylor, and Linda Ronstadt. *Harvest*. Reprise, 1972.