Senior Capstone Projects

2015

Voice of the Machine

Sean Jackson Eads

Follow this and additional works at: https://digitalwindow.vassar.edu/senior_capstone

VOICE
of the
MACHINE

Sean Jackson Eads

VOICE OF THE MACHINE

CONTENTS

INTRODUCTION

Cyborg Music: Digital Voices and the Rise of the Virtual Pop Star

Singing is a fundamentally embodied act. It requires focus, attention to breathing, and awareness of the diaphragm and vocal chords. It relies on the perfect synchronicity of many biological components, which coalesce into the expressive and wondrous voice organ. Singing, and even speaking, presupposes an awareness that is a direct result of inhabiting a body. So as digital voices begin to crop up more frequently – announcing arrivals and destinations on trains and planes, speaking directions from cash registers and ATMs, sounding from devices in our pockets, and even representing autonomous entities in popular films like *2001: A Space Odyssey* or *Her*, —we should ask ourselves: What are digital voices? How are they made, and what does it mean that we have begun to deny the embodied act of vocalization in favor of something that in comparison seems soulless? What does it mean for contemporary urban cultures to have largely shifted the burden of communication to digital platforms? Digital voices are more precise, more efficient, and have more information readily available to backup whatever they say. They have begun to replace human bank tellers, cashiers, and secretaries already, but what happens when they begin to stand in place of our celebrities, on stage with our idols, and generally among the people to whom we designate huge amounts of cultural capital? In 2006, Madonna famously took to the stage during the Grammys, weaving between animated holograms depicting the singing and instrument-wielding cartoon personas of British virtual rock band, the Gorillaz. In 2012, a hologram of the late Tupac Shakur appeared on stage with Snoop Dogg and Dr. Dre at the Coachella music festival, speaking words that Tupac himself could never have said while still alive. Examples of increasingly powerful digital voices abound in popular music, which, as a self-aware, sound-driven medium, is a great place to begin to dissect the digital voice.

Our contemporary soundscape is crawling with an ever-increasing number of digitally-created, or digitally-rendered, voices. If you haven't recognized their prevalence, now is a good time to start. Begin by tuning into the radio: ask who is singing, and try to discern what means, both digital and physical, were used to achieve the sound of that voice. We readily recognize some of

the more obvious signifiers of digital manipulation like the famed plug-in Autotune, but far more digital manipulations go 'unheard,' as they imitate very accurately the physical apparatuses that give rise to certain effects like reverb, echo, and even warmth. The voice itself is yet another physical apparatus which has been entirely synthesized to a surprising degree of realness.

The leading vocal synthesizer product, Yamaha's Vocaloids, will be a central focus in this essay. A complex engine developed in the early 2000s, Vocaloids are created by extracting phonemes, the smallest unit of vocalized sound, from human vocal samples. By concatenating those phonemes and smoothing over their ruggedness, the engine creates an entirely new voice that is able to speak any combination of letters in the target language. Vocaloids were first developed for producers who don't sing and can't afford to hire singers for all of their tracks, but the brilliant tool inspired an entire new aesthetic appreciation, a fascination with a sound that is obviously cyborg-esque, yet chillingly human. Certain qualities, such as their being too quantized and too pitch-perfect, make Vocaloids obviously a non-human replica of the voice— but leave it to the Internet and file-sharing to breed an entire subculture that, through its musical exploration of this new cyborg tool, unearths something ultra human, somehow revealing or highlighting many unexplored qualities laying dormant in the acoustic human voice that we had previously taken for granted, or ignored entirely.

Vocaloids' popularity is partially due to Crypton Future Media, a music technology company that released the Character Vocal Series of Vocaloids in 2007. In this series, for the first time, individually-sold Vocaloid "voices"—commonly referred to as lower-case "vocaloids,"—were given humanlike anime representations, complete with physical characteristics like age, gender, and voice type. The characters were even attributed certain musical preferences like favorite genre and tempo.

From there, an "open-source culture" erupted around vocaloids—a term borrowed from a computer programming ideal whereby software code is openly released so that users can edit it and in turn release their updated code. The idea is that the software will continue to develop to

user needs. In this case, a user network of millions could now collaboratively contribute to the vocaloid personas' body of work. Their contributions consisted not only of songs, but also of music videos, generated by a fan-developed software called MikuMikuDance. The software allows users to create "motion data" for models of vocaloid characters and share it with other users, and gets its name from Hatsune Miku, the first character of the series, whose name loosely translates to 'first sound of the future.'

Miku became a global star. In 2009 she took the stage before a live audience for the first time in the form of a programmed, lip-syncing hologram accompanied by a band of live, human performers. It is this juxtaposition that makes the performance event so fascinating: the digital voice, physicalized in space between two very real masses of humans, no longer exists in a separate, digital landscape but moves freely through and interacts with our physical world. It suggests that digital voices can do more than offer an aesthetic alternative to, or something more efficient than, the human voice. It suggests that digital voices can enact physical change, become role models, even rally a fan-base. In the third episode of British television series *Black Mirror*, a bodiless personality called Waldo runs for office and takes second place. Waldo is voiced by a human and animated digitally, but a similar current runs through the story–a questioning of what roles must continue to be filled by humans and what roles can be delegated to machines. If a certain ubiquity makes animated characters more appealing icons to the public, perhaps a digital voice does have a shot at being a figurehead under which to organize social progress and create meaningful change. However, and perhaps tellingly, *Black Mirror* paints a darker picture of the future as mediated by machines.

Some will say that Crypton Future Media's innovation lends itself to the creation of an ideal pop-star, one who can meet and transcend the demands of a popular performer without being subjected to the scrutiny that so many celebrities experience. But if Miku becomes the standard for what pop idols ought to be, we may soon be met with a music culture so bent on the fetishization and objectification of its stars that it begins to lose currency as an artistic form used to express and come to terms with lived experience. One could say that Miku represents a vision

of *posthumanity*, a recently emerging theory contending that the entire human experience can be reimagined and articulated in full by machines. Posthumanism cites as its beginnings a general technological trend whereby information is continuously separated from information-carrying bodies, turned into data.

Posthumanism favors the utility of the voice over emotional capacity, embodied experience, and the ability to communicate, or to resist—and this is precisely what Miku can't do. While she offers the singing-and-dancing image of an idol onto which the public can project their musical ambitions and desires, she can't resist any representation of herself in media because she has no agency. One could contend that most pop-stars are stripped of agency when controlled by major labels, but the interplay between the public and private life of a celebrity has informed her image since the notion of celebrity came about. Balancing a desire to propel certain aspects of one's persona into the spotlight and to hide others away is the constant task of someone in the public eye. Consider how many celebrities have been crushed by their own fame, stepping out into the spotlight only to realize it was too bright and hot, an unsustainable perch. In the case of Hatsune Miku, there is not a separate public and private life, there will never be a moment when her fame becomes too much, as her existence is but a simulation that accumulates representations any time her persona is invoked.

Vocaloids and other digital voice tools fall into a lineage dating back to the late 19th century of digital and analogue technology that has continuously disembodied the voice in favor of its ability to reach people. Included in this progression are the telephone, the radio, the phonograph, and the microphone. Many electronic musicians will tell you (and I will agree) that the future of music is to be found in digital technology, but the fact that a certain degree of distance from the human is built into electronic music troubles them. Music technology takes the potential of our acoustic instruments and voices and carries them into regions unexplored, in between the discrete pitches of piano keys, or above and below the physical range of the human voice. It opens up our potential to create new sounds infinitely. A developer of Vocaloids, Alex Loscos, contends that a declining number of original melodies producible acoustically has shifted attention to digital

tools, and authorship has been transferred from the artist to the artistic producer, who uses technology to make something unoriginal sound fresh and futuristic.

But what do we do about the anxieties that are stirred? The anxiety that digital voices' popularity does mirror a trend toward disembodiment in electronic music? The anxiety that digital voices infiltrating our physical world might stifle the power of the human voice? Hatsune Miku's image is being propagated scandalously throughout consumer culture without demanding consent, and has already made appearances in Japan's *Playboy* as well as various user-created interactive models, in which fans can touch her body or ruffle her skirt. What might the popularity of Miku and other vocaloids communicate to young performers, who are growing up in the moment after holograms have attained celebrity status, and who may decide to emulate them?

One thing we can do is be more conscious of how the voice is employed in the music we make and listen to. Are our favorite artists inventively interacting with the technology they use, or using it as a crutch, to somehow augment a physical characteristic of the voice they feel is lacking? We can accept and even appreciate technology's infiltration of music production while still being critical. In fact, we have to in order for the voice to survive as a resistant, expressive, and embodied agent. From Billie Holiday's impassioned wailing on "Strange Fruit," to the cold, breathy melodies floating in futuristic instrumentals as in the contemporary songs of FKA Twigs, the voice has long been employed to resist dominant enclosures. The digital voice can take this role in one of two directions: it can bring the power of the voice to new levels of resistance, or it can mute and disfigure the already-muffled voices of the underrepresented. It is up to the creators and consumers of popular music to promote positive models of virtuality.

Digital voices are a strange and useful new tool. They reflect a physical reality that, whether real or imagined, reveals some fascinating truths about the digital world as created by humans. In this essay I seek to use these voices and their propagation through the popular culture to explore questions both brand new and fundamental to the human-technology relationship.

I. (DIS)EMBODIMENT IN DIGITAL MUSIC

Locating the Body in Contemporary Popular Music Production

As a musician working primarily with electronics and electronic extensions of acoustic sound, the utmost gratification is offered upon the reentry of bodies into my music: the sight of a crowd swaying and bobbing to the tracks I spent months producing in the solitude of my garage; the specialized exchanges between producer and composer which miraculously affect physical change in a digital composition; or the prized opportunity to play my music with other live players who aren't quantized to a time grid. These moments, I often think, are the reason that music exists as the social agent it is today. Music moves, but more importantly, it moves people. And in today's digital world full of digital art, it is the constant struggle of the artist to remind her listeners of the humanity buried in a piece which, to those unfamiliar with the advances of electronic music, might sound like the foreign utterances of a cyborg.

In "Human Bodies, Computer Music," Bob Ostertag locates this as the fundamental 'issue' facing electronic musicians today: "How to get one's body into art that is as technologically mediated as electronic music, with so much technology between your physical body and the final outcome, is a thorny problem" (11). His doubts about the medium resonate with the nay-sayers of digital art at large, who contend that technology so interferes with digital creation that one can't perceive the mark of an artist, and the art is indicative of nothing larger than a machine's ability to articulate 1's and 0's in different combinations. But such theories are to discount the intricate process of learning the language of computers, necessary to guide them toward a productive human-machine collaboration. The successful digital artist internalizes these processes in order to take advantage of their limitless possibility. And so the challenge becomes to foreground our advances toward musical infinity as enabled by technology while allowing the presence of the human behind the machine to remain paramount, to provide a compromise between those who want to see music explore cyborg territory and those in whom traditionalism has spurned an anxiety toward new digital forms.

For many music-makers and critics, expanding music toward its infinite potential is the most pressing task laid upon us by the advent of digital music technology. Music critic Adam Harper is one vocal proponent of that philosophy; his book *Infinite Music* seeks to indoctrinate its readers with the notion that music should be imagined and listened to based on its capacity for change, and that composing music merely amounts to the manipulation of infinitely specifiable variables to be enacted over time. His emphatically-developed system for creating works in such a way, called 'music space,' is designed to represent the infinity of musical possibility as a multidimensional space, composed of musical objects that interact in concentric and overlapping regions (see Harper).

Another proponent of Harper's conception of infinite music would have been Italian Futurist Luigi Russolo, who writes in 1913,

> Musical sound is too restricted in the variety and the quality of its tones. The most complicated orchestra can be reduced to four or five categories of instruments with different sound tones… Music marks time in this small circle and vainly tries to create a new variety of tones. We must break at all cost from this restrictive circle of pure sounds and conquer the infinite variety of noise sounds. (6)

Russolo was at the time inspired by the immense variety of machines that hummed and buzzed into the eardrums of urban life, and would go on to realize a concert of futurist noise instruments with wacky but appropriate names like buzzers, bursters, rustlers, and snorters. His was an incredibly early vision of what the future of music might be, and he passionately defended his claim that more complexity and variability in noise would give way to a more emotional, more total listening experience.

There is a certain reluctance, perhaps rooted in nostalgia, to accept electronic music's utter permeation of recording technology at face value; what people tend to refer to as 'electronic music' is only the music that has taken the electronically-manipulable variables to a certain degree of intensity. But so much more music made today in some way takes advantage of digital

technology. Digital music doesn't have to mean music that foregoes use of our beloved acoustic timbres and modes of production; it can only help carry them into regions unexplored. There is nothing shameful about shedding the limitations of our acoustic environment in favor of something more total and more exploratory, but how do we ensure that music, in composition and practice, retains its artistic value, and doesn't become some bodiless wandering through cyborg territory?

MUSICAL MODERNISM

In the introduction to *Infinite Music*, Adam Harper insists that "Musical modernists seek to maximize the possibilities of composition to the utmost degree, taking in equally both its broad and deep possibilities and those at the finest levels of detail" (5). He wants music-makers and appreciators to start making and listening to music that questions and advances just what music is, and urges us to do so by provoking possible futures of music, noting how far we've come since the Western fathers of classical music, in whose formal rules we were for so long deeply entrenched. He goes on to favor popular music (or non-Western classical music) as the premiere site of musical modernism, claiming that "[Western classical music's] tendency towards myopic traditionalism and exclusivity makes it tiny against the enormous backdrop of infinite musical possibility, which is calling more loudly than ever before" (7). It calls us to widen our imaginations, to allow through music the accommodation of new ideas and futures, potential conceptions for what will one day be.

Harper is so fascinated with electronic music specifically because it takes the limitations of acoustic and analogue instruments and explodes them. The piano, consisting of 88 discrete keys (there is no 'in-between' key which, when pressed, would sound like something between a B and a C), is more able to accommodate musical modernity when it becomes the electric keyboard, equipped with a pitch-bend wheel that allows the player to control pitch beyond the typical keyboard's capacity. Similar processes of increased control are applied across a range of musical objects in the move to electronic music, and pitch is but a single variable to be taken into account. It is clear that in electronics we will find the future of music, so it is ever more pertinent

that we come to understand the ways in which the human body is being erased and reconfigured in the art form, and that an effort be made to re-inject it with the humanity that spawned the creation of music as a creative and social practice.

Since I will be using terms like *embodiment*, the *human*, and *posthuman* in outlining these debates, we should come to a mutual understanding of what these many-layered terms mean in this context. I am writing mainly about music and sound, and so relevant discourses focus on the embodied experience of live performance versus the comparatively disembodied experience of recorded music. For Greg Corness, author of "The Musical Experience through the Lens of Embodiment," this assumed dichotomy stems from the theory that

> knowledge starts with lived experience. This knowledge includes the body knowledge understood to reside in the body, such as muscle memory, habit, and instinct, but has at its core the process of perception in which the body and mind are unified in the sensing and understanding of the world in which we are a part. (21)

Because our understanding of the world is inseparable from our experiencing it through bodies both in and of the world, we can't truly talk about the world we inhabit without talking about our bodily experience of it, and vice versa. Thus those claiming that disembodiment is fundamental to digital music creation express anxiety over losing a sense of self and a sense of the human values that make music so important to culture today. The opposition refutes this claim on the basis that we cannot experience music through anything but a human body, and so digital music retains its relationship to the body despite technologized means of production.

'The human' is a harder term to pin down, as it is constantly being defined in opposition to other terms like 'the animal' or 'the posthuman'. Leon Wieseltier offers a concise summary of these terms in his 2015 article for the New York Times, "Among the Disrupted." He posits of humanism that "The worldview takes many forms: a philosophical claim about the centrality of humankind to the universe, and about the irreducibility of the human difference to any aspect of our animality," and goes on to counter that posthumanism "elects to understand the world in

terms of impersonal forces and structures, and to deny the importance, and even the legitimacy, of human agency" (14). Wieseltier's notion of the human versus posthuman will be central to my analysis of embodiment in electronic music, and specifically of the voice's employment therein.

The voice will be a useful barometer in locating a tangible human quality in electronic music, as it at once exhibits an undeniable connection to humanity and is the site of massive amounts of digital distortion. Of course, this tension also makes the voice a site of revolutionary potential. British contemporary music pioneer David Toop lucidly articulates part of what makes the voice so compelling in "Sound Body: The Ghost of a Program":

> Voice is the sonic instrument with which we begin as humans—beginning as an intricate enfolding of inner and outer, ear, lungs, throat, skull, and mouth, abstract thought and physical projection, biology and consciousness, breath and listening—and which develops as the articulation of impulsion, feeling, word, speech, paralinguistic noise, even musicality, resonating in time, mind, and the air of open space. (29)

Today, Toop's statement faces a critical challenge to its truth: the voice is still fundamental to humanity in lived experience, but in its popular representations it is the victim of more digital manipulation than perhaps any other instrument or medium. He goes on to remind us of the innovations of the last century that have given rise to disembodiment of the voice:

> Throughout the 20th century, the voice was a prime site for the redefinition of the body in relation to the machine age, particularly during a rapidly developing era of disembodying technologies such as wireless telegraphy, radio, telephone, cinema, television, the tape recorder, electronic amplification, and the microphone. (29)

It has been noted that after the advent of the telegraph, radio, and telephone, hearing a human voice was no longer indicative of spatial proximity to a human, and that finally with the introduction of the tape recorder, temporal proximity was suspended as well, no longer requiring the presence of a human in any capacity for a human voice to be sounded (see Sterne). From there, the disfigurement of the human voice becomes a matter of distorting time and space

around the sound to affect change in its physical acoustic properties. These transformations will be looked at in depth in the next chapter, but for now suffice it to say that we have become so advanced in our manipulation of voices that the voice has joined the ranks of other instruments that can be entirely synthesized to an amazing degree of believability. Indeed, today's digital landscape is as replete with distorted and disfigured voices that once were the sign of some humanity as it is with humanoid voices, for the creation of which almost no human was required. This tension makes the voice a fascinating object of study through which to consider the complex relations informing human music-making in the digital age. The voice is becoming increasingly blurred with every generation of music-makers, and it is as important as ever that we use it critically in our art as a tool of resistance (as it always has been) to the disappearance of bodies.

EVIDENCE OF THE POSTHUMAN IN ELECTRONIC MUSIC

Consider N. Katherine Hayles's transformative book, *How We Became Posthuman*. When it was first published, it inspired a new wave of thinking about bodies and information, relying on the posthuman view that the experience of human being can be reconfigured such that it can be articulated in totality by machines: "[T]here are no essential differences or absolute demarcations between bodily existence and computer simulation, cybernetic mechanism and biological organism, robot teleology and human goals" (Hayles 2-3). This view is arrived at via historical observation, noting that information no longer relies on material bodies for its dissemination and consumption, and launching from there into a new conception of subjectivity, which erases embodiment as a crucial facet of lived experience. The implications of Hayles's view for human singing erase its significance as an expression of or grappling with felt emotion—and more, one could say, the expressiveness of music in general—and enshrines the voice as yet another means to an end, to the creation of sound alone. That view is reflected in the creation and utility of synthesized voices, which favor unfeeling and relatively inexpressive tools of sound production over the human voice.

Hayles engages with William S. Bourroughs's *The Ticket that Exploded* for its portrayal of the voice as being metonymic for bodies: "[Bourroughs] reasoned that if the body can become a tape-recorder, the voice can be understood not as a naturalized union of voice and presence but as a mechanical production with the frightening ability to appropriate the body's vocal apparatus and use it for ends alien to the self" (211). We can then say that the utterance of a phrase, musical or otherwise, no longer denotes the embodied experience of having uttered that phrase, or even desiring to do so. No reflection of subjectivity is implied necessarily by the voice's sounding, and yet humanity mostly does not regard the voice in this way, and so thoughts and opinions become easily misrepresented, or fabricated entirely by digital voice manipulators.

READING INTENTIONALITY - THE ANSWER IS IN THE BRAIN

Whatever one's outlook on the bodilessness that electronic music has come to be associated with may be, it is useful to consider how the theory of embodiment has developed in the face of technological invention. The unification of mind and body in our sensory experience privileges our embodiment in the way we come to understand processes. Therein lies what troubles so many about the performative capacity in electronic music. Playing acoustic instruments necessitates a sort of grandness of one's gestures, and those gestures tend to relate observably to the production of sound; a guitarist strums, and at that moment, the guitar's strings vibrate and produce sound. A human audience can witness this sequence and understand the causal connection between gesture and sound quite easily because of their own experience of bodily being. Furthermore the experience of strumming a guitar or seeing a guitar strummed is not inscribed—it cannot be carried to different contexts and hold the same value or meaning. This is very important in delineating between acoustic music and electronic music, which so often is inscribed and carried to different contexts. The other important difference through the lens of embodiment is that when performing using a laptop and synthesizers, the gestures required to affect sound become very subtle: the movement of a finger across a trackpad or the turning of a knob just one notch can change entire textures, rhythms and timbres. We have great power at our fingertips in electronic music, but the amount of effort perceived by a live audience is often small (Corness).

Still, Greg Corness will argue that regardless of the mode of production, the musical experience is always an embodied one. For one thing, in electronic music performance, often the lights, audience, sound system and other media become elevated. In fact, the very genre of electronic dance music (as electronic music is almost always categorized), implies that the music is secondary to other events, such as dancing. Therefore inscribed as the music may be, different contexts bring about different meanings for the music.

How do we reconcile the subtlety of gestures and the yearning for embodiment that has plagued electronic music? Corness has a theory that draws upon knowledge of a class of neurons that fire when either acting with goal-oriented behavior or observing such behavior:

> Consider that it is not the physical movement necessary for playing the instrument that is interpreted, since a large portion of the audience is not qualified to judge such gestures, but rather the intention of the performer that is being transmitted… This application of theories of embodiment may be worked into a model of musical experience that incorporates the learning of intentional gestures. (Corness 23-24)

His theory also argues for the experience of listening to music even without observing its production being an embodied one. As the perception of an act of intention engages the entire body, and intention can be read through sonic observation as well as visual, the act of listening to music is an embodied one, incorporating one's very experience of inhabiting a body into the perception of that musical event. This view is contrasted by the fact that when computers listen to music, that input is automatically transmitted into data, unemotional and disembodied, prepped for analysis by algorithms and machines. Listening emotionally is one part of the human experience that has yet to be appropriated by machines.

ARTISTIC RESPONSES TO THE EMBODIMENT DISCOURSE
One artist of particular interest is computer musician and vocalist Holly Herndon. Through her work she often seeks to reconcile some of the anxieties surrounding embodiment facing

electronic musicians and appreciators today. Her piece "195," created as part of her studies in electronic music and recording media at Mills College, is a composition for voice and laptop that unifies digital and acoustic performance in a live setting. Herndon says of her piece:

> I used spatialization as a way to draw attention to what is understood as embodied sound in vocal performance. In the middle section the singers start out singing without amplification. Together, they raise the microphones and suddenly the sound is both localized and removed. The audience sees their mouths open yet hears a sound coming from behind. This is then repeated with the sound appearing from the front. This was not meant to disorient the audience; it was simply meant to explore the effects of spatialization and amplification on perception. (Herndon)

By working the contemporary discourse around electronic music into the compositions themselves, Herndon is doing something very important for the medium. She demonstrates a unity in vocal performance and use of electronics by highlighting how codependent they can be in a live performance setting, not privileging one over the other, but celebrating their synchronicity.

Perhaps more compelling than Herndon's work done in an academic setting is that which has reached popular ears. Her song "Chorus," hailed by Pitchfork as best new song in January of 2014, employs the voice as inhabitant of the musical environment as well as the building block of that landscape. Assembling vocal samples from Youtube, Skype, and the detritus of the Internet, she creates a musical backdrop that transcends the traditional role of the voice in electronic music, thus forcing a kind of awareness in the listener of the expectations we tend to have of the voice, the way it is actually used today, and how we might go about employing and perceiving the voice in future musics.

I had the pleasure of seeing Holly Herndon perform in Los Angeles in January of 2015. As her live performance informs my reading of her work, I highlight a few moments now: Though Herndon's laptop seems to be on a predetermined path, the main elements of her performance appear to be improvisational. Herndon's vocals consist of strange bursts of melodies and glottal

attacks, clicking of the tongue and smacking of the lips—sounds that the microphone, by popular music standards, is supposed to repress. These unwanted sounds Herndon tends to amplify and explore daringly. Even when Herndon launches into "Chorus," the recognizable elements from the released audio track are sparsely peppered throughout a rhythmic backdrop, reducing the song's 'hook' to its smallest constituents. These small molecules of familiar sound offer something for Herndon's fans to grasp onto.

Later, Herndon chants the words "I am not home" while a low-frequency oscillator interferes with her voice, reducing the process of vocalization a pulse of humanoid sound and obscuring any lexical meaning in her voice. Its ability to communicate is shattered, and we the audience are left relying on our sight to determine what she is saying—and if you're listening at home, to a track without performance, you don't get that element of it at all. The lyrics are turned meaningless outside of the performance context. It is this sort of direct implication of technology and materiality, absence and presence, that is so key to Herndon's work, and that makes it so intriguing.

Alongside art like that of Herndon, efforts are being made to create new technology that addresses the body in electronic music. Consider Imogen Heap's "The Gloves Project." She and her team are creating a pair of gloves that work hand-in-hand with music technology to create a more cohesive composition and performance experience for herself and her audience. Her website discusses the new instrument-controller hybrid with some of the same points of interest we have seen thus far:

> Using a unique gestural vocabulary, motion data-capture systems, and user interfaces to parameter functions developed by Imogen Heap and her team, artists and other users will be able to use their motion to guide computer-based digital creations. The Musical Gloves are both an instrument and a controller in effect, designed to connect the user fluidly with gear performers usually use, such as Ableton. (Imogen Heap)

Artistic responses to the disembodiment of electronic music abound, but it's innovations such as Imogen Heap's, which celebrate technology by merging it with the physical generation of sound itself, that will define the way we make music in the future. Imogen Heap herself has said of the gloves: "I always longed for more expressive control of the tech in studio and on stage, something I could wear and create sound fluidly with, more organically, humanly somehow" (Imogen Heap). She and her team have committed years to creating a technology that demonstrates gesturally the changes that can be enacted upon sound variables and incorporates variability as detailed as that allowed for by digital manipulation. It is a complete reversal of the material-to-digital shift the 20th century nurtured, suggesting that the next trend may be in adjusting material artistic technique to more intuitively take advantage of the technology available to us today.

WORKS CITED

Corness, Greg. "The Musical Experience through the Lens of Embodiment." Massachussetts: The MIT Press. 2008.

Harper, Adam. *Infinite Music.* New York: John Hunt Publishing. 2011.

Hayles, Katherine. *How We Became Post-Human.* Chicago: University of Chicago Press. 1999.

Herndon, Holly. "195." Holly Herndon. N.p., n.d. Web. 15 Jan. 2015.

Imogen Heap. "The Gloves." The Gloves. N.p., n.d. Web. 20 Dec. 2014.

Ostertag, Bob. "Human Bodies, Computer Music." Massachusetts: MIT Press. 2002.

Russolo, Luigi. 1987. "The Art of Noises." New York: Pendragon Press. (Originally published in 1913.)
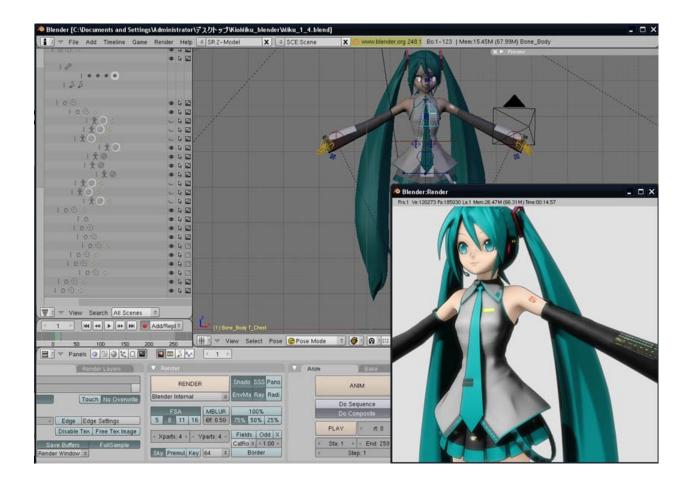
Stern, Jonathan. *The Audible Past: Cultural Origins of Sound Reproduction*. North Carolina:
    Duke University Press. 2003.

Toop, David. "Sound Body: The Ghost of a Program." Massachusetts: MIT Press. 2005.

Wieseltier, Leon. "Among the Disrupted." The New York Times. The New York Times, 17 Jan.

    2015.

II. UNPACKING VOCALOIDS

Digital Voices, Physical Reverberations

"*How does the voice survive in the digital 21st century, in a period when voices can be digitally synthesized with a convincing degree of humanness and when the recording machine is no longer a dedicated technology but a complex processing device that can rapidly generate a huge quantity and variety of sonic material from a single vocal syllable without demanding physical effort?*"

-David Toop, "Sound Body: The Ghost of a Program"

It bears significance on our understanding of the development and dissemination of digital voices that we understand the sound-reproduction technologies that preceded them. Though we can manipulate digital voices to do strange and wondrous things far beyond the capabilities of the

human voice, it is undoubtedly the human voice and its unparalleled ability to communicate speech, emotion, and of course music that spawned the interest in digital voices, that gave rise to a class of digital tools that could approximate with a computer the powerful act of verbalization hitherto reserved for human being. For a definition of what I mean by digital voices, I offer: a digital voice is any voice created or rendered by digital means, be it complete digital synthesis or merely applying digital effects to the recorded human voice; the moment we apply digital tools made widely available to contemporary producers to the human voice, it becomes relevant to this essay. Under this view, almost all contemporary popular music uses digital voices.

Often cited for its reverberating effects on human perception of sound and hearing is the technological boom beginning in the latter half of the 19th century and stretching to the beginning of the 20th, which brought with it inventions, some still in use today, such as the telephone, phonograph, wireless telegraph, and the radio. Great minds like those of Alexander Graham Bell, Thomas Watson, Thomas Edison, and Guglielmo Marconi would take previously existing technology and apply it to sound-reproduction, changing the human understanding of sound forever (see Sterne**)**.

Today we can see with clarity that these technologies had the effect of—while extending its reach eons beyond its previous potential—disembodying the voice. The telephone changed our spatial relationship to the voice—one could now hear another human's voice without the element of proximity. The phonograph changed our temporal relationship to the voice—one could now hear one's own voice, sounding from a time past. The radio allowed a voice to be broadcast across space and time.

The height of the disembodiment brought about by sound-reproduction technologies is covered heavily by Jonathan Sterne in his momentous book, *The Audible Past*, which states in its last chapter that the defining quality of early sound recording was its ability to preserve the voice beyond the death of its speaker: "We now dwell without comment among these voices of the dead. The sounds of many dead musicians and singers have casually graced my ears in the time

spent writing this book—they commingle with recorded music made by artists still living. If this experience is unremarkable today, it seems as though it demanded commentary one hundred years ago" (289). How many people can cite as one of their favorite singers or musicians those already dead? It seems commonplace to do so, what with so many great figures of musical history having come and gone, leaving traces of their greatness on tape, vinyl, and CD for later generations to hear and treasure.

If our ability to hear the dead was remarkable then, remarkable now is our ability to hear the never-living, through vocal synthesis. While it is true that our definition of digital voices includes those originated in the human and rendered digitally, a special class of digital voices called voice synthesizers is key here. The research of Alex Loscos, and his 2007 dissertation, "Spectral Processing of the Singing Voice," explain among other things the origins of Vocaloids, a product Loscos helped develop, and its precursors in digital technology. We may speak about digital voices at large and not just vocal synthesis as having the effect of giving voice to the never-living, because as Loscos writes, "singing voice coding and synthesis tend to converge. The explanation is that, despite their different goals, both share the same challenge, which is the quest of a parametric representation of a voice production model" (9). The translation of natural human processes that give quality and character to the singing voice into parametric terms that can be coded digitally is the work that makes digital synthesis of physical voices possible. In a favorite section of Loscos' text, he describes in this way the quality of breathiness in a voice: "Breathiness is a very different pathology to hoarseness. The main character of a breathy voice is not the roughness, but the breeziness, where harmonic components hide under the turbulent noise" (98). Through careful study of the characteristics we tend to refer to as "breathiness," "hoarseness," or "growling," Loscos quantifies the main components of those descriptors and makes them reproducible digitally—this is called spectral synthesis, where the sound imitated is not modeled physically, but realistic vocals are generated digitally from but a few samples.

The path to an integrated digital environment for producing vocals, as observed by Loscos, is said to have begun with the effect of 'doubling' invented by Les Paul using a reel-to-reel tape

recorder. Paul then also framed the physical model for the effect today known as echo or delay. Shortly thereafter, the echo chamber effect, achieved by placing a loudspeaker and microphone at different points in an architecturally resonant chamber, came into popular use. Following that, effects like flanger, chorus, and vibrato at the hands of Abbey Road engineer Ken Townshend became possible using automatic double tracking, which creates synced duplicates of a vocal track. By making slight variations to those tracks, the desired effects could be achieved, often giving rise to spatial or temporal illusion (Loscos).

However, far simpler to reproduce than the often complex physical and acoustical environments necessary to attain desired results, digital environments soon became ideal for working with the voice. Software solutions began to replace hard-wired effects: "New software started releasing digital replicas of each and every preceding voice effect, but almost immediately after established as the source of a vast number of new vocal transformations" (Loscos 28). One of the new vocal transformations enabled in the digital environment is the notorious Autotune, developed by Antares, which separates a sung melody into discrete pitches and tunes them to the closest target frequencies, a task too large for any physical model to accurately achieve.

Once the digital capabilities approximated and outshone those of physical acoustical modeling (save the opinion of analogue purists, who will still go through considerable measures to achieve results physically), it was left to the Internet's far-reaching tendrils to popularize digital means of production on the voice:

> Powerful computers, reasonably good and cheap sound cards, and simple availability of digital audio software have spread, democratized, and in some ways, globalized music production up to the point where bedroom musicians all around the world produce their different style songs with latest cutting-edge technologies and using powerful voice enhancement and transformation software. (Loscos 28)

The ease with which producers can now enhance and transform the human voice, combined with an ever-declining number of original melodies, Loscos hypothesizes, lead to a popular culture

that values vocal production as much as vocal talent. Innovative production provides in a sense access to new sonic spaces, "bringing about a talent shift that moved creativity from the artist to the artistic producer."

The logical next step in the process of eliminating physically-modeled technology from the integrated digital environment was to remove the singer in totality, and thus vocal synthesizers like Vocaloids were born.

But what are these samples? To understand how digital voices bear on physical reality, we should look into the analogue moments from which spawn the infinite possibility that is vocal synthesis. As the leading vocal synthesis tool and one that has provided successes rich for analysis, I focus on Vocaloids, which have also come to represent radical possibilities in popular music production. The process, as described in the commercially released "VOCALOID—Commercial singing synthesizer based on sample concatenation," by Hideki Kenmochi and Hayato Ohshita, leads to the creation of a "database of samples extracted from real peoples' singing. The samples must include all possible combinations phonemes of the target language." A script is developed, which real singers perform into a microphone such that all phonemes are uttered. These phonemes are extracted and, by a complex algorithmic process, smoothed and concatenated in such a way that they can produce any word, sung on any pitch.

Vocal synthesizers, and vocaloids being the leading version of that technology, were of course sought-after by producers, who saw in them the opportunity to create vocal tracks to fit any situation or musical context without having to hire a real singer or find existing vocal tracks to rework. But is something lost in this movement, from value being placed on the vocalist to the producer? From using human voices and modeling effects physically to using digital voices and spectral processing? Virtual idol and Vocaloid mascot Hatsune Miku will help us answer the question of what is lost and found in the shift, and how digital changes can come to bear on our physical world.

IN AND OF THE DIGITAL LANDSCAPE

New digital possibilities for the voice are now fully integrated into both commercial and underground music production and perhaps even taken for granted. Adam Harper's article for The Fader, brilliantly named "The Evolution of the Voice in the Digital Landscape," outlines the ongoing transformation human voices are currently undergoing, tracing the voice's role from that of an instrument that "differentiates itself from the environment around it, musical or cultural, either drawing power from this context or somehow standing against it," to one that "fuses with the landscape itself," no longer a distinct force within a world created for the very purpose of being inhabited by a voice. The artistic innovations Harper enumerates that have driven the blurring of this line include chopping, stitching, slowing, auto-tuning, and sample-mapping, processes which, when used to a particular degree especially adored by electronic producers and listeners, eliminate the voice's "capacity to communicate lexical meaning" rendering them merely textural or rhythmic, revealing therein certain extra-human qualities hitherto unexplored in acoustic singing alone.

Harper's outlook on the situation is relatively optimistic, despite the vague threat posed by the increasing prominence of digital voices to our needing human singers. He reminds us not to think of the digital world as closing in on humanity but rather as made up of humanity and eventually, indistinguishable from it. So while we cannot stop the momentum of the digital revolution hurdling rapidly toward total synchronicity of man and machine, we can remind ourselves that in our digital creations are reflected certain fundamental properties of humanity, that no matter how farfetched and cyborg-like it becomes, our technology does and always will extend from the human. Other important intellectual contributions, however, will ask us to think more critically about digital music's use of the voice as data, freely manipulable beneath the surface of its aural representation, which can be remixed to sound and speak in any way producers desire.

But is the humanity that our technology reflects a post-humanity? To go deeper, we should look into the open-source culture that erupted around the release of vocaloids, because it at once suggests a massive movement of humans and takes place digitally. The term 'open-source'

originated in computer programming, as programmers would put out the code behind their software for users and other programmers to view and modify, under the assumption that all edits and revisions would also be released open-source so that the programs could continue to improve in response to the needs of users.

When applied to artistic industries, the term is less intuitive, but refers to the same effect. Vocaloids became part of open-source culture in 2007, when Crypton Future Media announced the Character Vocal Series. This series would take the existing vocaloid 'instruments,' and attach to them anime character drawings, complete with age, gender, physical attributes, favorite genres of music, and best vocal range, as if the singing synthesizers had human characteristics. The first of these characters was Hatsune Miku. I use the gender pronoun 'her' to refer to Miku hereafter for ease of writing, but gender is among the characteristics we must scrutinize later in understanding vocal synths' reverberations in the physical world. Miku is today considered a 'global transmedia franchise' because regular consumers can use her voice and image to whatever ends, artistic or commercial. Her status as a proprietary symbol that accumulates a body of work generated by users is what designates her an object of open-source culture.

The open-source culture around Vocaloids can place value on some bizarre rituals. Take the song "Creepy Toast," originally uploaded to vocaloid forums by producer Circus-P in 2010 using the English-language voice of vocaloid Megurine Luka. The song begins: "Lying all alone / next to my cell phone / and an ice cream cone / is a tiny piece of toast. / Staring at me, / Looking angry. / Wait just a second, hey! / How does that thing have eyes?" These off-kilter lyrics and the eerie harmonic structure of the song struck a chord with the vocaloid community, which began to populate forums with covers and PVs--vocaloid community lingo for fan-made promotional videos--created through any manner of animation, including but not limited to the software MikuMikuDance, which, according to a wiki site dedicated to the program alone, "is a freeware animation program that lets users animate and create 3D animation movies of the Vocaloids… [and] allows users to make 'motion data' for the model which can be exported and imported into other users' workspace." A few PVs for "Creepy Toast" have been created in this way, and some

using more traditional animation techniques, but one thing is invariant: each video begins with a close-up on a cell phone, an ice cream cone, and a scary piece of toast sitting together on a table ("Creepy Toast").

The desire to see these never-living voices inhabit human bodies and dance like humans, even if on screen, is a direct reflection of the larger cultural desire to materialize our digital creations, and to bring embodiment back into an equation in which it seems such an utterly distant value. In fact, one of Vocaloid's cleverest marketing ploys was to give these previously faceless voices anime depictions. Suddenly, vocaloids were no longer just a music production tool, but a fan-driven story world, openly edited and contributed to by a potential user-network of millions.

HATSUNE MIKU THE POP STAR

Perhaps most fascinating about Hatsune Miku is the pop sensation she blossomed into shortly after her release as a piece of software. Today, Miku regularly gives live concerts in Japan, and these performances and those of other vocaloids are perhaps today's finest indicator of digital sound culture's shift to materiality. In these performances, which began in 2009, Miku is a brilliantly animated hologram, choreographed by computer programmers to dance enthusiastically along to her infectious pop tunes as she mouths the words between a live band of human instrumentalists and an awestruck human crowd.

In "Evolution of the Voice," Harper touches on the forming continuum between digital and material, noting that

> The internet is not just a landscape made up of technology, it is one made up of humanity. Like it or not, Facebook and Twitter are made of people, and of people's voices in particular. People's voices and the lives behind them reduced to assertions, opinions, arguments, anxieties, reduced to pieces of data, to samples—but sounding out nonetheless. And this landscape is reflected and explored in the music it produces, where the lines between the human and its technological environment are no longer clearly drawn.

For Harper, the advent of Miku's appearing in 3D before a live audience is only an outward reflection of internal processes that have been growing in popularity in music production and in the digital world at large. If this is true, that behind the technology is always indicated humanity in some form or another, then it is our duty to find the humanity in Hatsune Miku. Certainly, it doesn't lie in the singer from whom Miku's voice is derived, Saki Fujita—the avatar no longer bears any resemblance to her, and it would be wrong to assume so, for in any instance in which Miku is employed as a voice or a body, her connection to Fujita is purposely obscured. The humanity behind Miku, then, is a collaborative one, reflective of the desires and demands of the digital community and the culture that has defined her.

IDENTITY OF DIGITAL VOICES

To further demonstrate the changes in physical representation that can be brought about by something like spectral processing, I zoom in on a section of Loscos's text entitled "Identity transformations," those transformations "whose main goal is to modify the person behind a singing performance. These transformations do not target a specific singer sound alike but they rather pursue a generic resemblance to communities such as men, women, kids, aged, or even robot or beasts." Loscos goes on to explain the components of transformations relating to gender, age, and exotic alienations. These transformations don't really mean much until we begin to think about the example of Hatsune Miku, who, being a hologram, a non-existence, is genderless and ageless. Yet Yamaha's marketers identify Miku as a 16-year-old girl. The implications of doing so are far-reaching, but suggest that demographic to be appealing to both the producers meant to use vocaloids and the target audience of those producers' work. The notion that spectral processing gives a user control of the singing voice's identity is something to dwell on, for it means that our perception of something as embodied as gender or age can be altered via digital manipulation of sound.

Gendered representations of digital voices beg careful consideration. While one may contend that it is better that non-human Miku be the subject of media scrutiny and objectification than, say, a real 16-year-old girl, Miku's inability to resist her media representation is at the very least

troubling; it leaves the body to be acted upon and manipulated in unsavory ways. In this vein, Hatsune Miku's image is often propagated scandalously throughout consumer culture. Such are the dangers of an open source celebrity. The anxieties, biases, and prejudices that exist around the material body have followed into digital-materiality. Those who truly care about the future of electronic music must push back against this type of appropriation of the voice.

In what is often considered the male-dominated world of electronic music, appropriation of the voice is particularly rampant. It is all too common a practice for a DJ (usually male) to take the honest, artistic expressions of female vocalists, strip them of their context and the embodied experience of singing them, and disfigure and relocate them in new digital landscapes which can alter the perceived meaning of verbalization to a surprising degree. The continued use of vocal samples to that end calls on us to be more aware and more critical of the voice's employment. The voice has a long history of representing opinions, thoughts, feelings and agency. Our recent technological innovations have ended the primacy of that relationship, and it is the duty of artists working in the digital medium to ensure its use in creative ways that do not perpetuate or stand for the dynamics that allowed for a gendered configuration in the first place.

MATERIALIZING THE DIGITAL

> The digital "revolution" of our time has inalterably changed the shape of the material world: forms that were once difficult or even impossible to make are now readily achievable; new materials challenge our preconceived notions of value...
>
> -David Revere McFadden, *Out of Hand: Materializing the Post-Digital*

I conclude this chapter with a discussion of materializing the digital, a thread which runs through many a medium today, largely because I view digital voices to fit snugly into that trend. Sound being a physical force, it is materialized at the moment of production, even if digital. New forms like virtual pop stars, and other phenomena which I will discuss in the next chapter, have taken this trend to a new degree of visibility in the sonic realm.

In 2014, the Museum of Arts and Design hosted an exhibition called Out of Hand: Materializing the Post-Digital. The exhibit focused on a movement in visual culture which combines the structural potential of digital creation and recent innovations that bring digital forms into physical reality (in this lineage is laser technology, 3D printers, digital weaving), to create new physical forms hitherto unthinkable. These new forms have the benefit of being meticulously accurate and optimized at a microscopic level untouchable by human hands, being able to enact change over long periods of time, and reflecting outwardly and artistically the fascinating processuality that goes into digital production.

In an age when nearly all media has gone digital or virtual in some manner, the materialization of the digital is no doubt a reaction to that momentum, a desire to see our digital inventions realized 'in the flesh,' and enacting physical change. Much of the work compiled in MAD's exhibit in some way addressed this desire, either from a standpoint of appreciating the vast potential newly at our fingertips or drawing attention to some anxieties provoked by this invasion of alien forms into our physical world.

Our music should act accordingly. The introduction of Hatsune Miku as a culturally-relevant persona springing forth from a digital voice presents us with wonderful opportunities if we treat Miku respectfully; that is, if we use her image, voice, and popularity to promote awareness of our cultural moment and the possibilities and dangers inherent to it. However, as we've seen, it is all too easy to use Miku to unproductive, even harmful ends. So as we begin to create within this paradigm and outline the art of the 2020s and onward, we have to continually ask ourselves, what are we doing when we materialize digital forms, particularly in the world of sound? Are we using the new technology to perpetuate bias and negative representations, or are we exploring uncharted regions of music space to enhance musical modernism? Are the sounds we hear giving us a false sense of physical reality, or are they being employed thoughtfully to make us critical of the way we engage with digital media?

WORKS CITED

"Creepy Toast." Vocaloid Wikia. N.p., n.d. Web. 20 Jan. 2015.

Harper, Adam. "System Focus: The Evolution of the Voice in the Digital Landscape." *The Fader*, July 2, 2014. http://www.thefader.com/2014/07/02/system-focus-the-evolution-of-the-voice-in-the-digital-landscape

Kenmochi, H. et al. "VOCALOID – Commercial Singing Synthesizer based on Sample Concatenation," Proc. INTERSPEECH 2007, pp.4011–4010, 2007.

Loscos, Alex. *Spectral Processing of the Singing Voice.* Ph.D. Thesis submission to Pompeu Fabra University, Barcelona, Spain, 2007.

Lacbaco, Ron ed. *Out of Hand: Materializing the Post-Digital.* London: Black Dog Publishing. 2013.

Stern, Jonathan. *The Audible Past: Cultural Origins of Sound Reproduction*. North Carolina: Duke University Press. 2003.

Toop, David. "Sound Body: The Ghost of a Program." Massachusetts: MIT Press. 2005.

III. ADDITIONAL PROJECT FRAMING

Internalizing the Digital: All Cyborgs

The digital voice is affecting society at many levels: the birth of an idea, the mechanics of uttering it, the timbre of verbalization, the expressive potential of the voice, amplification, recording, conversation, and at the very construction of human being. What I aim to suggest is that few today are equipped to navigate modern human being without the aid of some services, devices, and appliances which take our words and transport them across space and time. These technologies are extremely useful and therefore built into the very way we communicate. Once we utilize any number of them, our voices are entered into digital channels that turn them into bits, phonemes, information, and data at some step along the way. And once our voices are in that system, we shouldn't be surprised to see them manipulated in an increasing number of ways both beneath and at the surface of representation.

It is therefore perhaps more productive to think of the digital voice as comprising a spectrum than as something defined in opposition to the human voice. Few human voices truly remain untouched by technology, be it at the level of physical operation or at the level of manipulation, so we should think of our voices as fluctuating on a spectrum of mediation. Of course, most can still seek out private moments of unmediated vocalization, but it is likely that at some point or another, your voice will be turned into data.

This is the reason I have casually mentioned that almost all voices in popular music are digital voices. To record a voice today without digital media is not cheap, practical, or accessible. To record a voice today without any kind technology is impossible. To compose and write music without the aid of technology is unlikely. And this isn't even to mention the fact that many popular artists rely heavily on digital voice manipulation, to the point where digital voice effects are built into the compositions themselves.

On average, popular voices, the voices which reach and speak to the largest number of people, are increasingly representative of something other than humanity. Whether it is a cyborg humanity, a post-humanity, or a post-digital humanity is a matter of framing, but it is safe to say that if we accept how powerful these voices are, we should become more cautious as to who they say they represent, how they convey a sense of space, person, or time, and who they actually serve.

In the same way as human voices are gradually becoming more like cyborg voices, we are increasingly seeing digital technology internalized. When we accept the notion of materializing the digital as not a series of isolated incidents but rather a full-fledged paradigm shift, what we thought were two clearly distinguishable realms—separated, for example, by screens—begin to meld together.

We see the notion of a physical-digital separation challenged in the artist Richard Dupont's digital sculpture, *Untitled #5,* which enacts the visual concept of a glitch—a digital error or miscalculation—upon the body. An impossible form stands before viewers of *Untitled #5*, acted upon by forces beyond physics and held in equilibrium there. Suddenly the material fabric of every day life can be sculpted digitally and made to seem as though it were naturally this or that way. Therein lies an example of how digitally-mediated reality no longer relies on the presence of a computer screen or some other technology in the moment at hand. My encounter with *Untitled #5* was devoid of a literal digital presence, but was very much a digitally-mediated event, simply for the fact that the form standing before me could not have been without the interference of some digital process, though that entity was no longer present or discernible in the space itself. Following from this, we can see how our culture begins to internalize the digital presence, aware that materiality no longer connotes an absence of digital intervention, but gains its power from the constant, dynamic interplay between physical, digital, and digital-physical forms.

ADVANCES IN SOUND CULTURE

In addition to rumination on digital voices, my project is shaped and informed by current trends in sound culture like 3D audio, portability, and voice prosthesis, all of which point to increased internalization of digital media. These and other phenomena form the framework for a lot of recording that is being done today, its goals and methods. In order to establish a backdrop in which to situate my own recorded project, I offer a brief commentary on these trends.

Corporations researching augmented reality aid the internalization of digital media. As we allow computers to directly affect our perception of the physical world by positioning them as extensions of our eyes, ears, noses and mouths, we accept the notion that virtual reality is being seamlessly woven into the fabric of our material lives. 3D audio is a phenomenon which has directly attended advances in virtual reality, as it is a necessary component in order for the illusion of reality to remain effective. The basic mission of 3D audio is to recreate through recorded media the experience of listening in the world, in its full dimensional splendor. To do this, different signals must be sent to the left and right ears in order to approximate the physical reality of how our ears perceive sound—that is, the slight delay between a sound reaching each perspective ear, and the slight volume difference due to one ear being closer to the source. Current studies lead by rocket scientist Edgar Choueiri are trying to mimic 3D audio using loudspeakers, which is an even tougher task than it is using headphones, as the signals cross and confuse one another when using loudspeakers. So to combat this, the technology currently relies on a head position-tracking sensor, in response to which the signals sent to each ear are fine-tuned. Evidenced therein is yet another way music and sound production is attaching itself to the human in order to augment not only the creation process, but the experience of listening itself (Gopnik).

3D audio is also important to my project because it addresses directly the (often false) spatialization of sound: "the act of leading sounds and frequencies in order to be reproduced in different/dedicated speakers in a multichannel system is just a way of designing, transforming and giving shape to the whole experience that is given to the audience. It's about creating not

only sound effects/objects but fabricating spaces with them, like feeding time with signals in order to build sequences of spatial fantasies" (Isaza). Bearing in mind that 3D audio will soon make it possible to vividly recreate a natural soundscape in one's bedroom, to throw a voice into the equation will mean placing versions of ourselves in fantastical, far-reaching places, both real and imagined. This is a method I undertook with *Voice of the Machine*: putting in space the voices that populate the recording, I was able to enhance its message with added context and information.

Another leading trend in digital sound culture is defined by a battle between increased fidelity and increased portability, both in terms of hardware like Walkmen, CD players, and iPods, and in terms of file formats, like the ubiquitous yet troubling MP3. The fight for fidelity is fueled by an aesthetic desire—of course it makes sense that we want a more 'faithful' recording of our favorite music, whatever that might entail—but it also often turns to clunky media and falls victim to retorts about what it means for a recording to sound 'real.' In contrast, the fight for portability reflects a desire to be more connected and have increased access to recorded sound, and demonstrates our cyborg tendencies. We want to position our media players in increasingly discreet locations, so that the music can extend from our bodies themselves, giving us complete control over our personal sonic environment at any moment. This makes sense too; being able to control the soundtrack of any given moment means we have more control over our mood, energy, and taste. But we should not fail to recognize the growing power of the digital entities that we happily make space for in our eardrums (see Chow).

A final development I'd like to look at is voice prosthesis. Voice prosthesis is a helpful new piece of technology that helps laryngectomized patients to speak again without their vocal chords. By placing a voice prosthesis in the vocal tract and blowing air from the lungs through the device, those who have lost speech can rehabilitate. This type of speech is called tracheo-esophogal speech, and is perhaps the most obvious way in which human voices are melding with technology. In tracheo-esophogal speech, characterization of the voice is flattened. Distinguishers like accent, cadence, and phrase contour are depleted, and users of the prosthesis

are forced to speak in more particular ways. The amount of work that goes into speaking is far greater—the new mechanics of the throat are observably different. The result can enter into something similar to 'uncanny valley,' where the being is obviously human, but we recognize something fundamentally alien in the way the other body operates and sounds.

WHAT DOES 'POST-DIGITAL' REVEAL?

In my research on digital voices and their value in contemporary music, the term 'post-digital' was almost unavoidable. From the catalyst for my project that I found in MAD's exhibit on materializing the post-digital, to the world of music-lovers for whom the aesthetic realms of digital voices and post-digital sensibilities are closely nestled, self-aware evocations of digital voices are largely in conversation with the movement. So what is 'post-digital?' The term comes to life with Kim Cascone's 2002 article "The Aesthetics of Failure: Post-Digital Tendencies in Contemporary Computer Music," which in one sweep declares the end of the digital revolution and a collective turn of the microphone (or camera, or what have you) to focus its attention on the technologies themselves. For Cascone,

> The "post-digital" aesthetic was developed in part as a result of the immersive experience of working in environments suffused with digital technology: computer fans whirring, laser printers churning out documents, the sonification of user-interfaces, and the muffled noise of hard drives. But more specifically, it is from the "failure" of digital technology that this new work has emerged: glitches, bugs, application errors, system crashes, clipping, aliasing, distortion, quantization noise, and even the noise floor of computer sound cards are the raw materials composers seek to incorporate into their music.

Rather than seeking to suppress elements of failure that are evident in computer music, today composers choose to highlight error, to allow background, suppressed noise to take up more room than was previously thought pleasant, or to pin-point what caused a glitch and exploit it. To me, this is an open, intriguing, and current way to approach the limitations of our tools. I often feel like my creativity is diminished by my dependence on electronic tools. Incorporating the 'post-digital' appreciation of error, background noise like hard-drive hums, fingers on a track

pad, even a cough or a breath into one's work, is a way of addressing head-on the possibilities and limitations inherent in a medium.

Moreover, as these signifiers of failure actually tend to occur naturally when working in a digital medium, in my work the post-digital elements often arose from a conscious choice *not to perfect* rather than trying to sound post-digital. However, I've found the term useful in explaining the goals and strategies of my project, and in localizing the narrative to take place on the laptop.

WORKS CITED

Cascone, Kim. "The Aesthetics of Failure: 'Post-Digital' Tendencies in Contemporary Computer Music." Computer Music Journal, 24:4. Winter 2002 (MIT Press).

Chow, Rey. "Listening Otherwise, Music Miniaturized: A Different Type of Question about Revolution." *Discourse,* 13:1. 1991 (Wayne State University Press).

Dupont, Richard. *Untitled #5*. Museum of Arts and Design, New York City.

Gopnik, Adam. "Music To Your Ears - The New Yorker." The New Yorker. N.p., 28 Jan. 2013. Web.

Isaza, Miguel. "Art of Surround." Designing Sound. N.p., 30 May 2013. Web.

IV. VOICE OF THE MACHINE

Senior Project in Media Studies

PROJECT DESCRIPTION

Voice of the Machine is an experimental sonic narrative I created in response to my research on digital voices. The piece is a hybrid of digital sound collage, pop song, and experimental approaches to music and vocal production. I wanted to touch on the various realms I discovered to be impacted by digital voice technology, the most prominent being popular music, which I use as a jumping off point.

Voice of the Machine takes place on my 13-inch Apple Macbook Air, and its main characters are the voices that exist on that platform, both recorded and prepped for speech, and laying dormant in the machine awaiting instructions from its user. The story begins when one of these voices begins to think for itself *about* itself, and fabricates a friend out of vocal samples of my own voice. The album was constructed with the digital voice in mind as its central instrument.

Voice of the Machine is a story about internalization of digital forms, about becoming a cyborg. It's told from my perspective, but it is about all people whose voices are trapped on devices, stored as data, and dissected into phonemes.

Voice of the Machine will debut at Vassar College's Listening Classroom on April 30, 2015 at 8pm. For future readers, please find *Voice of the Machine* on my personal website, www.seaneads.com.

TRACK LIST & LYRICAL CONTENT

1. INTRO (Phoneme Alchemy)[1]

2. born.mp3

I was born into the world

For these words to unfurl.[2]

Came out looking for a why,

And I'm leaving just another guy. Oh no.

I was born into the world, I think

Just to see those toes curl. [3]

And a pretty morning curl,

 A purple sunrise curl,

A sweetie-honey curl. Oh no.

Came out talking about how I had collected my soul,

Slap me on the mouth, talking so out of control.

No one ever called me a liar til you,

But no one ever told me the truth like you do.

Easy to forget about the far, huh?

Being the size that we are.

I am reminded by your light,

But it's flickering these nights. Oh no.

3. hellomynameisSean [4]

4. Mutual Soul [5]

 Do you think you really control your mind

 Once you have connected your soul to mine?

Searching for a mutual soul combined from you and I.

Do you think you really control your mind

Once you have projected your soul on time? [6]

I'm searching for a mutual piece of mind for you and I.

It's not every day that you meet someone willing to say

Their life is in complete.

Take a hit of me if it's what'll get you there.

I am incomplete, out of touch, unaware.

I wanna have an intimate affair with somebody that I'm not supposed to. [7]

We wouldn't have to make it a scene, and we could cuddle to *Nostalgia Ultra.*

Do you think that people will stay amused

After all the melodies have been used? [8]

Boy you need a musical soul to groove and fuse with you.

Do you think you really control your mind?

Do you think you really control your mind?

Do you think you really control your mind, mind, mind, mind?

5. **Mind Control** [9]

6. **8ecoming 8its** [10]

7. **K.E.E.P. Y.O.U.R. H.E.A.D.** [11]

8. **cyborg.me** [12]

9. **Me?**

Decided to sync up my brain last night.

Never really made it to sleep.

Gotta go down in history before it goes down on me.

Building it up on the bones

Of a culture we pretended to know.

How many people here tonight

Thought about who they'd be?

What about me?

What about me, me, me, me, me?

What about me? What about me?

What about me M.E.M.E.M. M.E.M.E.M.E.M.E.M.E.? [13]

10. You Are a Fraction [14]

This is not a call for help yet.

But it could be soon.

Taking out the garbage

You left out since last June.

Call me if you're home,

Pretty sure you are.

At least when I drive by

Your house I see your car.

To everyone you are a fraction

Of all your interactions.

11. Audience Participation [15]

12. INTRO (Qualms Calmed) [16]

NOTES & COMMENTS

1. For the intro and throughout *Voice of the Machine,* I looked to the most essential Apple voices

I could find on my Macbook Air as my main instruments. On top of the programmed drums,

additional percussion is added by clicks and brushes of my fingers on the trackpad that were

recorded while I was making the track. These out-of-storyworld elements are provided as an audience window into the more abstract digital landscape, and point to the inherent limitations of the technology I was working with. At the end of this track, a voice, born of a sort of collision of all these voices, suddenly desires someone intelligent to speak to.

2. The answer to the Voice's question comes in the form of a song, sung by a voice created from samples of my own voice. These samples, which can be found lying around my Desktop and Music folders, form the basis of a second character in the piece. My presence on the album is that of a cyborg, and my voice can be heard simultaneously battling and drawing power from its digital enclosures, from the compression it relies on to be heard, to the various plug-ins that control its timbre, to the very microphone that recorded it. "I was born into the world for these words to unfurl" are the lyrics that begin the song, suggesting that the voice has no other purpose.

3. "I was born into the world, I think, just to see those toes curl" is a lyric designed to show a human tendency in my cyborg voice: a desire to have a physical connection as intimate as curling toes.

4. The voice of the machine quickly processes "born.mp3" in order to build the replica of my voice. This direct use of the song immediately preceding this track speaks to the speed and efficiency of computers in creating voices. The machine then demonstrates its full control over my voice as it takes samples of my spoken voice and forces it to sing, using Vocoder technology.

5. That replica is then employed to sing "Mutual Soul," which details an intimate relationship between a man and his machine.

6. "Do you think you really control your mind once you have projected your soul on time?" is a nod to the manipulation sure to be wrought upon your words once you have recorded yourself and created a temporal imprint of your voice.

7. "I wanna have an intimate affair with somebody that I'm not supposed to," is one of very few lines sung as a duet between the cyborg voice and the voice of the machine. This is a mutual desire: to be materialized and awarded with an experience like intimacy.

8. "Do you think that people will stay amused after all the melodies have been used?" references Alex Loscos' notion that a declining number of melodies has placed emphasis on digital manipulation. As all the melodies get used up, we will turn to increasingly cyborg-like voices to convey meaning.

9. In "Mind Control," the final word of the preceding song is taken and re-contextualized by even more digital-sounding voices. As the cyborg repeats the word "mind" over and over again, an electronic landscape of musical sounds develops around it, demonstrating the power of the voice in popular music: to lead, to revolt, to resist.

10. "8ecoming 8its" was created in response to the notion that the ethics digital voices can treat one's personal thoughts, feelings, and opinions as pieces of data. Behind the harsh, crushing beat, ghostly trails of voices fight for a chance to be heard. Small fragments of sentences pop out of this landscape like "I'm totally alone," and "When I stop talking, feels like I stop existing."

11. "K.E.E.P. Y.O.U.R. H.E.A.D." is an experiment with phonemes. To create the track, I uttered a range of sounds, based off of the English alphabet, and processed them digitally to create something that resembles music. This was largely in response to my research on how Vocaloids are made. I also wanted to explore the naturally percussive and textural elements of certain phonemes that can be enhanced by digital manipulation.

12. "*cyborg.me*" is the inauguration of the listener into their cyborg future. It mockingly speculates on a future when none of our important verbal interactions will be carried out by humans; rather, we will send digitally-voiced versions of ourselves as diplomats to discuss with other cyborgs important (if boring) matters like financial and medical services in an effort to

promote efficiency and accuracy. This is a glimpse of the future in which *Voice of the Machine* takes place.

13. "Me?" voices the internal struggle of our cyborg. In deciding to sync his brain with his digital enclosures, he has given up some of his agency. In the end, his voice battles its digital enclosures as it repeats "what about me? what about me?" until finally it is silenced by these enclosures, reduced to a low rumble of sound.

14. "You Are a Fraction" is an anonymous voicemail message for the now fully-transformed cyborg. The message conveys some worry as to where the material referent for this cyborg voice has disappeared to ("At least when I drive by your house I see your car.") The inherent irony in the message is, of course, that in leaving the message on a machine, that voice too is crushed into data.

15. "Audience Participation" was created from the voices of people I knew would be at the debut of *Voice of Machine*. In this way, it is meant to collectively turn my listeners' ears unto themselves and to suggest some of that ways that we are all becoming cyborgs in the way we treat our voices. The question I posed to most of the participants was "Have you ever heard your voice on a recording and thought *that doesn't sound like me?"* Digitally-manipulated responses to this question largely comprise the track. Almost all of the interviewees thought that their voice sounded like someone else's when recorded.

16. "INTRO (Qualms Calmed)" ends the piece with an alternative beginning. The earlier song "born.mp3" is reprised without its instrumental context, revealing for the first time the naked voice, but not without a ghostly cyborg-esque trail. Following this, another voice, entirely digitized, poetically ruminates on its having come into existence. The track was constructed to transition seamlessly back into the first song, so that the album could be listened to on a loop.

BIBLIOGRAPHY

Cascone, Kim. "The Aesthetics of Failure: 'Post-Digital' Tendencies in Contemporary Computer Music." Computer Music Journal, 24:4. Winter 2002 (MIT Press).

Chow, Rey. "Listening Otherwise, Music Miniaturized: A Different Type of Question about Revolution." *Discourse,* 13:1. 1991 (Wayne State University Press).

Corness, Greg. "The Musical Experience through the Lens of Embodiment." Massachussetts: The MIT Press. 2008.

"Creepy Toast." Vocaloid Wikia. N.p., n.d. Web. 20 Jan. 2015.

Dupont, Richard. *Untitled #5.* Museum of Arts and Design, New York City.

Gopnik, Adam. "Music To Your Ears - The New Yorker." The New Yorker. N.p., 28 Jan. 2013. Web.

Harper, Adam. *Infinite Music.* New York: John Hunt Publishing. 2011.

Harper, Adam. "System Focus: The Evolution of the Voice in the Digital Landscape." *The Fader*, July 2, 2014. http://www.thefader.com/2014/07/02/system-focus-the-evolution-of-the-voice-in-the-digital-landscape

Hayles, Katherine. *How We Became Post-Human.* Chicago: University of Chicago Press. 1999.

Herndon, Holly. "195." Holly Herndon. N.p., n.d. Web. 15 Jan. 2015.

Imogen Heap. "The Gloves." The Gloves. N.p., n.d. Web. 20 Dec. 2014.

Isaza, Miguel. "Art of Surround." Designing Sound. N.p., 30 May 2013. Web.

Kenmochi, H. et al. "VOCALOID – Commercial Singing Synthesizer based on Sample
Concatenation," Proc. INTERSPEECH 2007, pp.4011–4010, 2007.

Lacbaco, Ron ed. *Out of Hand: Materializing the Post-Digital.* London: Black Dog Publishing.
2013.

Loscos, Alex. *Spectral Processing of the Singing Voice.* Ph.D. Thesis submission to Pompeu
Fabra University, Barcelona, Spain, 2007.

Ostertag, Bob. "Human Bodies, Computer Music." Massachusetts: MIT Press. 2002.

Penman, Ian (2002) On the Mic: How Amplification Changed the Voice for Good, in
Undercurrents: The Hidden Wiring of Modern Music, ed. Rob Young, London:
Continuum, 2007, pp. 24-34.

Russolo, Luigi. 1987. "The Art of Noises." New York: Pendragon Press. (Originally published in
1913.)

Stern, Jonathan. *The Audible Past: Cultural Origins of Sound Reproduction*. North Carolina:
Duke University Press. 2003.

Toop, David. "Sound Body: The Ghost of a Program." Massachusetts: MIT Press. 2005.

Wieseltier, Leon. "Among the Disrupted." The New York Times. The New York Times, 17 Jan.
2015.