TUNING IN SITU:
ARTICULATIONS OF VOICE, AFFECT, AND ARTIFACT IN THE RECORDING STUDIO

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

William Owen Marshall

January 2017

TUNING IN SITU:

ARTICULATIONS OF VOICE, AFFECT, AND ARTIFACT IN THE RECORDING STUDIO

William Owen Marshall, Ph. D.

Cornell University 2017

ABSTRACT

This dissertation examines practices of digital vocal tuning in the fields of audio engineering and music production. Drawing on an ethnographic study of two Los Angeles recording studios, it accounts for the recorded voice, and the tools used to shape it, as a collective sociotechnical accomplishment. It begins by tracing how recording engineers have historically thought about and worked with "emotion," specifically with respect to its relationship between the studio's technological artifacts and the voice of their client. It goes on to examine two major digital tuning tools – Auto-Tune and Melodyne – in terms of their initial development and the different technological prescriptions they brought to the technological problem that would eventually come to be known as "pitch correction." It first traces the ways that recording engineers developed for working with, problematizing, and *repairing* digital tuning techniques understood in terms of a "covert" and "corrective" approach. It then looks at the various ways that pitch correction was repurposed as an "overt effect," primarily by non-engineers such as songwriters, performers, and listeners. We then "zoom-in" an instance of corrective tuning as an interactional accomplishment that draws upon techniques of inscription, repetition, and inflection. The remaining chapters take as their starting point the distinction between "place" and "time" theories of pitch perception in psychoacoustics and signal processing, in an attempt to thicken and ethnographically respecify the ways that place and time are structured and performed

in-studio. This respecification provides a more adequate understanding of how the objects of digital tuning are practically constituted in two senses. The first of these is the *topical* structure of the recording studio, or the way that studio life is performed through various acts of place-making. The second concerns the *temporal* structure of studio life, taking up the various ways of keeping, making, and marking time in the studio. In the Appendix I develop the concept of the "*shibboleth*" as a complementary concept to boundary objects and boundary work, as well as a new way of thinking about the mutual production of social and perceptual difference in STS.

BIOGRAPHICAL SKETCH

William Owen Marshall studied political science and global studies at the Barrett Honors College at Arizona State University, where he also worked at the Consortium for Science Policy and Outcomes and served as the Music Director for KASC 1260 AM. After college he produced radio stories for Chicago Public Radio's Vocalo.org project. He began his graduate work in Science & Technology Studies at Cornell University in 2010. Following his PhD he will work as a Postdoctoral Scholar in The Science and Technology Studies program at UC Davis. He plays music. He grew up in Utah.

Dedicated to my mother, the musician, and my father, the engineer.

ACKNOWLEDGMENTS

This dissertation would not have been possible without the thoughtful, generous, and patient support of my committee members, Trevor Pinch, Rachel Prentice, Mike Lynch, and Steve Jackson. For their feedback and support I would also like to thank Karin Bijsterveld, Joeri Bruyninckx, Cyrus Mody, Alexandra Supper, Mara Mills, Viktoria Tkaczyk, Peter McMurray, Myles Jackson, Melle Krumhout, Tim Choy, Joe Dumit, James Griesemer, Tim Lenoir, Colin Milburn, Mario Biagioli, Alexandra Lippman, Dierdre Loughridge, Alexis Walker, Kasia Tolwinski, Danya Glabau, Enongo Lumumba-Kasongo, Hansen Hsu, Park Doing, Malte Ziewitz, Ron Kline, Paul Nadasdy, Suman Seth, Sara Pritchard, Stacey Stone, Debbie Van Gelder, Ryan Heichel, Trina Harrison, Vinny Ialenti, Tyson Vaughn, Marit McArthur, Jennifer Stoever-Ackerman, Patrick Brown at Sociology Compass, Eliot Bates, Samantha Bennett, Gretchen Gano, Lori Hidinger, Clark Miller, Dan Sarewitz, Dave Guston, Ira Bennett, Derek Anderson, the staff of the Consortium for Science Policy and Outcomes at Arizona State University, the Center for Science, Technology, Medicine, and Society at UC Berkeley, The Max Planck Institute for the History of Science, the staff of the Gannett Counseling and Psychological Services, WFMU, WRFI, KASC, KDVS, WBEZ, Vocalo.org, and Antares Audio Technologies.

I would not have been possible without the help of my friends and family. For their love and support I would like to thank Clare Casteel, Merrill and Bill Marshall, Cayla Marshall, Judd Anderman, Jeff Valla, Bahar Akyurtlu, Heather Fae Speaker, Adam Southard, Andris Balins, Jötunn Ymir Ginnungagap, Andrew Smith, Marty Royce, Tedd Roundy, Andy Roundy, Carl, Harold, Lucille, Bob, Seth, Barry, Hunter Davidsohn, Bubba Crumrine, Jason Muffler, Michael Hagblom, Illya Riske, Wasef El-Kharouf, Tobie Milford, Dave Gironda Jr., Valerie Lu, Chris Kasych, Luana Hakes, Scott Bransford, and Grey Ferran, This project was supported by an NSF doctoral dissertation improvement grant (award #1455647), a grant from the Institute for the Social Sciences at Cornell's Innovation Creativity and Entrepreneurship project, travel grants from the Cornell Graduate School and Dept. of S&TS, and a Sage graduate fellowship. Thanks finally to anyone omitted by accident, my anonymous reviewers, and to you for reading this.

LIST OF FIGURES

TABLE OF CONTENTS

INTRODUCTION

The digital pitch correction software Auto-Tune and I have at least one thing in common: we both got into music in the late 1990s. I can still remember when and where I first heard it: twelve years old, sitting in the passenger seat of my parents' car in a grocery store parking lot, listening to Top 40 radio. My childhood obsession with stage magic was fading (not a moment too soon) and I was in the beginning stages of what would be a prolonged interest in weird-sounding music.[1] The sound of Cher's voice during the second word of the pre-chorus of her song "Believe" ("I *caaaan't* break through!") seemed to bridge those two interests readily. Though I did not much care for the song itself, I was captivated by that particular sound as an instance of sheer production. It reminded me of the sound of a dial-up modem, or the sounds you hear when you call a fax machine by mistake. It was like listening in on the conversations machines have when humans are not around.[2] It sounded like a voice being possessed by a demon or infected by a computer virus, or like the musical equivalent of the scene in *Terminator 2* where Arnold Schwarzenegger peels the fake skin off of his forearm and reveals a metallic skeleton. It was, I think, my first encounter with the concept of something sounding "wrong on purpose." At any rate, it seemed like quite a trick, and I wanted to know how it worked.

[1] This would eventually lead to an interest in the practice of circuit bending, or the deliberate short-circuiting of electronics for purposes of soundmaking. For an analysis of how circuit bending's material intervention resonates with work in Science and Technology Studies (STS), see (Pinch 2016)

[2] Conversation between non-humans is a topic that is unfortunately underdeveloped in STS, with the notable exception of the postscript to Woolgar's "Configuring the User" (Woolgar 1990)

I had to wait a few years to get my first clue, which came in the form of a friend explaining to me that a local band that had recently gained national attention were actually sell-outs, in large part because their singer used a computer program that fixed mistakes in his voice. Gradually I came to understand that these two formative moments could be traced back to the same piece of technology. Years passed and, contrary to many people's expectations, Auto-Tune would not seem to go away. By the end of high school I was hearing it in hip-hop tracks. By the end of my undergraduate studies, Jay-Z was rapping about the "Death of Auto-Tune" and the band Death Cab For Cutie were wearing blue ribbons to the Grammy's in protest of "Auto-Tune Abuse." People on the Internet were Auto-Tuning news footage. The effect had become a household name.

On the one hand Auto-Tune came to seem ubiquitous, unavoidable, and annoying. On the other hand, one had a sense that its use as an over-the-top effect was just the tip of the iceberg. Before long, Auto-Tune was claiming the title of best-selling software plug-in of all time, a piece of studio equipment nearly as standard-issue as the Pro-Tools production software that hosted it. (Walden 2015) Moreover, it was clear that engineers were not paying hundreds of dollars on a piece of software just because it could make their clients sound like late-career Cher. They were using it as 'intended:' as a way of correcting the intonation of their clients' voices. Auto-Tune was operating simultaneously in the foreground of popular culture as an index of technological sophistication and in the background as a way in which the culture got produced behind-the-scenes.

It is this protean combination of superficially faddish and deeply infrastructural qualities that make auto-tune such a compelling object of sociotechnical investigation. Whether it will prove, with hindsight, to be an enduring fixture of 21st century culture or a trend that just happened to saturate the pop culture of my generation's formative years is a prediction I will not attempt here. Instead, this dissertation is an attempt to engage with digital tuning as a way of understanding the ever-changing relationships between technology, culture, and the bodies of listeners and performers. If digital pitch correction offers young radio listeners a way to tune-in to the sounds of machines, it is also a way for practitioners of STS to tune-in to the socio-teechnical production of voices (literal and figurative) and the voicing of technological assemblages as elements of social action. In the remainder of this introduction I will argue that auto-tune is usefully analyzed in terms of how it "breaches" commonly held ideas about the relationship between voice, emotion, and technology. For this reason, I argue that it offers a particularly useful case for studying in detail precisely how the emotionally expressive voices of popular music are accomplished *in situ*.

Auto-Tune as Breaching Experiment

Pinch and Bijsterveld have argued that historical cases of newly emerging technologies[3] can be understood as "breaching experiments" which, by disrupting the taken-for-granted practices which structure everyday experience, provide insight into how social order is locally constituted. (Pinch & Bijsterveld 2003) The phrase "breaching experiment" comes from the field of ethnomethodology, wherein researchers would breach mundane social conventions (e.g. by responding to formalities, such as "how are you?" with clarifying questions or detailed descriptions, or performing illegal but apparently useless moves in a game of chess) so as to reveal how everyday social life is constituted and repaired "at all points." (Garfinkel 1986)

---

[3] Though they focus specifically on musical technologies, such as the player piano and synthesizer, they suggest that this approach may be fruitfully applied to the study of technology in general.

Among several examples from the history of music technology, Pinch and Bijsterveld show how the *intonorumori* noise making machines of early 20th century painter Luigi Russolo challenged conventional ideas about tonality and rhythm, while also providing a way to practically "capture" randomness and chaos within the conventions of compositional western art music.[4] The *intonorumori* both provoked the interwar European art world with noise and occasioned a process by which that noise was gradually made socially and technologically accountable. This dissertation's topic is the articulation of voice, affect and artifact in the recording studio, and Auto-Tune is the impetus of a "natural" breaching experiment into that topic.

I argue that the insights afforded by the natural breaching experiment approach come at the expense of two major blind spots. The first is the apparent problem of accounting for the emergence and shaping of the technology that is doing the breaching. When a technological artifact is being deployed as a conceptual "probe" into social order, it is liable, like any well-functioning tool, to perceptually disappear into the task at hand.[5] The second, more problematic blind spot has to do with the problem of conceptual "capture." Just as studies of scientific controversies put researchers at risk of getting "drawn into the fray" (Scott et al. 1990) as one side begins to make use of their work, so does the natural breaching experiment make the research prone to capture by, for example, a contemporary tech industry deeply invested in narrative of innovation-as-disruption. (Lepore 2014)

While this makes the natural breaching experiment approach, in terms of matching actor's and analysts categories, uniquely well-tailored to accounts of late capitalist technological

---

[4] Emily Thompson makes a similar observation in her discussion of how experimental music of the mid-20th century involved a domestication of noise. (Thompson 2002)

[5] Here I have in mind Heidegger's observation that tools in use, or ready-to-hand, take on a certain phenomenological transparency, at least until they break.(Heidegger 1996)

change, it runs the risk of recycling entrenched narratives while missing others. This problem is especially relevant in the case of Auto-Tune, the founder of which is something of a self-fashioned ‘serial disrupter.’ Jonathan Sterne, drawing on Bourdieu, has advocated the practice of an “epistemological break” from the “common sense” of technology, whereby inquiry begins with an attempt to identify and think beyond the ways in which the technological system has been “pre-constructed.”(Sterne 2003a) In an attempt to balance the conceptual affordances of the breaching model while also maintaining a reflexive stance with respect to the widely deployed narratives surrounding Auto-Tune, I will venture an approach based on “partial” breaking and a counterstrategy of repair – both as an epistemological orientation and a topical concern. I will keep the vehicle, but check the blind spots.

Checking the first blind spot means that, in addition to using the case of Auto-Tune to examine how the producers and consumers of popular music think, feel, and work with technologically mediated voices, I will approach Auto-Tune as a “black box” that needs to be opened. This process of black box opening can begin with a prefatory note concerning terminology. Even at the level of common word usage, Auto-Tune contains multitudes: in addition to the many versions and material instantiations of the actual brand name product produced and marketed by Antares Audio Technologies, “auto-tune” has become shorthand for a diverse array of digital tuning tools. These include software and hardware such as Celemony Melodyne, Waves Tune, iZotope Nectar, Synchro Arts Revoice Pro, Melda Production’s MAutoPitch, Cakewalk/Roland’s V-Vocal, Mu’s Mu Voice, Steinberg Pitch Correct, Logic Pitch Correction, TC Helicon VoiceTone, Boss VE-20 Vocal Processor, and smart-phone apps like Smule’s I Am T-Pain or Atonality’s Tune Me. I will use the uppercase “Auto-Tune” to refer to the Antares product. The lowercase “auto-tune” will refer to pitch correction or digital tuning in a broader, more generic sense. "Big-A" Auto-Tune is a particular brand. "Little-a" auto-tune is the broader technological complex of artifacts and practices of digital vocal tuning, pitch correction, and any interesting empirically observable conflations thereof. To wit: auto-tune has also come to stand for (or at least get routinely conflated with) earlier vocal manipulation technologies such as the vocoder or the talk box.(Weheliye 2002; Tompkins 2010; Smith 2008; Dickinson 2001) As I will show, this already multiple, "small-a" auto-tune also contains a wide array of histories, practices, concepts, investments (emotional and financial), copyrights, and controversies. Auto-tune is not one black box but several, and (*pace* Winner (1993)) they are anything but empty.

Checking the second blindspot means accounting for auto-tune's "breaching" effect on music production while also keeping in perspective the ways in which this narrative was itself produced and circulated, the effects it may have had, and alternative narratives that might have been overshadowed. Auto-tune did not simply appear fully formed like the monolith in Stanley Kubrick's *2001: A Space Odyssey*, though it can seem that way to the producers and consumers of pop music. To borrow Annemarie Mol's terminology, it was born "multiple" (Mol 2002) and continues to be enacted and reenacted across a wide range of contexts. Most importantly, auto-tune cannot be understood fully without taking into account how it has become a point of sociotechnical articulation, or "joint,"[6] between practices of technological voicing and emotional experience. Emotion, as show in Chapter 1, is a central and contested topic for engineers, who have historically approached it both as a problem of production and transmission. Auto-tune

[6] The concept of a joint as a point of sociotechnical articulation comes from (Latour 2004) and has been further developed by (Prentice 2005) alongside the concept of "mutual articulation," which brings bodies and artifacts together in a relation of skilled perception.

offers an empirical entry point into the sociomateriality of the voice, understood as both a medium for emotion and an object of technological intervention. I use it here to trace the ways in which voice, affect, and artifact are articulated together through practices that implicate and produce social as well as material arrangements.

Jarzabkowski and Pinch have recently identified two key concepts widely used in investigations of "sociomateriality."(Jarzabkowski & Pinch 2013) The first is the concept of the "affordances" of material arrangements, which constrain and enable particular forms of action (Gibson 1979). The second is the concept of sociotechnical "scripts," which are embedded within and deployed alongside technological artifacts, thereby encouraging particular scripted sequences of action over others (Akrich 1992). They suggest that these concepts, by focusing on materials and intentions respectively, distract from the ways that people "accomplish" social actions with materials. They advocate increased attention to the "three Rs": *repurposing* (finding new affordances within objects), *reinscripting* (revising the social scripts of objects), and *repairing* (making objects work in response to moments of breakdown and contingent inadequacies of an object's affordances and inscriptions over the course of everyday life.)

In Chapters 2-4, I take up these three Rs as ways of teasing out and weaving together the ongoing tangle that is auto-tune-as-breach. Under the heading of repair, (which implies damage, or at least *disrepair*) we can consider three moments of sociotechnical articulation. The first of these concerns the way in which auto-tune both crystalized a problem in studio – a perceived zero-sum relationship between emotion and technical accuracy – and provided a way to fix it. The second concerns how auto-tune came to be perceived as entailing a deskilling of both engineer and vocalist, but ultimately led to a proliferation of new skills related towards the repair of dislocated social relationships in-studio. The third moment encompasses how auto-tune came to be perceived as inadequate to its professed task – the covert correction of out-of-tune performances – and the ways that it was refashioned by engineers and software designers in response to this breakdown of expectation and affordance. Sociomaterialities of repair involve cases where phenomenological transparency of use breaks down and has to be re-negotiated. In the case of auto-tune, breakdown and repair can be located at the levels of the corrigible voice, the skilled relations among studio workers, and the functionality of the software artifact itself.

Under the heading of "repurposing" goes the emergence of "overt" tuning practices among user groups outside of the domain of recording and software engineering proper. By using settings and techniques not originally intended by the software designers, these groups – disproportionately composed of black and female users - found new affordances in automatic pitch correction algorithms. Songwriters used it as a tool for generating ideas and communicating them to their client. Vocalists like Cher and T-Pain incorporated it into their artistic personas. The "consumers" of pop culture made critical and creative use of the effect, extracting and recombining the latent vocal melodies of news anchors, politicians, infants, and other unwitting singers.

Reinscription occurs in dialogue with repair as well as repurposing. While digital tuning tools were each pre-inscribed with particular modes and assumptions of use, these scripts were revised in response to the ways in which users fit the technology to their particular arcs of work. Along the pathway of repair, for example, Antares began to include new features and new scripts (in both technical and marketing materials) that responded to perceptions that the effect made voices "inhuman" or "lifeless." Melodyne – Auto-Tune's primary competitor - was increasingly inscribed as a tool for automatic correction as opposed to post-hoc composition, through the introduction of features such as the "correction macro." The repurposing pathway, meanwhile, found artists like T-Pain reinscribing auto-tune as "the T-Pain effect" and promoting his own social scripts concerning how such an effect is "correctly" accomplished. Antares, meanwhile, also worked to reinscribe its products in ways that facilitated and took ownership of the overt effect in their own way.

The breaching experiment framing, combined with attention to the sociomaterial renegotiation of vocal production in the wake of auto-tune, opens up two further lines of inquiry. The first is into the detailed local practices by which an engineer actually carries out covert tuning in a way that is accountable to the voice's broader publics (e.g., vocalists, labels, imagined commercial audiences, and fellow engineers.) Decisions regarding the relative "correctness" of a voice are performed locally and according to particular socially cultivated methods. Just as the repurposing pathway saw non-engineers playing with the boundary between everyday speech and technologically stylized singing, so do engineers turn "rough" voices or "raw" takes into hit singles by way of mundane conversational and inscriptive practices. To illustrate this, I present in Chapter 5 an instance of covert vocal tuning as an interactional accomplishment between the engineer, the pre-recorded and technologically mediated voice, and myself as an auditor of the process. Zooming-in on the act of digital tuning is useful in that it demonstrates the way that the tuned voice comes into being as a mundane interactional accomplishment. By showing how an engineer working in the context of digital tuning breaks down a performance into its various tones, timbres, rhythms, and intentional objects, we gain an understanding of the tuned voice as the result of a skilled procedure.

Showing how the voice is decomposed into corrigible pitch-objects, however, raises the question of how things like vocal pitch are sustained in the studio prior to and well after they have gone through the tuning process. It is this question that guides my inquiry into how tuning is related to the topical and temporal articulations of the studio as a domain of work. Psychoacoustic and signal processing science accounts for pitch perception in terms of a wave's topological and temporal characteristics. Pitch makes itself known to the listener's body both in terms of the *place* on which it is felt, as well as the *time* it takes to repeat itself. Pitch makes itself known in a waveform in terms both of the shape of the wave (i.e. its harmonic structure) as well as its periodicity, or timing characteristics. I offer an ethnographic elaboration and specification of these narrow technical conceptualizations of place and time, arguing that to understand how vocal pitch comes into being in the actual context of audio production practice we need to attend to the various ways that time and place are performed in-studio. Drawing on ethnographic observation of two Los Angeles studios, and contemporary accounts from recording engineers, I offer in Chapters 6 and 7 a thick description of studios as topically and temporally organized spaces within which voices are collectively produced.

An act of breaching is epistemologically useful precisely because it makes boundaries visible by animating the work that sustains or refashions those boundaries. Auto-tune is a useful site for investigating the social construction of voice, emotion, and technologies which articulate them relationally because its advent required that these relationships be re-articulated, both in the sense of re-stating that which has become unspoken (or, to use a more theoretically loaded term, "tacit"[7]) as well as in the sense of crafting new boundaries and relations in place of old ones. If

---

[7] For my analysis I specifically avoid relying on the concept of tacit knowledge, as introduced by Polanyi (1962) and developed by Collins (Collins 1975) as a way of referring to knowledge that is either inexplicable or simply unexplicated within a certain domain of knowledge production. As Thornton has argued, tacit knowledge is rooted in a *via negativa* argument whereby it is defined solely in terms of what it is not. (Thornton 2013) To me this means that tacit knowledge is not an explanatory concept but a phenomenon in need of explanation. Early on in my fieldwork I attempted to explain to one of my informants that I was interested in the aspects of

the body is often thought of and treated as a border between an interior self and an exterior world, the voice is often taken to be one of the most important crossing points between these worlds. Technologies like auto-tune, which claim to work at the nexus of voice and emotion, trouble the boundary that is the (figurative and literal) voice.

By making the voice legible to algorithmic representation and available to digital manipulation, auto-tune functions partially as what Rachel Prentice has, in the context of surgical simulation technology, termed a "body object." As Prentice notes, body objects might be considered a type of "boundary object" (Star & Griesemer 1989) though they are better understood specifically as "bodies or body parts that have been engineered to inhabit computers." (Prentice 2005, 863) Boundary objects "inhabit several intersecting social worlds… *and* satisfy the informational requirements of each of them." They are

> "plastic enough to adapt to local needs and the constraints of the several parties employing them, yet robust enough to maintain a common identity across sites. They are weakly structured in common use, and become strongly structured in individual-site use. These objects may be abstract or concrete. They have different meanings in different social worlds but their structure is common enough to more than one world to make them recognizable, a means of translation." (Star & Griesemer 1989, 393)

---

his work that could not be written down. Naturally he wanted to know how I planned to write about something that cannot be written. I realized that this was a very good question. Collins, in his book Tacit and Explicit Knowledge, has attempted to answer it in the form of a typology, which eventually leads him to a position that he terms "social Cartesianism" which erects a boundary between a type of tacit knowledge that can be socially held and the kind that can be explicated algorithmically. (Collins 2010) Though for the purposes of this project I am interested in boundaries between human and non-human knowledge as actor's categories (the ways, for example, that emotional creativity and technical accuracy are parsed out through acts of pitch correction) Collins' formulation of this sort of distinction as an analyst's category seems superfluous. While Susan Schmidt-Horning has examined studio techniques in terms of tacit knowledge (Horning 2004) I would argue that her cited examples such as mic placement and various "tricks of the trade" are better described in terms of their situated articulations than by a framework which takes their "tacitness" to be of primary importance.

Prentice's body objects do indeed traverse social worlds – specifically those of medicine, engineering, and computer science. They also appear at the boundary between human and computational comprehension, providing representations to both the computer and the user in ways that are respectively legible.(Prentice 2005, 847) However, for this reason, they depart from the concept of boundary object insofar as they are not "loosely" structured, especially at the level of common use. Where a boundary object facilitates "collaboration without consensus" in the form of tacking back and forth between the ill-structured common construal of the object and its well-structured local form (Star 2010, 604-5), a body object (as I read it) is rather a way of facilitating collaboration by way of a thoroughly elaborated material-semiotic apparatus of consensus- a tightly woven fabric of human-ready and computer-ready representations.

Digital tuning technology can be understood, in part, as both boundary and body object, though it also troubles the definitions of these concepts in important ways. Auto-tune resides at the intersection of several social worlds – particularly those of the engineer, the artist, the listener, and the software developer. Like many audio signal processing technologies, auto-tune contains algorithmic representations of parts of the bodies of singers and listeners[8] alike. Auto-tune is a "bizarro" boundary object, a negative image of how a boundary object ought to work, or a figure-ground reversal of the concept of collaboration without consensus. It is an example of a type of boundary object that I describe as a "shibboleth" – a concept which I develop more fully in the Appendix. Instead of facilitating cooperation between social groups in the absence of

[8] Telephone systems that employ linear predictive coding, for example, contain algorithmic representations of the human throat as a way of reducing the amount of information required to transmit the sound of a human voice. Even the ubiquitous mp3 can be construed as a variety of body object, as a result of the perceptual coding algorithms – mathematical representations of the human auditory system – that it employs to compress files. (Sterne 2012)

consensus, a shibboleth affords a way of collaborating that reinforces the divide between social groups. Boundary objects occur when people have specific things in mind but come together on the basis of a simplified version of that thing. Shibboleths occur when social divisions are deepened by differences that, outside of the context of collaboration, would seem unimportant. Shibboleths are, in short, boundary objects for doing boundary work. They are common in the recording studio, and play an important role in the production of different domains of music-productive skill and the credible creative and technical voices that these skills imply.

Articulating Voice, Affect, and Artifact

Recording engineers of the early 21st century care about many things, but "emotion" is one of the main ones. If you asked an engineer what the most emotionally important element of a piece of recorded music is, they would probably say the voice. If I were allowed only one word for the technology most deeply implicated in how these recording engineers conceptualize and work with emotion and the voice, that word would be "auto-tune." If you talk to audio engineers for long enough, ask enough strange questions, listen to the stories they tell each other, read the things they write in books, magazines, and message boards, you will find that one of the few things they agree about is that what ultimately matters in their line of work is the emotion associated with a recording. Though, in absolute terms, more time, breath, sweat, saliva, and keystrokes might be spent on hashing out concerns like analog vs. in-the-box summing, microphone placement, or room treatment (all of which will have a place in my account), the generally agreed-upon end game for all of this is the emotion that seems to wind up on record.

It is precisely because of its close association with emotion that the human voice, among other musical instruments, takes on a privileged role in the studio. The vast majority of popular music recordings feature singing voices, and feature them centrally. The voice typically occupies the loudest, most prominent point of a mix. If a project is named after a single member of an ensemble-based recording project, that person is very likely to be the singer. The voice is such a dominant element of modern sound recording and music that even in cases where a voice is not literally present it still has a powerful effect on how people think about and work with sound. The studio's best microphones are those designed specifically with "presence peaks" that emphasize particular frequency ranges of the human voice. A polyphonic synthesizer is defined by the number and quality of its voices. A skilled pianist or guitarist will know the multiple "voicings" of any given chord name. Voice, as "the original instrument"[9] takes the form of an exemplar for emotional expression as a site of articulation between objective skill and subjective richness.

The privileged role of the voice in sound reproduction has a historical basis. Jonathan Sterne has argued that one key historical pre-condition to what he calls "the possibility of sound's reproduction" was a conceptual shift in the 19th century from sound understood primarily in terms of *vox* or voice, to a modern understanding of voice as a particular instance of sound. (Jonathan Sterne 2003, 23) While this is a good description of changing scientific and philosophical approaches to sound prior to the development of sound recording, the modern recording engineer still works in a sound world where voice has not been fully subordinated to vibration. As I will argue, the work of an engineer in studio is largely seen as a negotiation of the boundary between voice as significant sound and voice as just another collection of vibrations at various frequencies. Because the literally and figuratively creative voice of the client is the

[9] I take this phrase from pioneering experimental vocalist Joan La Barbera. (1976)

engineer's main priority in studio work, and because the voice-as "just another signal" is, for the engineer, a key technical commitment, the studio is often a place where Sterne's *longue durée* epistemic inversion between vox and vibration is replaced by multiple local oscillations.

For recording engineers, the significance of the voice, and thus the point in engaging in the oscillations that produce it, is its unique ability to convey emotion. The voice, as an object of engineering work, is spoken about and worked with as a means of producing and/or transmitting emotion. Emotion is an important actor's category for recording engineers, one that I will distinguish as such from the analyst's category of "affect." By affect I mean the social and embodied practices by which emotion is expressed. Though this dissertation turns to the words and actions of engineers for evidence of how emotion is understood and enacted in-studio, my topic is distinct in that it brackets the ontological question of emotion as a repertoire of natural states or an object of psychological research, attending instead to affect as a broader topic of how feeling is socially accomplished in the particular context of recording. My examination of "affect" should be understood as an "ontography" of emotion (Lynch 2013) – an examination of how answers to questions of how emotion is made or defined (i.e., emotion as a "matter of fact") and how and why people care about it (emotion as a "matter of concern") are arrived at locally and historically through mundane practices. My interest in affect does not, for example, involve its philosophical explication as an "autonomous" reaction or remainder. (Massumi 1995) I am instead interested in affect as a social practice and as a topic that encompasses the way recording engineers conceive of and work with emotion.

Though it is centrally important to the recording process, emotion has no formal or technical definition among recording engineers. Its lack of technical specification is part of what makes it such a productive basis of creative evaluation among recording engineers. Its loose technical structure is bounded, however, by a broader popular understanding of emotion in modern Western culture. "Emotion" is often talked about as an expression of one's subjective interiority, something that is first felt and then expressed (or, perhaps, repressed.) It is increasingly thought of as a collection of natural types – anger, joy, disgust, sadness, etc. – which dwell inside of us. In the popular imagination one variation of this "internal expression" model of emotion is well illustrated by Pixar's *Inside Out* (2015) wherein personifications of particular emotions collaborate to shape human behavior. This is a picture of emotion as the product of a social relationship, not between people, but between the "little voices"[10] in people's heads. Unfortunately, neither my informants nor I have been able to gain direct access to any of these little voices and the emotional essences they denote. For that reason, what follows will be limited to the considerably messier question of how voices are contingently articulated and emotional experiences are crafted by recording professionals, their clients, and the technological complex within which they work.

In order to unpack pop-psychological understandings of emotion as an expression of internal feeling by way of the voice, I will draw on work that has shown how emotion and voice are categories of practice that vary across time and place. I approach emotion partially as "stance," or as a "situated practice entailed in a speaker's performance of affective stance through intonation, gesture, and body posture," among other techniques. (Goodwin et al. 2012,

[10] The film, which was produced in consultation with experimental psychologists from UC Berkeley, has as its tagline an invitation to "meet the little voices inside your head." (Keltner & Ekman 2015)

16) Following the example of Prentice's study of surgical training, I specifically attend to how engineers "construct an emotional stance" consistent with their professional position, while cultivating their own emotional experience as a resource for the performance of skilled perception, or becoming "articulate" with respect to the voices of their clients. (Prentice 2013, 38) I seek in part to locate emotion within the particular "moral economy" of the recording studio, or the way it might function as a "balanced system of emotional forces, with equilibrium points and constraints… a contingent, malleable thing of no necessity" that nonetheless "has a certain logic to its composition and operation."(Daston 1995, 4) as well as its "affective ecologies" or ways of being shaped by "pleasure, play, and experimental propositions."(Hustak and Myers 2013, 78)

As with the breaching experiment framing, the concept of an economy of emotional force and affective experience reflects certain actors' categories almost too well. Recording professionals care about emotion because emotion is understood to be a major component of a recording's value. The recent growth of digital distribution practices such as online streaming and file sharing has decreased the emphasis on pieces of music as physical objects, leading iconoclastic engineer Steve Albini to declare that music is "no longer a commodity, it's an environment, or atmospheric element." (McDuling 2014) That the solid artifacts of music[11] seem to be melting into air does not mean that the atmospheric element of music will not be thoroughly economized, however. Grammy-winning pop star Taylor Swift, writing in the Wall Street Journal, has placed her hopes for the future of the music industry in the ability of a recording artist to "break through on an emotional level and end up in people's lives forever."

[11] Promises of the enduring vinyl record "revival" notwithstanding. (Huet 2015)

(Swift 2014) As the economized unit of music moves from the physical recording to an individual stream of a song, the emotional resonance of a particular artist-voice has increasingly become the focus of music production. My approach here is the opposite of the one that says that emotion and voice are substances of internal subjectivity, the expression of which can be understood in terms of transmission and fidelity. In transmission models, artifacts are akin to noise, or unwanted byproducts of the transmission process. In order to make sense of how engineers go about crafting voices as affective objects, I instead take voices and emotions to be fundamentally artifactual things. As artifacts they are socially constructed, suspended within complex socio-technical systems of power, and articulated through local material-semiotic practices (Bijker et al. 2012; Latour 2005; Haraway 1999).

The oscillating way in which recording engineers approach voices as denoting either signals or singers needs to be understood as part of a broader relationship between instrumental and practical modes of evaluation in engineering practice. Peter Dear has argued that technoscientific knowledge production is characterized by two ideal-typical modes. On the one hand, science has an instrumental rationale; it *works*. On the other hand, it has a natural-philosophical rationale; it is supposed to accurately *represent* nature. Dear notes that these ideal types are in what he describes as an "invisibly" circular relationship, which he analogizes to a gestalt figure which can be seen in two distinct ways (though never simultaneously) without the picture itself ever actually changing (Dear 2006, 6). This invisible circularity between two ideal types itself works as a powerful rhetorical device: it works because it's true; it's true because it works. The instrumental/natural-philosophical invisible circle is so mesmerizing to consider that it is easy to ignore, for instance, our everyday encounters with "true" beliefs that don't work and "false" beliefs that do (e.g., geocentric astronomy as a means of navigation).

Recording engineers, like most people who call themselves engineers[12], do much of their work within something resembling Dear's invisible circle. Instrumental and naturalistic rationales constitute, in part, what Mulkay terms the "evaluative repertoires" of their work (Mulkay 1976, 643). Recording engineers also make careful but effective use of creative and artistic repertoires. They are engaged in the production of popular music recordings where emotional impact and disposable novelty are valued primarily, which means that pretenses to technoscientific prestige are often taken as fundamentally suspect. At the same time engineers occupy the instrumental-rational and objective role within a technologically complex production process, and thereby base their professional pride on their ability to perform this role well. The invisible circle of the recording engineer is complicated by the fact that the instrumental-rational goal of their work is in large part taken to be irrational insofar as it is emotional.

Technologies of the voice, such as digital vocal tuning software, are tangled knots of social relationships and material practices which cater to this double bind of technical and emotional evaluation. In this sense they echo the hybrid qualities of voices taken as both windows of subjectivity and objects of mechanical reproduction. The voice is so often deployed as a metonym for internal emotional identity, the possibility of interpersonal communication, and political agency that it is easy to lose sight of its everyday materiality, its basis in the movement of hot air through individually specifiably assemblages of wet flesh. The media technologies that have emerged to bridge these corporeal and conceptual domains are places where the material

[12] For more on the historical relationship between engineering, technology, science, and technology as "applied science" see (Kline 1995)

and the social bleed into one another and come into contingent but durable arrangements. This is to say that they are negotiated within particular social and material structures of meaning.

An iconic example from art history may help illustrate the kinds of negotiations which I aim to locate within the field of modern vocal production. Michael Baxandall argued that 15th century Italian paintings were products of culturally and historically specific modes of perception and artistic production, appearing as "among other things, fossils of economic life."(2) Details of a commissioned painting, ranging from the posture of the figures depicted to the qualities of the pigments used, arose at the intersection of contractual social negotiations between painter and patron, and the play of social conventions which formed the context of these transactions. The development of vanishing point perspective, for example, had much to do with shared ways of "gauging" the world in terms of geometric planes, a perceptual habit rooted in the everyday surveying practices of quattrocento mercantile life. With his concept of the socio-economically conditioned "period eye," Baxandall built on the work of Ernst Gombrich, for whom perception was learned through a culturally conditioned dialectic between schema and correction.

The recorded voices of mass-market popular music at the turn of the 21st century can be thought of in similar terms. They are, to borrow Baxandall's phrase, "deposits of a social relationship," (Baxandall 1988, 1) one that includes but is not limited to clients and signal processing practitioners. If Renaissance painting was indeed "too important to be left to the painters" (Baxandall 1988, 3) then the voices of modern pop music are today certainly considered too important to be left to the singers. As Engineer Sylvia Massy puts it: "Only one person typically gets in the way of a great vocal performance: the singer." (Massy 2016, 69) Much about the precise ways in which voices are currently "not left" to vocalists can be captured in Ernst Gombrich's framework of schema and correction (Gombrich 1977). Gombrich argued that the work of a painter involved a tacking back and forth between act of representation and attempts to correct representation by way of comparison with the thing being rendered. Vocal takes, which are taken to represent one or more aspects of the performer's (literal and figurative) voice are routinely subject to "correction." Pitch correction, in particular, has over the last two decades become a standard practice in computer-based music production (i.e. the vast majority of it.) Recording engineers have increasingly directed their efforts towards repairing vocal performances digitally, either during a singer's performance or long after they have left the vocal booth. The oscillation between how the voice is rendered and corrected through the apparati and practices of the studio is where we need to look if we want to find how voices are tuned-in *in situ*. The "zero-crossing," or obligatory passage point[13] of these oscillations is the recording studio, so into the studio we now turn.

The Field Sites: Studios A and B

In addition to a wide array of textual sources regarding the historical and contemporary practices of recording engineering, I draw on an ethnographic study of two recording studios in Los Angeles during the latter half of 2013. I will refer to these as Studios A and B. Both studios are in the middle-range in terms of size and cost-of-service. This means that they are considerably more well equipped than the home studios that have proliferated with the increasing

[13] A zero-crossing is a point at which a waveform intersects the X axis, or a point at which Y is equal to zero. Finding zero-crossings is one widely used technique for approximating the pitch of a signal via "feature extraction." I read this signal processing metaphor alongside the Actor-Network Theory concept of the "obligatory passage point" (or OPP: itself a term drawn from modern military parlance (Callon & Latour 1981)) in order to suggest that recorded voice is itself the effect of a network of actors and actants arranged in space, as well as a phenomenon (like a waveform) which only comes into being as temporally emergent movement.

availability of inexpensive computer-based recording technology. However, they are not as large or expensive as more famous studios such as Capitol, Henson, or Paramount, which draw financial and infrastructural support from large media conglomerates. Most of the important early studios (e.g. RCA, Atlantic, Columbia, and Motown) were originally owned by major broadcast and publishing corporations. The major label system is currently composed of three organizations: Universal Music Group, Sony Music Entertainment, and Warner Music Group. The recent contraction of the recording industry has seen a pattern of consolidation among major labels and a decrease in in-house label studios, though some persist. Capitol Studios, for example, is owned by Universal, which is owned by Vivendi Media Group.

The midrange studio has recently taken on greater importance as major labels have sought independent facilities for their projects and artists who started in home studios have sought more professional facilities. Hip-hop artist Kendrick Lamar, for example, recorded his acclaimed platinum-selling debut studio album Good Kid, M.A.A.D. City at four different studios. These included his mother's house, the suburban home studio of his label CEO, an independently owned mid-range studio in Atlanta, and Dr. Dre's No Excuse Studio, which is owned by Interscope, a subsidiary of Paramount, itself a subsidiary of UMG. The distinction between major label, home, and midrange studios maps onto the local geography of Los Angeles. Studios A and B are both located in Downtown LA, the Arts District of which has become something of an articulation point between the established studio system based in Hollywood to its northwest, the socioeconomically disadvantaged neighborhoods of South LA, and the rapidly gentrifying neighborhoods of Echo Park, Angeleno Heights, and Highland Park. Downtown's central location and relatively low costs of living have made it a hub for do-it-yourself (DIY) musical communities including the flourishing home-studio hip hop scene of areas like Compton, the punk scene associated with the now-defunct venue The Smell, and the hipster neo-psych scenes rooted in Silver Lake. These latter two communities have been recently bolstered by Bay Area musicians seeking more affordable housing.

Studios A and B often work in collaboration with projects based in part in larger studios, and the engineers and proprietors of these studios both spent a good deal of time working in the major label system before striking out on their own. They are also relatively young studios, each having been open less than a decade at the time of my fieldwork. The owners of studios A and B each had over a decade of experience working as in-house and freelance engineers at a number of studios in and around Southern California prior to starting their own studios. Each serve as chief engineer and proprietor of their respective studios, while also employing assistant engineers, runners, interns, and project specific guest engineers as needed.

Studio A, whose owner and chief engineer I call Carl, is the smaller of the two. Carl specializes in the genres of rock, pop, and hip-hop. Much of his studio's work involves bringing in projects that were begun or may continue to be worked at other studios. Because the tracking area is not large enough to comfortably host an entire live ensemble simultaneously, larger bands will often bring Carl sessions consisting of tracks recorded elsewhere, and Carl will work to finish them up by adding overdubs, editing, and adding effects after the fact. Studio B, whose owner I call Harold, is larger both in terms of physical space, employees, and number of clients. In terms of genre, Studio B specializes in heavy metal, experimental, "indie rock" and pop music. Studio B also occasionally hosts recording sessions for advertising music, which are usually helmed by non-resident engineers who simply want to make use of the equipment and recording space. Most of my time was spent working directly with either Harold or Carl. Accordingly, they formed the two major nodes within a single tight-knit and largely overlapping community of recording professionals. This community included regular interaction with 15 recording professionals at various career stages (including technicians, assistant engineers, mixing and mastering engineers, runners, and interns.) I conducted extended one-on-one interviews with eight of these engineers and worked on sessions involving 35 in-person clients. Three additional clients passed through Studio A in the form of pre-recorded material needing additional work, meaning that I never properly met them but nonetheless spent hours listening to and helping Carl work on their voices.

Between studios A and B, with considerable overlap as a result of their physical proximity, I spent five months in-studio, beginning in August of 2013 and continuing through the end of the year. I entered studios A and B as an intern. This was for reasons both substantive and logistical. On the one hand I was interested in investigating both the infrastructural work behind music production and the accultural processes that recording professionals encounter in their earliest career stages - both of which would be most apparent at the lowest level of the studio hierarchy. On the other hand the intern role offered a solution to a key question of access and reciprocity. For a studio owner, having an intern in-studio can be both a luxury and a liability. Even the sharpest interns have some sort of learning curve, and the long hours and close quarters involved in studio work can lead to strained relationships. As an intern, I was in many ways a risky prospect for Carl and Harold. Unlike most of the candidate pool I had not attended a formal recording school, and would thus require more on-the-job training. Secondly, I was openly engaged in pursuing a career other than engineering in the long term, meaning that I might not be fully focused on the job at hand and would not necessarily build their informal networks by going on to work in other studios as a product of their mentorship. Third, I had come to know Harold and Carl through mutual friends, and had gotten to know them both in other contexts prior to my fieldwork - this could have complicated what would have otherwise been a more straightforward relationship between studio owner/engineer and intern. Finally, I was proposing to study them ethnographically, which was of course an unusual thing for an intern to be doing. All of this is to say that Carl and Harold were extremely generous to bring me into their studios and to allow me to attempt to show others at least a portion of how they go about their work. The following section takes up the particulars of how the show will go on.

Chapter Overviews

Chapter 1 examines how engineers think about and work with "emotion" specifically with respect to its relationship between the studio's technological artifacts and the voice of their client. Chapter 2 looks at two major digital tuning tools – Auto-Tune and Melodyne – in terms of their initial development and the different technological prescriptions they brought to the technological problem that would eventually come to be known as "pitch correction." Chapter 3 traces the ways that recording engineers developed for working with, problematizing, and *repairing* digital tuning techniques understood in terms of a "covert" and "corrective" approach. Chapter 4 looks at the various ways that pitch correction was repurposed as an "overt effect," primarily by non-engineers such as songwriters, performers, and listeners. Chapter 5 zooms-in on an instance of corrective tuning as an interactional accomplishment that draws upon techniques of inscription, repetition, and inflection.

Chapters 6 and 7 take as their starting point the distinction between "place" and "time" theories of pitch perception in psychoacoustics and signal processing, in an attempt to thicken and ethnographically respecify the ways that place and time are structured and performed in-studio. In these chapters the goal is to provide a more adequate understanding of how the objects of digital tuning (i.e. the tuning technologies as objects and vocal performances as objects of tuning work) are practically constituted. Chapter 6 zooms-out to examine the *topical* structure of the recording studio, or the way that studio life is performed through various acts of place-making. Chapter 7 concerns the *temporal* structure of studio life, taking up the various ways of keeping, making, and marking time in the studio. In the Appendix I draw on observations of studio life in order to develop the concept of the "*shibboleth*" as a complementary concept to boundary objects and boundary work, as well as a new way of thinking about the mutual production of social and perceptual difference in STS.

CHAPTER 1: EMOTION AND THE RECORDING ENGINEER

This chapter traces how ways of thinking about and working with emotion as an element of recording engineering has changed over the last century while remaining an enduring and central topic. It focuses, on the one hand, on the emergence of the engineer as a professional category and its changing meaning with respect to emotional expression over time. On the other hand, it tracks the shifting balance between understandings of emotion as something that is actively technologically produced as opposed to something that is transmitted more or less faithfully by equipment and the engineer. It traces the interplay between these two conceptual strands over what Edward Kealy has termed the course of the craft-union, artistic and entrepreneurial modes of engineering work. It finally argues that the categories of transmission and production have been rearticulated through perceptual practices which locate emotional production in the sound of the recording process as a transmission medium itself.

Production and Transmission Models of Emotion

Luigi Russolo developed his noise instruments as a solution to what he saw as the problem of the inability of normal musical sounds to evoke emotion in the modern listener. In 1916 we wrote in his futurist manifesto The Art of Noises, that "machines create today such a large number of varied noises that pure sound, with its littleness and its monotony, now fails to arouse any emotion." (Russolo 1986, 5) For Russolo, if machine noise was the problem, it could also be the solution. His new orchestras, composed of mechanical noise generators, would "produce the most complex and newest sonic emotions"(Russolo 1986, 11) Upon hearing "the harmonized combined pitches of the bursters, the whistlers, and the gurglers, no one remembered autos, locomotives or running waters; one rather experienced an intense emotion of futurist art, absolutely unforeseen and like nothing but itself." (Russolo 1986, 15) Russolo viewed the relation between emotion and sound technology as one of production.

Around the same time that Russolo was trying to convince people that his devices could furnish them with new, futuristic emotions drawn from the same mechanical sounds that had deadened them to the music of the past, Thomas Edison was hard at work convincing people that musical emotion in the machine age was alive as ever, and could be summoned on demand through his phonographs. As early as 1916, Edison's "realism tests" were combining the genres of sales pitch, scientific experiment, and training regimen in order to teach would-be phonograph owners how to listen properly to recorded music such that they may experience "same emotional re-action"(Katz 2010, 23) they would have to the real thing. Edison would later enroll Walter van Dyke Bingham, experimental psychologist and architect of the U.S. Army's first foray into intelligence and personality testing during World War I, to design and facilitate a mass campaign of "mood change" tests in order to scientifically analyze "Your Mental Reactions to Music, as RE-CREATED by the New Edison, 'the Phonograph with a Soul" (Hui 2012, 611) Edison saw emotion and sound technology in terms of a problem of transmission.

Edison and Russolo were both working during the emergence in the US and Europe of the concept of "technology" as a singular, inclusive term for all things mechanical and instrumental (Marx 1994, 18) as well as an increasing interest in "emotion" as an alternative to the romantic Victorian concept of "spirit" (Stearns & Stearns 1985; Dror 2000). They were also working with sound – a sensory register that had long been associated with feeling and emotion (as opposed to vision's connotations of objectivity and intellect) in Judeo-Christian culture.

(Sterne 2003b) Their respective approaches to the problem of emotion and sound technology serve as conceptual anchor points for what would be an ongoing concern with the proper relation between the technological capture and production of sound and the way it relates to emotional experience. Russolo's narrative cast mechanical technology as both the cause of our degraded emotional relationship with music and also the means by which we can discover new emotions. Edison's narrative was simultaneously more optimistic and conservative than Russolo's. It was one of naturalistic recovery of musical emotion through sufficiently transparent means of technological reproduction. For Russolo, there was no going back to the garden (or at least no way of getting the machine back out of it) so we ought to get to work making a music for a mechanical world. In Edison's narrative, it was possible for the sound of machinery to destroy emotion, but only by way of unfaithful representation (an excess of technological noise in the reproduction process) or the failure of a listener to experience it correctly.

All of this is to say that emotion has been closely associated with the problem of sound's technological (re)production from the earliest days of the profession that we now refer to as recording engineering. The discourses of emotional realism and emotional creativity are still largely current, but have also recombined in significant ways. In many cases the ideal of transparency in sound recording has mingled with the idea that emotion can be achieved technologically in ways that go beyond the "natural" musical sound. This has led many engineers to pursue a production aesthetic that is a "more perfect" (or "too perfect" as some would argue) emotional representation of the performance than the "raw" takes. Many engineers take the opposite approach, wherein the audible "artificiality" of the recording process is taken to index the emotional intimacy of the performer with respect to the listener. Mark Katz has shown how recorded music, made widely available through playing the phonograph, was initially understood as offering "even the most hardened man a way to express deeply felt emotions and explore his latent musicality."(Katz 2010, 67) Katz has further argued that new musical techniques – specifically the use of vibrato by violinists – emerged in response to the problem of making one's playing style recognizable and conveying emotion in the absence of performer's visual and bodily presence:

> Mischa Elman's "throbbing," for example, can be easily distinguished from Jascha Heifetz's "nervous" vibrato; and no one would confuse Fritz Kreisler's omnipresent shake for Marie Hall's essentially decorative use of the technique. And within a single work, any of these artists might choose to emphasize certain notes or phrases with added vibrato, or to communicate increasing or relaxing emotional intensity by changing the speed or width of vibrations. A distinctive and nuanced vibrato, then, could go some way toward restoring some of the "poetry" lost in the recording process. (Katz 2010, 107)

The last line of this passage conveys the other side of the emotional promise of sound recording. Even as the possibility of recorded sound was perceived as a revolutionary new way of conveying and experiencing emotion, the recording process was itself understood as a threat to the very emotion it was supposed to capture. Listeners' bodies had to be trained to be moved by recorded music. Musicians had to develop new ways of playing in order to make sure their emotional signature survived the recording process. The facilitators of the actual recording process – the engineers – took on an ambiguous role with respect to the project of emotional transmission. As operators of the recording equipment, they were both the gatekeepers of new networks of emotional expression and the ones thought most liable to ruin it. As Susan Schmidt-Horning has vividly captured in her book *Chasing Sound*, early recording engineers often seemed to treat their clients like pieces of equipment. The engineer's relation to the musician was largely one of adjustment and manipulation: "Some musicians might get carried away in the emotions of the performance and forget these instructions, putting recordists in the awkward position of manhandling them" (Horning 2013, 16). Gramophone recordist Max Hampe would push and pull singer Frieda Hempel towards and away from the microphone depending on the volume and pitch of the notes she was making. Violinist Maud Powell, saw the engineer as a sort of opponent to her own emotional expression:

> The artist who makes a record sings into the horn and sees nothing else except a bare wall and the face of the operator at a tiny window. Facing this cold, indifferent prospect, which is ominous in its scientific aloofness, one must dig into one's own soul for the impassioned touch that is afterwards to thrill one's hearers, and— it is something of an ordeal (Horning 2013, 19).

Restrictions on time and movement, imposed by the engineer on behalf of the relatively unforgiving apparatus of the early recording studio, were taken to stand in opposition to the emotional qualities of the performance even as they supported clarity and technical accuracy of the resulting record. In one characteristic account, the time-consuming and time-restricting aspects of the recording setup appear as impairments to emotional expression:

> Ponselle echoed the sentiments of other dramatic vocalists of her time when she recalled, in 1976, that what she found least satisfying upon hearing her early recordings was the absence of emotion: "I couldn't give as much as I would have liked to give. Emotionally I would have taken more liberties, but what can you do in three or four minutes? Besides, all the cuts they had to make— cut this bar out, and cut that bar out to save a second or two. I knew that I was limited to a short amount of time, and that sort of handicapped me, put a harness on me. (Horning 2013, 21)

This necessary but antagonistic relationship between the experience of musical feeling and the technical basis of its reproduction was of a common thread with broader attitudes toward emotion and science in early 20th century America. Historians of emotion (Stearns & Stearns 1985) have pointed to a shift from a passionate Victorian emotionology to a more economical one they term "American cool." Brenton Malin has drawn a parallel shift in modes of speech training by the 1920s, from the highly gestural and emotionally exaggerated techniques of "elocution" to the more conversationally "natural" and emotionally controlled style of "public speaking." (Malin 2011) Otniel Dror, meanwhile, has observed a major shift in the Anglo-American world during this period with respect to the relation between emotional experience and technical, medical, and scientific knowledge. Dror argues that by the 1930s,

> predominantly laboratory- and clinic-based developments transformed emotion into a modern object of knowledge and introduced a new era in the history of emotions. The novel interactions between scientific observers, "disinterested" machines, and emotions generated new knowledge and new kinds of relationships: instruments supplanted personal interactions in retrieving intimate emotional knowledge; the interior body superseded language as the primary medium for expressing emotions; the boundaries between private and public, subjectivity and objectivity, and inside and outside were blurred, as experience was exteriorized through laboratory procedures; class, gender, and race were reified through new quantitative emotional measurements and gradations, replacing Victorian qualitative categories; and physiologists and clinicians exhibited and exchanged graphs of their own emotional experiences--producing an economy of emotions inside the laboratory and clinic. The new scientific images of emotions differed significantly from previous artistic depiction in their mode of production, form and style of representation, method of interpretation, and use. Their production and interpretation resonated with and mediated broader historical, epistemological, and cultural transformations of the late nineteenth and early twentieth centuries. These latter include the rise of the graphic method; the emergence of modern forms of knowledge, and, in particular, the changing definitions of objectivity and of the "image of objectivity"; the modern "appropriation of subjectivity"; the shifting representation of the body (i.e., from a "spiritualized" Victorian to an emotionalized modern); the modern embodiment of emotions; the emerging challenge to the boundaries that separated science from art; and the advent of a late-nineteenth-century "evidential paradigm" (Dror 1999, 3).

One specific cultural transformation that Dror does not mention but nonetheless resonated with these changes was the birth and consolidation of the commercial recording industry. The development of electrical recording techniques, the professionalization of recording engineers, and the opening of the first modern recording studios took place alongside a massive refashioning of the relationship between voice, emotion, and the technologies that tied them together. As engineers, recording professionals saw themselves as part of a scientific order, which was newly able to engage with emotional experience at the interface of technical equipment and the body. Again, however, this new legibility of the performing body for the technical apparatus became a site of tension, with engineers often caught in the middle. Limitations on recording time, for example, caused problems with artists like Frank Sinatra, who in 1947 refused to change a song's tempo, lest it "kill the feeling" even though it meant it would not fit on the three minute window of the 78rpm record. (Horning 2013, 108)

The professional structure of the recording industry also conditioned the possible economies of voice and emotion. Legendary British engineer Joe Meek, for example, was well-known for his hand-built and idiosyncratically wired "black boxes" with which he could get uniquely strange and expressive sounds in studio. Meek found it more difficult, however, to get an emotionally compelling performance out of the session musicians who he worked with at Lansdowne studios in the late 1950s.

> [Meek] would try all means of coaxing more feeling out of them, sometimes even going so far as plying them with alcohol. But it really shouldn't have been all that surprising that the musicians didn't pour their emotions into every song that they recorded. Top-drawer session players worked lots of sessions, sometimes from ten o'clock in the morning until well into the evening, often at more than one studio. They were routinely expected to sit down cold in front of a sheet of music or a chord chart and be doing a fine job of playing the tune a few moments later. (Cleveland 2001)

Similarly, the era of what Edward Kealy terms the "craft-union" mode of audio engineering that emerged after World War II meant that the relationship between musicians and engineers was often extremely formal and impersonal – with the engineers (often former military radio operators) rarely being musicians themselves – and guided by a decidedly "utilitarian" aesthetic.

(Kealy 1979) This strong institutional separation between the engineering and the performing professions, coupled with the limitations of early recording setups, meant that the emotion of a performance was understood as largely being the purview of the musician. Engineer Glyn Johns, who would later make his name recording The Beatles, relates how his concept of musical emotion is strongly tied to the practice of recording all of the musicians simultaneously:

> Having started when everything had to be recorded at once, I have never lost the value of musicians interacting with one another as they play. This can be so subtle, and invariably is nothing more than a subconscious emotive reaction to what others are playing around you, with what you are contributing having the same effect on them. When a musician overdubs his or her part onto an existing track, this ceases to be a two-way reaction. With only the musician who is added being affected by what he or she is playing to. Recording equipment was originally designed to capture the performance of a piece of music. Now it influences the way music is written and performed. (Johns 2014, 169-70)

The barriers between engineers and musicians at unionized corporate owned studios were by the late 1950s and early 1960s being challenged more "entrepreneurial" mix engineers operating independently of the major label system. These engineers took advantage of newly available and relatively inexpensive recording techniques such as magnetic tape and catered to audiences (typically younger and from marginalized racial and socioeconomic backgrounds) that did not value the "high fidelity" aesthetic of the large studio productions. The engineer of this period became "more like a service worker who must please his clients without the benefit of appeal to a set of craft standards enforceable through his union." (Kealy 1979a) In the absence of craft standards, the experience of the client and the satisfaction of their particular expressive goals became increasingly important for the engineer. The mid to late 60s saw yet another shift, this time towards what Kealy calls an "art" mode of production, where engineers (increasingly with musical backgrounds of their own) began to assert themselves as artists with their own aesthetic vision. The service ethic was replaced by one of creative collaboration. The engineer in this case was charged with shaping the emotional statement of the recording, or even contributing emotional content to it, rather than simply capturing it in a detached and hopefully non-destructive way.

Kealy's periodization is both picked-up on and partially challenged by that of Chicago-based engineer Steve Albini, whose engineering career began roughly where Kealy's account leaves off. According to Albini:

> Over a fairly long period of the evolution of recorded music, the people who worked in studios recording music were a different class of people from the people who were in bands and performed music. And there developed an aesthetic where the recorded sound of music - in particular, the recorded sound of a drum kit - was not representative of the organic live sound of a drum kit in a rehearsal room or a club or anything like that. And that culminated in a period in the late 70s, early 1980s, when producers and engineers were willfully replacing drummers in bands with drum machines because drum machines were more predictable and they more closely approximated the idealized sound that these producers and engineers had been striving for in this disassociated studio aesthetic. There's been a paradigm shift where people who work recording music a lot of the time now come from a background of having played music, having been in bands and having been fans of music and being steeped in the exposure to live performed music. And so the aesthetic has changed and gradually supplanted the old aesthetic of making studio recordings that were uniquely studio sounding and not necessarily representative of a live sound, to studio recordings that were intended to evoke a naturalistic or live sound... Starting with the punk era [late 1970s] and carrying on to now. (Albini 2015a)

Kealy and Albini's professional narratives agree to the extent that both note a decoupling between a "realistic" and "studio" sound by the late 1970s. Where Kealy attributes this to a blurring of the line between engineers and musicians, however, Albini sees it as a continuation of the sort of class-based separation Kealy limits to the craft-union period. We can account for this disagreement partly in terms of genre orientation. Albini locates himself within the "punk era" which was getting underway when Kealy's article was first published. Writing in 1982, Kealy categorizes punk in the "entrepreneurial" rather than the "art" mode. This means that, from the perspective of a punk like Albini, the "art-mode" mix engineer would represent a "different class" from musicians in the emerging punk scene. This dissonance highlights how community-specific understandings of the proper relation between emotional expression and musical production could be. For Kealy's "artist" engineer, the ability to cultivate an idealized studio aesthetic provided an opportunity for emotional articulation in what had previously been a purely technical and representational profession. Albini, by contrast, locates the site of emotional production in the everyday sound of the rehearsal room or live venue. In his account, then, a return to sonic "realism" in the punk era in fact meant a more emotionally resonant recording style.

New Orientations towards Emotion and Application

By the early 21st century the craft-union mode of recording engineering, at least in the context of non-film or television music production, had been largely displaced.[14] Even at major label-owned studios engineers are no longer unionized, and the ideal of the completely technical and detached engineer (let alone one with expectations of job security) is no longer seen as viable. At the same time, however, the "art mode" mix engineer is also a poor fit to the current landscape of recording engineering. Engineers, when they are working under the title "engineer," at least, are better described as entrepreneurial. When an engineer ventures into the "art" mode, they are increasingly likely to be credited as a "producer,"[15] and with that title come logistical responsibilities (booking studio time, managing the recording budget, providing overall guidance to a project, arranging for musicians and hiring engineers, etc.) that go beyond the job

[14] This is especially the case in Los Angeles, where the ethnographic research for this dissertation was conducted. Nashville studios make much more frequent use of session musicians, which makes the craft-union ethos somewhat more current in studio work.
[15] For a thorough analysis of the role of the producer of popular music as a "mediator" between production and consumption, see (Hennion 2013)

description of the engineer. There may be an increasing number of producers who, to save time and money, take on engineering responsibilities, and an increasing number of engineers who take on producing roles in projects.

Still, the engineering role and the producing role are still understood as distinct (even if one person may choose to perform both) and with it remains a distinction between the emotional contribution expected from the producer-client and the modicum of un-emotional objectivity expected from the engineer. Among the most influential public figures in the audio engineering community, it has become a commonplace that the role of the engineer is fundamentally about capturing and manipulating emotion. Portland, Oregon-based engineer and founder of Tape-Op magazine Larry Crane writes that, “Capturing art, expression, and emotion – in any way possible — is the goal.”(Crane 2016b) Dave Pensado, a Grammy-winning LA-based mix engineer and the host of a popular online show catered to engineers, articulates the emotional goals of the engineer in terms of affective escape:

> We talk all the time about the emotion in a record. We’re not selling our EQing chops, we’re not selling our abilities to run a console. What we’re selling is four minutes of escape from the real world. And the way you get them to forget about the real world is by serving them up a heavy dosage of feeling, vibe, and emotion. So, we’re kind of emotion manipulators in way, when we mix. And the way you create emotion, it can be with a plugin it can be with analog gear it can be with a washing machine. It doesn't really matter what you use, it’s the end result is somebody listening and forgetting about the war, forgetting about the election, forgetting about their parents, forgetting about life. (Pensado 2012)

Eric Sarafin, LA-based mix engineer and author of widely read books including the Mixerman Chronicles, writes in his book Zen and the Art of Mixing: “the only thing that matters is the music itself. If that music elicits an emotional response, then we’ll appreciate it; if it doesn’t, we won’t.” (Mixerman. 2010, 25) He goes on to highlight that the engineer’s work is defined in large part by the tension between emotion as a matter of subjective taste and the possibility of crafting it in an objective way:

> The problem is that there's really only one way to personally measure emotional impact, and that's with your ears and brain. Further confusing the issue is how a song's impact will differ from person to person. How a song affects any given individual is beyond our control. It's how a song affects overall human consciousness that offers prima facie evidence regarding its overall effectiveness... At some point, if you can separate yourself from your own personal likes and dislikes, you can begin to understand when you're working on a song that is on par with the great ones— it manipulates our emotions. (Mixerman. 2010, 3-4)

The tension between emotion's irreducible subjectivity and its status as the major goal of engineering work is mediated in the domain of specific practices, or tricks of the trade that allow the gap between technical procedure and emotional effect to be overcome. The importance of emotion and its unique technological intractability comes across in an exchange between Crane and Bay Area engineer Scott Fritz:

Fritz:	If the emotion is not coming out of their voice and they're not feeling it, it's very apparent.
Crane:	And harder to fix than a bad band.
Fritz:	You really can't EQ that out can you?
Crane:	Compress it.
Fritz:	EQ in some emotion. (Crane 2000)

Crane and Fritz are speaking ironically here. It is not taken as possible to simply "compress" feeling into an uninspired vocal, and you likewise cannot simply "EQ in" some missing emotion. The fact that emotion is taken to be so complex and, almost by definition, beyond the technical domain, is what makes it such an enduring interest for recording engineers. Emotion is not easily defined, pinned down or reproduced. It is emotion's durability as a sociotechnical puzzle, and the possibility that it might eventually be solved, that animates how some engineers imagine the future of their profession. Engineer and recording software developer Steven Slate describes his reaction to a new online tool that allows for rapid automatic mastering of stereo mixes:

> I had to let go of my ego. I'm an audio engineer, and here is a system that attempts to replace an engineer. Artificial intelligence is the new reality. Stop being dismissive and afraid. Yes, mastering is an art. But do you really think that in some years' time, some software process won't be able to analyze emotion? That's what AI is. (Hong 2014)

The concept that engineers must let go of their ego is a familiar one, but it is usually understood as a need to subordinate one's own tastes and priorities to those of the client. The possibility that future tools will be able to analyze and work with emotion directly suggests that engineers will have to subordinate their own abilities to deal with the client as an emotion source to a more efficient technological system. The engineer's dream of perfect control over the emotional expression of their client during the recording process, in Slate's vision, seems to promise the twist ending of taking emotional agency away from the engineer as well. Imagined futures where AI will compete with the emotional skills of the engineer should be read alongside current ways in which engineers are encouraged to cultivate an unemotional disposition so as to better manage the emotional qualities of their recordings. Sarafin writes:

> You don't serve anyone well by getting emotional, no matter how hysterical those around you become. Not that you won't occasionally get sucked into the drama— you will. You're human, and hopefully you're passionate about what you do. But the better you become at ignoring the emotional outbursts and instigations of a drama queen, the more effective you'll be as a mixer, and the less problems you'll have completing a challenging project. (Mixerman. 2010, 203)

Accordingly on of the most important of Sarafin's self-described techniques is the judicious participation (or abstention) in what we might term the diegetic and non-diegetic[16] emotional

[16] In film theory, diegetic elements are those that are depicted as occurring in the world being depicted, while non-diegetic elements occur at a separate level associated with the particularities

energies of the studio setting. For instance, while Sarafin and others advocate strategies of emotional provocation in the interest of getting a singer riled up for a take, he specifically recommends a "Zen" approach to actual crises:

> It's best to be unemotional when dealing with a mix catastrophe; otherwise your frustration will surely reveal itself. This will only serve to make your client defensive, or worse yet, emotional. Emotion evoked by any source other than the music itself only clouds communication. (Mixerman. 2010, 218)

Here Sarafin is carving out a specific aspect of the emotional experiences possible in-studio and marking it as the proper object of the recording process. The engineer is not supposed to get "emotional," but neither is the client, exactly. It is instead a problem of accomplishing the emotion required by the project. The recording process, in its approach to the voice, has become a way of commodifying or otherwise *economizing* emotion, both in terms of controlling its circulation and in terms of making it available as a saleable object. In-studio drama, though commonplace, is considered undesirable because it is a *waste* of emotion. For Sarafin, emotional impact is something that an engineer can learn to accomplish by way of "concrete techniques" (Mixerman. 2010, 5) which render up emotion at the interstices of social and technological action. As Seattle-based engineer Randall Dunn puts it, the emotive aspect of a recording is a matter of artisanship:

> The emotive aspect of the recording is directly related in a syncretic way to how they're hearing the music, and the choices you make as an engineer should back up everything that the record's trying to say emotionally. It takes anyone a very long time to get the colour right, and you don't always get it right either. I just think there's an expertise there and an artisanship, so to speak, that is lacking in a lot of modern music. (Kostelnik 2016)

---

of the film-as-production. A film's score is typically non-diegetic, for example, while the music on the radio in a character's car *is* diegetic.

Dunn's description of the relation between the recording's emotive aspect and the way clients hear the music as "syncretic" gets at the way that studio work brings together technological systems and social worlds in highly localized and provisional instances of what actor-network theorists have termed "heterogeneous engineering."(Law 2012) Emotional and musical articulations are cobbled together in-studio from a wide range of phenomena, including but not limited to sensory registers, shared cultural references, local material conditions, and interpersonal relationships. Engineering work is marked by the reconciliation of diverse beliefs about what is going on in the recording process and how things ought to proceed. Not least among these potentially conflicting beliefs are those drawn from the technical domain of the engineering tradition and those drawn from the aesthetic and musical domain of the artist.

In linguistics, syncretism refers to the combination of inflectional forms in the creation of a new language. The term "inflection" must in itself be understood syncretically. For an electrical engineer it might refer to a change in sign (or positive or negative direction) in a signal plot. For a linguist, inflection means the alteration of a word in order to express different grammatical forms (e.g., verbs are inflected via conjugation – to *be*, I *am*, or we *are*.) For a singer, voice over artist, or speech therapist, inflection means a change in intonation in order to convey a mood or feeling (e.g., an upward inflection might be used to indicate a question.) In a studio setting, all of these meanings and more are potentially fair game. Engineers like Dunn take emotion to be both practically accomplishable and technically unspecifiable because of the diverse range of inflections that it incorporates.

Even as new techniques and technologies are elaborated in the pursuit of emotion on-record, the anxieties of Russolo and Edison are still in evidence. Dunn suggests that skilled attention to the emotive aspect of music is lacking in modern music. This is a common sentiment among many engineers, and can be understood in part as a continuation of the anxieties about how recording techniques risk sacrificing emotion. With every technical refinement or elaboration of a recording technique seems to come a wave of concern that the displacement of established approaches entails a loss of emotion. Where early recording techniques raised concerns about the loss of emotion because the apparatus was so restrictive, the trope has become that recording techniques have become overly permissive, allowing engineers to indulge in the pursuit of perfection at the expense of emotions that would otherwise come to life through minute imperfections. Pioneering engineer Al Schmitt provides one articulation of this feeling:

> Unfortunately, because we can do all these things to manipulate the recordings now, we are fixing every mistake and everyone is singing perfectly, and sometimes we make records that are just too perfect. We take the emotion out of them. I mean, nobody ever had to tune Frank Sinatra.[17] His pitch wasn't perfect every time, and y'know, if there was a little mistake on something, it added to the emotional effect of the record. If somebody in the band played a wrong note, but the emotion was right there, you kept that take, and that was the one that would come out. (Massy 2016, 67)

The risk that something will be "too perfect" and therefore lack emotion leads engineers to rehearse the trope of "mistaking something for a mistake." Engineer Andy Selby makes a practice of listening to the vocals he edits "over and over again just to make sure that I am not missing the emotional impact of something being technically wrong."(Tingen 2015) Albini draws on his background in punk music to avoid making things overly perfect.

> I come from a punk rock background and a lot of the punk records that are really important to me and that are touchstones and things that I cherish, if you listened to them in a sort of analytical way they are total fu***** disaster. The sounds are horrible, the playing is really bad and nothing about it is "correct", but that record changed my life so where would you improve something that has been a critical life changing experience. I

[17] A 2016 advertisement in Tape op for Bock microphones make similar use of Sinatra as an icon of imperfect intonation and perfect emotion. The copy reads: "Capture a Classic / When Frank serenaded New York… there was no pitch correction."

> think I have to allow for the possibility that I just don't understand something, I refer to it as listener error. Like if I'm listening to something and a band does something and it sounds off to me and I play it for them and say 'hey I want you to check this one spot, it sounds odd.' And I play it and the band says yeah that's normal, then that was listener error on my part. Like I mistook something for a mistake. (Albini 2015b)

The service ethic that Kealy describes as characteristic of the entrepreneurial mode is increasingly intertwined with an image of engineer as custodian of recorded emotion. Albini is well known for his habit of wearing auto mechanic coveralls while working, a uniform associated with his decision to always be credited as an "engineer" rather than "producer." Regarding the rise of digital vocal tuning in particular, on online commenter in a recording engineering forum draws an explicit connection to his earlier experience as a janitorial worker:

> I am reminded of when I worked at a service station in high school just cleaning up around the place, where I learned a valuable lesson. One person's mess is another man's job security. All in all the more dependent music is on technology the more work there is for us right? So why should we complain about it! Here's to job security! (Fallforward 2009)

Tape Op editor Larry Crane, in an article titled "Welcome to the Service Industry" draws an explicit connection between the decade he spent working in restaurants and his approach to engineering.(Crane 2013) By some accounts, the recording engineer's service ethic and role as emotional custodian increasingly resembles that of a therapist or a bartender. Engineer Harold describes his approach to Hip Hop sessions as "like a therapy psychology session in that you're taking these nuances and kind of creating this atmosphere to create, for everything to be fine." Engineer Bob takes a somewhat more jaundiced view, where the devaluation of the recorded product has emphasized the emotional connection between the engineer and client:

> It's a service industry… I can purchase PBR [Pabst Blue Ribbon beer] at any bar in LA, the same exact price, you know what I mean? But why do I go to these certain places? Because I like the person doing it, I enjoy them doing it. I like certain bartenders giving

> me that PBR. You know? It's the same crap. It's just mp3s. At this point it's shitty mp3s we're talking about.

In an attempt to secure an emotional performance, engineers police their own pursuit of perfection. They also work to make sure that the performer does not fall into the trap of being "overly perfect." This is especially the case with vocalists, who are considered particularly prone to dwelling on intonation at the expense of emotion. This is problem is often spoken of in terms of the singer being "stuck" in their head or throat. Engineer Sylvia Massy is well known for cultivating compelling vocal performances in her clients. In her book Recording Unhinged, she explicitly opposes the conscious pursuit of perfect vocal technique to the accomplishment of emotional take:

> "the nature of singing is to convey emotion. Emotion is not perfection. When conveying a story about heartbreak, I want to hear that vulnerability in the singer's voice. This doesn't happen when the vocalist is concentrating on pitch and timing." (Massy 2016, 66)

Here we see a negative definition of emotion: we might not know what emotion is, but we know that it is not the same thing as perfection. In order to keep perfection from eroding emotion, Massy makes use of a variety of techniques designed to distract and prime the vocalist:

> If a singer does too much warming up, they will not give you that raw emotion that you may want. Sometimes you want a performance that does not have control – a performance that has vulnerability, emotion, pain, and fear. Those first few takes of the day more often hold this type of feeling… you may want a performance with anger or bewilderment or deep sadness. I suggest you prepare your singer for that type of performance by forcing them to do something they don't like, or reminding them of a sad story. (Massy 2016, 59)

Massy also recommends time-tested techniques such as forcing the singer to stand on a chair during a vocal take[18] and coaching them in their bodily posture.[19]

> Many singers are self-conscious about their voice not working properly, concentrating too hard on the feeling in their throat. And who can blame them? It's not like they can change the strings and retune, or try a different instrument – they are attached to their instrument… I find the best treatment is to get them out of their head… I give them so many other jobs to do during the vocal recording that they are completely distracted from thinking about their throat! (Massy 2016, 59)

Sarafin, meanwhile, recommends the technique of compiling a single vocal track from multiple takes (a technique known as "comping"), being sure to exclude the singer from this editing process:

> In my opinion, it's better for the singer to let someone else comp the vocal and then listen fresh to the comp as a whole performance. This prevents the singer or artist from becoming oversaturated, and consequently unable to adequately judge the comp as a whole. Aside from that, singers tend to get very preoccupied with pitch over performance, and good producers are interested in performance over pitch. (Sarafin 2002, 169)

Massy and Sarafin carry on the tradition of Max Hempel in the ways that they wrangle their performers. Where Hempel was concerned about the distance of the vocalist's body from the recording apparatus, however, engineers are now directing their handling strategies towards the more nebulous goal of getting a good, i.e. emotional, performance. The domain of emotional experience has become an area of involution (Hustak & Myers 2013) for engineers, clients, and equipment – a question of how best to *involve* performers and the studio apparatus with one another. Involvement means that the investment of feeling in equipment and technique has

---

[18] She quotes Shelly Yakus' endorsement of this strategy: "He's not thinking about the song anymore, is he? He's thinking about trying not to fall of that chair!" (Massy 2016, 61)

[19] Massy describes techniques she learned from her mother, who was a professional opera singer: "Stand like you are getting ready to take a punch in the abdomen, like a fighter's stance. One foot slightly in front of the other. Rotate your pelvis and clench your sphincter!" (Massy 2016, 64)

deepened to a point where it is not enough to talk about the proper application of something as separate from the feeling it evokes. In conversation with engineer Tony Maserati, Pensado observes a shift from an “application phase” to an “emotion phase” in the way that people relate to studio equipment and the sounds it affords:

Pensado:	When someone says ‘make it warm’ what does warm mean? It’s a word that’s lost its meaning to us on this side of the word warm. When somebody says the word warm I think, ok um, I don’t know. It’s not dull, it’s not just hook up a bunch of analog crap. What does warm mean anymore, have we lost the meaning of that word?
Maserati:	You know I think you and I probably interpret that, um, quite well because we have navigated a couple generations of audio
Pensado:	Well warm was easier in ’89 than it is now. Just grab a lo-pass filter and yank it up around 6k and print [laughs] and that was warm.
Maserati:	And now of course it means what kind of compression are we using? Down to, what is my attack setting? Because we are in our ears either trying to mimic tape or we’re trying to mimic an old microphone… I mean to our ear.
Pensado:	Now warm has moved out of the application phase to the emotion phase. It’s like what people think is warm is an emotion now rather than an EQ thing, which, I love that concept. (Pensado 2013b)

What Pensado is describing here is a shift in meaning over the course of his and Maserati’s careers as engineers. Meanings and perceptual examples of terms, such as the oft-invoked “warmth” of a sound, change over time. Where he once understood warmth in terms of a particular application – specifically the sound of a low-pass compressor rolling off the frequencies about 6khz - Pensado now finds that warmth is understood in a more “emotional” way that encompasses a broader set of connotations. Describing a sound in an emotional way is, to borrow Dunn’s term again, a syncretic activity. The emotional impact, or evocativeness of an object have become key ways of evaluating studio equipment. Gear reviews, especially those appearing in publications like Tape-Op or Sound on Sound, often articulate the particular emotional affordances of, for example, compressors that “center and stabilize a vocal, even in a dense mix, without all of the emotion of the delivery (as heard in the dynamics of the breathiness and glottal sounds, etc.) being lost." (Hong 2013) A compressor might, for example, be said to work well "if you really need to slam a sound but you want to retain some emotion." (Hong 2008) As Geoff Stanfield observes in his review of a Pulse Width Modulation Compressor:

> Writing about audio gear is an interesting exercise in trying to capture the emotional impact of an inanimate object. I think that it is less important how something works than how something makes you feel. When a piece of gear allows you to transfer a performance to a fixed format with emotion intact, it is a piece of gear worth owning. (Stanfield 2014)

Technical concerns of sound recording and signal processing are increasingly deeply intertwined with concerns about the experience of the client, their ability to express themselves effectively and have that expression come across in the final product. Paradoxically, recording engineers have gone from being able to dismiss the client's emotion as beyond their concern because of its subjective nature, to being primarily concerned with the production of the client's emotional experience as something objectively accountable and thus subject to commercial circulation. Recording engineers working in the late 20th to early 21st century can be characterized by their interest in emotion as an objective phenomenon. In terms of this project's broader relevance to STS work, it offers an example of how the boundary between subjective and objective experience is socially managed and contingently shaped through histories of the work that makes emotion its object. In the chapters that follow, I will consider how emotion and voice are articulated within one domain of studio technology in particular, that of digital tuning.

CHAPTER 2: COMPETING PRESCRIPTIONS IN DIGITAL TUNING

This chapter, and the two chapters which follow it, are structured respectively around the themes of reinscription, repair, and repurposing. It begins by focusing on the context within which two of the most important digital tuning tools – Antares Auto-Tune and Celemony Melodyne – were developed, and the sorts of "prescriptions" with which they were brought to market. In characterizing how these technologies were initially conceived, it provides a basis for two subsequent chapters which take up how it was re-inscribed through use. Chapter 3 takes up the theme of "repair" by looking at how recording engineers took up digital tuning technologies in dialogue with their previously developed ways of working with the voice, and how they developed new skills around what they perceived to be the limitations of digital tuning. Chapter 4 examines the "repurposing" of digital tuning by groups outside of the software developer and engineering communities, specifically musicians, listeners, and songwriters.

Here we begin by looking at the mundane objects on the desks of Auto-Tune and Melodyne's respective inventors, using them as a way to open up questions of how their particular social worlds informed their technologies and framed their approach to digital tuning. It goes on to situate Auto-Tune and Melodyne's particular prescriptions for digital tuning within the social milieus of their developers. These inscriptions take the form of two "origin stories" or marketing narratives deployed by Antares and Celemony, each of which establish specific bases for the proper use and evaluation of their products as, respectively, corrective and compositional tools.

Evocative Objects

On a table behind Dr. Harold Anson Hildebrand's desk, between his MacBook Pro and a stack of external hard drives, sits a large wooden model yacht. It's a J class Ranger, with cutter-rigged sails. Fresh out of college he had worked on sophisticated inertial navigation systems for the U.S. Navy, but at 70 he is now more interested in sailing. He and his wife are considering a move to Puerto Rico and are in fact part of an online network – The 20/22 Act Society – designed to inform and connect people who want to immigrate to take advantage of the commonwealth's recently passed tax incentives. On his bookshelf are awards from his previous work in petroleum geophysics, as well as a half dozen thick technical reference books and a couple of stacks of old rack-mount hardware modules for his company's flagship product, Auto-Tune. Four guitars are distributed around the room, some leaning on chairs or walls. An ergonomic kneeling office chair is tucked under the table.[20]

Almost six thousand miles away, Peter Neubäcker's desk is flanked by a collection of stones, lined up across a raised shelf so as to catch the light coming through the angled top-floor windows of his Munich home office. A nearby shelf, lined with green felt, holds a conch shell and selection of small Platonic solids in wood and glass. A well-used copy of Kepler's *Weltharmonik* (Kepler 1978) is ready to hand. Set apart from the other stones, just behind the computer monitor, is a tapered translucent stalagmite in pale green-gold plastic. It is the result of Neubäcker's year-long investigation of a plucked open A note on a trichordo bouzouki. The base of the sculpture indicates the note's initial attack, with the decay and release over time corresponding to the tapering point. It is a painstakingly crafted piece of frozen time. The

[20] For a picture of Hildebrand's office, see (Crockett 2015)

method used to translate the information from the bouzouki's string to the shape of the stalagmite is the same as that which Neubäcker's software, Melodyne, uses to translate recorded sounds into visual representations that can be reshaped in terms of time, pitch, and timbre.[21]

Sherry Turkle has used the concept of the "evocative object" to describe how the various useful and aesthetic things that furnish everyday life can serve as "companions to our emotional lives or as provocations to thought."(Turkle 2007, 3) Evocative objects underscore the "inseparability of thought and feeling in our relationship to things" and illustrate how we "think with the objects we love [and] love the objects we think with."(Turkle 2007, 5) The objects Hildebrand and Neubäcker keep close while they work provide an entrée into the local and socio-historically situated practices which brought two of the most widely used digital tuning technologies to market. There is something of Hildebrand's yacht and textbooks in an Auto-Tuned voice. There is something of the plastic stalagmite and Johannes Kepler in a Melodyned note. Auto-Tune and Melodyne have come into wide use in the production and circulation of the some of the most widely heard recorded voices on earth – those of modern pop music. Their respective constellations of algorithm, interface, material hardware, inscription, ideology, and practice constitute objects that are evocative in the sense that they are globally deployed in the production and circulation of singing voices. They are evocative in a way that gets at the word's etymological roots, being many engineers' preferred tools for calling-forth (*evocationem*) and bringing-out voices (*ex-vocare*), which, through wide-ranging and heavily capitalized networks of music production and consumption are meant to evoke feelings within listeners.

[21] For video of Neubäcker's office see (Schönherr 2011)

While the figure of the evocative object helps us think through the material bases and consequences of affective sociotechnical experience, it tends to frame this experience around the relationship between a particular person and a particular object. A rather different situation arises when we apply the method of de-scription, which Akrich defines as an "inventory and analysis of the mechanisms that allow the relation between a form and a meaning constituted by and constitutive of the technical object to come into being" (Akrich 1992, 210) While Turkle describes objects as evocative insofar as people have "spun a world" (Turkle 2007, 10) with them, Akrich highlights through the concept of inscription the dialectic between "*the world inscribed in the object* and *the world described by its displacement*." (Akrich 1992, 210, italics in original) Reading these aspects of the object together invites us to consider how the evocativeness of an object is produced through a process of social negotiation similar to that of in-scription. If artifacts acquire scripts through (in-, pre-, or de-)scription, they become evocative through processes of "voicing." If sociotechnical "scripts" put forth particular patterns of interaction with respect to a technological system, "voicings" evoke particular articulations of thought and feeling between users and artifacts. To account for studio equipment in terms of processes of in-voicing and re-voicing, alongside those of in-scription and re-inscription, would be to provide a thicker description of the affective relations between people and objects without being limited to either a textual or evolutionary metaphor.

When, as in the case of the recording studio, the intentional object of a sociotechnical system is "expression" (as opposed to, for example, diagnosis, duplication, education, or testing) the question of how one comes to speak with or through a technology becomes especially relevant. In-voicing comprises the work of marketers and salespeople, who craft and shepherd the personality and aura of a product, and users who cultivate their own affective stances and expressive possibilities with an object or system. Because this work is about how one feels about and with an object, it takes place in the touchy-feely yet ruthlessly pragmatic idioms of ad copy, packaging lingo, brand identity, and trade show innuendo. Looking at how Auto-Tune and Melodyne came to market and mutated across contexts of use, the line between catchy product narrative and nitty gritty production account becomes quite blurry. One of the most important stories of Auto-Tune and Melodyne thus concerns the ways in which these stories changed over time.

A comparison with an earlier, canonical case of music technology will help set the scene. In many ways, Neubäcker and Hildebrand are to digital tuning software what synthesizer pioneers Don Buchla and Bob Moog were to electronic music instruments.[22] The synthesizer's revolutionary role in the popular and academic musics of the mid 20th century is paralleled by the transformative effect of software like Auto-Tune and Melodyne beginning in the late 1990s. Like Moog, Hildebrand was a talented electrical engineer who would go one to develop both a new musical technology and a brand name that became synonymous with it. Buchla and Neubacker, by contrast, have comparatively countercultural backgrounds, with Buchla providing the music for Ken Kesey's 1960s acid tests and Neubäcker spending his formative years wandering the Black Forest and frequenting theosophical hippy communes across Europe. Neither Buchla nor Neubäcker studied engineering formally, opting instead to study music and astronomy (and, in

[22] For a thorough account of the Moog synthesizer and its relation to the Buchla synthesizer, see Analog Days. (Pinch & Trocco 2002) Pinch and Trocco show how Buchla and Moog's synthesizers emerged in a mutually constitutive relationship with their respective social contexts. Moog, for example, developed a keyboard for his synthesizer, in response to a "listening strategy" which identified and responded to the practices and interests of musicians for whom the piano keyboard was a common technological paradigm. Buchla, by contrast, eschewed the traditional keyboard as a way to encourage more experimental uses, engaging with the "new music" and psychedelic musical communities.

Neubäcker's case, astrology.) Melodyne and Auto-Tune were developed in parallel and largely independently of one another, reflecting their inventors' markedly different social contexts. Though their respective products have come to be thought of in roughly comparable and complementary terms by audio engineers, they brought divergent philosophical and practical approaches to digital tuning – a category of artifact and practice that retrospectively appears to us as a single technological problem.

Scripts and voices have styles. Thomas Hughes coined the phrase "technological style" (Thomas Parke Hughes 1983) to analyze how, at the level of nationally-directed infrastructure, technological problems are pursued in different ways by different countries. British and German electrical grids, for example, were developed in ways that emphasized local and centralized governmental authority, respectively. Wiebe Bijker has developed the concept of "technological frame" as a way of taking Hughes' key insight –that engineering is a creative process that must be accounted for as more than economic or scientifically determined – and applying it to the role of "current theories, goals, problem-solving strategies, and practices of use" (Bijker et al. 2012, 167) at the level of the social group. Hansen Hsu, more recently, has used the concepts of technological style and frame to examine the work of software developers. Style, in particular, takes on a wider range of meanings for computer programmers, where "coding style" – community-specific ways of writing and reading code - appears as an important actor's category. (Hsu 2015, 273) Coding style may make little or no difference to the computers that compile the programs, but it serves as a key way of articulating group identities among programmers. The ways that Auto-Tune and Melodyne differently articulate the practice of digital tuning cncompass these diverse concepts of style and frame.

Auto-Tune: "It's About Time"

Antares Audio Technologies is a small company located in Scotts Valley, California. When I visited them they were in a nondescript suite in a small tree-lined office park – 213 Technology Circle, just off of Highway 17 as it emerges from the winding mountain pass that separates Silicon Valley from Monterey Bay. The office was spare, with no need for a receptionist or front desk. When I wandered down the hallway looking for someone, VP of Marketing Marco Alpert greeted me cordially. I was a bundle of nerves. He offered me a beverage, walking me to the office kitchen. I accepted his offer and awkwardly poured myself half a plastic cup of room temperature ginger ale from a two liter bottle. We sat down at the office's conference table. The walls bore framed posters and publications from the company's history. Their brand is currently based around a blue alien-looking character. I asked Marco what the blue guy's name is. He does not have one, apparently. Earlier campaigns were based around retro sci-fi art, coupled with the slogan "Where the future is still what it used to be." One wall bears an advertisement for Kantos, an early vocal effects plugin with a sinister Giegeresque interface. On another wall hangs the newly designed Auto-Tune guitar, which has a built in string-specific pickup and processing system that allows the user to play in any tuning at the press of a button.

I learn that Harold Anson Hildebrand, whose title is Antares CEO and "Chief Scientist," drew upon a deep well of experience when he set about crafting Auto-Tune. Born in Coronado, California in 1946, he was recording as a session flautist by the age of 16. He studied systems science at Michigan State University from 1965-69, and worked for the following two years on inertial navigation systems for the Naval Ships Engineering Center in Washington, D.C. He went on to complete his PhD in electrical engineering, with a focus on random processes and estimation theory, at the University of Illinois at Champaign-Urbana in 1976 under the instruction of control theorist Abraham Haddad. His dissertation was on estimating density of weevil populations in alfalfa fields, using nonlinear filtering techniques. By his own account, he was able to put himself through graduate school using a form of what would come to be called the Black-Scholes financial investment model, making thousands of dollars using it on the Chicago Board of Exchange shortly before the actual Nobel Prize-winning paper was published in 1973.(Sjuggerud 2015) During his graduate research he worked for the University's Office of Long Range Planning and Analysis, "performing studies of student and staff flow in the University, and projecting budgetary requirements." (Hildebrand 1976, 119)

Upon graduation, Hildebrand took a job with Exxon's Production Research team in the middle of what was proving to be a particularly tumultuous period for the oil industry. The series of "oil shocks" beginning in the early 1970s saw major increases in oil prices, oil company profits, and investment in geophysical exploration research as companies sought oil reserves outside of OPEC control. As predictive models were taking over the domain of financial speculation, so was the haphazard "wildcatter" approach to oil speculation being overtaken by more highly rationalized model-based strategies. Hildebrand left Exxon in 1979 to start an independent geophysical consulting firm called Cyberan, which developed a microcomputer-based seismic analysis workstation for oil company Seiscom Delta United. By this point he was working specifically in the field of seismic geophysical analysis. Seismic geophysicists, Hildebrand explains, "listen to sounds on the surface of the Earth (or in the ocean), listen to reverberations that come up, and, from that information, try to figure out what the shape of the subsurface is. It's kind of like listening to a lightning bolt and trying to figure out what the shape of the clouds are."(Crockett 2015)

In 1982 he formed Landmark Graphics, which focused specifically on computational tools for seismic prediction and analysis. In 1983 OPEC began lowering its prices in order to slow the development of competing energy sources, leading to a price war and a precipitous decline in prices globally. Oil companies jettisoned their research initiatives and much of their exploration personnel, and the ranks of exploratory geophysicists contracted rapidly. Landmark, meanwhile, was one of a handful of companies that introduced new PC workstations for seismic analysis, which could be operated by a single person – inaugurating the so-called "era of the workstation" in exploratory geophysics (Lawyer, Bates, and Rice 2001, 200) and a boom period for Landmark. Landmark's database management system, called OpenWorks remains the industry standard in geological surveying. (Halliburton n.d.)

Business was so good that Hildebrand retired from exploration geophysics in 1989 and spent a year studying composition at Rice University's Shepherd School of music. Part of his coursework involved composing using computer-sampled instrument sounds – taking, for example, a collection of separately recorded violin notes and assembling a new piece. He found existing tools for sample manipulation restrictive, in large part because they were not able to loop samples seamlessly and thus required large amounts of computer memory to process. He designed an audio looping program of his own to solve the problem. In 1991 he started a company called Jupiter Systems in order to sell his looping software, which he called Infinity. This would be his first foray into commercial digital signal processing for music.

Hildebrand's career narrative is characterized by innovations that disrupt and reshape entrenched and perhaps stagnant markets. Just as the Landmark PC workstation capitalized on early 80s petro-politics and represented a new mode of oil exploration, so did Infinity provide a new mode of computer music production. He claims that Infinity "bankrupted the Los Angeles

Philharmonic"(Crockett 2015) by making it possible to produce film scores using looping sampled instruments. He similarly accounts for his Landmark work in terms of disruptive innovation:

> At the time the seismic contractor community was beginning to acquire three-dimensional data for geophysical interpretation. That data was extremely time consuming to interpret, so myself and three partners created this company to build a workstation that would do that very quickly. So we took interpretation down from a year and a half to a month. It allowed a lot more places to be researched before drilling, and it opened up how we acquire oil. Nowadays, we're able, because of that technology, to acquire oil in very small plays. At that time America was only providing 30 percent of their oil. The other 70 percent was being imported from OPEC. That technology changed how things worked and now we're supplying close to 70 percent of our own oil and only importing 30 percent. And I was involved with some other projects in geophysics, but it was a profound change to a limited resource, which is a geophysical interpreter's time. (Sjuggerud 2015)

Shortly after Infinity he developed a product called Spectral Shaping Tool (SST). SST was one of the earliest means of speaker "room correction," a now well-established technique for tuning acoustic spaces by analyzing the resonant qualities of the room and design compensating filters to cancel it out. This process effectively erases the room's acoustic qualities for the purposes of sound playback. SST can be understood, in part, as a technology of testing. It was meant to turn a room into a suitable test-site for the multiple acoustic situations within which a recording would later be played.

The Salesman's Wife

The oft-told story of Auto-Tune's inspirational moment finds Hildebrand in the midst of his work on SST and having lunch with an Antares salesman and the salesman's wife, who suggested he should make a box that could make her sing in tune. As Alpert recounted, Hildebrand "looked around and everybody was looking down at their plates, and were so embarrassed for her 'cause that was such a bad idea… they thought it was impossible, that what she was suggesting was so preposterous." As the story goes, Hildebrand's genius was to look beyond the preposterousness of such a box, and to see that the tools of geophysical analysis could be readily applied to the problem. In 1996, the year that Landmark was acquired by multinational oilfield corporation Halliburton,[23] Hildebrand debuted Auto-Tune at the annual meeting of North American Music Merchants (NAMM.)

The techniques of exploratory geophysics and vocal tract modeling had in fact been deeply intertwined well before Hildebrand brought them together in Auto-Tune. From as early as the mid 1960s researchers were using similar tools to analyze geophysical and speech time series data (Gray 2005). The technique of linear predictive coding, which interprets signals in terms of resonant spaces with particular points of constriction, originated partially in seismic strata analysis and later found its way into vocal coding algorithms that would later facilitate the earliest forms of the Voice Over Internet Protocol (VOIP). Stanford computer music expert Julius O. Smith explains the connection between vocal modeling and seismic analysis in this way:

> You drill a hole and you put dynamite in the hole and you set it off and you get all these reflections everywhere. And then you take the impulse response of the earth - your dynamite response - and then you fit a ladder filter[24] to it and so you find delays between layers and you're finding the reflection coefficients between the layers. And that allows you to predict what kind of rock it is and how hard it is. You know, the speed of sound

[23] At the time the New York Times reported that "The purchase offers the clearest sign yet of the direction that Halliburton's new chief executive, Dick Cheney, the former Defense Secretary, wants to take. In the oilfield services industry, companies like Halliburton want to rely increasingly on computerized brains instead of educated hunches and roughneck brawn. "The most exciting developments are occurring in the information technology field, and that's what Landmark specializes in," Mr. Cheney, who took over Jan. 1, said in an interview." (Myerson 1996)

[24] See (Gray 2005) for an insider account of the history of LPC, speech coding, and exploratory geophysics. For more on the social history of the ladder filter, developed by Robert Moog in the 1960s, see (Pinch & Trocco 2002).

> and the reflection coefficients and the impedance. So reflection coefficients are the function of the impedance change and the impedance change is the function of different kinds of rock and so they do underground modeling with this stuff. And that was first, as far as I know, that was the beginning of it. And then it made its way into linear prediction of speech and made its way up to processing. (Smith 2015)

The harmonic reflections of acoustic waves off of layers of underground rock, mapped as a function of time over a span of the earth's crust, in many ways resembles the harmonic patterns of a vocal formant frequencies as a function of time. Both geophysical and vocal acoustic data can be analyzed in terms of their autocorrelation, a way of measuring periodicity (or the prevailing length the period of one cycle of a signal) by taking the correlation of the signal with itself at a given time lag.[25] This method can help characterize the composition of geological strata (e.g., the presence of oil or natural gas) in a tract of land. It can also quickly and accurately estimate the average periodicity (or pitch) of a vocal signal over time. In a study of interpretative practices among petroleum reservoir geologists, Petter Almklov and Vidar Hepsø have described how geologists are trained to employ

> a type of visual intelligence whereby the geologist applies a set of templates that organize sets of marks into a body of significant signs. What was once mute stone becomes significant as the rock reveals its latent language to the trained eye. (Almklov & Hepsø 2011)

The suggestion that geologists' visual intelligence involves extracting language from "mute stone" takes on a new resonance when we consider how the paradigm of the reverberant geological stratum afforded Hildebrand a model for the vocal tract. Hildebrand's formal training and social context help explain how he came to approach and technologically reproduce the

[25] A signal will, for example, have a perfect correlation with itself at a time lag of zero. The correlation will then decrease as the time lag increases and the signal falls out of phase with the lagged version of itself. The correlation rises and peaks again, however, once the signal starts to repeat itself. The time lag values at which autocorrelation peaks occur can be used determine the prevailing periodicity, or wavelength, for even noisy signals.

voice as the product of a reverberant medium, the vocal tract, which was transmitting vibrations to a receiver – the listener – who was tuned to the particular context of the performance. He specialized in extracting hidden regularities from complex time-varying systems – e.g., alfalfa weevil populations, the earth, the market, and the vocal tract. Control and correction were key conceptual resources for Hildebrand. With SST, he had recently dealt with the problem of correcting the irregularities of acoustic spaces, identifying their particular resonant signatures and canceling them out so as to let the originally intended recording be heard without the distortions of place. Room correction tools typically get marketed as ways of ensuring that the mix will be able to travel well from one listening context to another. It is a tool for turning the studio into an acoustic "view from nowhere" (Haraway 1988). In the world of control systems theory, the fundamental randomness of the universe is only knowable and surmountable in terms of correction and control:

> Majoring in signal processing I majored in estimation and control theory. Control theory is what guides submarines under the water. Control theory is what guides space navigation and orienting your space craft at dock. Control theory governs controlling a rocket when it's been launched. There are control vanes inside the jet flow. Part of that control theory is these devices to save weight, money, and, you know, weight is the big deal in many cases. Have limitations. You get to a point where you're exercising your control surface and you can't go any farther. So a sub-class of control, and also gaming theory, is what happens? What should you do? (Sjuggerud 2015)

Hildebrand's formal training was situated at the center of what Peter Galison has referred to as the "Manichean Sciences" – or the areas of command, control, and prediction research spawned during World War II (Galison 1994). Galison argues that this historical formation, which comprised the fields of operations research, game theory, and cybernetics, constructed a distinct and influential "servomechanical" model of human and organic-mechanical behavior, based largely on a preoccupation with the need to anticipate and engage a technologically mediated enemy. This "Manichean" ontology involved the displacement of an "Augustinian" notion of evil-as-incompleteness with a model of an opponent with its own agenda, capable of resorting to deception if necessary in order to accomplish its goals. In Galison's analysis of the early cybernetics work of Norbert Wiener, he argues that this Manichean paradigm produced an opaque, black-boxed vision of the other, understood as a feedback system with its own teleological parameters. The telos driving the servomechanism, the control signal in its uncorrected form, was exactly the signal that Hildebrand had learned to dig out of the noise.

For Wiener, it was the autocorrelation function of an enemy pilot's evasive path (the correlation of the signal with itself over time) that allowed one to track him down with an anti-aircraft predictor. (Galison 1994) Galison argues that the autocorrelation function in particular served, for Wiener, as a way of getting at the human element of the human-machine assemblage. For Hildebrand, autocorrelation was useful as a way of quickly identifying the periodic pitch regularities of a singing voice, and thus the best way to track and correct its pitch. Auto-Tune was designed to function as a corrective prosthesis for a poorly functioning feedback-based (or "servomechanical") vocal control system.[26] Though it should not be overstated, reading Hildebrand's use of autocorrelation as an outgrowth of the Manichean Sciences helps highlight the implied models of human action which make Auto-Tune workable. The idea behind Wiener's autocorrelation-based predictor was that the intended path of an enemy pilot could be extracted from the noise of in-air maneuvering. The prevailing flight path was composed as the intentional object of the anti-aircraft prediction system, such that it could be used to guide the trajectory of anti-aircraft shells. This invites an analogy between the singer and the anti-aircraft gunner, who

[26] For a more specific discussion of the heterogeneous intellectual and practical terrain of cybernetics, see (Kline 2009; Kline 2004; Kline 2015; Dunbar-Hester 2010)

are both following a path (a melody or in-air maneuver, respectively) and must correct for the noise in the system that makes pitch tracking and expression possible.

In addition to control and correction, Hildebrand had a conceptual commitment to standardization – conventions which facilitate communication and uniform analysis and intervention across time and space. Auto-Tune's software interface proudly claims that it is the "world standard" in pitch correction. For Hildebrand, randomness is overcome either by external control or by the feedback systems within a particular system (e.g., missiles, commodity prices, or singing voices.) These modes of containment are not natural but conventional and context-specific, meaning that standards are necessary for this sort of control to be possible. Hildebrand was well accustomed to uniformity as a practical accomplishment within a particular industrial setting. The scarce resource that Auto-Tune serves to extract was not taken to be anything like a perfect vocal or musical essence, but simply a usefully captured voice as a function of studio labor time. Because of this, Auto-Tune's object of simulation and corrective intervention was not initially the sound or music *per se*, but rather the thing that produced it – the resonant cavity of the throat. This was a model of a voice made commensurable with electrical signals[27] – the cords functioning as vibration source and the throat and mouth as a series of bandpass filters. It was the vocal tract as an algorithmically thinkable and producible object. This model came in part through the 20th century history of telecommunications engineering (Mills 2011; Mills 2012), through which techniques such as linear predictive coding were developed for the voice by Bell laboratories and other research centers in order to aid bandwidth compression (Sterne 2012) and

[27] For more on the history of "sound as signal" in early 20th century America, see (Thompson 2002).

permit the recognition and synthesis of speech.[28] Working within this tradition of telecommunications engineering and speech signal analysis, then, Auto-Tune analyzes the throat of the singer in roughly the same way as room correction algorithms analyze studios or other acoustic spaces. It tries to separate out the original vibration source from the resonances of the space (voice or room) based on a generic model of that space that can have certain measurable coefficients fitted to it on a case-by-case basis. Once it has a guess as to the dimensions and resonant characteristics of the space in question, it can compensate accordingly and reverse engineer the signal into its original source vibration using filtering algorithms. In this way, Auto-Tune aims to excavate the fundamental frequency of the vocal cords from its resonant context so that it can be brought into accordance with a standard frequency and then fed back into the original formant.

The software's marketing slogan – "It's About Time" – has to do both with its overarching concern with time efficiency, as well as its paradigmatic concern with vocal performances as time-series phenomena. Hildebrand casts the process of vocal takes in terms of the effective capture of emotion – fixing things like pitch errors after the fact – is inscribed right into Auto-Tune's patent and technical documentation. Here emotional expression is defined in terms of transmission, which varies as a function of the in-tune quality of the performance within the musical context. Following Akrich, we might identify the following excerpt from the user's manual as one of the most important pre-inscriptions for Auto-Tune:

---

[28] This model is known as the Kelly-Lochbaum source-filter vocal tract model (Kelly & Lochbaum 1973). It models the voice as a periodic impulse generator feeding into a series of tubes of varying diameter and volume. This model makes salient the constrictions of the vocal tract as the key parameters of a singing voice – the control vanes in the jet stream; For an ethnography of more recent computerized speech recognition research in Germany, see (Voskuhl 2004)

> “Voices or instruments are out of tune when their pitch is not sufficiently close to standard pitches expected by the listener, given the harmonic fabric and genre of the ensemble. When voices or instruments are out of tune, the emotional qualities of the performance are lost. Correcting intonation, that is, measuring the actual pitch of a note and changing the measured pitch to a standard, solves this problem and restores the performance.” (Antares Audio Technologies 2000a)

Standardization and correction, in this understanding of the voice, are put in service of emotional expression, and this expression is understood primarily in terms of transmission efficiency.

Recently, Hildebrand has taken pains to distinguish his work in geophysics from his work on Auto-Tune, pointing out that over a decade passed between his work in geophysics and his development of Auto-Tune. This relatively recent claim should be read against earlier statements where he explicitly linked the two competencies of seismic and vocal data processing, as in the following:

> “Seismic data processing involves the manipulation of acoustic data in relation to a linear time varying, unknown system (the Earth model) for the purpose of determining and clarifying the Influences involved to enhance geologic interpretation. Coincident (similar) technologies include correlation (statics determination), linear predictive coding (deconvolution), synthesis (forward modeling), formant analysis (spectral enhancement), and processing integrity to minimize artifacts. All of these technologies are shared amongst music and geophysical applications.” (Hildebrand 2009)

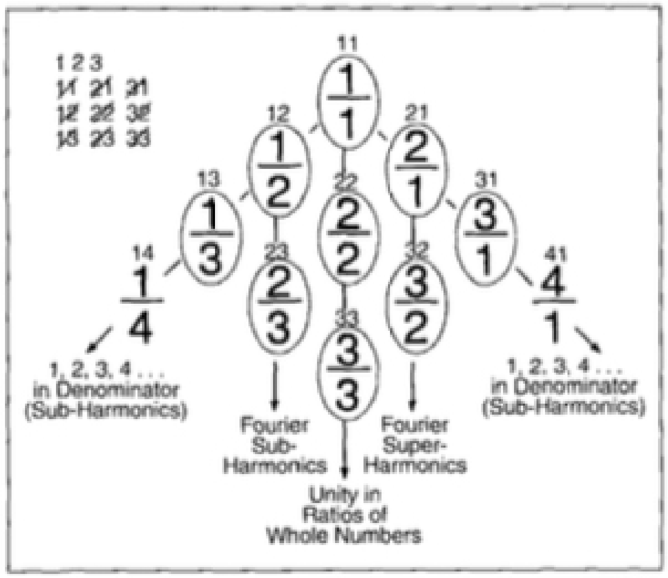
To say that Auto-Tune was simply an application of geophysical analysis techniques to singing voices would be an oversimplification. To say that his earlier work had nothing to do with how Hildebrand approached the problem of vocal tuning, however, would be to ignore his own comments and many of the rationales through which the program was marketed and taken up in the first place. Between these two inadequate accounts of the initial construction of Auto-Tune is an account of resonant styles that persist across Hildebrand’s career and the communities of practice that it strings together. The framing concepts of scientific rationality, standardization, correction, time-efficiency, and emotional expression which are articulated in the design and deployment of Auto-Tune, come into sharper relief when read alongside those of its competitor, Melodyne.

Melodyne and Peter Neubäcker's Harmonic Pythagoreanism

Melodyne, by contrast with Auto-Tune, was developed from a much more haptic, material, and object-oriented way of understanding and working with the voice. Its approach derives from a conceptual orientation rooted in Neo-Platonic and German Romantic thought, as well as astrological and lutherie practices. In 1994, Melodyne inventor Peter Neubäcker gave a talk in which he compared the study of harmonics to the fictional "Glass Bead Game" described in Hesse's novel of the same name (Neubäcker 1994a). Published in 1943 and set in the 25th century, the novel tells of a complex "game of games" played and understood only by an elite monastic class. The game consists of drawing dense and elegant connections between aspects of the shared eternal corpus of human knowledge and the universal relationships that undergird it. It would hardly be an exaggeration to say that this story captures a crucial aspect of Neubacker's general philosophy and deeply informed how he designed Celemony's flagship product, Melodyne.

Neubäcker was born in 1953 in the village of Buer in the northwest of Germany. His parents, Rudolf and Eva, owned a radio and television shop. He had two younger brothers. When he was 14 he and his family moved to the Black Forest, where he graduated from high school in 1973. Rather than go to university he "began to roam the Black Forest with only a sleeping bag and a book about edible herbs" (Schönherr 2011). After getting involved in the Freiburg hippy scene, where he played guitar in a group called "The Release Music Band," he hitched a ride on a VW bus that took him to a Christian-oriented spiritual community on lake Thun in Switzerland. During his three year stint at the commune, he came across a citation of

Viennese musicologist Rudolf Haase's "History of Harmonic Pythagoreanism." (Haase 1969) This work got Neubäcker interested in harmony, and he took up correspondence with Haase. In 1977, while hitchhiking back to the Black Forest to visit his mother, Neubäcker found himself at Lake Constance, where he took up an apprenticeship with a luthier. Building and maintaining stringed instruments provided a practical version of the theories he had encountered through Haase. Since then, by his own account, his philosophy has been strongly rooted in the Harmonic Pythagoreanism movement, founded in the 1920s by Hans Kayser, a student of Composer Arnold Schoenberg. (Schönherr 2011) Kayser drew primarily on the writings of Pythagoras, Kepler, and Leibniz and developed a sort of theory of everything based on the intervals of a monochord and a numerical chart known as the "lambdoma." (B. Hero & Foulkrod 1999) Taken from the writings of Nicomachus (100 BCE) and attributed to Pythagoras (ca. 500 BCE) the lambdoma is a matrix of ratios with one axis consisting of increasing integers and the other axis consisting of fractions with increasing denominators. Modern proponents of the lambdoma, the most vocal of these being artist and mathematician Barbara Hero, argue that it represents the mathematical relationships embedded in many natural phenomena, including but not limited to the harmonic relationships which structure musical perception.



1. The basic Lambdoma matrix, named after the Greek letter "lambda." The basic V shape is the first written indication of the Lambdoma, which was called a "Lambdoid" in a footnote in one of the books written by Nichomachus in 100 AD. This work was translated by Iamblicus a century later. Modern notation is added to the original ratios, with references to Fourier, supraharmonics (musical overtones), and subharmonics (musical undertones).

Figure 1: The Lambdoma Matrix (B Hero & Foulkrod 1999)

For example, a guitar string tuned to a "fundamental" frequency of A 440 Hz, will tend to vibrate not just at 440 cycles per second but also (with varying amplitudes) at integer multiples, or "overtones" of the fundamental frequency (880, 1320, 1760, etc.) This phenomenon is known as the harmonic series, and forms parts of the lambdoma's structure of ratios. Writing in a 1999 issue of the journal IEEE Engineering, Medicine, and Biology, Hero and her co-author Robert Foulkrod make specific claims about the effect on emotional experience by lambdoma-generated music produced by a BASIC program running on an IBM PC Jr. computer. They cite, for example, the testimony of a study participant:

> "Since listening to a Lambdoma harmonic matrix tape, I have had new feelings of bodily excitement and awareness. Sometimes the feelings are sort of part orgasmic, and sometimes a sense of mental excitement, which I haven't experienced since l was 20,

sometimes an over all happiness. I am sure these feelings are brought on by the tape, even though some of them come several days after use. Once, five hours after 1 used the tape, a prop plane went over the house and set up an astonishing resonance inside me." (Hero & Foulkrod 1999, 70)

The authors go on to hypothesize that the effects of lambdoma-based music, "because of its well-defined mathematical construct, are that of emotional and physical harmony (cenesthesia.)"(B Hero & Foulkrod 1999, 72) Around the same time as Hero's research into Lambdoma music, Neubäcker built a physical version of the lambdoma, across which he installed moveable monochord bridge that could dial in any of the given ratios it represented. In a promotional video for Melodyne he demonstrates it and goes on to explain his own understanding of the relationship between emotion (or, as it is translated in the video, "feeling") and the mathematical structure of music.

> "Music actually consists of numbers. Of numerical relationships. For me that is what constitutes music in its essence. When you sing a melody an interpretation is added to it, perhaps as the expression of a feeling. The actual musical element, however, is already there before the feeling. And I can perhaps make use of an especially well-suited melody to express a particular feeling. That which is musical in the melody, however, is more comprehensive, and earlier in origin than the feeling for which I'm making use of it." (Schönherr 2011)

Earlier, in his essay on Harmonics and the Glass Bead Game, he had written more specifically about emotion. Responding to Leibniz' claim that music is the "hidden mathematical endeavor of a soul unconscious it is calculating," he explained that because such a claim

> "sounds somewhat prosaic to our ears; those with a rather more emotional relationship with music who are told that music is in fact the subconscious calculation of the soul are unlikely to concur at all readily. What is meant, however, is that music and mathematics have roots in the same soil and that mathematics for this reason offers the same aesthetic and even religious edification as music - which, of course, is not something any true mathematician needs to be told."(Neubäcker 1994a)

Emotion, in this Leibnizian view, takes on a decidedly different relationship to music than it does in the Auto-Tune inscription. Thinking with the monochord and the lambdoma, Neubäcker and his colleagues Hero and Foulkrod suggest that music, and its emotional possibilities, somehow pre-exist expression in the form of mathematical relationships. A "well-suited" melody may be used to express a feeling, but the melody pre-exists its expression. The feeling does not appear as a message in the mind of the performer, but as something imminent to the mathematical ratios that can be selected and expressed.

In 1979, Neubäcker was put on probation for dodging national civilian service. Shortly thereafter he moved to Munich to work as a luthier. There he studied astrology, homeopathic medicine, and built an alchemical laboratory in his basement. His second wife, and Celemony co-owner, Hildegard Sourgens, was a physician who shared Hildebrand's interest in Renaissance and baroque music. (Sourgens 2007) While in Munich Hildebrand opened a publishing house for astrological texts, and wrote regularly for the astrology magazine *Meridian*. He worked with a community music school called the Freien Musikzentrum in Munich starting in 1984, where he presented a series of lectures on harmony and eventually founded a forum on experimental music and computers. It was an interest in creating visualizations for astrological calendars that eventually led him to begin working with computers. While Hildebrand was developing software for visualizing geophysical data, Neubäcker was using an Atari computer to visualize and sonify astrological data.

> "Around there somewhere: 84, 85. I had a publishing house at the time for astrological tools, offering things like calendars. Until then I had always prepared the print material for the calendars by hand but realized then it would be far better to program a computer to do it. You could show relationships on the screen such as the course of the planets or fractals. And the Atari also had a midi interface, with which I could translate movements

> into notes, and output them in musical form. That was also very creative." (Schönherr 2011)

In all this time Neubäcker had virtually no formal training in either math, physics, or computer programming, considering himself primarily as an astrologist and luthier. Celemony, a portmanteau of "Celestial Harmony" illustrates these dual commitments for Neubacker.

The Sound of a Stone

Melodyne's approach to pitch-shifting, which Neubäcker refers to as "local sound synthesis," analyzes and reproduces the voice in a way very different from that of Auto-Tune. Melodyne approaches sound spatially, as an object, as opposed to Auto-Tune's time-based predictive approach. Instead of reverse engineering sound in terms of a source-filter resonant model, Melodyne samples it wrapping it around a mathematically modeled cylinder so as to prepare the parameters of time and phase position for separate intervention. This unusual approach came to Neubacker, according to an oft-recounted story, when he was drinking ouzo and pondering the question of "what does a stone sound like?"(Schönherr 2011) Neubäcker arrived at the cylinder-based sampling method, in part, because it provided a way to engage a sound as a three dimensional object. Whereas Hildebrand made his software's unit of analysis the voice of a singer, Neubäcker focused on each note as an object in itself as something that could be frozen (or, perhaps melted) in time. Auto-Tune tries to guess what the singer is going for, where the melody is supposed to go, and nudges it into place. Melodyne is tuned into each individual note as a discrete thing, not defined so much by its position in an intentional melodic movement (its expressive *telos*) but rather its own note-ness. This object-orientation was for Neubäcker a prior aesthetic and philosophical commitment based on his engagement with what a Keplerian cosmology of the music of the spheres, which was itself engaged in a Pythagorean paradigm based on the model of the monochord – a single-stringed musical-scientific instrument.

Neubäcker originally sketched out this cylinder-based sampling approach by marking sampling points around a cardboard tube (Thielemann 2010). By “coiling up the sound” in this way, he was able to produce something resembling a literal sound object - an icicle-shaped three dimensional musical note molded in orange plastic, suitable for use as a paperweight. In an interview he shows off the note:

> “The question originally raised by the stone was how can I translate a given sound into a three dimensional form? Here I’ve arranged the individual sampling values of the sound indicated here by 1, 2, 3 and so on, in a spiral. And it turns out that if you interpolate between the points, a landscape emerges that also represents the individual cross-sections in the sound.” (Schönherr 2011)

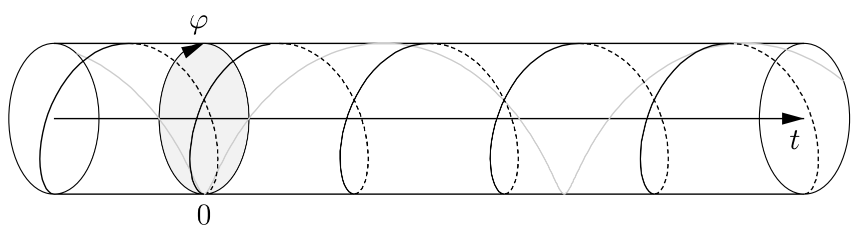


Figure 2: *The cylinder we map the input signal onto (black and dashed helix) and where we sample the output signal from (grey).*

Figure 2: The Cylinder Sampling Model (Thielemann 2010, 2)

Where Hildebrand concerned himself with the conventional production and use of standards, Neubäcker instead worked with the concept of naturally “just” harmonic relations. In writing about just intonation, which he regards as a more mathematically pure mode of tuning, he defines his terms:

> “Is a tuning system simply a cultural convention or is it something prescribed by Nature?

> I would like perhaps to anticipate at this stage by putting it this way: In the area where the term 'correct' is used, we are dealing with a system that stands in a context of culture and convention - this has a subjective character. In the area where the term 'just' is used, on the other hand, we are dealing with a system that stands in a context of auditory physiology and the physical and mathematical fundamentals of music - something that is given by Nature, in other words, and has an objective character."(Neubäcker 1994, 2)

The difference between Auto-Tune and Melodyne's approaches came into sharper relief as Melodyne debuted its "dynamic note access" (DNA) feature, which allows the user to track and change the individual notes of a previously recorded chord. His characteristically gnomic account of DNA's development included the explanation, widely quoted in recording engineering trade publications, that "For theoretical reasons I thought it would be impossible, but the more I pondered the subject, the more I began to see that what doesn't work in theory can still work in reality." (Assar 2009) Auto-correlation provides a quick way to identify the fundamental periodicity of a signal. It assumes that it is being given something monophonic, or originating from a single sound source such as a voice or flute. It approaches the sound as a path pursued by a single vibrational source. Melodyne, by contrast, does not take into consideration the source of the sound but rather tries to identify individual note-objects.

Conclusion

Emily Thompson has examined the modern historical shift from sound as a spatial phenomenon – understood originally in terms of architectural acoustics - to "sound as signal" which made sound commensurable with electrical current, and thus thinkable and controllable in the absence of any particular reverberant space (Thompson 2002). In his essay on artificial reverberation technologies, Jonathan Sterne has built upon Thompson's thesis by arguing that a sound without space is not a "schizophonic" (Schafer 1967) separation of sound from some original source, but the separation of sound from a part of itself.[29] This task, he argues, is possible only retrospectively, as an effect of particular engineering projects that aim to produce "detachable echoes." He argues that,

> "all artificial reverb is based on the founding construct of detachable echoes: that sounds have an essential dry interior structure separate from the spaces in which they emanate. But this dry, echoless condition can exist only if it is carefully manufactured. The dry interior condition of a sound "in itself" is thus an aftereffect of manipulation and not the sound's true essence. And yet, to engineer sound in this way it is useful to describe sound as if it could exist outside of space" (Sterne 2015, 118)

Neubäcker and Hildebrand cultivated distinct worlds out of various tools, exemplars, and heuristics that they had learned from their respective backgrounds. Hildebrand lived in a post-cybernetic world of signals, which needed to be filtered from noise. His object was the time-signal, which he approached as a practical question: how do we get at it? Neubäcker, by contrast, was invested in the play of ratios as something aesthetically compelling in their own right. For him, the performance of music was almost incidental because it served as an occasion to approximate the perfect mathematical relationships, which he took to constitute the most basic forms of aesthetic beauty. Neubacker's aforementioned "stone" was constructed from the sound of a string being plucked. He points to the triangular shape of its cross-section as evidence of its "fifthness" or the predominance of the fifth in its overtone series.

Without any formal training in electrical engineering, let alone the distinctive cybernetic[30] American control systems variety which shaped Hildebrand's thought, Neubäcker

[29] Piekut and Stanyek offer the Deleuzian term "rhizophonia" to refer to this decentralized and proliferative potential of modern sound technologies. (Piekut & Stanyek 2010)
[30] (Aumann 2011) has shown that, while cybernetics did gain currency in West Germany, it was of a very different 'flavor' than the servomechanical command and control variety that predominated in post-war America. The differences Aumann describes are subtle and various,

had no such paradigmatic commitment to the vocal tract model, standardization, or correction in developing his pitch detection and manipulation techniques. His paradigm was a neo-Keplerian concept of the music of the spheres, and a desire to render sounds into objects that seemed to persist, stonelike, outside of time. He did not encounter music as a problem of coded transmission, but as "musica mundana" - a set of inherently meaningful mathematical relations that were only elaborated and ornamented by the musical codes of "musica humana." He insists in one interview that "Melodyne was never intended to be a correction tool. The idea behind it was more about the vision to break up the recording process, to meld frozen audio up again and to work with it freely." (KVR 2013)

Auto-Tune and Melodyne represent distinct technological styles of digital tuning. They also arose from distinct styles of research. Melodyne arose from a lengthy material engagement with a particular note, painstakingly translated into a three dimensional object, rather than a conceptualizing of a vocal signal as the product of an unknown time-varying resonant system consisting of randomized data with buried intentional regularities. The meanings that Auto-Tune and Melodyne had for the people that developed them would differ greatly from the meanings they would have for the thousands of people who actually put them to use.

---

but suffice to say that, whereas Hildebrand was clearly working in the servomechanistic control theory tradition developed in the "first wave" of cybernetics (Hayles 1999), it is not at all clear that cybernetic ideas had any deep bearing on Neubacker.

CHAPTER 3: TUNING AS COVERT CORRECTION: REPAIRING AUTO-TUNE

The previous chapter examined how the distinct technological frames of Antares and Celemony produced two different ways of approaching the problem of digital tuning. Once digital tuning tools came into wide use they were reinterpreted, repurposed, and repaired by a wide variety of social actors. The following two chapters are concerned with how these two products were re-articulated along two distinct pathways[31] of practice. The first of these paths – which will be the topic of this chapter - concerns the question of how to use digital tuning in a covert and corrective way. Here questions of transparency, naturalness, and humanness became central to the social construction of both Auto-Tune and Melodyne as tools for corrective tuning. The second path, which we will take up in the next chapter, concerns the use of digital tuning as an *overt effect*, wherein digital tuning's potential for producing new and highly artificial sounds, as well as the meaning of these sounds with respect to established sociocultural conventions, became salient.

---

[31] It is tempting to replace "pathways" with "trajectories." As Bijker and Law argue, however, one must be "cautious of explanations that talk of technical trajectories or paradigms. Even if we can identify a paradigm, this does not mean that we have thereby identified what it is that drives the way in which it is articulated. And even if we can observe a trajectory, we still have to ask why it moves in one direction rather than another."(Bijker & Law 2012, 8) The term "pathway" is meant to complicate the Newtonian metaphor of a frictionless social space by invoking both the concept of path dependence – or the historically situated and contingently emergent quality of technological innovation (Garud & Karnoe 2001) – and Tim Ingold's concept of "wayfinding" as "a skilled performance in which the traveller, whose powers of perception and action have been fine-tuned through previous experience, 'feels his [sic] way' towards his goal, continually adjusting his movements in response to an ongoing perceptual monitoring of his surroundings." (Ingold 2000, 220)

This chapter begins with a discussion of the use of pitch correction techniques prior to auto-tune and other digital pitch correction tools. It then shows how Auto-Tune was first received by recording engineers, and the ways that it complicated the way they recorded and worked with the voice. We then look at how Antares responded to the perceived limitations of their product, introducing new features inscriptions. We then look at how engineers developed ways of working with auto-tune in a way that they considered covert and corrective, as well as the way Melodyne reshaped itself into a more corrective mode in order to better agree with these techniques. Finally we look at how engineers come to be able to hear what they would describe as covert tuning.

Pitch Correction Before Auto-Tune

Auto-Tune's initial framing as a time-saving, emotion-preserving covert correction tool was well-tuned to the concerns of the late 1990s to early 2000s major label music industry. This period was defined largely in terms of popularization of two other technologies – Napster and Digidesign Pro-Tools. The former was the first widely used peer-to-peer filesharing platform, which made it possible for people to share music for free. The latter was the first widely used digital audio workstation (DAW), which would eventually overtake magnetic tape as the industry standard approach to sound recording. Auto-Tune's development was closely tied to that of Pro-Tools. Antares VP Marco Alpert had been involved in the founding of Digidesign, and when Pro-Tools first started supporting plugins (programs that could be integrated into Pro-Tools' signal chain, as opposed to running externally from it) Antares began developing their software within the plugin framework.

Napster, meanwhile, signaled the beginning of an ongoing crisis in the music industry regarding the value of musical recordings, their legitimate means of circulation, and the resources available for professional music production.[32] With the shift from magnetic tape and twenty dollar CDs to DAWs and freely circulating MP3s, there was a widespread perception that the major label music industry, professional recording studios, and music retailers were in trouble.[33] As record labels tightened their belts, there was an increasing focus on maximizing the efficiency of studio time. Auto-Tune and other digital tuning platforms co-emerged and found success largely in terms of newly pressing emphases on time-efficiency. A pre-lapsarian moment in late 90s-early 00s recording budgets is humorously captured in the semi-fictional memoir The Daily Adventures of Mixerman, penned anonymously by a mix engineer who later revealed himself to be Eric Sarafin. Sarafin narrates the downside of what was to shortly become an unheard of phenomenon, the unlimited recording budget:

> "I was thinking about the opportunity (or curse) of recording a band with an unlimited budget. It's a well-known principle that you always spend up to your budget. So if there is no budget, what do you spend? Will we start recording guitar chords one string at a time? Will we record the same album three times?" (Sarafin 2002, 193)

Most of the engineers I interviewed had come up later than Sarafin, and their stories were of scarce gigs with meager budgets and an overwhelming demand for speed. They frequently cite time efficiency as the most valuable skill in the studio. Engineer Harold notes that he was able to get an edge over other engineers because he could "fly a hook" in a hip hop track before the singer was even done laying down a verse. Engineer Barry explains that this was the major difference between his formal and on-the-job training:

[32] For more on file-sharing and the history of the mp3 format that helped make it possible, see (Sterne 2012).

[33] Some recording engineers point to 1999 as the beginning of the recording industry's decline, with the Sept. 11 2001 attacks really marking the tipping point. Seth, who was working at a studio in upstate New York at the time, recalls, "After 911 hit the phones just stopped ringing."

> “School’s very much focused on you know as much as you can put on sessions and record bands and you know put up orchestras in school you’re not really working towards anything it’s all very academic you don’t really. The stakes are low. And so coming in here it was about particularly speed, I think was a large thing. Skills that pertain to being faster, as much as it’s important to be good at something if you can’t do it fast, at music you can’t do it. So it’s really quick problem solving.”

In engineer Bob’s experience, a demand for quick turnaround and a pragmatic attitude often took precedence over more ideal artistic or technical considerations:

> “I’ve done recording, I’ve done mastering gigs, I’ve done live sound, live recording… where you have literally an hour to get it all done. And you have no time no resources, at the last moment they’re like the date’s changed but we still need it in a half hour… A lot of people are like “well that’s not me, so I didn’t do it” no it’s what they want – they want something high quality, quick and ready. Can you do that or not? Who cares how it happened, just do it. That’s how I operate.”

With the possibility of digitally correcting technically flawed performances, the emotional resources of younger vocal talents became increasingly available for capture and circulation. Underlying assumptions about pitch accuracy and emotional expression as existing within a sort of zero-sum temporal economy were already well-established prior to Auto-Tune’s arrival, such that Sarafin was able to make the pre-Auto-Tune observation that “singers tend to get very preoccupied with pitch over performance, and good producers are interested in performance over pitch.” (169) Prior to Auto-Tune, pitch correction was possible using varying tape speeds or digital harmonizers from companies such as Eventide and Publison, but this was difficult to accomplish and used sparingly. In 1975, for example, David Gilmour of Pink Floyd described slowing down the tape playback speed when tracking vocals for “The Machine Song” so as it make it easier to sing. "It was a line I just couldn't reach so we dropped the tape down half a semitone and then dropped the line in on the track." (Cooper 1975) As engineer Seth recounts, this way of tuning before auto-tune involved tedious handiwork and socio-technical forms of management.

S: And then you had the harmonizer. The [Eventide] H3000 that could do pitch shift… and that was real hard. You had to set it and just record that one word. But back then the singers, you made em do it more, they trained a little harder they tried a little – I don't know if they tried a little harder but you know. So you might only have to do one or two. But that could still take an hour to do a couple of them. You know? Not like today.

O: Couldn't ask for that sorta consistency today?

S: Well there's two things, and this is what I tell new artists that I'm working with that are singing and like, there's the using the tools as tools but not using them as a crutch. So there may be times when I'm competent but I'm like "that's a little sharp or flat, I'm gonna fix it". I'm not gonna ask you to do it again 'cause you were close enough that I can do it without it being noticeable and we don't have enough money to spend the time to do it again when I can do it right now in 30 seconds. It's a lot of times just balancing the budget, is what it is. When it's being done properly.

Epistemologies of impairment and prosthesis, as Mara Mills has shown, are deeply intertwined with the discourse and practice of audio engineering (Mills 2012; Mudry & Mills 2013). Logics of normative ability inform the ways of evaluating and intervening with recorded voices in-studio. It is frequently considered poor form to use tuning as a "crutch" because to do so implies an inability, a lack of creative agency, or a failure of expression. Concerns about efficiency, precisely because they are perceived as external constraints on the creative process, have come to be considered legitimate reasons to resort to tuning among engineers. Just as practical contexts are subject to this sort of management, so is the artifact itself. The next section attends to how the voice was negotiated at the level of the signal processing technology as an artifact and system.

The Initial Uptake of Auto-Tune as Covert Correction Tool

Auto-Tune was marketed directly at this growing interest in studio time-efficiency. It was not so much meant to produce perfectly in-tune performances, as it was meant as a tool for time saving. As Auto-Tune's inventor Andy Hildebrand puts it:

> "[Auto-Tune's] largest effect in the community is it's changed the economics of sound studios... Before Auto-Tune, sound studios would spend a lot of time with singers, getting them on pitch and getting a good emotional performance. Now they just do the emotional performance, they don't worry about the pitch, the singer goes home, and they fix it in the mix." (Hildebrand 2009)

Part of the push for efficiency was the increasing use of younger, less expensive and less experienced singers. The association between an absence of technical experience and the potential for an emotional connection with an audience was forged in part through this logic of pop music production and artist promotion. Engineer Carl recalls:

> "[I noticed increased use of Auto-Tune] especially when I was working with people in the Disney crowd, it's younger people singing and they don't have as much of a trained voice. Their audience are young people and they hear these people singing and they think people sound like that, that people always hit the pitched center, it doesn't go sharp or flat at the end of the note. And I really think people, especially, like I said, in the Disney crowd when they go to see Miley Cyrus and they hear tuned vocals, that there's nothing wrong with that. That's how pop music is supposed to sound… For Disney stuff it's all tuned. The artist has less control over their sound because it's controlled by the machine (laughs). Disney has a whole corporation that does a bunch of things and it isn't so much about artist integrity as it is about getting the kid's face up on the screen and, you know, graduating them from being a Mouseketeer to being a tween hearthrob."

Engineer and Modern Rock Producer Howard Benson was an important early adopter of both Pro-Tools and Auto-Tune. He cites his background in electrical engineering as a major reason for his interest in computer-based recording and, later, Auto-Tune. In an interview with engineer Dave Pensado, he recounts his first encounter with the program:

> "What happened during the [1999] recording [of Less than Jake's album "Hello

> Rockview"] was something even more magical. I got a call from them [the manufacturer of Protools] and they said we're gonna send you this program from this company that makes stuff for the military, called Antares, and it's they do Fourier transforms - and I knew what that was from going to engineering school. And we can change the pitch on things. And I was like oh really? And they sent me a floppy drive and it's Auto-Tune 1.0. And I remember putting it into the computer and the singer Chris DeMakes was probably the first rock singer ever to sing through auto-tune because it didn't exist before then. And I remember putting it through there and I remember thinking oh my god, because at that time as you guys remember, we tuned vocals using Publisons [the Publison "Infernal Machine"], [Eventide] harmonizers, whatever we could do to change pitch. And it was a horrendously archaic way of working - well at the time it was the forefront of working… [On "Hello Rockview"] we even tuned the horns through Auto-Tune, which was a huge mistake, because they sound like square waves on that record. It was a different sound because I put it on everything. I was like hey this'll tune everything. I put the guitars through it. It didn't work on everything…"(Pensado 2014b)

For Benson, Auto-Tune and Pro-Tools together offered a way to get good recordings quickly out of bands that were not necessarily seasoned studio musicians. After Less Than Jake he used it on the San Diego-based rap metal band P.O.D's 2001 album *Satellite*:

> "That was the beginning of my career really taking off because the next year I did [San Diego-based rap metal band] P.O.D. and when they came to me, that was a record that was really done in post. A lot of the tuning and editing. That was a band with a huge amount of energy but not really any recording chops. And that was a record that would've taken a year or two to make. It took me two months to make, because we were in the computer."(Pensado 2014b)

Though Benson identifies the software in terms of its use of fourier transforms, in fact Auto-Tune's approach from the beginning was considered innovative largely in terms of how it replaced the more common Fourier approach to pitch tracking with a more efficient and accurate autocorrelation algorithm. It is also telling that Benson, from early on, thought of Antares as both "magical" and involved in military research. Though Antares has never (as far as I can tell) done any work for the military, Hildebrand had done inertial navigation research in affiliation with the US Navy early on in his career. Pensado's response to Benson crystalizes the unease that Auto-Tune and Pro-Tools produced among engineers in the late 90s and early 00s. Benson locates the emotional source of the recording in the discretion his larger team of engineers.

Pensado: Sometimes I miss the humanness of records. You seem to find a balance between keeping the humanness but massaging some of the things that got a little un-human, let's say. How does that work?

Benson: I think it's 'cause I take a more aerial view of the project at times I don't - I'm not and I think this has rubbed a few artists the wrong way… I let other guys on my team bring their humanness to the project. And I believe in their emotional connection to the work. So as opposed to me being a one man band where I'm micromanaging every single thing I'm not doing that.(Pensado 2014b)

Benson's early and influential advocacy for computer-based and correction-heavy production techniques was not about displacing the human element, but rather delegating and multiplying it across a broader team of engineers. He found in young artists such as POD and Less Than Jake a store of charisma and emotional energy that, while it lacked "recording chops" or the ability to perform in a reliably technically adequate way (i.e., in time and in tune) could be easily polished up and shipped out with the help of new digital techniques. Mastering engineer Heba Kadry, by contrast, associated pitch correction with a broader shift towards plugins, a move that she sees as having compromised emotional expression in music production:

> "I remember reading an article about the origins of music and how engaging in it was an indication of a more advanced society; it has allowed us to compartmentalize our emotions and therefore the ability to relate to others. The need to capture and preserve sound is sort of an extension of that… I have succinct memories of important turning points and phases in my life based on the records I was listening to at the time, like all the way down to triggering specific smells and emotions. … We don't use plugins, so we use the digital medium more like tape, and we run it though an analog console. What it's done, is required musicians to step it up. We want to work with better musicians. So the quality then starts with the musician and not letting them get away with things that we could just pitch correct, because for me that's when all the emotion and the passion – all those things that we were really in tune with 20 years ago that kept getting dampened by things like automation and plugins – are lost. I have to say that even working with protools when I did, it was never faster than tape. It just allowed musicians to be perfectionists and try to achieve something that they thought in their minds helped their ego. A record doesn't sell better just because it has been pitch corrected.(Kadry 2015)

The new economy of tuned voices has its share of discontents. Bollywood, which had for decades relied upon a division of visual and auditory labor between its actors and a small number of equally famous playback singers, has faced disruptions as star actors like Salman Kahn use Auto-Tune to make their singing, leading to objections from well-known playback singers (Saxena 2013; Tiwari 2016; Dasgupta 2015). Even as it was exploding in popularity, Auto-Tune came under criticism for doing exactly the opposite of what it was meant to: being overtly noticeable and removing humanity and emotion from the voice. Singers have come out against Auto-Tune, framing it as corrosive to musical integrity (Dovey 2009). Pop star Willow Smith expressed a characteristic perspective when she complained that "putting a lot of Auto-Tune on it really covers up the essence of you and is like putting a cloud over real true emotion." (Scott 2015) These concerns with pitch correction's destabilizing effect on established distributions of vocal and emotional labor would become key sites of negotiation as additional versions of Auto-Tune and Melodyne were developed.

Humanizing Auto-Tune

We can partially trace the interplay of engineering knowledge, design, and the recording studio's economy of skill and affect in the evolution of the features Auto-Tune offered in an attempt to make its tuning sound more "human." The initial distribution of human and technological agency – what Latour has referred to as "delegation" (Latour 1992) - between the Auto-Tune servosystem for that of the performer, was only the beginning of a longer, more complex negotiation over how the quality of human-ness can best be written into Auto-Tune's algorithms. This negotiation, or one side of it at least, is visible in the changing features and accompanying documentation of Auto-Tune's various versions. It took a while for concerns about Auto-Tune's transparency and sonic artificiality to be reflected in the program's technical documentation: in the first three versions of AT, the problem of making the effect sound more "human" or "natural" goes entirely unmentioned in the user manual. By version 4, however, we begin to see the first of several attempts to make the effect better suited to negotiating the natural/artificial boundary routinely managed by the recording engineer. In this edition Antares included a "variation" parameter for their vibrato generator, a feature which:

> "Sets the amount of random variation that will be applied to the Rate and Amount parameters on a note to note basis. Useful for humanizing the vibrato by adding random "errors." (Antares Audio Technologies 2000b)

If users sought a more human-sounding Auto-Tune, in version 4 they received one wherein the human was imagined as a 'random vibrato error generator.' The error-based model of human-ness recalls what David Noble termed the "ideology of control" that underwrote the design and deployment of early numerically controlled (and servomechanistically stabilized) machine tools, wherein the human element is "construed purely in terms of error" (Noble 1978)[34]. The ideology of control resembles a sort of productive inversion of the axiom "to err is human," quietly implying in the process that "to be human is to err" and, by extension, "to be not human is to not err." Embedded within this ideology is a sort of applied "sociology of error" which delimits social action *per se* to the domain of the "mistake." David Bloor has pointed out the limitations of such an approach, namely that it fails to account for the fact that all systems of knowledge production (in this case, for example knowledge concerning the relative correctness

---

[34] Randomness has a long history of troubling and/or reproducing the distinction between human expression and mechanical entailment, particularly in technologies of music production. Don Buchla recounts that, when his synthesizers were used in the Columbia Electronic Music Center in the 1960s, composer Vladimir Ussachevsky would tape over the random voltage generator (dubbed the "source of uncertainty") in the undergraduate studios so that he did not have to evaluate any randomness in student work. (Buchla 2007)

of a musical note) are socially constructed in the first place.(Bloor 1991, 14) While it is perhaps unfair to evaluate a somewhat tongue-in-cheek software design choice as comprehensive sociological claim, the critique of sociologies of error highlights the “variation” parameter’s peculiar and partial enactment of what it means to “humanize” a musical performance. At any rate, the variability parameter’s humanizing effect remained sufficiently in doubt that version 5 introduced the more on-the-nose solution in the form of a “Humanize” knob. This feature was accompanied in the manual by the following, somewhat grudging, rationale:

> “Humanize: One of the criticisms occasionally (if, we feel, unfairly) leveled at pitch correction is that it results in performances that are “too perfect.” Frankly, it’s been our experience that for the skillful engineer, Auto-Tune has always offered the means to perform its magic in a way that makes it virtually impossible for the listener to hear that it has been used (without, of course, comparing the result to the original performance).” (Antares Audio Technologies 2000b)

The “Humanize” knob works by distinguishing between short and long notes, retuning the shorter notes more quickly than the long ones so as to preserve the original vibrato of the voice. Auto-Tune Evo took a step further by including a geometrically modeled human vocal tract parameter called “throat length,” which allows users to alter the distribution of frequency peaks (or “formant,”) as well as the pitch. Here Antares offered up the presence of a model vocal tract as an index of humanness in response to users’ demands. With the introduction of a feature called “Flex-Tune,” Auto-Tune 8 offered yet another way of dialing-in the “natural” sound of tuned vocals. It controls the threshold for how close to a given scale-specified note an actually-sung note needs to be before the program retunes it. With a moderate setting, it tunes notes that are close but not quite right, while leaving the more deviant notes untouched, “allowing you to let through whatever amount of pitch gesture you choose, resulting in more natural sounding vocals.” (Hildebrand 2015) The Flex-Tune model is based on the idea that the threshold between real and artificial sounding vocals is to be found at the boundary between small (and presumably involuntary) and larger (voluntary and expressive) deviations from the selected scale frequencies.

Death and Auto-Tune: The Role of Melodyne in Late-Period Michael Jackson

Another way of tracing the negotiation over the transparency or natural-ness of digital tuning is by looking at cases where it was implicated in a broader-reaching controversy. Public discussion of Michael Jackson's posthumously released recordings, and questions of authenticity which surrounded them, offers just such a case. Jackson has released eighteen new songs since his death in June 2009. Several of Jackson's family members suspected early on that some the vocal performances on the record *Michael* were not actually Jackson, but rather those of a young R&B singer named Jason Malachi. Leaked vocals-only versions of the tracks in questions began making their way around the internet – the *a capellas* from the song "Breaking News" faced particular scrutiny from Jackson's fans and family members alike.[35]

Skeptics pointed to specific sonic qualities of the recording: First, the signature Jacksonian nonverbal sounds (Owww! Unhh!) seemed repetitive and out of place, as though they were dropped in from other, more genuine recordings. Second, whereas Jackson would clap his hands and stomp his feet during vocal takes, in the acapellas for breaking news there was just silence between the lyrics. Third, the vibrato (slight oscillations in pitch over time) was rapid and

[35] For many, the lyrics alone were enough to spark conspiracy theories:
"Everybody wanting a piece of Michael Jackson
Reporters stalking the moves of Michael Jackson
Just when you thought he was done
He comes to give it again
They could print it around the world today
He want to write my obituary"
(Jackson et al. 2010)

prominent, similar to that of Malachi, whereas Jackson's vibrato was generally slower and less pronounced. Upon hearing this, Jackson's nephew TJ (son of Tito) wrote on twitter: "Sounds like Jason Malachi to me too. The vibrato is a dead give away that it's not my uncle."(Jackson 2010)

In response, *Michael* producer Teddy Riley claimed in an interview that the voice was, in fact, "the true Michael Jackson" and that if it sounded strange it was because they had to use pitch correction software to bring certain recordings up to scratch:

> "When they first had it they did the processing and they did Melodyning because there were a lot of flat notes, um, because Michael was going in to do the the *ideas*, you know. He would never consider it being a final vocal. But, because he's not with us he cannot give us new vocals, so what we did was we utilized the Melodyne to get them in key, and uh we didn't go with the uh Auto-Tuning, we went with the Melodyne 'cause we can actually move the stuff up. Which is the reason you know why some of the vibrato sounds a *little*, you know, um, off, or, uh processed, over processed. Um, we truly apologize for that happening but still in all you're still getting the true Michael Jackson."(Reuters 2010)

For the purposes of tracing the discursive construction of Melodyne as a corrective tuning tool and Auto-Tune's shift towards a stylized effect, we can bracket the "imposter" question and attend to Riley's apology at a performative level. He deploys an admission of artifice – the use of pitch correction - in order to secure a deeper claim towards authenticity – that this is really Jackson's voice, even though it sounds strange. Riley's working relationship with Jackson, mediated through vocal processing techniques, thus came into conflict with familial closeness as a criteria for knowledge about "what is going on inside of Michael" and being able to say whether a recording is a true expression of that interiority. Later in the interview Riley situated the decision to in terms of his personal relationship with Jackson, saying that following his passing, "he was hope for me and he still is you know 'cause he's given me hope, he's given me a second life in the music industry."(Reuters 2010) Though his intentions were to return the favor by giving Jackson a second life in the music industry through an act of digital repair, in doing so Riley came up against a set of strong associations between vocal tuning and death. Auto-Tune had been declared dead by Jay-Z a mere three weeks before Jackson's passing. Prior to the Jackson tuning controversy, Riley was himself a well-established connoisseur and trendsetter of technologically mediated voices in pop music. He is widely credited with creation of "New Jack Swing," a 1980s subgenre that made prominent use of vocoded voices and that Riley claims "gave R&B a new lifeline."(Williams 2012)

This context helps us understand why Riley was so careful to specify that he had used Melodyne instead of Auto-Tune, which had already become strongly associated with the robotic vocal aesthetic made famous in the late '90s work of Cher.[36] In this story the practice of digitally correcting vocal performances in particular appears as one of the several ways of managing the affective and perceptual boundary between life and death in the recording studio. Experiences and performances of deadness and liveness, as sound studies scholars have noted, have long been key material-discursive resources for recording engineers (Piekut & Stanyek 2010; Sterne 2003b). Rachel Prentice has written about the tension surgical students conducting dissections experience between experiencing cadavers both as assemblages of body parts and as former living people. Through "tactical objectification" and cultivation of affective relations with patients, surgeons must learn to "mutually articulate" their own bodies with those of their patient. (Prentice 2013) Indeed, surgical metaphors are commonplace in the recording studio. This is largely as an echo of earlier days of the profession that frequently involved the literal cutting and splicing of magnetic tape, though the surgical imagery discloses the decidedly somatic qualities

[36] A pre-echo of Riley and Jackson's concerns with the figurative afterlife, Cher famously used Auto-Tune to ask whether there was "life after love."

that engineers perceive within the recordings of people's voices. The tape has been replaced by the scrolling audio regions of the computer's protools display, though cutting up and carefully reassembling the digital inscriptions of embodied performances still makes up the a large part of an engineer's day to day work. Carl describes a colleague skilled in vocal editing a "surgical ninja."

Because the vocal track is widely considered the most emotionally expressive and thus most important part of a pop music recording, it generally occupies the loudest, most stereo-central place in the mix, and it receives the greatest amount of attention in the production process. Digital vocal tuning's popularity is accountable in large part because it allowed an unprecedented degree of control over the vocal performance and thus intensified this pre-existing conceptualization of the voice as a the quickest route to the singer's personality. A repeated sound, from either the body of an intern conducting a microphone check, or a pro-tools session set in a loop, is both more and less than its component elements. Spoken phrases or random sounds reveal inadvertent melodies, musical phrases break down into their component sounds – they become collections of production choices. Auto-Tune's inverse law of emotion and technical accuracy is not so much an accurate description of reality of how voices and emotions work as it is a collectively held heuristic - a particular rule for generating and interpreting vocal recordings in practice. Exceptions to this rule are common, however, and the negotiation of these exceptions is in large part how qualities of emotional expressiveness and in-tuneness get articulated in practice. Engineers take great care in tuning voices so as to preserve what they regard as the emotional qualities of the performance. The problem is that what counts as "in tune" is always changing from case to case – it needs to be negotiated and performed relationally between the engineer and the client. They need to be tuned-in to one another.

Making Tuning Inaudible: Developing Covert Tuning Techniques In-Studio

The growing field of engineers specializing in vocal tuning, meanwhile, had to develop specific sociotechnical skills that would allow them to walk – and, occasionally, draw – the fine line between technical correction and creative discretion. Seth developed his tuning chops working with Jazz and Blues musicians:

S: when I was assisting a lot of the old school engineers I got to work on a lot of cool blues and jazz and stuff, and the producers, I was telling them, you know I can tune that note and that note, just. They're smart guys obviously if they're still doing it, and I wouldn't have to ask they'd say tune that right there that one note. And then later on, oh tune that note! And that'd be it and I'd slowly build learning why are they tuning this one and not this one – well just listen to the song.
OM: Yeah, context tells you whether it's appropriate?
S: Yeah. You gotta let the artist be artful.

The injunction to "let the artist be artful," circular as it is, requires the tuner to decide whether any given moment of the performance should be treated as an instance of technical error or emotional expression. More complicated still, moments that even the performer might have in-the-moment regarded as a mistake can be accounted for later on as a fortuitous deviation that actually works better than the originally intended performance would have. The grid and automatic correction features of tuning software afford engineers a degree of what Daston and Galison have termed "mechanical objectivity" by allowing them to refer to a standard embodied in an established technological procedure (Daston & Galison 1992). Depending on the relationship between the client and the recording under consideration, however, recourse to this technique of "tuning by the numbers" can easily be taken as a slight against either the technical skill or creative decisions of the performer. Seth explains the importance of understanding whether a note is flat "on purpose."

> I think one of the reasons they liked me was ‘cause I didn’t just go by the numbers. Because I spent a lot of time training. And I was lucky enough when I moved out here I worked with a lot of old school artists, producers and engineers. And you know they as Protools came in I’d be like oh I can fix that one note. And sometimes they'd be like is it hard? I’d say no. Sometimes it was like “no no no that note should be flat. It adds to the song, it adds to the vibe. She sang it like that on purpose.” [laughs]

Tuning work requires diplomacy and sensitivity on the part of the engineer. Though engineers occasionally resort to a rhetoric of mechanical objectivity, qualities of technical incorrectness or artistic expressiveness do not ultimately inhere in anything like objective technical criteria or authentically in-the-moment intentions of the artist. In concert with the flexible materials of the artist, the technological system, and the “voice itself,” these qualities are assembled and reassembled collaboratively over time. In the absence of clear and universal tuning guidelines, effective tuners must listen to and manage their client’s expectations as well as their recorded voices. For Seth this involved a practice of prefacing, or articulating his role in advance of the actual work:

> I’m always huge at prefacing – I’m not doing it by the numbers. If there’s anything that has to be straight up just let me fix it real quick for free. And there’s always a couple. You know – this one I didn’t like, this one needs to be tuned better, and I’ll just snap it right to the number. Every once in a while. You’re not gonna be right perfectly in sync with the producer, with the artist all of the time.

It is especially when people fall out of sync that an economic rationale gets evoked in justifying tuning decisions. When I ask who ultimately gets to decide whether things get tuned, Seth explains:

> “Well first and foremost it depends on who’s paying. [laughs] And I always use, some people, even the private people who’re paying I’ll be like no you’re tuning too much. You’re sucking all your work out of it. I’ll be like pick some notes you can’t live without and I’ll fix those up, but the rest let me make ‘em sound real.”

In addition to these negotiations over what sounds "real" or "artful," engineers frequently cast tuning decisions in terms of its effect on the emotional qualities of the performance. Engineer Harold complicates Auto-Tune's framing of emotion as something that is communicated, successfully or otherwise, according to genre-appropriate accuracy of intonation. He articulates the relationship between emotion and tuning in terms of the artist's feelings of attachment to their own recorded voices within the song's emotional context. The problem is less one of effective transmission to the listener, and more one of preventing the feeling of alienation between the client and recorded representations of their embodied action. Eliciting and preserving this emotional relationship in the context of tuning takes place within a complex, contingent, and implicit economies of skill and trust, the negotiation of which requires careful staging, as Harold describes:

> "If you're still a little bit detached and want to use Auto-Tune because that's technically what you've been trained to do, or technically that's what makes it right, pitch-wise, that's not necessarily helping the emotion of the song. Maybe having that note out of tune really makes the song more emotional. Maybe the singer all of a sudden feels detached from his performance all of a sudden by having auto-tune on it."

Tuning-in to the needs and desires of the client is crucial for being able to discern what counts as "in tune" and, in so doing, making a compelling final product. As engineer Lucille articulates, it is sometimes necessary to balance one's own sensibility with that of the client in order to make one's work sufficiently "invisible" to let the music come through.

> "Yeah, [tuning] is something I do a lot. But I would say it completely depends on the artists. I did the last [Band Name] record and [Engineer] worked on it and [Carl] was mixing it. And um, like I tuned it like where I thought, man there's like this is like invisible tuning and like it still doesn't sound like great. It like sounds good but it doesn't like – it still has the kind of Nirvana singing out of tune sort of thing vibe to it. Like there's no artifacts you can't be like 'oh that was like I just heard the tune happen' or like I can hear that it was manipulated sort of thing."

It is rarely a question of being simply in or out of tune: being in-tune is always context-dependent and a matter of convention in the recording studio, and there are varieties of out-of-tuneness. Lucille draws a distinction between being “artistically” and “annoyingly” out of tune:

> “So I spent like 2 days on this one vocal…. like artistically out of tune, not annoyingly out of tune. But they came in on Monday were like ‘this sounds way too good’ and they were like pissed about it. And I was like bypass bypass bypass bypass, back to playlist back to playlist. And then I tried again the next weekend and made it even like more artistically out of tune. And that was like I’d put some things in like better tune, but I’d pull things farther out of tune to try to like make it sound more like a fluid thing. ‘Cause the guy wasn’t a very experienced singer… Um. And so it like there wasn’t a lot of artistry to the out of tuneness it was just like out of tune.”

This distinction between being artistically and annoyingly out-of-tune, in Lucille’s case, became accountable in terms of both the experience of the singer and the quality of self-consistency that his intonational deviations seemed to have. The young singer’s voice required tuning, not because it was ‘objectively’ out of tune, but because it did not seem to convincingly follow any signature stylistic pattern:

O: Did it ever occur to you to, was it ever an option not to tune at all?
L: No ‘cause it was way too far…
O: But you were pulling stuff out of tune sometimes…
L: Overall it was too far. Um, and it wasn’t like consistent in any way, like he’s sharp so it sounds like he’s still singing well. It was kind of like this [gestures] like he was too sharp and flat on different words.

Because they are required to rely on relative as opposed to absolute standards of pitch, tuners learn to identify distinctive patterns in the intonational deviations of particular singers and to adjust their tuning techniques accordingly. Negotiations can become even more complex when considerations of song structure, other engineers’ opinions, and the authority and perceptual skill of a disagreeable client come into play. Lucille continues:

> “It took me like three tries and then when we got to mixing [engineer] was like this is so out of tune. And he was like you can’t have the first lyric of the song out of tune. And I was like, sorry pal that’s staying in that way yeah [laughs] he was like are you serious? Like the first line? He was really bargaining with me and I was like, there’s nothing we can do about this that they won’t notice.”

What counts as “in tune” varies widely from genre to genre, band to band, singer to singer, moment to moment. Carl recalls working on a session from a hard rock band and realizing that, outside of the context of the rest of the music, the singer’s vocals sounded flat:

> “we soloed up [singer’s] vocals and were like, woah, those are *under*. Like the entire time. I’m like holy shit I had no idea. You know but once you double it and put it in with the music, it sounded fine to me, all grown up… you want to hear them strain their voice, and if they fall off a little bit you identify, you identify with the passion more so than the actual pitch.”

Based on these experiences of assembling and disassembling vocals, tuners develop schema for how to evaluate and intervene with different voices. These sometimes take the form of rules of thumb that often serve to reinscribe certain collectively-held vocal stereotypes: generally speaking, Carl says, “men go flat, women go sharp.” They also reflect the characteristic styles and habits of both the tuner and the vocalist. Just as approaches to tuning vary across clients and engineers, so it is with tuning software. Auto-Tune and Melodyne, along with other tools such as Waves Tune or ReVoice Pro, each bring their own characteristic sound and style to the vocal signal. One particularly common workflow, described here by Lucille, involves using both Auto-Tune and Melodyne together:

> “Actually on this last record that I did, I went through and tuned everything – the lead vocal – I went through and tuned it all to some artistic out of tuneness in Melodyne, and then I did the same thing in Auto-Tune and then I had one that was completely out of tune, and then I did a new comp out of the three of those because there were some parts where the Auto-Tune pulled it nice and there were some parts where Melodyne had like a smooth thing that was going on. It was nice and was just finally ended up being way

> more awesome just to like spend the time on it even though like if it's like technically taking – I guess artistically like… uncompromised I guess is the word I would use."

Carl elaborates on this approach:

> "Auto-Tune has a sound to it, even when it's used correctly, [for example] the main difference between Auto-Tune 5 and Auto-Tune Evo, they changed the sound a little bit. In fact, when I'm on Melodyne and I get done tuning and I say, okay I want to commit and go print this, the vocals sound a little glossier and they lose some maybe high-end EQ. Not sure why, probably because of the processing wanting to get rid of digital artifacts that may be noticeable. But yeah, the point I'm trying to get at is both those things have a sound… When I was tuning with [project] it would be Melodyne first and then Auto-Tune after it, so you could tune faster and Auto-Tune would kind of smooth over all the glitches that you missed with Melodyne. Actually using Auto-Tune in auto mode, tweaking parameters and automating it depending on how noticeable it was for certain phrases. And I've heard of other, especially Swedish producers doing the same thing, actually using both, just 'cause they have a certain sound to them."

Even within the relatively small community of professional vocal tuners, cross-cultural traditions are apparent to experienced engineers, such that there can be a distinctively "Swedish" way of tuning vocals.[37] As Auto-Tune continued to pass from the Antares drawing board and into the hands and ears of the recording engineer, a number of its key assumptions came under revision. Where it was supposed to correct pitch without producing audible evidence of its use, engineers found that it had a sound. Where it was meant to correct pitch and restore emotion, engineers found it put a performance's emotion at risk. Where it was meant to be automatic, standardizing and corrective, engineers found multiple ways of tuning and applied their own discretionary repertoire to its use.

[37] The existence of a Swedish style in this case might also have something to do with the disproportionate number of Nordic engineer-songwriters such as Arnthor Birgisson, Kristian Lundin, and Jorgen Elofsson, in the field of elite pop music production. See: (Karp 2014; Seabrook 2012)

Melodyne's Pitch Correction Macro – tuning in emotion and creativity

While the versioning of Auto-Tune involved adding features that allowed the user to dial-in at their own discretion parameters that were initially left by coarser controls to the correction algorithm, Melodyne's later versions incorporated features that allowed for automatic pitch correction. Interestingly, Melodyne's movement towards automatic correction was largely directed towards the same goals as Auto-Tune's movement towards fine-grain discretion. Melodyne's pitch correction "macro" slider, by providing a means of automatic correction, gave engineers a way to better articulate their own corrective work from the emotional work of the singer. This is shown in the following passage, which examines how recording engineer and vocal tuning specialist Carlo Libertini carries out what he refers to as "professional" tuning technique in Melodyne 4. In this example we can see how he specifically articulates two distinctions. The first is between the "emotion" and the technical accuracy of the vocal performance. The second is the distinction between "corrective" and "creative" tuning. After importing the vocal take into Melodyne and checking to see whether it has correctly identified the key of the song, Libertini begins attending to specific passages within the song. He highlights the first part of the song's opening verse and opens up the "pitch macro" – a new feature for Melodyne - which (very much in the style of Auto-Tune) allows the user to dial-in how much the selected notes are automatically aligned to the pitch grid. The macro window offers two parameters: "pitch center," which controls how close the average pitch of a note blob is to the precise note of the scale, and "pitch drift" which controls how much any given part of the blob is allowed to deviate from its center. He dials in pitch center to 100%, locking the center of each note to the pitch grid. With pitch drift, however, Libertini is more circumspect:

"because her pitch drifting is really good - her emotion is good - pitch drifting is the

> equivalent of staying in your lane while you're driving. Well it's the same with singing, she's not wavering too much. This may affect some of the emotion so maybe I'll tweak it up only about 20 percent."(Libertini 2016)

Here Libertini is articulating a connection between the emotion of the singer and the quality of her pitch drift. He has no problem bringing the overall melody into precise agreement with the scale at the level of tonal centers – this is merely a matter of technical accuracy on the part of the performance. The way that each note moves around its tonal center, however, is taken as an instance of emotional expression. Crucially, it is not an identical relationship between drift and emotion, but rather an assessment of the drift's quality as an index of emotion as opposed to lack of control. If – to use his metaphor – the singer were swerving in and out of the lane, the drift would become a source of trouble and thus something requiring a much stronger degree of correction. The engineer's decision is guided by a sense of whether the singer intended to use that amount of vibrato on the note. The affective qualities of the vocal performance are procedurally accomplished within the context of genre, distribution of labor, and the relationship between the engineer and the vocalist, through the specific and situated use of the pitch drift slider. The slider provides a way to tune-in what the engineer takes to be an appropriate relationship between accuracy and emotion, which the slider places into a relationship of exchange.

It is also significant that Libertini is using the automatic pitch macro in this case rather than the more "Melodyne" approach of adjusting the note's characteristics by hand. By using a tool that automatically nudges the tonal center and drift characteristics in place, the engineer is making use of the mechanical objectivity (Daston & Galison 1992) associated with the program in order to keep from imposing his own discretion on the emotional expression of the note. Had he gone in and edited the drift himself, Libertini would have been intervening in the very aspect of the performance that he has identified as its singer's emotional signature. The macro settings allow for all of the selected notes to be adjusted in a uniform manner, thereby preserving part of their relationship to one another even as they are all being transformed by the algorithm. Previous versions of Melodyne, which did not include the "macro" tool would have required the engineer to decide how much of each note is adjusted on a case-by-case basis, a time-intensive procedure which runs the risk of appearing to put too much of the engineer into the performance.

Shortly thereafter Libertini encounters a place where the singer's intentions need to be recovered at the level of pitch center and note distribution, instead of simply the drift of each note. He selects a passage where the first note appears to have an intended but unarticulated note hidden within it in the form of a 'plateau', which strains at the edges of the blob:

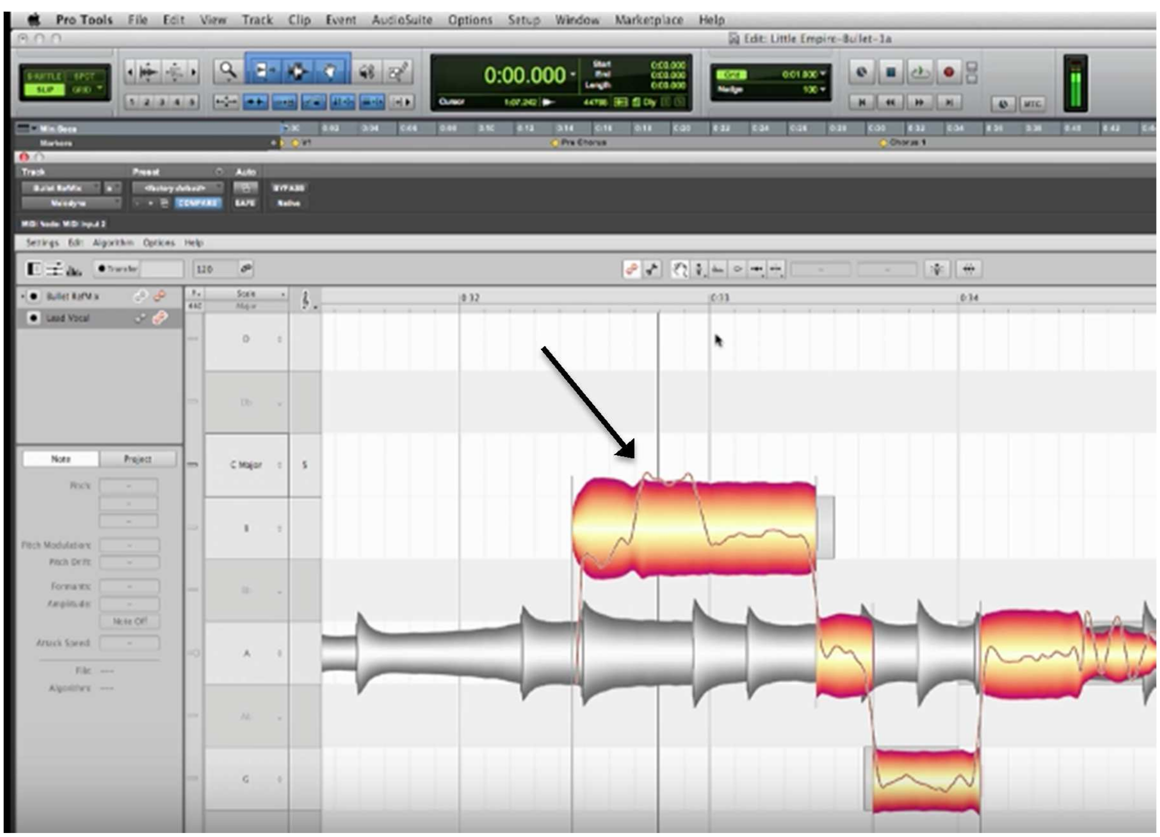


Figure 3: An intended but unarticulated note, as indicated by a 'plateau.' I have added an arrow so as to indicate the note's location. (Libertini 2016)

> "Now I can see that the artist had what I call an 'intended note' here but didn't quite articulate it… I'm just gonna scrub this audio and listen to what I mean [drags cursor through the blob] So when you see a pitch line extending like this, it means that the artist didn't quite articulate what they were intending. So now I'm gonna take my note separation tool and I'm gonna divide that note where the pitch lines change. And sure enough there's enough sonic energy to denote two notes: an E and a C…"(Libertini 2016)

By assigning a new boundary to the blob, creating three notes where there was only one before, Libertini has carved out a new note in the melody, which he can then correctively nudge into its proper tonal center. The note appears as an "intentional object" of the tuning process, something that the singer and the tuner accomplished collaboratively but in a temporally disjointed and technologically mediated fashion:

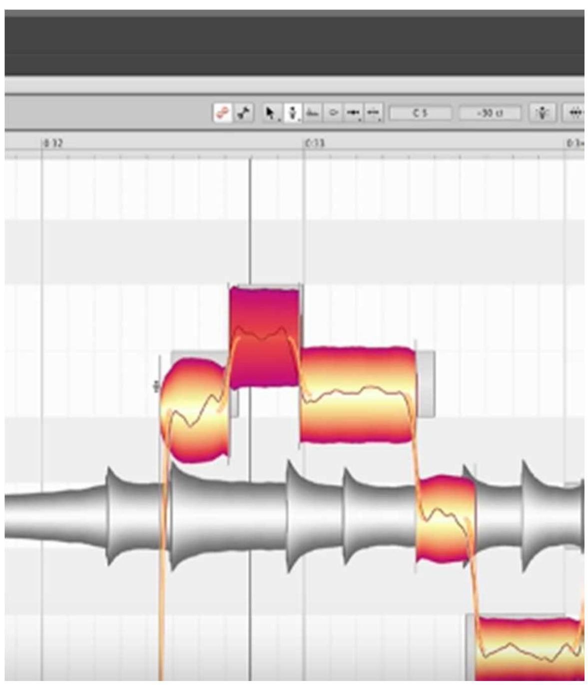


Figure 4: The intended note (dark red) is articulated as it is released from the blob. (Libertini 2016)

Later in the process Libertini encounters a different variety of intentional object in the vocalist's performance. In this case he marks a transition between a corrective and a creative intervention. Whereas the above correction involved bringing-out a note that was unable to escape the gravitational pull of its neighboring notes and was instead registered as an unusually large drift within a single note, the next intervention concerns a sustained note without any significant internal drift.

Libertini:	"now let's do something creative, something that the artist may have intended but didn't quite come out. Melodyne is also a wonderful tool to communicate ideas too, so keep that in mind.
Vocal:	"Don't open my ↑eye-"
Libertini:	It's that word "open" right here I'm hearing. Alright let me take my separation tool and separate it right there. Take the pitch tool and [moves new note up a half-step] Let's take a listen to this performance – now, I literally changed the performance, let's see if it works.
Vocal:	"Don't ↑*op*↓en my ↑ey↓e↓e↓e↑es"

(Libertini 2016)

Libertini listens back to the passage a couple of times, comparing the edited and unedited passage, "yeah that's kind of creative, I actually like that! I'm gonna keep that."(Libertini 2016) In absolute terms of transformation of the vocal signal, this change is not much more drastic than any of his previous interventions, but here the engineer notes that he has moved from the corrective to the creative mode by virtue of having put a new note that he "heard" as something that the singer might have intended but that does not appear to have been actually *attempted* in the source take. Libertini still frames his intervention as addressing an "intention" on the part of the vocalist, but also claims it as a creative act. He does this, in part, by framing it as a way to "communicate" an idea to the singer, situating it within an extended creative conversation with the client. If this is something of a boundary case whereby the engineer's corrective and creative mandate is performed, he moves on to more clearly creative tuning in the form of creating new harmonies during the chorus.

A distinction needs to be drawn, at this point, between what Libertini calls "creative" tuning and what I will, in the next chapter, describe as "overt" tuning. The creative tuning Libertini performs in the above passage is still done with the goal of sounding like it was accomplished by the original vocalist in a relatively naturalistic way. The singer may notice that Libertini has changed the note, by virtue of the fact that she has never sung it that way before, but a listener outside of the context of the production is not supposed to be able to tell that it was tuned. Covert tuning, by contrast, is tuned-as-such. If a voice has been overtly tuned, the effect should be relatively obvious without having to hear the original un-tuned version. Even the most creative instances of covert tuning, however, typically require a performance of skilled listening, or professional audition in order to be heard-as-tuned. Because the transparency of the tuning is a function of the skill required to tell that something is tuned, a key skill in repairing the artificiality of pitch correction tools is precisely the ability to "tell" tuning when you hear it. The next section examines how this skill is performed and adjudicated among engineers.

Learning to "Tell" What Inaudible Tuning Sounds Like

As part of learning to carry out covert tuning work, engineers develop ways of telling whether tuning has been carried out in other people's recordings. Online forums frequented by audio engineers feature lengthy threads with titles like "Auto Tune: Tell Tale Signs of its Use" and "How Can You Tell if a Voice is Auto-Tuned?" Among engineers, being able to "hear the tuning" is both a safeguard against sloppy tuning work and one of many informal tests of the sensitivity of one's ears and the extent of one's professional experience. One way to tell whether someone is a "real" engineer is by whether they can tell whether something has been tuned (and, by extension, how easy it is to hear artifacts in their tuning). Skilled ears are gained through practice. Seth, for example, has gone to great lengths to cultivate his ears, including listening to "golden ears" ear training CDs which present the listener with different frequency spectra in the same way one might learn vocabulary words in a foreign language. It is this personal investment in his listening ability that allows him in part to tell whether an engineer is truly "high end."

S: For some stuff, which some of the high end engineers they have better hearing than I do. So there's been, when I started working at [Studio] I'd go on, 'which one do you like better this or this or this or this?' because honestly 2 tenths of a dB I can't really hear

O: Yeah but some people can? Have you ever thought that maybe they can't but they want you to think that they can?

S: Oh no 'cause it's always me switching and they can tell the difference. When they know what they're doing and when they can really hear the difference they're not afraid to do a double blind, or even a blind [sic], they'll tell ya. They yeah. Now I've run into the other when I was assisting when I have engineers that don't really know what they're doing. They'd go and they wont hit the insert button so the piece of gear isn't in and you'll see them you know adjusting it for 15 minutes, and I'll be sitting there thinking should I tell him that he can't hear that? So I just let him keep adjusting the knobs.

The use of the word "tell" is significant here, as it helps capture the way that listening ability is tested and enacted relationally in-studio. The good engineers are the ones who "can tell," while the bad ones raise the dilemma of whether the assistant should "tell them they can't hear that." "Telling" talk also highlights the uncertainties inherent in the practice of hearing tuning-as-tuning. One is rarely well positioned to "know" definitively that something has been tuned, but one might still be able to "tell." The only real way to be sure that something has been tuned is to have done the tuning yourself. Even then, with the various takes, vocal comps, and plugins that get knotted together over the course of a project, the question of precisely which parts have been pitch-corrected, by how much, and for what reason, can become genuinely unclear to even the most fastidious engineer. Being able to tell whether or not something has been tuned by someone else is often the most that can be expected.

Beyond its use as a verb, "tell," meant in the sense familiar to poker players,[38] is also the object of a particular heuristic activity. Picking up on a "tell" is a way of knowing something which may be right in front of you but is nonetheless (perhaps intentionally) obscured as such. In its noun form, a "tell" is closely related to the term "shibboleth" as I develop it in the appendix. A shibboleth (an object or practice which marks a social boundary) can be thought of as a kind of "tell" that works by articulating and reinforcing boundaries. The terms are not identical, however. Being able to tell that someone has used tuning does not necessarily mean, for example, that an engineer thinks of the person using tuning as an outsider. Often quite the opposite is the case, and, as Lucille related, engineers may be tempted to tune something against their client's wishes in order to save face in front of their engineer colleague. On the other hand, having people be able to tell that you yourself have been tuned may well mean jeopardizing your status as a "real singer." Singer-songwriter Neko Case made such a gesture in a 2006 interview:

> When I hear auto tune on somebody's voice, I don't take them seriously. Or you hear somebody like Alicia Keys, who I know is pretty good, and you'll hear a little bit of auto tune and you're like, "You're too fucking good for that. Why would you let them do that to you? Don't you know what that means?" It's not an effect like people try to say, it's for people like Shania Twain who can't sing. Yet there they are, all over the radio, ****ing saccharine all over you. It's a horrible sound and it's like, "Shania, spend an extra hour in the studio and you'll hit the note and it'll sound fine. Just work on it, it's not like making a burger!"(Lapatine 2006)

Among audio engineers the question of whether you can "hear" auto-tune when it is being used covertly, oft-posed though it may be, is fundamentally paradoxical. From an

[38] There is considerable overlap between the skills of recording engineers and poker players. These include patience, an ability to read people, and a good poker face. Many engineers are also poker aficionados: legendary Chicago-based engineer Steve Albini is known for being an avid poker player. Grammy-winning engineer David Reitzas organizes an annual recording industry charity poker tournament held at The Village studios in LA.

engineer's perspective, any worthwhile use of pitch correction is going to entail some change in the way the recording sounds (otherwise what is the point?) In this sense, all tuning is audible in principle. In addition to the presumption that any instance of tuning is going to be technically "audible," it is generally accepted that it is often possible – even effortlessly so— to hear tuning-as-tuning on a track. More persistent questions concern 1) whether one could ever do a tuning job that would be completely undetectable as such, and 2) how one goes about hearing tuning-as-tuning even in cases that begin to approach non-detectability. Online forums provide an important site for engineers to discuss these and other foundational questions of their trade. The seemingly most common answer to the "hearing tuning" question involves a claim about the circular relationship between audibility of tuning work and the quality of the work itself. A Canadian engineer participating in the Steve Hoffman audio forums wrote in July of 2009, for example, "if you know what you're doing with good pitch correction software, you cannot "hear" it being used. If you can hear it, it either means that the artist wanted you to hear it (as a vocal effect), the engineer/editor did some shoddy work, or they used some crappy software."(Puddin 2009) The quotation marks around the first instance of "hear" in this passage indicate that the term is potentially problematic in the aforementioned technical sense. In an April 2009 thread on a separate audio engineering message board, one user acknowledged this apparent circularity of the "hearing tuning" question: "Truth is sometimes you can't hear it, or so I am told... I have no first hand experience... because I can't hear it --- catch 22."(D 2009) Also common are "know it when I hear it" claims, which root the ability to recognize tuning in one's own experience rather than any collectively held empirical criteria: "when I use it, I know when and where I am using it, so... I 'hear' it."… "All I can say is, I know when I 'hear' it... it has a certain characteristic, not unlike compressors, you are not suppose to hear them either, but you do or can I guess is more accurate."… "It is like anything else, if you know the gear you know the "sound" or be able to get close to what 'it' is anyway."(D 2009)

Others are so attuned to the sound of pitch correction that they admit to hearing it even when they know it is not there. As one engineer wrote, "I have tracked singers where I heard the auto tune artifacts, but [the singer] did it naturally somehow." To which another responded, "yup heard it too. I've also been sat at orchestral recitals and heard edits!! That's why you can get away with quite a lot - because much of it CAN be replicated live."(Narcoman 2009) A 2010 contributor to the Stack Exchange Sound Forum similarly reported that

> There are new singers, having grown up listening to pop singers that relied heavily on auto-tune, actually frighteningly imitate the "perfection" of the autotune glitches - that is freaky to hear! This was actually an issue with a singer in the Polish version of the X-Factor/So you think you can Sing shows. The "Cowell" positioned judge asked the contestant to sing it a capella. Holy crap if she didn't nail it! (Drordy 2015)

In a forum hosted by Pro-Tools manufacturer Avid, one user had presented an audio file in order for people to figure out whether it had been tuned. One San Diego-based user reported being surprised at hearing what sounded like a tuning artifact in one of their own un-tuned vocal performances:

> I definitely hear an AT artifact. But that does not mean it was used. It's very possible to get an AT effect with a uneffected [sic] vocal. I know this as it happened to me on my own vocal track. I heard it and said, how the hell did I do that?!? I actually really liked it for that spot in the song, so I fired up my AT plug-in to emphasize it, and AT did change it. It was as if it was already run through the plug!! So, AT may or may not have been used... (Wiesman 2003)

In puzzling over whether a particular voice had been tuned or not, one engineer described it as having "an extremely unique tonal character which borders on nasal, yet has a certain disarming richness to it...all of which can be somewhat mistaken for AT artifacts at times." Similar experiences of uncertainty regarding tuning take the form of the "accidental bypass" where one hears an effect without realizing that it is not actually incorporated into the signal path. In hearing whether or not something has been tuned, engineers draw on aspects of the vocal sound itself as well contextual and comparative resources. The first mode can be thought of as a way of listening *for* one or more tells or signs within the recording. The latter mode proceeds as a listening *against*, which works by reference to the circumstances of the recording and the experience of listening to other recordings in contrast with the one in question.

An important quality that is listened-for is an excessive quantization of the notes. This is basically a way of listening to the “retune” parameter of the pitch correction process, which controls the amount of time a note is allowed to spend away from a scale-specified pitch before it is nudged into place. This process may result in held notes that are overly smooth sounding and transitions between notes that are not smooth enough. As one Brooklyn-based engineer and forum contributor puts it, “There are two primary things to listen for. The first is the transition between notes. If this is too quick it sounds keyboard-ish. Also if a sustained note is too steady, it will also give a synthetic sound. In the extreme, it sounds like a vocoder.”(Osumosan 2009) This sound often gets described as an unnaturally rapid transition, or “warble” between two notes. The term “key-boardish” implies that the notes seem to arise from the action of discrete keys rather than a continuous change in pitch. Notes are said to “snap” or “lock” into place. One user refers to the sound as “digital stair step flutter,” while another uses the term “shelf” to describe the transition from note to note as choreographed by the pitch correction algorithm:

> When a singer sings a note, and slides to the next note, usually stepwise, you can hear a little "shelf" between the two notes. That shelf is where the first corrected pitch ends, and the second corrected pitch begins. Without auto-tune, the slide will simply begin on the first pitch and end on the second pitch. With auto-tune, you'll actually hear the note change from one perfect pitch to the next. A skillful engineer that's going for a truly transparent sound will know how to mask or "ease" these pitches with respect to the amount of correction so that the transitions sound more natural and less shelf-like.

(Lining 2015)

Some engineers point to characteristic timbre of tuned notes. Timbre is one of the more difficult-to-define aspects of musical sound (Fales 2002), indicating a quality distinct from the sound's pitch or intensity. Whereas pitch is described in terms of a metaphor of vertical position, timbre involves metaphors of color and texture. If a saxophone and a guitar each play a note of identical pitch, intensity, and duration, the difference between the two resulting notes are their respective timbres. As one New York-based engineer explains, it is possible to hear corrective tuning through timbre precisely because pitch correction algorithms assume that pitch and timbre are distinct when, in the context of actual performance, they are deeply intertwined:

> This is the thing that people who have played around and edited voices a lot know -- when a singer is "off," wavers around a notes pitch, or does vibrato that is not right, **the actual timbre of the voice changes.** So with Autotune, you have a note where the pitch is holding steady but the timbre is changing. This would normally not happen! Timbre changes occur hand-in-hand with pitch changes (vibrato). (Ahl 2009, emphasis in original)

This effect is particularly audible when the tuning program treats the note as a single object as opposed to a continuous movement. Pitch tracking algorithms are apt to treat a deliberate "slide" up into a target note as a case of inaccuracy, as one engineer puts it, the result is a slide that has been straightened out or flattened:

> Many times when a vocalist approaches a note, they begin the note below the target pitch and slide up to it. In Melodyne, this slide is included in the body of the note to be pitch corrected. If you flatten the slide to one constant pitch, you'll hear an artifact because the timbre of the singers voice will change but their pitch won't, if the slide was flattened. (Lining 2015)

Flattened slides are especially noticeable when the musical genre is one that makes characteristic use of sliding vocal technique, as in the case of vocal "twang" in country and western music. As one music journalist put it in a 2014 review of a live George Strait record that had apparently been tuned in post-production,

> George Strait sings at times with a cadence that Auto-Tune can't keep up with. Many times he purposefully sings by beginning a note out-of-tune to eventually bend it back into place in an attempt to squeeze the emotion out of a lyric. This is called "Twang," and is a critical part of the George Strait experience that someone in a studio decided to destroy because of some silly notion that George Strait's singing must be perfect. Performers who regularly sing with the aid of Auto-Tune like Rascal Flatts, they know what to do to make sure to not send the program into overdrive. When you add Auto-Tune on to a live performance after the fact, it almost always results in obvious artificial electronic sounds that erode the authenticity of the listening experience. (Coroneos 2014)

Whereas an engineer would be more apt to describe the sound of a flattened twang in terms of a change in timbre without its accompanying change in pitch, here the music critic uses the evocative image of bending a note as a way of "squeezing emotion out of a lyric." It is through the play of pitch and timbre together that twang enacts emotion for the listener.

The distinctive timbre of covert tuning may also appear to a sophisticated listener as a difference in instrumental "temperament." Some instruments, including pianos and guitars, make use of fixed pitches in the form of keys or frets. Instruments such as the violin or the voice, by contrast, have no built-in or fixed pitches and are able to change pitch continuously. Because they can change pitch continuously, these instruments are able to play in "just intonation" where the relationship between notes has to do with low whole number ratios. Fixed pitch instruments need to be tempered, otherwise a set of notes that will be in-tune with respect to a particular key will sound out of tune when played in a different key. A style of temperament is basically a decision as to which musical keys will be the most "in tune" and which will be the most out of tune. Since the early 20th century most fixed pitch instruments are tuned according to "equal temperament" wherein each musical key is equally out of tune with itself, which makes it possible to modulate from key to key without some sounding better or worse than others.

This (admittedly esoteric) bit of musical knowledge serves as a resource for articulating the differences between tuned and non-tuned vocals. As one engineer explains:

> The voice is a justly intonated instrument in that it contains natural harmonics along the natural order, which vary in pitch exponentially the further up the series you go when compared to the tempered scale that we use. Auto tune can't allocate for these "naturally" out of tune harmonics and when overused gives an 'un-natural quality' to the vocal aside from the obvious cher, kc and jo jo, r&b special effect use of it (Fallforward 2009)

Here the reference to the harmonic series, which is the structure of resonant frequencies or "overtones" which shape a note's timbre, serves as a way of distinguishing "natural" from "unnatural" timbres in a tuned voice. Another commenter extends this concept to include the effect of the backup instruments and even the resonances of the performance space on the way that a singer "tunes" her voice in a natural setting.

> One way you can tell, is simply that their vocal lines, every single pitch and inflection of their voice is perfectly in tune. And by that, I mean, perfectly in tune, and tuned to equal temperament. When a singer sings live, their tuning will take on the characteristics of their backup instruments, reverb of the hall, etc. These reflections allow overtones of their vocal harmonics and harmonics of the instruments to "ring" and allows the singer to slot their intonation in that overtone series. Some notes are flatter than equal temperament (i.e. flatted 7th in barbershop music is "just" intonation. If in equal temperament, or you're listening to auto-tuned barbershop music, those 7th chords will be too sharp). So when I hear a singer that's completely, perfectly, inhumanely in tune, I assume its auto-tuned or at least pitch corrected manually. (Lining 2015)

The point here is essentially that, because the voice offers so much flexibility in intonation, it will always bear the traces of what counted as in-tune at a particular place and time. One is never "in tune" full-stop, but is always in tune *with* something, or with respect to a set of circumstances.

Hearing tuning-as-tuning need not be so deeply rooted in technical minutiae, however. The sound of covert tuning is often described in more impressionistic language. Tuning is sometimes said to have a “shimmery,” “metallic,” “cheap,” or “grating” sound. It may be articulated as a particular “feel” or “signature” that one has to learn to hear. One commenter described it as “Phasy, smeary. not right sounding like vocals recorded in the old days i.e. Patsy Cline, Nat Cole, Frank Sinatra, McCartney, etc.” (Vernier 2009) Another describes it as sounding like an artificial larynx. Yet another describes hearing the sound of consonants and other non-periodic vocal sounds (e.g. “pickup” breaths, sibilance, or vocal-fry) being “crushed” (Lining 2015) by the correction algorithm.

In addition to listening for tuning in the sound of the recording, engineers also listen comparatively or against certain examples and considerations of what tells you a voice is tuned. In the same way the T-Pain has become an icon of overt tuning, certain artists, such as Maroon 5 or Britney Spears serve as exemplars of covert tuning. One commenter describes hearing “Cher, Tim McGraw, and Carrie Underwood all at once” when listening to a tuned voice. Inferences may also be made on the basis of certain overt stylistic indicators (As one engineer jokes, “most of the time auto tune can be heard if the song's lyrics contain "baby" a bunch of times.”) or judgments about the integrity and intentions of the people involved:

> the way that you hear tuning is usually when 1. the person doing the tuning is a non musician. 2. The person doing the tuning is lazy. 3. They are deliberately using the quantizing of the tuning and all of the other artifacts as an effect, which seems to be pretty common these days in what passes for R&B, or its due to either 1 or 2 (Gimenez 2009)

Inferences may also be made based on the year that the song was recorded:

> If it was recorded after the mid 1990s, you can almost bet 100% there are auto tuned vocals. It is such a standard procedure in recording studios that only extreme purists won't use it... (TLMusic 2009)

At their most extreme, era-based inferences of this sort might take into account the make and model of tuning software that was available at the time of the recording, each of which implies a particular workflow and style of tuning:

> All of this depends on the version of Auto-Tune (we like to call it Outta-Tune), as well. The most recent 2 updates allow for live tracking - requiring either a completely synced backing track or a live keyboardist who must adhere to the melody and/or to play in those sections where the choreographer or producer know there will be issue without it.(Drordy 2015)

Recordings made before Auto-Tune became available, or in a presumably non-tuned context may also serve as a basis for comparison:

> One easy way to tell is if you hear a recent recording by someone who's been around since before Autotune was invented and you're very familiar with how his or her voice used to sound. If the vocals on the new recording sound strangely nasal, robotic, or otherwise unnatural to you in a way that the simple passage of time couldn't account for, then it's Autotune. Once you hear it, you can spot it anywhere. A good example I can think of is Donna Summer's *Crayons*. Talk about someone who doesn't need Autotune, but boy, it sure is in that album. (Raf 2009)

Even the most recent and highly polished pop recordings made before digital tuning may, in retrospect, seem rough hewn to ears that have become tuned-in to auto-tune:

> I just listened to *2 become 1* by the Spice Girls (don't ask...) and the lack of autotune was extremely apparent. The vocals weren't exactly out of tune, but they were wobbling around and missing the mark in a natural way (the Spice Girls weren't the greatest singers), which just doesn't happen on 2010 pop music. (Earwicker 2010)

For some artists, the fact that their own voice had been tuned becomes only apparent in retrospect. Rock band The Killers' vocalist Brandon Flowers commented on this experience and its role in the decision to not use the effect on the group's sophomore album:

> I think [my voice on the new album] sounds better, yeah, but that could just be because I've been singing everyday for the past three years. I love it. We didn't use too many vocal effects. On the first album, we used auto-tune, and I didn't even realize what was going on with these machines and the computer. I was adamant about not using it this time. You really hear what my voice sounds like, for the first time. (Scaggs 2006)

Telling differences may also emerge between an artist's live and studio performances, even in sophisticated televised productions. One engineer noted seeing country western artist Brad Paisley on Jay Leno, "singing way off pitch. No auto-tune, but painful to hear."(Turnaround 2009) The context of a "live" performance is in itself no guarantee that tuning was not used, however, as realtime tuning hardware from companies like TC Helicon and Tascam can tune a singer's voice before it reaches the venue's public address system. Even if ears or performative context fail to tell the difference, some engineers will make use of spectrographic software that can make inaudible tuning visible. As one discussant in an audio engineering forum describes:

> In a frequency plot of a vocal track, you should be able to see that there is no data or little data in frequencies that aren't exact notes. In a waveform where frequency is plotted as if it were amplitude, the track would look like a series of lines at different heights; without autotune, there should be variable curves between lines of different heights (representing the singer shifting from note to note). With autotune, there will be straight vertical lines (no shifting, just jumping to the note) or "perfect" curves (deterministically-produced curves to decrease the unnaturalness of the jump, with no variability.) (Read 2010)

Claims about being able to "see" the tuning in the spectrographic representation locate artifacts of tuning in the inter-note transitions, which would take on visibly artificial straight lines or "perfect" curves. Recourse to the visual, in this case, carries an imprimatur of objectivity and scientific authority for engineers insofar as it seems to reveal the entirety of the selected audio in a static form that seems to put it outside of time. As studies of visualization in STS have shown, however, even the most scientifically "representational" images must be crafted in ways that incorporate aesthetic as well as epistemic judgments.(Lynch 1988; Knorr-Cetina & Amann 1990) Visualizations such as spectrography, pitch tracking outputs, and waveforms, though they are often deployed by engineers as uniquely artless ways of encountering the sounds with which they work, are best understood as ways of what Janet Vertesi has termed "*drawing as*" or theory-laden representational practices. (Coopman et al. 2014; Vertesi 2015)

As a way of bringing out the contingencies of representation involved in visualizing a recorded voice as-tuned, and investigating the work that goes into making the above claim about "seeing tuning" come true, we can apply spectrographic techniques to digital tuning software in the context of an epistemic studio ritual known as a "shoot-out." Shoot-outs proceed by way of side-by-side comparison of various objects in the studio. One example of a shoot-out would be having the same vocalist sing the same vocal part through several different microphones, so as to establish which mic sounds best. A shoot-out is similar in format to a blind comparative test, but can be distinguished to the extent that it involves the comparison of several different items as opposed to just two. Shoot-outs have become increasingly common in online contexts, allowing many different listeners to witness the conditions of the test and weigh in with their opinions.

Russ Hughes, UK-based engineer and writer for the website "Pro-Tools Expert," performed a shoot-out of three different digital tuning tools: Auto-Tune 8, Melodyne 4, and Revoice Pro 3. (Hughes 2016) The audio clip included in the original article presents, in random order, a vocal track tuned with each of these programs, as well as an uncorrected version. Below the audio I have produced spectrograms of these tracks in order of appearance. The reader is invited to see if they can guess which is which 1) based on the sound of each version and 2) by looking at the spectrogram and following the heuristics discussed above:

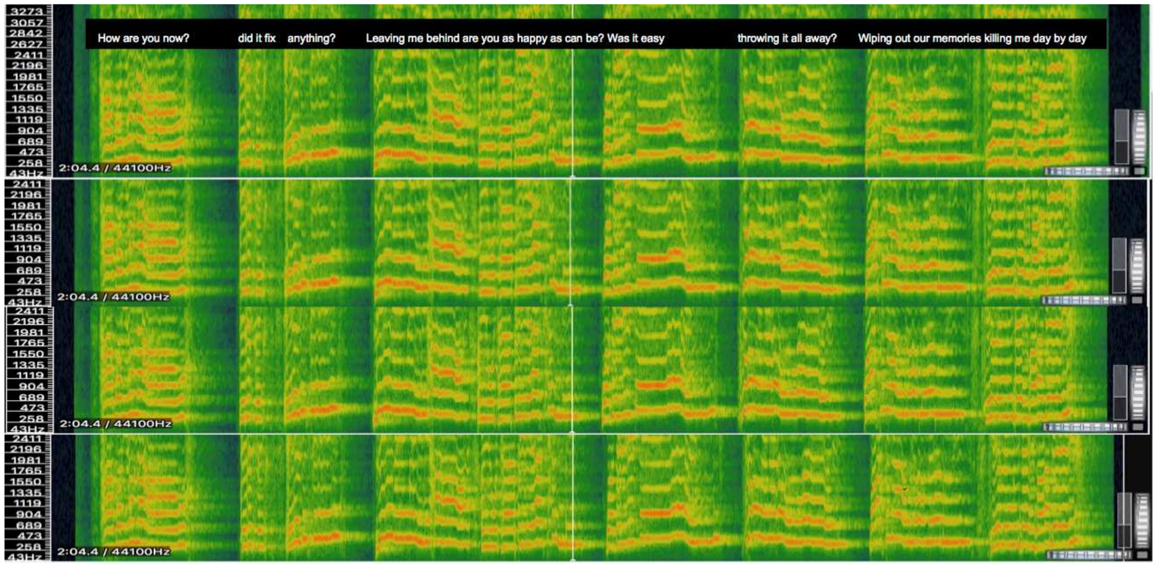


Figure 5: Spectrographic Representations of pitch correction examples 1-4, arranged from top to bottom. Visualization produced by the author using Sonic Visualizer software. Three versions of the same vocal track, through Melodyne, Auto-Tune, Revoice Pro, and non-tuned (not necessarily in that order.) from an online poll (Hughes 2016)

It should first be said that neither the various versions of the vocal take, nor the spectrographic representation of them, were produced "automatically." The engineer behind the shoot-out tracks claims to have attempted the best tuning job possible as afforded by each piece of software. The shoot-out has been staged in such a way as to downplay the personal preferences and tool-specific skills of the tuner, who goes unidentified. What has been highlighted is the identity of the source vocal, which is presented as a basis for rendering audible the effects of the tools under consideration. Similarly, in addition to the frequency distribution of the shoot-out's audio, the spectrograms represent a number of decisions on my part. These include the selection of the visualization software (in this case the free program Sonic Visualizer (Cannam et al. 2010)), the decision of where to crop each spectrogram so as to make them all simultaneously visible without losing the most important frequency data (many of the upper frequencies are not shown, for example), the degree to which time and spectral range have been dialed-in using the controls on the lower right corner of each row, the decision to include the lyrics, and the way that they are aligned with the spectrograms. Applying the above instructions on how to "see" the tuning involves a particular way of drawing our eyes across the image and drawing observations from it. The presence or absence of curves between notes, and the "perfection" of those curves if they are present, are orienting questions that locate signs of tuning as intentional objects. (Lynch 1984, 67) Listening to the recording while looking at the graphs, moreover, it is possible to note how aural and visual attention are shaped in terms of one another.

Of course, engineers usually have little need to identify production decisions visually. Visual representations typically serve as ways of orienting oneself in a mix, rather than a way of "experiencing" the recorded material in qualities that will be relevant to the final product. I passed the shoot-out test along to Carl (minus the spectrogram) to see what he made of it. His response shows how intricately taste and technical evaluation are woven together in the construction of a vocal track. Certain things were treated as obvious. For example, he does not even bother to say explicitly that the first track is the untreated one, even though this was not specified in the test itself. He was able to discern this based on noticeable deviations in pitch on this version where there were none in the other versions. The presumption that 1 is "real," moreover, structures how he evaluates the other versions:

> I prefer 3. It sounds the most natural to me. It's not perfect but its seems real. It has a very slightly thinner eq than 2 and 4, and the closest in eq to the og [original vocal]. That's a big reason why I think its AT. That and the natural quality of it not being snapped perfect pitch. So much of this is taste though. (I tend to tune things a couple cents under if I need to sound more natural, and a couple cents sharp if its a chorus vocal that needs to pop out a bit etc)

The similarity of EQ between the "og" (version 1) and version 3 combines with a personal habit

– discretionary deviation from the actual tonal center depending on the place in the song – in forming his stance toward the recording. I also happened to like version 3 the best, and I joked with Carl that his taste must have rubbed off on me. However, where he thought it was Auto-Tune I thought it was Melodyne. Version 2, which I took to be Auto-Tune because of how its notes seemed “snapped” to the pitch center, struck Carl as Melodyne or Revoice because of its “duller” sound:

> 2 is very pleasant and sounds a little more like the pitch centers of each word is snapped right on to the pure note. "anythING" is pitched correctly even on the tail end of the word. It is slightly duller than the original and this is one reason why I think its Melodyne or ReVoice Pro.

Version 4 meanwhile seems to be characterized by a certain lack of attitude, which Carl connects with his own experience with the (to him) less familiar Revoice software:

> 4 is cool but kind of the middle of road for me. The pitch peaks and valleys are a bit rounded off. Feels a bit safe and vanilla. This sounds like what I would be happy with if working in Revoice... largely due to my lack of expertise lol.

Carl also notes what he hears as pitch detection artifacts in each of the tuned examples. These appear as low end pop glitches and a “watery sound.” He also notes that what he takes to be the singer’s natural vibrato could be mistaken for tuning:

> They all have a slight low end pop glitch due to bad pitch detection in the software. 2: "Now" "Memories" 3: "How" "Me" "Memories" 4: "How" "Memories" [the low end pop glitch in] "Memories" is consistent [across examples] but the other words aren't. It's noticeable that something is affected. A boss would Izotope RX that shit. Also a couple watery sounding words. "Now" "Easy" "Our" memories. The vibrato is noticeable in the og but someone might hear that and assume its tuning. It's rough out there.

In the same way that he articulates his recognition of Revoice in terms of his relative lack of ability with the program, Carl hears the noticeable glitch in terms of the affordances of the program Izotope RX, wielded by a hypothetical “boss.” The unusual listening structure of the shootout highlights the comparative gesture at the heart of all critical listening among engineers.

End-listeners of pop recordings will rarely if ever be listening comparatively with different versions of the same song (at least not by the same performer) but engineers listen in terms of what their tools are able to accomplish, what another engineer could or would do with them, and what others might be able to notice about their own work by listening to it later on.

Digital tuning in the covert mode emerged along two pathways of repair. The first was the way in which pitch correction was taken to be a way to repair imperfect vocal takes. Artists with ample energy but lacking recording "chops" became objects of automatic repair. Where the relationship between the engineer and the artist became unclear, tuning techniques were employed to rearticulate the boundary between technical and emotional skill in new ways. The second involved cases where the tools of tuning themselves broke down and had to be repaired in order to remain accountably covert and corrective. In the covert mode, the breakdown of auto-tune meant that its use became perceptible as such. In response the "natural" and "human" qualities of auto-tune were repaired through acts of versioning and the development of new features accompanied by new forms of documentation. Users of auto-tune repaired their own perceptual skills by tuning themselves into its particular sound, the way that a skilled carpenter might get a feel for a hammer or saw. Paradoxically, the elaboration of "transparent" tuning tools and techniques over time involved the cultivation of recognizable sounds and styles. This leads us to the next chapter, which examines how an overtly audible use of digital tuning became a standard in itself.

CHAPTER 4: TUNING AS OVERT EFFECT: REPURPOSING AUTO-TUNE

While software manufacturers and the recording engineers went about repairing the perceived shortcomings of cover-corrective tuning methods, other social groups began to explore the ways in which auto-tuning could be used to produce creative new sounds. Like Russolo, who sought to seize machine noise as a mode of emotional production rather than mask it in a project of emotional fidelity, artists outside of the fields of audio and software engineering went about amplifying the artificiality of tuning algorithms in the pursuit of a new sounds. This chapter looks at how pitch correction software was *repurposed* as an overt effect. In the hands of pop stars, songwriters, and critically creative audiences of popular music, the sounds of overdriven tuning plugins were woven into afro-futurist and feminist musical traditions. As covert tuning inscribed itself in the standard practices, included features, and public conversation around auto-tune, engineers and recording artists sought new ways of making voices together.

Overt Tuning as a Critical Practice

Scholars in the fields of cultural studies and critical theory took an early interest in how auto-tune was deployed as a conspicuous effect in pop music production, especially by black and female artists. Kay Dickinson drew on the idiom of Cyborg Anthropology (Downey et al. 1995) in her reading of Cher's pioneering use of the effect in her 1998 hit "Believe," suggesting that it provided a way of thinking about "who and what is reproduced (And by what sorts of technologies) when a 'human subject' is recognized."(Ibid) Dickinson highlights the "camp" aspect of Cher's repurposing of vocal processing, arguing that "[b]y pushing current (largely straight male) standards of pop, perfection, fakery and behind-the-scenes mechanisation in unusual directions, a vocoder, like other camp objects, might complicate staid notions of reality, the body, femininity and female capability." (Dickinson 2001, 345) She also recovers the fact – easily forgotten in a post-T-Pain era, that for several years the so-called "Cher effect" was primarily employed by female artists, noting that "no long- standing male artists of the stature of Madonna (no Bryan Adamses or George Michaels) have used the device, nor even have the more 'feminised' male performers such as the contemporary boy bands" (Dickinson 2001, 341)

Cher had long been considered an icon of empowerment for the LGBTQ community, for which camp's project of "delight in the inauthentic" had proven, as Dickinson puts it, "inspirational in its survivalist hints." Upon the release of her album Believe and its titular single, Cher had recently been awarded the Gay and Lesbian Alliance Against Defamation's (GLAAD) Vanguard Media Award. She had also become a key public voice for bodily self-determination, famously asserting that her decision to use plastic surgery is "nobody's business but my own."(ABC 2002) She took a similar stance towards her vocal processing. Warner Music executive Rob Dickins reportedly wanted the effect taken off the recording prior to release, to which Cher replied "Over my dead body. Don't let anyone touch this track, or I'm going to rip your throat out."(Clayton 2016, 27) Overt vocal processing was taken to highlight the artifice inherent in acts of vocal expression and provided a way to play with the boundaries between male and female voices, as well as that between human and machine. Dominic Pettman has noted the critical affordances of Auto-Tune in its ability to disrupt and erase ways of assigning identity to voice, such that "the culturally gendered binary of the larynx is revealed to be a rather ambiguous modulation of the sonorous spectrum: a woman can sound like a metallic man, while a man can sound like a fiber-optic woman." (Pettman 2011, 151)

Anthropologist Nick Seaver has compared the way that Auto-Tune troubles conventions of vocal skill with Judith Butler's analysis of drag performance as a way of challenging entrenched preconceptions and aesthetics of gender:

> Drag provokes unease in those who are invested in traditional gender roles, Auto-Tune in those invested in traditional vocal skills... Where drag dislocates external appearance from biological sex, Autotune dislocates the ability to hit notes from the ability to sing. Now, in its most obvious "Cher effect" incarnation Auto-tune does more than just make someone who can't sing hit the right notes. It conveys a human-machine fusion, a matching of precision with imprecision, and it does so in a way that is legible. You know that T-Pain is singing through Auto-Tune, and you won't mistake him for a very precisely melismatic singer. In drag, the signifiers of female (or male) gender are performed, but in a way that highlights their performance. The legible use of Auto-Tune can be so discomforting to those invested in vocal skill because it denies the need for skill, but also because it always refers back to the intended use of the tool. (Seaver 2009)

The analogy between digital tuning and the practice of drag performance highlights how the topic of legibility is reconfigured between covert and overt tuning. Questions of legibility in the context of covert tuning concern whether or not the algorithm is able to correctly parse (and thereby correctly correct) a signal. Auto-Tune is in-scripted so as to make voices legible to it. Legibility for the engineer means preparing a recording within the context of the DAW and tuning software so as to accomplish the legibility of the tuning act. Legibility is accomplished as demarcation between what is supposed to be there and what needs to be fixed. Seaver points out that, in the context of overt tuning, what becomes legible is the tuning *per se*, and that this move has the subversive effect of decoupling criteria of technical pitch-accuracy from received understandings of skill, while also drawing attention to the means by which they are usually articulated together in the first place.

Osvaldo Oyola, in his essay "In Defense of Auto-Tune" has similarly suggested that digital tuning implies a critique of "the authority of the arbiters of talent" in the same way that Benjamin saw mechanical reproduction as a challenge to the "aura" of an artwork:

> Mechanical reproduction may "pry an object from its shell" and destroy its aura and authority–demonstrating the democratic possibilities in art as it is repurposed–but I contend that auto-tune goes one step further. It pries singing free from the tyranny of talent and its proscriptive aesthetics. (Oyola 2011)

Critically resonant readings have also emerged from scholars working traditions of African and Afro-American Studies and Black critical theory. Alexander Weheliye read overt pitch correction (which he called the "vocoder effect") as part of a broader practice of technological mediation of the voice in black popular music, including the use of talk boxes, and "telephone voice" effects. Overtly technologizing the voice, in his account, becomes a way of creating layers of meaning and challenging tacitly racialized conceptions of the voice. Nielson draws specifically on Weheliye's observation that many of these sonic tropes have the effect of aestheticizing surveillance tactics such as wiretapping as part of the musical text. (Nielson 2010, 1264) He locates the overt use of auto-tune within a larger tradition of "Black signification through recontextualization using existing elements (words, technologies, fashions, etc.) in unique ways to create new meaning." (Nielson 2010, 1262) The practice of "sampling" in hip hop, for example involves

> an invocation of another's voice, sometimes dozens of voices from dozens of times and places in a given piece, it effectively works to deemphasize individual rappers and producers by delocalizing their presence across a vast continuum of place and time. In effect, rappers and producers willingly undermine their presence with sampling by immersing themselves in a sonic tradition that is far greater than the individual. (Nielson 2010, 1263)

Citing what he takes to be an instance of vocoding in Tupac Shakur's 1996 track "Can't C Me," he emphasizes the way that the effect is used as an "obscuring force" that simultaneously complicates the identification of the individual voice and reinforces a broader Black musical tradition of repurposing and redistributing vocal identities. Rayvon Fouche has argued specifically for an increased attention towards practices of Black signification in STS and the history of technology, developing the concept of "black vernacular technological creativity" in order to "describe the ways African American people interact with material forms and affects of technology."(Fouche 2006, 641) Fouche builds on Joel Dinerstein's concept of the Techno-dialogic, or the way in which ""the presence (or 'voice') of machinery became integral to the cultural production of African American storytellers, dancers, blues singers, and jazz musicians,' … how, during live artistic performances, black 'musicians brought the power of machines under artistic control and thus modeled the possibility of individual style within a technological society.'"(641) One example of black vernacular creativity is the practice of "scratching," or the manual manipulation of a record on a turntable. Fouche writes:

> When DJs began scratching, they subverted the fundamental meaning constructed for record players as well as for that of the LP records. What is significant about this basic maneuver is that it drastically diverges from the principal meaning embedded in the technological network associated with records and record players: to listen to prerecorded sound/music. DJs were thus able to creatively reconceive the technological products associated with recorded music and the knowledge associated with their functions based on their own black/ethnic musical sensibilities. (Fouche 2006, 655)

In his attention to the "technological vernacular" as a way of recovering the critical and innovative practices of black communities, Fouche is also in conversation with Henry Louis Gates Jr.'s theoretical explication of the African-American discursive practice of "Signifyin(g)." To "Signify on" something is to engage in a form of linguistic deep play which multiplies, layers, and rearranges meanings, often to a subversive effect.(Gates, Henry Louis 1988) In his genealogy of the term Gates connects Signifyin(g) to the partially synonymous practice of "the Dirty Dozens:"

> "a very elaborate game traditionally played by black boys, in which the participants insult each other's relatives, especially their mothers. The object of the game is to test emotional strength. The first person to give in to anger is the loser."(Gates, Henry Louis 1988, 68)

“Doing the dozens” remains an enduring tradition, especially in online contexts. As Judy Isaksen observes in her study of Hip Hop and the construction of masculinity among young men, forums like BlackPlanet host threads dedicated to dozens and signifyin(g). In a 2009, for example, commenter A-Deuce asked his interlocutor Benson to “Tell me why yo mama so dumb she thought T-Pain was a medical condition.”[39] As with many instances of Signification, it is all too easy to underestimate the complexity of A-Deuce’s line. Behind the apparent insult against Benson’s mother’s intelligence (or at least cultural awareness) there is a riff on the sentiment, common by 2009, that T-Pain (whose legal name is Faheem Najm) was not so much an artist as a pathology. T-Pain’s work has been described using phrases like “the death of music.”(Matousek 2015) In his 2009 single “Death of Auto-Tune” Jay-Z turns Najm’s stage name into a verb as he calls on hip hop to “get back to rap, you T-Paining too much.” Katherine Turner, meanwhile, has used the concept of signfyin(g) to argue that a critique of T-Pain (and, by extension, overt tuning) based on imputations of laziness, ignorance, or lack of talent,

> fails to recognize this activity – the rejection of auto-tune’s literal application (that is, as a pitch corrective) – as sustaining a tradition of techno-black cultural syncretism that has attended hip hop since its inception. Such critics and communities succumb to the cleverness of T-Pain, who exposed their inability to perceive the non-literal and spotlighted the hubris of those unwilling to extend him the capacity for figurative expression. In short, they suffer the same failure as the Lion in the Signifying Monkey Tale… In combination, irony, Signifyin(g), and techno-black cultural syncretism reflect a single motivation to revise and resituate texts and technologies, to exploit the semantic space between literal and figurative expression, and to maneuver within this space in an effort to combat inequality, to challenge social imbalance and injustice, or simply to rebut harsh treatment by music critics. (Turner 2016)

Taking T-Pain’s vocal aesthetic as a practice of signifyin(g) – as a way of playing with voice, expectation, and meaning, perhaps in part as a way to “test emotional strength” – provides a

[39] Online commenter A-Deuce, cited in (Isaksen 2011, 159)

thicker, and in many ways more adequate description of how he has repurposed Auto-Tune. It explains, for example, why Najm uses the effect even though he actually does appear to possess a conventionally talented "natural" singing voice. (Kelley 2014) It also helps explain the professed depth of his engagement with the technology itself, particularly his knowledge of its development and the ways that moves with the voice:

> I can firmly say that nobody has looked into Auto-Tune the way I have. You know what I'm saying? Like I've looked into Auto-Tune. I've literally met the inventor of Auto-Tune. I've talked to him about the way that Auto-Tune was invented. Auto-Tune was invented by this guy who used to be an oil digger. This guy used to work on a oil rig, and they used to send sonar signals and tones down in the ground, and if it came back a different tone up to where your equipment was, then that means you, that determines if you've got oil or not. So you send a signal and if it comes back a different tone then it changes the tone like, oil will change the tone so it's like he used that same math to make Auto-Tune. And it's like you send a tone into ProTools and it sends you the right tone back. And a lot of math went into that shit and just some shit that's more complicated than — it would take us fucking a billion minutes to explain this shit to regular motherfuckers. But, like, I really studied this shit, and I know for a fact that nobody has sat down in the studio and studied this shit that much. Nobody has done that. Because it happened too fast. They didn't have time to do it. I studied Auto-Tune two years before I used it once. And I know it happened too fast. After I used it, n***as just started coming out of nowhere. So it happened too fast for them to know how that shit works. And I know I studied that shit and I know the technology. I know why it catches certain notes. I know why it doesn't catch certain notes. I know why the shit works the way it works and I know n***as ain't did that. For sure… (Todd 2014)

He re-tells the auto-tune origin story in a way that makes it his own. Hildebrand appears as a former "oil digger," and the functional connection between seismic analysis and the skilled use of auto-tune forms the expertise that Najm claims with respect to the technology. He re-inscripts both the narrative of auto-tune as well as the "manual" language for how it ought to be used. Where, from Antares' perspective, the T-Pain approach was incorrect, Najm argues that his own approach is the correct one with respect to his colleagues who misuse it. Among the colleagues singled out for their lack of skill with overt auto-tune is the rapper Future, and Kanye

West. T-Pain assisted on West's album 808s and Heartbreak, which made use of auto-tune throughout and marked the beginning of West's continued use of the effect. Regarding the tuning on Heartbreak, Najm was quoted as saying

> Kanye uses it, but he doesn't use it correctly. He makes great music with it, but the way that I use it and the way that I've shown Chris and Jamie to use it, he doesn't use it that way," he said. "He sings without it first, and then he puts it on it. ... You don't know how it's going to come out. You can't catch your mistakes before that happens. Sometimes he gets a little wobbly and things like that. (Branch 2014)

Where T-Pain went against auto-tune's prescribed corrective use, repurposing it as an overt effect, he continued to reinscribe it as well, labeling West's use of the effect "incorrect" because it is not monitored live (the very thing about T-Pain's use that Alpert had initially considered "incorrect.") For Najm, Auto-Tune was not a gloss that gets thrown on after the fact, but an instrument and technique that one could master. He claims a particular depth of knowledge of the software itself, both as a way of distinguishing himself from others who have adopted a similar technique but don't "really" know how to use, and as a way of situating himself in a broader history of technologically altered voices as a musical gesture. On the one hand he prides himself on having met Auto-Tune's inventor, on the other hand he claims (via the signifyin(g) techniques of naming and punning) a techno-aesthetic kinship with talk box pioneer Roger Troutman and nu-jack swing vocoderist Teddy Riley: "let me get a moment of silence, for the late great Roger Troutman/ Y'all n****s ain't holdin him down, so we had to put me at you/ B***h, I'm Teddy Pain, the son of Teddy Riley." (Najm & Khaled 2008) In a 2008 interview with the New Yorker, Faheem Najm (T-Pain's legal name) noted that he decided not to use Auto-Tune on a song about his son.(Frere-Jones 2008) Najm's son, who was later diagnosed with autism, would serve as the inspiration for T-Pain's *Stoicville: The Phoenix* project. The titular Stoicville, he explains, is a place "where everybody is stoic—where nobody has emotions. You don't get shit from *anybody* in Stoicville. You don't get people saying or doing fucked up shit to you. Everybody's just stoic. Nobody has emotions and everybody minds their own fucking business. That's the town for me. That's where I want to live." (Neyfakh 2014) Though the album has not yet debuted as of this writing, the first single, "Intro" features none of his signature vocal effect. The possible exception is what seems to be a melismatic and pitch-shifted female vocal line running in reverse, which echoes through the song's outro. (Najm 2014) Najm's repurposing of Auto-Tune as the "T-Pain effect" involved re-inscribing the popular narrative of auto-tune's development as well as his own development as a "rapper ternt sanga."(Neyfakh 2014)

Beyond the world of western major label pop music, overt tuning came into particularly wide use in the popular musics of Northwestern Africa. Mdou Moctar, a popular Tuareg guitarist and singer from Abalak, Niger, began using auto-tune after his recording engineer recommended it during a session in a Nigerian studio. His song "Tahoultine," featuring heavily distorted auto-tuned vocals, became a huge hit among young listeners in the Sahel region, who trade music with one another via the memory cards of their cell phones. Ethnomusicologist Chris Kirkley, who first brought Moctar's music to American and European audiences through a 2011 release titled Music From Saharan Cellphones on his Sahel Sounds Label, describes his initial reaction to Moctar's use of tuning:

> What I really like was its autotune, that crazy blown-out autotune. Like if Cher knew about that, she probably would've regretted it. She probably already regrets it. But what's very distinctive about that Mdou Moctar track is that it's Tuareg guitar music, but it has a drum machine and this crazy autotune and it sounds sort of like a robot or like a spacey anthem. (Isaacson 2013)

Music journalist Jace Clayton has further explored the popularity of auto-tune in Berber music (the Tuareg are a nomadic subset of the Berber ethnic group). Clayton points to the Moroccan singer Adil El-Miloudi, who "employ social-realist lyrics, chromed in Auto-Tune, to elevate the travails of the underclass without being depressing or didactic." (Clayton 2016, 37) He explains the popularity of overt tuning, in part, in terms of its aesthetic resonance with traditional melismatic vocal styles from the region:

> Melisma's swoops and dives are exactly the type of melodic movement that provokes Auto-Tune into extreme corrective mode, thereby producing its most unusual sounds. This, I believe, explains the software's mind-boggling success in North Africa. The region embraced Auto-Tune so early and so heartily because for more than a millennium audiences have been listening to— and for— those gorgeous, subtly rising and falling pitches. And they sound especially startling when processed through Auto-Tune. The familiar pitch slide gets activated by bizarre effects. A weird electronic warble embeds itself in rich, throaty glissandi. The struggle of human nuance versus digital correction is made audible, dramatized in a zone of heightened attention. Listening habits from the dawn of Islam helped Auto-Tune mean so much here. (Clayton 2016, 46)

Clayton describes the way that engineer El Hajj Ali Ait Bouzid incorporates Auto-Tune into the recording process at al-Maarif studios in Amazigh, Morocco.

> First, the backing band records their material. The engineer takes a few days to edit together the best selections and tidy things up, then the singer will come in to lay her raw vocals on top. Finally, the engineer will go in after that, to mix it all together, and that's when Auto-Tune creeps in. In a typical Amazigh combo, the banjo (a notoriously fickle instrument that is constantly going out of tune due to micro-changes in ambient temperature and humidity) sets the reference note, the rest of the band tunes to the banjo, and the singer, days later, adjusts her intonation to fit the band. Auto-Tune then gets tuned— or rather, detuned— to whatever results. I say detuned because in studios across Morocco, I watched time and again as the people who used Auto-Tune the most relied the least upon its interface. The software offers many ways to customize its effects, including settings called Arabic and Pentatonic— registers that wouldn't be hard to tailor-fit to Berber songs. Nobody bothers with any of that. Instead, people click a single knob— Pitch— and twist. (Clayton 2016, 49)

Within the heavily pitch quantized voices of female Berber vocalists, Clayton hears an unexpected resonance between auto-tune's radically modern technological sound and the traditional distribution of gender roles in Moroccan society. He writes:

> Auto-Tune activates deep-seated and conservative ideals of Berber womanhood by making those high-pitched voices cut even more keenly. The software amplifies old ideas of the rural and the feminine… Berber Auto-Tune acts something like a digital veil to protect— and project— a woman's purity as she makes her way in the modern world. (Clayton 2016, 55)

Setting aside the question of the accuracy and ethnographic authority of Clayton's analysis, we can take his account as an index of how a use of auto-tune that, in the American context, would come across as radically subversive of accepted musical norms can be entirely consistent within a musical tradition which locates its emotional resonance in the melismatic gesture more than the stability of pitch. It is suggestive that auto-tune may "amplify old ideas of the rural and the feminine" in Berber music while, in the American context, it carries an unsettling and futuristic connotation. Music technology researcher Ethan Hein writes, for example,

> The Cher Effect is used so often in pop music because it's richly expressive of our emotional experience of the world: technology-saturated, alienated, unreal… Here we have it, the perfect encapsulation of what it's like to be a feeling human being in a hypertechnological, hypercapitalist society. Auto-Tune gives that indefinable feeling a literal voice. No wonder it's so popular. (Hein 2015)

Responding to Hein, composer Greg Brown suggests that, considered in terms of Hip Hop's genre conventions, auto-tune actually highlights and facilitates emotional expression.

> I think Auto-Tune here is somehow protective for Kanye when he is expressing emotion in a genre where that is not really smiled on. I haven't quite put my finger on it, but I think the dehumanizing of the human voice is somehow a foil for the expression of inner turmoil. It's haunting… The hard part for me to wrap my head around is the fact that Auto-Tune is a filter, a dehumanizer, and it manages to make Kanye both closer and more human. (Hein 2015)

Critic Mike Barthel advances a similar theory regarding the emotional resonances of West's voice on *808s*:

> It wasn't the raw emotion of humans, but the synthesis of emotional impulses and mechanical restraint, a computer's inauthentic attempts at automatic expression which nevertheless sprung from a real human need to communicate. (Hein 2015)

Ge Wang, who helped develop the Smule App I Am T-Pain, suggests that it is the very restrictiveness of auto-tune, the way that it filters out all but the "right" notes, that allows people to let their guard down. Auto-tune, he suggests, has the ability to

> give people some freedom to be out there without so much inhibition, at the same time retaining a little bit of themselves and that's using Autotune of all things. And it kind of has, you know, say what you will about Autotune but it has this effect of kind of, like alcohol perhaps, loosening people up just a bit. Maybe that's why T-Pain sings so many songs about drinking. (Wang & Applebaum 2013)

Wang's analogy between pitch correction and alcohol is something of an echo of the emotional logic auto-tune presented with respect to its covert correction mode. Not having to worry about hitting wrong notes, he suggests, has the effect of "loosening" up the voice while still retaining some of themselves within it.

Overt Tuning as Songwriting Tool

As black and female performers overtly repurposed digital tuning as a way of refashioning and augmenting their own voices, professional songwriters and producers took up overt tuning as a songwriting tool. By Carl's account, it was the songwriting community who really spurred on the use of overt tuning, precisely because their work was not meant to be the definitive version of the song:

> When I started noticing (overt tuning) being used more and more was when writers would come in. You'd give them the beat, and they were supposed to write over it, and so they would demo their vocals on the beat, even though it was supposed to be a writing

> session for someone else entirely... The writer would just have you put Auto-Tune on his voice and just print it. Not even have it as an effect you can dial in and tweak later and make it more subtle, but I'd just print the shit right on there. And I'm like, 'you sure about that? because you can totally hear it,' and they're just like 'no no that's great.'... And they could get away with it because they'd just say, 'oh I'm using it as a writing tool, the actual singer that it's for is gonna come by and sing over it anyway...'

His account of overt tuning in songwriting is echoed in a *New Yorker* profile of pop songwriter Ester Dean, who employed tuning as a central part of her writing process:

> The first sounds Dean uttered were subverbal—*na-na-na* and *ba-ba-ba*—and recalled her hooks for Rihanna. Then came disjointed words, culled from her phone—'taking control . . . never die tonight . . . I can't live a lie'—in her low-down, growly singing voice, so different from her coquettish speaking voice. Had she been 'writing' in a conventional sense—trying to come up with clever, meaningful lyrics—the words wouldn't have fit the beat as snugly. Grabbing random words out of her BlackBerry also seemed to set Dean's melodic gift free; a well-turned phrase would have restrained it. There was no verse or chorus in the singing, just different melodic and rhythmic parts. Her voice as we heard it in the control room had been Auto-Tuned, so that Dean could focus on making her vocal as expressive as possible and not worry about hitting all the notes… After several minutes of nonsense singing, the song began to coalesce. Almost imperceptibly, the right words rooted themselves in the rhythm while melodies and harmonies emerged in Dean's voice. (Seabrook 2012)

When these syllables congeal into the words stored in Dean's blackberry, their semantic meaning appears less important than their qualities as bodily-shaped sound. A comparison with another case of the methodical production of nonsense syllables provides some perspective on the work Dean is doing with auto-tune here. The glossolalia (or "speaking in tongues") of charismatic Christians, writes anthropologist Thomas Csordas, "by its formal characteristic of eliminating the semantic level of linguistic structure highlights precisely the existential reality of intelligent bodies inhabiting a meaningful world." (Csordas 1990, 25) The lack of linguistic meaning, then, is not a drawback but its primary strength as a mode of communication. As Csordas puts it,

> the stripping away of the semantic dimension in glossolalia is not an absence, but rather the drawing back of a discursive curtain to reveal the grounding of language in natural

> life, as a bodily act. Glossolalia reveals language as incarnate… the utterance takes place at a phenomenological moment prior to distinction between body and mind, a distinction that is in part contingent on the objectifying power of natural language. (Csordas 1990, 26)

Dean is not working within the context of religious practice from which Csordas draws his insights. She is instead working in a context of commercial sound production. She is not channeling embodied emotion into an idiom of group exorcism or "resting in the spirit," but is instead trying to fit her voice "snugly" to the beat. She is tuning-in to the context of the backing track. Csordas' account, however, does usefully suggest that, even in Dean's case, the tentative adventures of her voice can be understood as a sort of pre-objective phenomenological moment. Her voice has the character of both expression and sensation, as she uses a set of stock phrases to feel around the beat for an effective melody and cadence. To borrow Csordas' words, Dean's songwriting practice "serves the cultural process of self-objectification and is not simply a dreamy state of meditatively emptied consciousness" (Ibid) The self-objectification (or, to use Prentice's more felicitous term, bodily object-formation) of the singing songwriter – the shaping of her voice into words, melodies and rhythms that other performers can take on as their own – is part of the process of how singing becomes song. Auto-Tune's prescribed form of use situated it at the end of the vocal production process: the singer would do their best and pitch correction would "restore" the performance by fixing the mistakes. Dean repurposed Auto-Tune by placing it at the beginning of the vocal production process, before the "real" singer even enters the vocal booth. She refashioned it as a way of presenting her voice as a model for her client, the artist.

Even Cher's *ur*-instance of overt tuning can be traced largely to the work of the six songwriters behind "Believe." Mark Taylor, one of these writers, initially led the readers of trade publication *Sound On Sound* to believe that the robotic yodel in Cher's voice was the product of a relatively obscure vocoder pedal called the Digitech "Talker" (Sillitoe & Bell 1999) Taylor's ruse even made it way into the pages of *Popular Music* (Dickinson 2001). Even as he pioneered the overt use of digital tuning, Taylor apparently felt compelled to retain some degree of covertness regarding how exactly it was accomplished. One way of accounting for this is that Taylor was attempting to protect a "trick of the trade" and retaining his monopoly on a popular new sound. It could also be argued, however, that Taylor's "Talker" story had to do with the fact that, at the time, overt Auto-Tuning was very much an instance of "misuse," one that might cast Taylor as having incompetently stumbled upon a mistake that sounds good instead of having crafted it deliberately. While Cher was in a position to take ownership of the effect as a purposeful case of self-making, the person who first put the tuning on the track had to be careful to avoid appearing either technically unskilled or careless with respect to the voice of his client.

As overt tuning's use as a writing tool began passing into the hands of engineers and artists, these groups had to figure out what to do with it. For engineers, the tuning seemed, at first, incorrect and itself in need of further correction.

> It would start happening that (replacing the hard-tuned vocals) wouldn't actually happen. Whether it was just backgrounds would be hard tuned and maybe Britney Spears didn't have time to sing all the backgrounds so they just left the writers' backgrounds in there and pretended it was her. Basically, I think it was a writing tool to use as a reference and then everybody just started going with it... People would use it and it wouldn't even be in the right key, they'd leave it chromatic, like, 'key chromatics, oh I love that!' you know? Not even tuned to any of the notes of the scale, just this warble between every note. When I was working with (another producer) later, he'd say, dude, that's not the way to tune it,' but by then we'd just go through and Melodyne it. And he has a great sense of pitch and he would just get real surgical and use Melodyne to fix it.

Once artists started using tuning on their own vocal takes, however, the layering of tuning work became more baroque. In the following scene, for example, Carl is working on a demo song written by an up and coming producer for a well-known Hip Hop artist. Carl knows the producer from a session he worked on at a different studio. The track consists of a beat with a chorus refrain, or "hook" which is tuned heavily and overtly. Given the style of the song – a somewhat sinister R&B slow jam with a minor key melody and lyrics that one might describe as seductive if they weren't so direct –the obvious tuning would sound entirely appropriate even if it was not just a demo recording. Carl had been in-studio with the producer the night before, and had begun polishing things up while everyone went out for a long dinner break. Now he is going over last night's work, making everything sounds good, and moving on to the parts that were giving him trouble. One line trails off indeterminately, so he makes the decision for the singer and resolves the melody in a way that catches the tuning quantization in a particularly strong way ("auto-tune at the end of the phrase, he'll love that!" he explains.) He listens to the next line, exhales audibly, and takes a beat to consider his approach. The singer's voice is moving in unpredictable ways, and after singing the hook to himself for a moment, Carl ventures a guess as to why:

C: As simple as this is, I mean, it takes a fuckin' long time
Daaah
[singing] "You know I wanna – oh – you know I wanna" [hums melody]
It's all over the place!
He's singing with Auto-Tune tune on, that's what it is. Shocking
O: You're tuning something that's already been tuned?
C: No, uh, he just had it on as he was hearing it on auto. When he was tracking.

Even though the vocal track that Carl was given initially did not have any tuning effects applied to it, the singer had apparently been monitoring an auto-tuned version of his own live voice while doing the vocal take originally. Its unpredictability, it turns out, has to do with the fact that it has already been predicted by a pitch tracker, corrected, and sent to the singer's headphones during the take, changing the way he sang the line. I ask Carl to clarify whether people sing differently when they are listening to a tuned version of their own voice.

C: Yeah, it does. It totally does! They sing to the tune sometimes…
You sing slightly flat and Auto-Tune will be more effect sounding.

Here auto-tune is having an audible effect on the singer's voice even though it is not actually being applied to, or "printed on" the track audio. Its effect is apparent not in the ways that it manipulates the recorded voice, but in how it caused a voice to move with itself in the original recording session. Why they did not simply print the effect on the track from the start is unclear, though it probably was because they wanted to option of not using it. Now, however, because he wants to tune a performance that was recorded according to some (presently unknown) set of tuning parameters, Carl is faced with the somewhat strange task of pre-tuning a non-tuned performance (that was nonetheless tuned at the time that it was performed, at least from the perspective of the vocalist) that it will feed more easily into the tuning plugin yet again. I attempt to clarify, and he complicates things further by explaining that the different versions of Auto-Tune are in fact relevant to how this line is going to turn out.

O: So you're tuning something so it can go into Auto-Tune better?
C: Yes. And in fact Auto-Tune Evo works slightly better.
I think I need to do Auto-Tune Evo. Auto-tune the regular one works better with tracking, uh, recording better on my setup because I'm running HD Native [version of Pro-Tools] I just found that it does, I have no idea. Evo is a newer plugin so it probably takes up more shit and so there's more latency, you know what I mean? So I have to use… even though I have Evo and some people like it better I have to use the Auto-Tune 5.

The problem, in this case, is that different versions of Auto-Tune work in different ways, and each works differently depending on the system in which they are being used, how much CPU power they are allotted by that system, and how they are positioned within the signal chain of that system. Carl has many audiences to think of: that of the producer of the song, that of the artist for whom the song is being written (as well as the A&R representatives from their label who will also want a say on which songs the artist chooses), that of the listening public who will hear the ultimate version of this song should the artist include it as a track on his album, and that of his colleagues (including myself) who will hear this work as his.

As engineers developed new ways of working with tuning software, effectively revising its built-in prescriptions about what counts as a correct voice, so did singers develop new ways of singing *with* tuning algorithms. New worlds of vocal possibility were opened through the practice of monitoring the tuned, rather than "dry" output of the pitch correction processor. As Antares VP Marco Alpert explains, this was something the designers of Auto-Tune specifically did not recommend:

> I write our user manuals, and one of our tips in the manuals pretty much from the beginning has been, when you're tracking don't monitor through AT, when you're performing live – you know we have these boxes for live performance – you do not want to monitor that because it'll throw you off. You'll be hearing something, you'll hear the post-corrected, and as you start reacting to that it's going to be really really inconven… anyways that was the accepted wisdom for a really long time. With Hip Hop use of the effect, a lot of those people they got used to sort of playing the effect, with their voice they knew what to do and they wanted to hear it while they were performing. And we supported that. There's two modes of Auto-Tune, there's automatic mode which is a realtime mode and there's the graphical mode which you can go in and tweak to your heart's desire afterwards, but and that's just continued apace. It's become more and more the standard recording technique is people want to do their tracking through correction or with the effect whichever they're doing.
> (Author Interview with Marco Alpert, April 29, 2014)

In 2008 Antares released its Auto-Tune EFX product. The box and manual included several references to "The T-Pain/Cher Style Effect." In 2011 Najm sued Antares for unauthorized use of his trademark. (Gardner 2011) Meanwhile Najm developed T-Pain branded pitch correction products through DSP developer Izotope and an app company called Smule (which includes many faculty members and students of Stanford's Center for Computer Research in Music and Acoustics.) The ability to monitor from a tuned signal, has in fact become one of AT's major selling points and a way of distinguishing itself from Melodyne. Alpert explains how the unrecommended tuned-monitoring feature has become an important selling point:

> "You can't do graphical mode in real time obviously. That's always after the fact. Um, And some of our competitors, Melodyne in particular, doesn't do realtime at all. It's just incapable of doing that, so that gives us an advantage."(Alpert 2014)

Anticipating and tuning the voice, in this case, resembled the task of coaching a boxer pitted against an invisible opponent. By taking up the corrected signal and incorporating it into their own singing style, singers appropriated and subverted the corrective logic of the pitch correction algorithm. They refused to be excluded from the tuning process, leaving the software and the engineer the final say on their voice. Of course, their voice still circulates and is subject to retuning down the line, but even when the effect itself is removed from the signal path, it remains in the tendencies of the "raw" vocal take.

The narrative of taking creative ownership over what was originally intended to be a covert, limiting, and standardizing technology, is common in the history of music technology.[40] (Pinch & Bijsterveld 2003). T-Pain took ownership of the effect, playing it like an instrument rather than letting it "correct" his voice. Moreover, he overtook the brand itself, building up a popular association between "hard-tuned" Auto-Tune and his own voice. Songwriters like Ester

[40] Examples of this include the Maestro Fuzztone pedal, which was originally advertised in the early 1960s as a way of making an electric guitar sound like various orchestral instruments, and ultimately helped form the basis of an entirely new distorted guitar sound. Don Buchla's 16 step sequencer design was meant to replace having to do 16 tape splices, though it became a tool for repetitive rhythmic composition.

Dean, similarly, used AT not in any effort to allow their personally "authentic" voices to shine through, but to allow the production of melodic terrains that other singers could inhabit.

Just as we traced the humanizing reconstruction on the engineering pathway, we find a similar process with this artificiality pathway. The Earliest editions of Auto-Tune made no mention of its robotic potential. By Auto-Tune 3, however, the manual makes two somewhat dismissive references to the "infamous Cher effect." (Antares Audio Technologies 2001, 11, 26) Cher's name is not written in Auto-Tune documents without the word "infamous" preceding it until Autotune 7 came out with a dedicated chapter on the "Auto-Tune Vocal Effect," which, the manual reminds us, was "First heard on Cher's 1998 mega-hit 'Believe.'" (Antares Audio Technologies 2011, 77) Again, following this reinscription of overt tuning as 'not a bug but a feature,' Antares saw fit to release several zero-retune specific plugins to fill the market niche.

Even more radical reappropriations have arisen from the creative and critical consumers of popular media. Hard tuning has been used as an absurdist filter for anything from recordings of crying babies to Martin Luther King Jr.'s "I Have a Dream" Speech. Perhaps the most high-profile such use is "Auto-Tune The News," a series of elaborately edited compositions of network news footage wherein people's voices have been re-tuned into melodies and composed into parodic musical revues of the week's current events. By 2010, auto-tune had become a widely used forum for deep play with voices both marginal and overexposed. As it became increasingly entrenched as a standard studio practice, it also came to seem like a parody of itself, not just taking "talentless" vocalists and turning them into singing cyborgs but strip mining melodies from everyday speech. It was as though Auto-Tune had fallen prey to its own logic: flooding the market with singing voices and making it difficult to tell what counted as a "correct" or "incorrect" voice anymore. The next chapter looks at the everyday work of vocal tuning as a different kind of translation of conversation into song. It examines the work of an LA engineer on a single vocal track.

CHAPTER 5: DIGITAL VOCAL CORRECTION AND CONVERSATIONAL REPAIR

"When in doubt, mumble." – Duncan's Law[41]

In recent decades, digital corrective tuning[42] has become a standard component of sound recording, particularly within the field of major label pop music production. Engineers regularly use software such as Antares Auto-Tune or Celemony Melodyne in order to correct pitch and timing problems, especially in vocal performances. While the technique has gained popular attention through the overt use of automatic pitch correction for an artificial-sounding vocoding effect, its more common "corrective" use within the recording engineering community continues as an open secret – something everyone knows about but is hesitant to acknowledge. This is due in large part to a broader controversy concerning the desirability and ethical-artistic implications of corrective tuning. While proponents argue that pitch correction increases efficiency of studio time and allows greater emotional expression by reducing the burden of technical skill among musicians, critics claim that the practice erodes artistic expression, depreciates talent, and homogenizes music. The popular auto-tune imaginary is one that oscillates between the trans-human technological sublime and an inhuman automatism. To listeners of popular music, digital tuning has come to seem both commonplace and fantastical.[43]

---

[41] One of the Audio Engineering Society's recognized "laws for audio engineers," attributed to Engineer Kent Duncan (Chinn 2014)

[42] For the sake of brevity, the terms "pitch correction," "intonation correction," "pitch shifting," "auto-tune" and "auto-tuning" (the lowercase "a" indicating the term's common synechdochic use) while each having distinct meanings and connotations, will be treated as roughly interchangeable unless otherwise specified.

[43] In an online video titled "The Auto-Tune Ascension," YouTube user "BarelyHuman11" weaves Auto-Tune into a dizzying array of esoteric narratives. He notes, for example, that the release of Cher's early auto-tuned hit "Believe" coincided with: "the sun's central-most alignment with the milky way galaxy. The galactic alignment has, for eons, been said to signal the age of humanity becoming in-tune. This is why Auto-Tune is made by Antares Technologies,

In response to this context of hype and controversy, this chapter's approach is critical to the extent that it is deliberately small-scale and mundane. Rather than participating in the ongoing polemic concerning auto-tune as a monolithic and autonomous technological force, entailing a particular set of "impacts" on society, I aim to account for it as a specific, contingent, and situated socio-technical practice. By closely examining an instance of vocal tuning in-action, I will show how a tuned voice is collectively accomplished as an instance of "invisible" infrastructural labor within a studio's arc of work. (Star & Strauss 1999) I take this to be a way of answering the practical question of how engineers come to experience a vocal performance as "incorrect" and how acts of "correction" are performed. Building on work in the sociology of repair (Henke 1999), specifically ethnomethodological approaches to repair and correction within domains of technoscientific work (Suchman 1985; Schegloff et al. 1977; Maynard 2013), I offer a thick description of auto-tune as a mode of everyday socially accountable action within the recording studio. Within this repair-oriented framing, I analyze digital tuning as an assemblage of activities which occasion a field of practice, or "topical contexture." (Lynch 1991)

I employ the tools of conversation analysis (CA) (Sacks 1992) as a way of highlighting the polyvocality and temporal emergence of digital tuning in-practice. Whereas auto-tuning is often assumed to be a unidirectional and deterministic application of a ready-made procedure (a pitch detection and correction algorithm) to a docile object (a live or pre-recorded vocal signal),

---

because Antares [the star] is the heart of the scorpion [the constellation Scorpio] whose tail points directly to the intersection of the golden gate and the silver gate, galactic center, the heart of the galaxy. Before being called Antares, the company operated under the synchromystically [sic] potent name 'Jupiter Systems.' Being another celestial entrainment with love and the pursuit of the heart. The heart of the galaxy is a black hole, and thus Antares is pointing to a singularity, a breakdown of all known governing principles of experience." (BarelyHuman11 2012)

CA makes specifically legible vocal tuning's turn-based interactional structure. Though it is possible to perform digital tuning "in real time" over the course of a recording's playback, it is generally the case that the tuner will apply a correction of some sort, and then listen back in order to assess whether the correction worked. This can be difficult to notice due to the rapid tempo of exchange and correction performed by a skilled tuner. Representing it as a conversation between the engineer and the vocal track being tuned helps to emphasize the practice's dialogical dimensions.

A tuning session is depicted here as a three-way conversation between myself (an intern-ethnographer who began working at the studio several weeks prior), the engineer carrying out the tuning, and a pre-recorded track of a singer's vocal performance, to which the corrective tuning is being applied. As a literary-dramaturgical conceit, setting the scene in this way has two key effects: The first effect is that I appear as a sort of Greek chorus – a correspondent for the unfolding action within the scene and a proxy for a broader audience. Over the course of the scene I echo, amplify, clarify, obfuscate, laugh, and "hmmm" along with the tuner and the tuned. By periodically asking after and helping to track the rationales that go into the task of tuning, I can be seen collaboratively bringing-out and making collectively accountable the reasonings and heuristics of tuning that would normally be used but not noted aloud.[44]

The second effect is to cast the tuner and the tuned in something resembling a ventriloquistic relationship, turning the vocal track into the tuner's (occasionally stubborn) interlocutor, even as the reader understands that the pre-recorded but mid-tuning vocal track is in a process of becoming, by way of the application of the tuning software, a second voice for the

[44] Harry Collins (Collins 2010) would categorize this normally unexplicated, though practically explicable skill, as "relational" tacit knowledge.

tuning engineer. The tuner-tuned interaction appears as an entanglement of two previously distinct vocalic agencies so as to produce the effect of a "corrected" voice. This emergent voice reflects the combined labor of the engineer and singer while also working to erase the retraceable effects of that labor. In short, this should not be mistaken for a case of an agential subject imposing its will on a passive object: the pre-recorded voice presents a kind of hybrid or second order agency by virtue of its entanglement with an actual client to which the tuning intervention will be held accountable. At the same time, the pre-recorded voice offers certain material resistances by virtue of its partial connection to an actual vocal performance. At times the tuner struggles with the vocal track and the tuning software, a technological complex that asserts and remakes itself in unpredictable ways. In these moments tuning becomes a "dance of agency" (Pickering 1993a) or a play of resistances and accommodations between multiple material-semiotic voices.[45] The following analysis concerns 105 lines of transcribed conversation excerpted from a tuning session conducted by Carl. In this particular session, he is working partly with vocal takes he recorded with the artist the evening before, and partly with rough vocal tracks recorded by another engineer at a studio across town. This is a relatively common arrangement within the music production community.

Inscription

Tuning, like much of the engineer's work, is largely a practice of eliciting and assembling inscriptions. "Inscription," here refers to a way in which an artifact embodies a particular mode

[45] This observation need not conflict with the ventriloquism analogy. As any ventriloquist will tell you (If I may ventriloquize ventriloquists) it is not a matter of simply making the dummy talk, but rather learning to talk *with* the dummy.

of use, or the broader production of enduring traces indexically or symbolically related to some temporally emergent activity. One of the most important of these modes of inscription in the context of vocal tuning is the tracing of the pitch of the recorded voice over time, as represented in the software interface. Most digital tuning plugins, like most digital audio workstation (DAW) programs within which digital recording work is done, share a common visual paradigm. Time is typically represented as running left to right along the X axis while pitch and amplitude are represented vertically, with higher notes corresponding to a larger value on the Y axis and larger objects representing greater energy. This model is particularly well represented in the "graphical mode" interface for Auto-Tune 8, where the Y dimension is overlaid with the a piano keyboard:

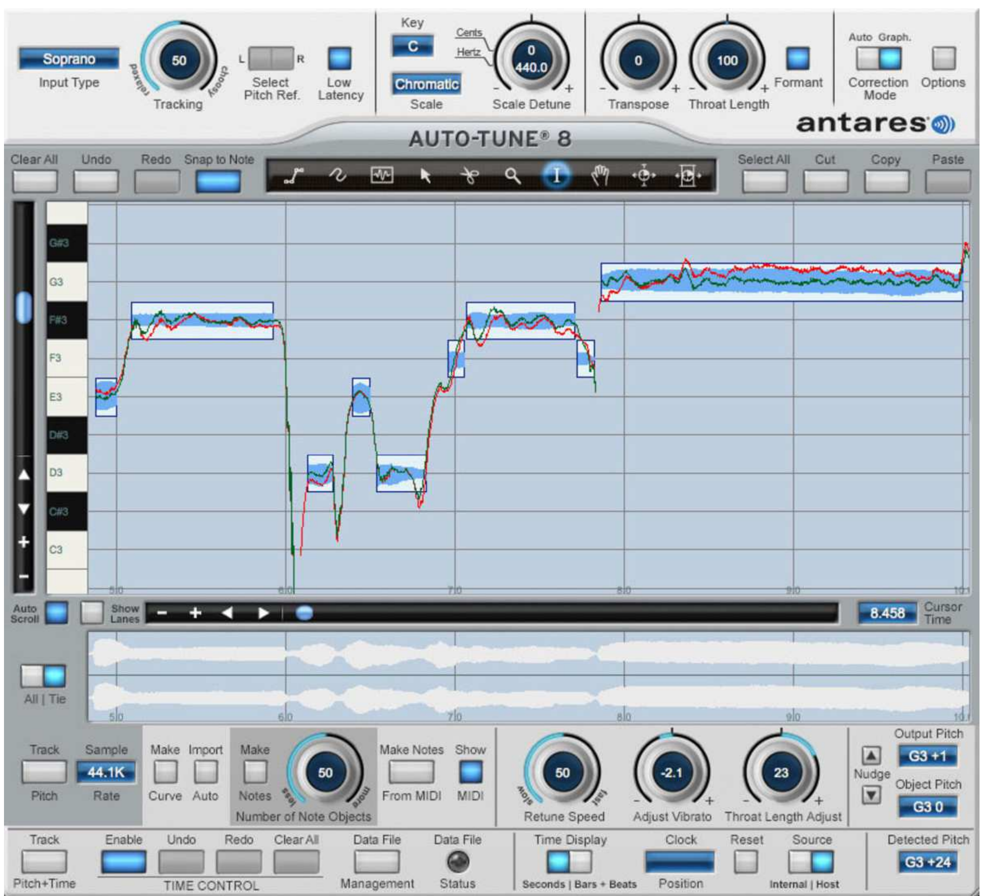


Figure 6: Auto-Tune 8 Screenshot (Walden 2015)

This model has its roots previous technologies of musical representation and reproduction, including the aforementioned keyboard, western musical notation, cylinder and magnetic tape recording, and player piano rolls. It also bears a family resemblance to a wide range of non-musical technologies of inscription, such as seismography, polygraphy, signal processing, speech pathology, and market forecasting. The role of this two dimensional grid representation in digital tuning is twofold. First, it produces visual representations of the pitch and timing of the musical performance subject to tuning. It also, crucially, offers a uniform basis of comparison in the form of a pre-selected musical scale. Even within this paradigm, there are numerous interpretive contingencies built into these structures of visual representation and comparison, which manifest in the various "technological styles" (Thomas P. Hughes 1983) of tuning tools available. Each tuning tool offers its own distinct pre-scription (Akrich 1992), or way of construing the task to which it is being applied, while also participating in broader shared contexts of meaning (standards of notation and signal representation, names of component functions, etc.)

Vocalists typically record in a separate "isolation booth," while listening back over headphones to pre-recorded instrumental tracks. Recording sessions are temporally structured around repeated "takes" or recorded performances of a piece of music. The "take" as a unit of the recording process can be traced to the practice of filmmaking, where an actor may do several takes of a scene, from which the director and the editor may later select and compile the best ones. The vibrations of the voice in the air move the diaphragms of microphones, the movements of which are converted into patterns of electrical energy. These patterns, in the case of computer-based recording sessions, are sampled periodically and translated into digital signals, which are saved within the computer's memory.

As the take is performed, the engineer listens and takes notes. Engineers frequently make use of a "comp sheet" – a piece of paper on which the song has been mapped out in terms of its various lyrical and structural components, with a number of columns available to note the salient qualities of each component take. It is called a comp sheet because it will ultimately be used for "comping" (short for "compiling") a final vocal track from the best pieces of each take. A basic comp sheet might take the following form (annotations vary, but a common shorthand would resemble something like: VG = Very Good, x = skip it, OK = okay):

SONG TITLE: *"Everything Around You"*

LYRIC:	***Take 1 :***	***Take 2 :***	***Take 3:***
(Verse 1)			
everything around you	*OK*	*VG but flat*	*x*
is going crazy	*x*	*x*	*excellent*
(Pre-chorus 1)			
everything around you	*VG(timing?)*	*x*	*good*
ooooooh yeah	*x*	*OK*	*quiet*
alright	*x*	*x*	*OK*

These sheets are generally prepared for the benefit of the engineer, to save them having to re-listen and evaluate takes without the immediate context of the performance to serve as a basis of comparison. Singers are rarely shown their own comp sheet. It is often expected that the best performances will come from the first few takes, though it may take more than that to come up with a completely technically polished performance. The first take is, so to speak, the deepest. There is a general concern with "freshness" – both of the musician with respect to the particular part being recorded, as well as of the engineer having "fresh ears" from not having listened to the part repeatedly already. Ears become less fresh as the takes pile on. Part of this process is simple bodily exhaustion, as singing listening both involve sustained physical exertion. Freshness of attention is also expended – a phrase or melody can easily lose its ability to surprise, move, and convey meaning over several repetitions. It can also hypnotize and inhibit critical listening as it takes on the quality of a mantra.

Some singers will have a lyric sheet in the booth with them, and they are often supplied with a music stand and a pen so they can annotate it as they go along. The idea behind this practice is that it allows the singer to focus on expression, rather than memorizing and recollection of the lyrics. Occasionally, however, the lyric sheet can be considered disruptive. In one vocal session, for example, Carl suggests that the singer should avoid looking at the lyric sheet while doing the take – "it sounds like you're reading" he explains. Voices bear the marks of their production in subtle ways; consider, for example, how a speaker's tone changes when shifting from speaking extemporaneously to directly quoting someone else. Carl's complaint in this case is that the voice coming through the microphone sounds less like it is coming from the singer and more like it is coming from the ink and paper resting on the music stand.

Carl, meanwhile, makes use of the comp sheet precisely because it allows him to distinguish his impressions of the vocal take as it occurs from his overall subjective experience. His practice of note taking is a way of delimiting his interpretive contribution to the recording. Just as the singer has to be careful not to seem to be singing on behalf of their lyric sheet, so does the engineer have to ground interpretive work through note-taking as one of many ways of performing objective accountability. Distinctions between expressive and technical labor get played out in how the recording gets written-out as it is being recorded and processed. The use of file naming conventions and track annotations – often entered both within the track-specific data fields of the DAW (in this case, Pro-Tools) and on a piece of notebook paper - is meant to map out the session file and make it legible both to the engineer currently working on it as well as anyone else who may need to work with it in the future. While there is only one singer of this song, there are textual conventions in place to allow for multiple engineers. Carl, for example, has appended "2n" (1) to one of the vocal tracks, which he explains indicates that it is a tuned track.

1 Carl: Two N. Which means it's tuned. ((*clack*))[46]
2 (Instead [of)
3 Owen: [mmh:mm

Following the invitation of the keystroke,[47] the pre-recorded vocal track makes its appearance. This particular take is an overdub, or double of the lead vocal line, but at the end of the second line it veers slightly from the main melody. The exact way it veers is, as yet, difficult to articulate. To my ears it is unclear whether the problem is the pitch of the notes or something closer to the way the singer changes the shape of his mouth over the course of the "youuuuuu" refrain. Before the actual repair process can begin, this portion of the vocal track needs to be made accountably *repairable*. Troubles only appear as repairables in terms of their relation to a broader situation. Carl explains that, while it is not necessarily out of tune with respect to the scale of the song, its position as a double of the main vocal necessitates correction because, in its current state, it would sound like it's "rubbing" (11) against the main vocal:

4 S:[48] *Everything around youu↓uuu*
5 *Everything around↓youuuuuu↓uu* [*uuuuu↑*
6 C: [If they doubled it exactly like
7 this we could probably ↑leave ↓it and it'd be fine.
8 It's just that (.) with all the other vocals

[46] I employ Jeffersonian Conversation Analysis notation (Heritage & Atkinson 1984). Parentheticals denote pauses, with (.) denoting a just noticeable pause. Arrows indicate upward or downward inflection. Horizontally aligned brackets and indentation indicate simultaneity of statements on adjacent lines. Quietly spoken portions are indicated as in: °*example* °.
[47] "((clack))" is my onomatopoeia for an audible stroke of the computer keyboard. This is often the sound of the spacebar, because this is the hotkey for starting and stopping playback. Because of its central role in the editing process, this key is struck frequently and emphatically, usually serving to announce the return of the recorded voice, or punctuate its interruption.
[48] S, whose contributions are highlighted, refers to the Singer's *pre-recorded* voice being played back in the DAW.

9		there's no way
10	S:	[*Everything around ↑yoooooouuuuuuuuuu↓uuuu*
11	C:	[ºIt just sounds like it's rubbingº(.) really badly
12	O:	↑rubbing? (.)↓yeah
13		(8.6)
14		((*clack*))
15	S:	*Everything around ↓youuuuuuuuu↓uu*
16		(3.5)
17		((>*clack clack*< *clack*))

The metaphor of "rubbing" casts the line in terms of a frictional play of textures within the scene of the composition. This haptic image, combined with the evaluative statement that it is rubbing "really badly," serves to bring out the rubbing as a trouble source in the recording. At this point, from my perspective, the way in which it is rubbing is largely undecided as I am still working to hear in a way that allows Carl's statement to make sense. How, exactly, does one hear something as "rubbing really badly?" Is it a literal "rubbing" sound that one might associate from everyday experiences of pencil erasers, windshield wipers on dry glass, or ill-fitting shoes? Perhaps he has a more technically specific image in mind: when two strings are tuned to almost the same note but not quite, too close to one another without being the consonant, their interfering vibrations produce a "beat" frequency that sounds rough or bumpy to the ear. I know this sort of rubbing from the practice of tuning a guitar. I work to imagine the sound in order to hear it properly, but neither of these strategies of hearing rubbing seem as yet to fit the problem Carl is describing.

Carl pulls the offending vocal line up in the editing window of Melodyne. Whereas most digital tuning tools, including Auto-Tune, process tracks in real time while running in the background like other plugins, Melodyne (at least the version Carl is using) requires the user to feed the selected audio portion into the program. As the vocal track plays into Melodyne's grid, it forms a series reddish-orange "blobs" (the actual term used in the Melodyne manual) of varying height (pitch), size (loudness or overall spectral energy), and shape (timbre, texture, or color). As the note-blobs form across the screen, the rubbing metaphor immediately becomes more visually literal, with the offending line visibly hitching downward and forming a new blob where it deviates from the other backing vocals.

The rubbing now appears as the kind you might experience if you tried to push your way through a crowd of people moving in the opposite direction. Perceiving the take as falling out of formation and jostling with its neighbors, initially a task of visual imagination, now becomes one of shared *observation*. Melodyne turns the notes into blobs, which are able to rub against one another. Whereas Carl is habituated to Melodyne's particular way of blobbifying sound, and the way those blobs can be seen to interact with one another, I am still learning to hear in terms of these blobs before they are made. By now, however, we can both see the blob that has fallen out of place. Carl is further able to see that it should not be a difficult job to tune this line:

18	S:	*Everything around ↓youuuu↓uu↑uuuu↓uu*
19	C:	This one seems ((*clears throat*)) ºahumº (.)↓pretty easy
20	S:	*↓youuuuuu↑uu↓uu*
21		(5.1)

Having dealt with the inter-blob domain, the intra-blob now comes into focus. Each Melodyne blob consists of a reddish membrane, which traces the amplitude – or total energy - of the note at any given moment. The red skin lightens towards the center of the note, which is positioned vertically according to the blob's average pitch. A red line snakes through the blob, following the pitch drift of the note. If the singer hits the note dead-on and stays there without any vibrato (i.e., minute fluctuations in pitch which are not large enough to be considered new notes in their own right) then the line will be relatively straight and will be contained within the blob's membrane. Notes performed in a vibrato-heavy, or otherwise warbly manner, form serpentine lines that may escape the membrane entirely. They transgress the inter-intra blob boundary, like springs piercing the upholstery of an old sofa or bone breaching the skin in a compound fracture. See, for example, the final note of this four note phrase:

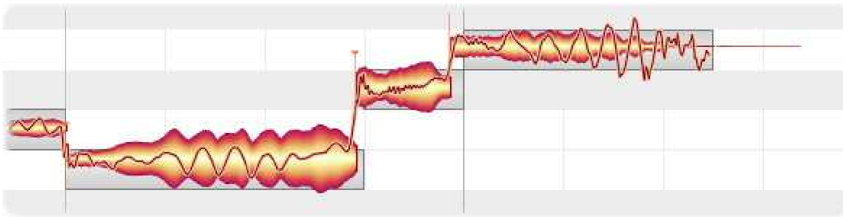


Figure 7: Melodyne "blobs" (Celemony 2005)

Carl proceeds by first "centering" the pitch of the blob, pulling in the drift of the inner-line and making it adhere more tightly to the blob's main frequency. He then shifts it down to match the main vocal melody:

C: Pitch center it (.2)
then change the pitch
S: *youuuuuuuu*

In-tuneness is a fundamentally relational quality. Things are only ever in-tune *with* other things: scales, other instruments, specific elements within a single instrument, rooms, keys, scales, temperaments, previous notes played within a phrase, the overall 'vibe' of a piece of music. Melodyne's interfae provides a way for the user to change the visualization and thereby transform the corresponding sound. Inscriptive strategies such as note-taking and computer-aided visualization are important for vocal tuning because they provide ways of making explicit or holding constant at least some of these tuning reference-points. The next section considers repetition as a different way in which in-tuneness can be established through the production and management of relationality within the tuning process.

Repetition

Whereas inscription is about making something temporal endure in space, or turning a practice into an artifact, repetition is about making something spatial endure in time, or turning an artifact into a practice. In the context of tuning, these modes of practice take on a relationship of exchange. A comp sheet employed during vocal takes allows the engineer to inscribe differences into a stream of repetition. The multiplicity of digital intonation correction tools would be entirely redundant if each one did not bring a particular inscription to the task of tuning. Techniques of inscription and repetition are ways of producing, assembling, and redistributing relations of similarity and difference. Over the course of the tuning session they serve as key methods for the interactional production of identifiable troubles and accountable interventions, what Lynch refers to as "turning up signs." (Lynch 1984)

Attending to repetition brings into focus the elements of the tuning practice that are often most easily taken for granted. My presence in the work thus far, for example, has largely proceeded by way of non-verbal affirmation (mmh:mm) or clarificatory echoing ([↑rubbing? (.)↓yeah). Though I am not adding much in the way of new verbal information to the actual content of the interaction, my presence is framing the broader interaction at least as much as the actual work being carried out. The repetitive, occasionally monotonous work, which constitutes this framing presence, is largely directed at the production and circulation of accountability. Peter Weeks, in a study of temporal coordination among members of an amateur string ensemble, uses the phrase "double accountability" to refer to how ensemble members are able to identify and repair troubles without their audience noticing (Weeks 2013, 216). In order to identify, contain, and repair troubles in the vocal track, it is necessary to produce, coordinate, and maintain multiple levels of accountability.

In the case of vocal tuning, I would replace Weeks' "double" with "multiple" (or "more than one and less than many") accountability (Mol 2002). As material-semiotic things, voices come into being already marbled with intersecting veins of error and intention. They occupy local "repair worlds" (Jackson et al. 2012) which practically configure their modes of use while conditioning their possibilities for re-use. The processes of recording and signal processing entailed in computer-based music production cause these meanings and repairables to proliferate. To a person not accustomed to hearing their own voice on record, the voice can even appear as, to borrow Homi Bhabha's phrase, "less than one and double."(Bhabha 1994) A recording of yourself is always already a potential source of trouble simply because it is the sound of your voice from somewhere which is usually thought of as 'outside.' For a novice to vocal recording, learning to hear your voice as your own is thus a social accomplishment. Turning up and tuning out a sign of vocal trouble, similarly, takes work of social coordination.

Accordingly, just as he uses inscriptions to make his work visible to himself and others that might encounter this session in the future, Carl is announcing his actions verbally in order to make them available to me. My echoes and affirmations are meant to indicate my continued attention and to encourage the continued explanation of his work. They fill in the spaces offered by Carl to check in with my understanding of what has happened thus far. If I failed to continue indicating my presence within the interaction, it would disrupt the process and force Carl to attend to the problem of my work of audiencing and apprenticeship. When performing listening together, keeping silent can in fact become disruptive. Carl, the singer's pre-recorded voice, the clack of the keyboard, and my gestures of attendance, form a rhythm of work over the course of the session. This rhythm is partially organized through a process of turn-taking, as the participants try to make way for and respond to one another.

Weeks observed that when members of a musical ensemble encounter moments of temporal disagreement, they often make reparative use of what he calls a "holding pattern." (Weeks 2013, 212) Holding patterns occur when someone does something unexpected – playing a wrong note or play the right note too early or late, for example – at which point other members of the ensemble proceed to alter their parts – usually by holding a note longer or shorter than usual or repeating a phrase. A holding pattern is a way of inviting a correction and making it possible to repair a trouble retroactively while keeping it from becoming accountable as such to the audience. Like an airplane making a second approach in order to compensate for a crowded runway, the players use techniques of repetition to collaboratively and improvisationally enact such situations as repairable.

Whereas the ensemble playing and holding patterns that Weeks describes take place in a sort of "all-at-once" temporality, tuning work is distributed over longer and more disjunctive timescales. Instead of the holding pattern, we find the repeated take and the use of looped playback as the primary modes of repetitive repair in vocal tuning. These two modes of repair have a sort of nonlocal entanglement in that they occur in well-differentiated spatio-temporal scenes, but nonetheless anticipate and refer to one another in order to produce the effect of a coherent voice unfolding in a single moment. Through the use of a comp sheet, vocal takes are gathered up, not like specimens from which an ideal single representative can be chosen, but more like salvaged parts that, with the right tools, can be made to work well together. The practice of getting vocal takes often appears as a mode of repetition specifically pre-conditioned by the possibility of digital repair after the fact.

As an improvisational salvaging activity, repetition-as-repair is also an opportunity for creativity. As a cliché in jazz improvisation puts it, "if you make a mistake, do it again!" An awkwardly played phrase, if it is replicated and incorporated into the continuing performance, can be imbued retrospectively with intention and creative agency. The playful repetition of an error works on several levels. To an audience savvy enough to notice the mistake in the first place, it allows the performer to acknowledge and retake control of their performance. To an audience for whom the mistake was not apparent as such, a repetition can make the error come off as "on purpose." The same phenomenon can be found in the world of improv comedy groups, where the guiding rule is to support one's co-performers through a strategy of unconditional acceptance – the proverbial "yes, and…" technique. When an improv performer misspeaks, a capable partner will seize on the strangeness of the statement and use it to guide the scene into a new direction. In the case of vocal tuning, the singer and the engineer may approach repair in a similarly creative way. Repairing an awkward take can just as plausibly involve copying and pasting it, as opposed to undoing the apparent transgression.

Post-hoc tuning work, meanwhile, is largely a treated as a practice of procedurally recovering a certain elusive feeling of spontaneity that is supposed to have been disassembled throughout the process of exploding the voice into so many individual takes. For example, a decision as to whether an event within the vocal track is "supposed to be there"(91) implies a back-formation of a coherent vocal *supposition* on the part of the vocalist. The question of whether there was really such a supposition in the first place is irrelevant, even as the question of what such a supposition might have been serves as a useful heuristic for the tuner. Tuning is usually directed only peripherally towards producing agreement between the movements of a vocal trace and the possibilities pre-determined by the pitch-time grid, because this agreement's adequacy as a criterion for "in-tuneness" is conditioned by, among other factors, an understanding of what the singer was "going for." (101-2) Where clarity regarding the in-the-moment intentions of the vocalist has been exhausted by the sheer proliferation of moments and intentions over the course of the vocal take process, the engineer may well resort to continuously scrutinizing and fine-tuning a single take. The looping function works like a lathe or a rock polisher. It is an ordeal through which the listener's ears perform a due diligence, wherein rough edges are removed, and a certain grain[49] and sheen of the voice are made to come through:

88	S:	*↑uu↓uu↑uu↓uu↑uu↓uu↓uuuuuuuuuuu (.) ↓uuuuu↑uuuuuuu*
89		*((click click))*
90		*↑uu↓uu↑uu↓uuu↑uu↓uu↓uuuuuuuuuuu (.) ↓uuuuu↑uuuuuuu*
91		*↑uu↓uu↑uu↓uu↑uu↓uu↓uuuuuuuuuuu (.) ↓uuuuu↑uuuuuuu*
92		*((click))*
93		*↑uu↓uu↑uu↓uu↑uu↓uu↓↓uuuuuuuuuuu (.)↓↓uuuuu↑uuuuuuu*
94		*((click click))*
95		*↑uu↓uu↑uu↓uu↑uu↓uu↓↓uuuuuuuuuuuuu (.) ↓uuuuu↑↑uuuuuuu*
96		*((click click))*
97		*[continues in this way for ~5 minutes]*
89		*↑uu↓uu↑uu↓uu↑uuuuu↑↑uuuuuuuuuu↓uuuuuu↓↓uuuuuu*
90		*((CLACK))*
91	C:	Sh(oo):↑t (3.0) °↓I'm not sure (it's) supposed to be (↑there)°
92		((*clack*))
93	S:	*↑uu↓uu↑uu↓uuu↑uu↓uu↓uuuuuuuuuuu (.) ↓uuuuu↑uuuuuuu*
94		*↑uu↓uu↑uu↓uu↑uu↓uu↓uuuuuuuuuuu (.) ↓uuuuu↑uuuuuuu*
95		*↑uu↓uu↑uu↓uu↑uuuuu↑↑uuuuuuuuuu↓uuuuuu↓↓uuuuuu*
96		((*clack*))
97		*↑uu↓uu↑uu↓uuu↑uu↓uu↓uuuuuuuuuuu (.) ↓uuuuu↑uuuuuuu*
98		*↑uu↓uu↑uu↓uu↑uu↓uu↓uuuuuuuuuuu (.) ↓uuuuu↑uuuuuuu*
99		*↑uu↓uu↑uu↓uu↑uuuuu↑↑uuuuuuuuuu↓uuuuuu↓↓uuuuuu*

[49] Roland Barthes' concept of "the grain of the voice," is often deployed in Sound Studies literature as a way of describing a particularly "rough" or "distinctive" vocal quality. This reading, while not incorrect, is incomplete in that it fails to capture the term's use for critiquing (or at least socially and historically situating) a "transmission" model of the voice, wherein coded emotion is conveyed as a message. As Barthes writes, "The 'grain' of the voice is not - or is not merely - its timbre; the *significance* it opens cannot better be defined, indeed, than by the very friction between the music and something else, which something else is the particular language (and nowise the message)." (Barthes 1977, 185) I use "grain" here precisely in this frictional sense, as something that is produced where music meets language (which I construe broadly to include the grammar of the tuning software in-use.) It is fair to assume that Barthes would likely balk at the idea that the "grain" as he heard it is ever able to survive the tuning process.

100 ((*clack*))
[continues in this way for ~2 minutes]
101 O: ss:were you not sure what he was ↑going
102 [for? (.)↓err
103 S: [↑*uu*↓*uu*↑*uu*↓*uuu*↑*uu*↓*uu*↓*uuuuuuuuuuu (.)* ↓*uuuuu*↑*uuuuuuu*
104 (1.6)
105 C: Yeah

In everyday conversation, repetition is frequently deployed for general repair purposes, and correction in particular (Heritage and Atkinson 1984, 41).[50] The same could be said of everyday repair of technical objects (e.g., rebooting a computer, cranking the ignition of a stubborn car engine.) The case of corrective tuning, by addressing a voice as both an object of technical intervention and as a source of subjective expression belonging to another person, straddles the worlds of conversational and technical repair. The invitation to repeat is simultaneously a way of turning up a sign of trouble and a way of creating a space within which that trouble might be managed. In this way, the repetitive techniques of vocal correction are also ways of *testing* a vocal track. Pinch has used the similar example of the microphone check to

[50] Work in CA has found that repetition serves a wide range of functions beyond repair, and that repair involves a wide range of techniques not limited to repetition. Schegloff, for example, has identified numerous distinct operations within the domain of self-initiated same turn repair, including replacing, inserting, deleting, searching, parenthesizing, aborting, sequence-jumping, recycling, reformatting, and reordering. (Schegloff 2013) The preference in CA for self-correction (as opposed to "other-correction") is also well documented. (Schegloff et al. 1977) Limiting ourselves to repetition as repair, however, we can identify certain common patterns. Echoing a problematic word or phrase, for example, is one common device for initiating repair:

A: my cousin's seven inches tall
B: inches?

It can also be used for self-correction:

A: my cousin's seven inches tall.
inches. feet!

Or as an acknowledgement of acceptance of other-correction:

A: my cousin's seven inches tall
B: seven feet?
A: seven feet

illustrate the basic gesture of "projection" involved in technological testing: a stagehand repeatedly checks the mic in order to project that mic's reliability once it is in the singer's hands (Pinch 1993). The practices of repetition employed during tuning is similarly projective: the engineer repeatedly tests the vocal performance for problems and subjects them to repair as a way of ensuring that the voice will "work" across multiple contexts of audition.

Inflection Part 1: A decision to not use Auto-Tune

Within this process of contingent repetition and rhythmic coordination Carl and I are improvising a sort of prosodic shorthand wherein our conversational inflections come to move with, make room for, and respond to the musical inflections of the singer. Again, the infrastructural invisibility of tuning work in-practice is apparent in the way it provisionally straddles domains of meaning. Though conversational prosody and sung musical pitch typically form distinct economies of sense, the specific context of showing someone how to tune a vocal recording produces a sort of "trading zone" between otherwise incompatible structures of meaning (Galison 1997). In this space of ambiguous prosodic coding, the affirmations, questionings, tentativenesses, pressings-on, and focusings-in of everyday speech become partially commensurate with the meanderings of the singer's voice as it moves along the terrain of musical key and melodic motif. They tentatively assemble makeshift economies of feeling, which allow the shared tasks of tuning, teaching, and learning to proceed.

Even when I am not looking over his shoulder, I occasionally overhear Carl talking with, or otherwise audibly responding to the pro-tools session he is working on. His interactions with the computer and the voices it summons forth are usually non-verbal or interjectional. He laughs at, curses out, makes frustrated noises with and otherwise invites action from the notoriously fickle Pro-Tools interface. Most of the prosodic exchanges taking place between Carl and myself have a basis in our habitual everyday interactions. These, of course, are not universal rules for how prosody might convey meaning in human speech. They are, rather, polyvocal ways of producing shared sense, and are inseparable from Carl and my personal histories of interaction. Learning to work with others in the studio means learning how they speak and how to speak with them. With Carl, for example, I quickly learn that when something needs to be done, or when something has not been done correctly, he habitually raises the topic with a quick rising and cheery "oh yeah!" which will generally be followed by an instruction. (e.g., "Oh yeah!... be sure to get receipts next time" or "Oh yeah!... how's the mic book coming along?) This prosodic habit, which has the effect of making Carl sound like even this thing that has been getting on his nerves has in fact just come out of nowhere and caught him by surprise, goes a long way towards maintaining a friendly informal atmosphere even when something is clearly at issue.

Some of our habits are more on the generic side: An upward inflection, connoting a question or inviting a response, is resolved by a downward inflection ([↑rubbing? (.)↓yeah) as a way of indicating that I am following along, that my understanding may have been snagged upon a newly deployed term ("rubbing"), but as the singer continues and I begin to hear what Carl means, I resolve my invitation to explanation with a downward inflection ((↓yeah)).[51] The barrier between our conversational and musical ways of listening and speaking is decidedly permeable, however. As the refrain's familiar melody rises, it seems to require once again resolution of a falling tone. In response to Carl's "centering" of the note, I respond by offering an possibly synonym, i.e. correcting the "drift" of the note from its center. My flat inflection mirrors

[51] Another way of interpreting this would be to say that I am performing an abortive other-initiated repair via repetition of a trouble source.

that of Carl's matter-of-fact procedural accounting (then change the pitch) as does the newly tamed, pitch-centered, and drift-corrected note:

25 O: The [drift=okay
26 S: [*Everything around ↓youuuuuuuuuuooo*

Having fixed that particular note to his satisfaction, we listen back to the complete phrase. Carl frames the playback by suggesting how he might have used Auto-Tune instead of Melodyne, possibly with a "weird"-sounding result. I imagine what the weirder possibility would have sounded like as I listen:

27 C: See with Auto-Tune on [that might sound weird ((*coughs*))
28 S: [*Everything around ↓youuuuuuuuuu*
29 *Everything around ↑youuuuuuuuuu↓uuuuu*
30 C: (clack)
31 S: *Everything around* ↓youuuuuuuu
32 *Everything around* ↓youuuuuuuu
33 *Everything around ↑youuuuuuuu*

We listen together to the newly tuned line, holding it between two shared imaginary moments: 1) the way the line sounded and felt prior to tuning and 2) the prospect that a different method (Auto-Tune, in this case) might have produced an outcome that may not have sounded right. We move on to the next line, which should be more or less identical to the previous one, but presents new difficulties:

34 S: *Everything around* ↓youuuuuuuu
35 *Everything around* ↓youuuuuuuuu[uuuu
36 C: [°↑*oooo*↓*ooo*°

Carl sings along, softly enough that I imagine it cannot be for my benefit. I realize that he is providing himself with a reference to how the line is supposed to go. He lets the line run and, again, compares his own sung version with that of the track:

37 S: *Everything around* ↓youuuuuuuuuuuuuu [uuuu
38 C: [°↑*oooo*↓*ooo*°

39 S: *Everything around ↑youuuuuuuuuuuuu↓uuuuu*

Instead of my usual hmm, I respond to the new difficulty with a "huhh", which is quickly answered by a severely off-pitch beginning to the next line:

40 O: huhh
41 S: *Everyth-*
42 C: Oops(.) .hh ha
43 S: *[[being sharply retuned]] ↑u↓uu↑↑uu*
44 (1.2)
45 *-thing around ↑youuuuuuuuu↓uuuu↓uuuuu-*
46 *((clack))*
47 C: º↑*uuu↓uuu↓uuu* º ((*clack*))
48 S: *-thing around ↑youuuuuuu↓u↓uuuuu↓uuuuu-*
49 C: º↑*uuu↓uuu↓uuu* º ((*clack*))
50 S: *-thing around ↑youuuuuuuuuuuuu↓uuuuu-*
51 C: º↑*uuu↓uuuu↓uuu* º
52 S: *↓uu↑uuuu*
53 *↓u↓uuu↑uu*

Carl hunts for the right note, but it seems to be stuck between two quantized options. As he searches for a note that works I keep up the rhythm, acknowledging the trouble with a laugh (55). Carl's intonation matches the uncertain searching of a vocal track in mid-tuning, rising and falling (56). I echo his inflection (57):

54 S: *uu↑uu (.5) ↓↓uu↑uu*
55 O: haha.
56 C: ºwelpº ↑any↓wa:ys (that's wha-)
57 O: haa. ↑some↓where betwe:en
58 C: uhn:ha↑
59 S: *↑u↓uuu- ((clack)) -thing around ↑youuuuuuu↓uuuuuu↓↓uuu*
60 *everything around ↓youuuuuuu↓↓uuuuuu↑↑uuu↓uu*

Though Carl still seems dissatisfied with the exact note he had to settle on, he decides to play it back in context of the lead vocal it is accompanying. Carl's voice becomes quiet with anticipation, and I clear my throat, as we wait for the main vocal to respond.

61 C: ºlet's=see if that passesº
62 O: ((*clears throat))*

63 C: ((*clack*)) ºt's uhh, this oneº ((*clack*))

[Singer comes in double tracked with backing vocals:]

64 S: *everything around ↓youuuuuuu↓↓uuuuuu↑↑uuu↓uu*
65 *everything around ↑youuuuuuu↓uuuuo↓ooo*
66 *everything around ↑youuuuuuu↓uuuuo↓ooo↑oo↓uu*

Apparently convinced that the tracks work well enough together that we can move on, Carl expresses his satisfaction, again in terms of the need to use Auto-Tune instead of Melodyne. In this way the tools at his disposal provide a key context for decisions about the adequacy of the intervention. I try to clarify what Auto-Tune has to do with it, if it's not being used on the track:

67 C: ↑Oka:↓y=↑I'm ↓okay with=that (.) No Auto-Tune this time.
68 ºbus (.) toº (1.5)
69 O: Does >Auto-Tune not let you do <the ↑pitch drift> thing? (.) Or.
[>'zit justa nother way of doing (.) things<
70 C: [Yeah. It might <u>fight</u> you (.) sometimes.
71 If it's a long note it might go (.) up and [down
72 O: [uhuhh
73 C: ºOr um (sometimes maybe (.) sound too robot)º
74 O: mmhmm
75 S: *everything around ↓youuuuuuu↓↓uuuuuu↑↑uuu↓uu*
76 *everything around ↑youuuuuuu↓uuuuo↓ooo*
77 *everything around ↑youuuuuuu↓uuuuo↓ooo↑oo↓uu*

The decision to use Melodyne instead of Auto-Tune here has to do with a concern that, given the length of the notes in need of tuning, Auto-Tune might "fight you" by going up and down or "sounding too robot." The decision is made in terms of an anticipated response to a projected intervention. The risk of a fight or a robotic sound are located inflectionally within the sustained notes of the vocal performance and Carl's own orientation towards how they should sound. Later in the piece, however, Auto-Tune becomes appropriate as the singer moves to a falsetto register and Carl's orientation moves towards one that is open to "weirdness."

Inflection Part 2: A Decision to Use Auto-Tune:

Later on, while working on the same song, we come to a falsetto[52] backing vocal for which Carl opts for Auto-Tune instead of Melodyne. He explains that, given the quality of the vocal and the type of tuning needed, Auto-Tune and its previously invoked potential for weirdness may be justified. He does this playfully, with an exaggerated command that we agree to say the weirdness is intentional:

78		*((clack))*
79	C:	Alright, well, Auto-Tune might help with that.
80	O:	Mmhmm
81	C:	And they're high vocals so they can be a little weird sounding.
82		(0.5)
83	O:	Mmhmm
84		(2.3)
85	C:	*[[with raised pitch and rough timbre]]* >jus=say it's on ↑purpose!<
86	O:	((*laughs quietly through nose:)* hfff hfff hff
87	S:	*uuuuuu↓uuuuuu↑↑uuuuuuu↓uuuuu*

If I transcribe part of the above passage as a musical score, certain aspects of the interaction become apparent. It becomes possible, for instance, to see that each of the first two exchanges find Carl inflecting downward and resting on the same note (approximately an F#2 on a piano keyboard) to which I respond with an "Mmhmm" on the same, or nearly the same note. Carl's third turn begins on his low resting note and quickly jumps up by the musical interval known as the perfect fifth, deviating both from the gradualness and downward sweep of his first two turns.

[52] Falsetto is a vocal "register" usually attributed to male singers venturing above their usual vocal pitch range. Prince, for example, is well known for his falsetto singing voice.

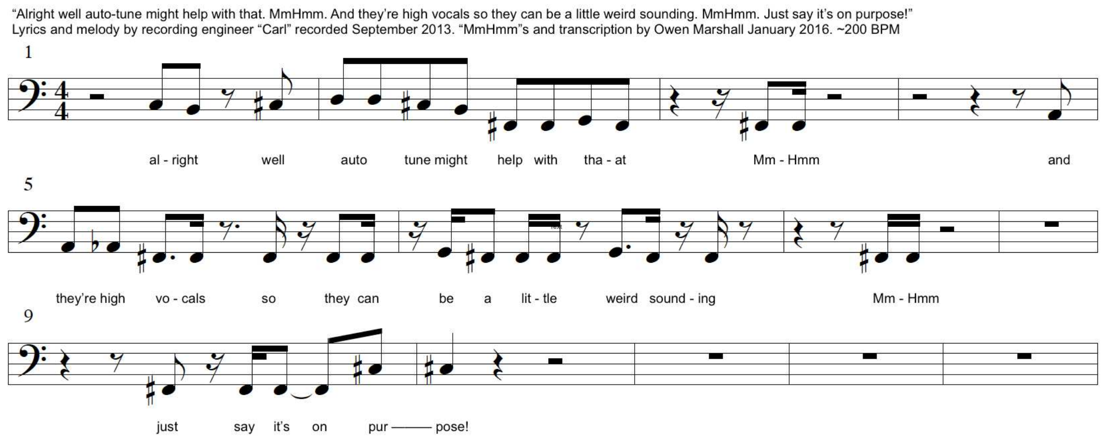
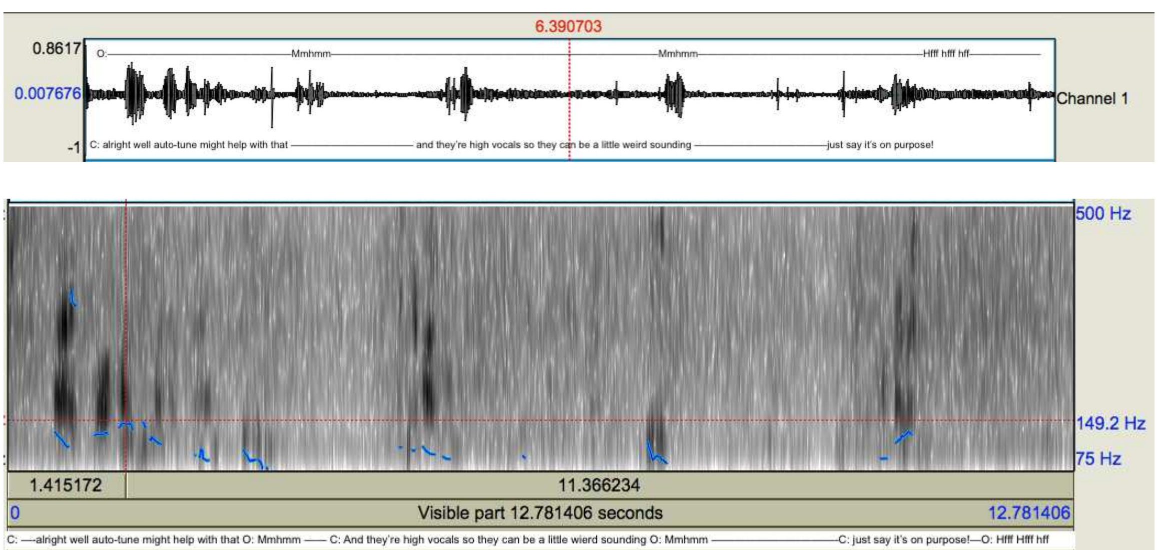


Figure 8: Staff notation of "… just say it's on purpose!" by author.

Spectrographic and pitch tracking analyses of this passage make visible certain aspects of this passage not included in the above transcription. They also, crucially, obscure elements of the passage, specifically the interactional rhythm with which each passage sits within the broader exchange. Different aspects are captured by the amplitude, spectrographic and periodic pitch representations of the above exchange over time:[53]



---

53 These visualizations were produced with the free phonetic analysis program Praat: (Boersma & Weenink 2013)

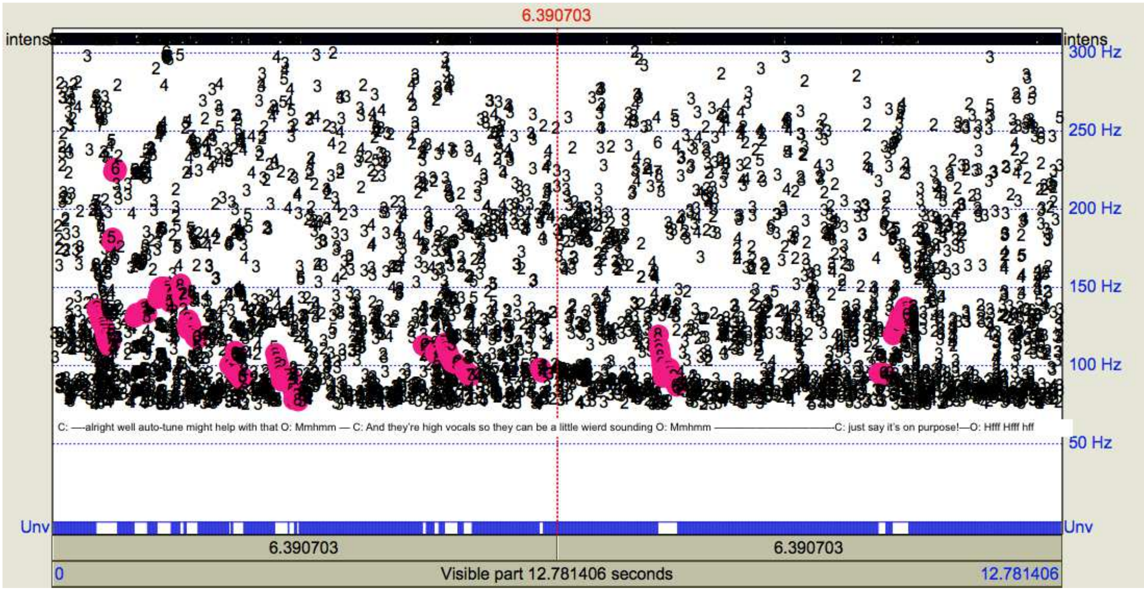


Figure 9: Spectrogram and autocorrelation of "…just say it's on purpose!"

Just as the raised falsetto pitch of the singer is departing from the typical intonational range of the rest of the song, seemingly indexing a certain shift in character and justifying a tuning technique that may be considered "weird sounding," Carl is raising his own prosodic pitch and applying a gravelly texture in order to enact the playfulness of his justification for using a weird sound – "just say it's on purpose!" From our everyday interactions I know this is the tone Carl uses when he's half-joking. It is the same tone he uses on talkback when faux-commanding the vocalist to do another take with a terse "again!" With the vocalist absent, and his audience consisting primarily of myself, himself, and the imagined future auditors of this vocal track, the implication is that the use of Auto-Tune in this case could be taken either as an instance of inexpertly attempted "invisible" tuning technique or, alternatively, as a deliberate creative choice.

In the context of his previous (and soon to be repeated) experience of uncertainty as to where and how the vocal needs to be tuned, the joke serves as an acknowledgement of that frustration and the possibility that a bit of “weird” tuning needs to be declared and agreed-upon as intentional in order to be clearly accountable as a deliberate decision. The injunction to “just say it’s on purpose!” serves a dual role, drawing attention to an ambiguity in the tuning job while also playfully distancing us from some imagined strict criteria or audit of the tuning job down the line. At this point, we might take Carl’s previous statement that high vocals are allowed to be a little weird sounding, as playfully implying a strict set of rules. My response of “mmhmm” (83) does not indicate a recognition of the play in this gesture, but my laugh two lines later serves to acknowledge the playfulness of the sentiment.

Conclusion: On Producing an Accountably Unaccountable Transcription

This chapter is not a transparent representation but a situated and discretionary interpretation of my own experience learning the craft of recording engineering. As such, it is connected to the “actual practices” of recording engineers less in terms of accuracy of reference than in terms of its accountability to various audiences. My accountability includes my readers and the system of scholarly production within which am working. This includes the terms of data collection and disclosure outlined in my Institutional Research Board approval: I have permission, for example, to use the material collected from my interactions with engineers, I can reproduce their words and details of their practices with the provision that I anonymize them. I do not have this kind of permission from their clients whose performances, and in particular their voices, are the object of the engineer’s everyday labor. Securing this sort of permission would be unlikely, and even the act of requesting this permission could jeopardize the reputation of the engineers with whom I worked.

As a pretext for institutionally accountable human subjects research, this distribution of ownership of voice, and conditions for the collection and textual representation of voice, should not (at least in theory) be a problem since I am interested in the work of the engineers rather than that of the musicians whose performances they capture and manipulate. The conceit comes into immediate tension with the fact that the two parties, by virtue of the interactions that serve as the key data for this project, are deeply involved in the production of one another as socially accountable actors. The distinction between technical and creative action is at its blurriest in the recording studio, precisely because this is the place where the work of generating that distinction gets carried out behind the scenes. It is not that emotional-creative and objective-technical roles are dispensed with in-studio (in fact, it is where they are dispensed) but that it is the site where they are rigorously and contingently negotiated. The exchange entailed by this negotiation gets projected back to an imagined pre-existing role as musician or producer, even as these roles are being actively constructed.

The practical methodological question becomes how to go about removing (or selectively preserving) these identifying marks, and thus eliminating the risk that their performance will be identifiable in its potentially embarrassing moments of becoming, without losing the aspects of the tuning practice that I am interested in elucidating. In working through this practical problem of representational ethics, I am also necessarily working through the problem of what makes the vocal performance accountably that of the singer vs. that of the tuner (meant here to refer to both the tuning engineer and the tools they are using). Producing an account that reveals the hidden work of the engineer and obscures the identity of the performer is similar to describing a mold-making and casting process without revealing the final statue. Better still, it resembles the exercise of subtracting an MP3 waveform from a lossless version of the same recording and revealing the ‘ghost image’ of what was compressed.(Sterne 2012) It is the production of something that is, for my purposes, ‘accountably unaccountable’ (or perhaps, to use CIA-speak, ‘plausibly deniable.’) This tells us something about what matters to the members of the studio ensemble, because it reveals what I need to take into account as someone taking part in and making myself accountable to this ensemble. Displaying the tools of obfuscation can, to some extent, remove the need for their specifically accountable application. Indeed, to produce a specific account of their application would open the door for the recovery of the things they were meant to obscure. As such, I will try to make transparent my reasoning process with regard to how to produce an accountably unaccountable transcription of tuning work.

What can stay in, what needs to be changed, and what needs to be removed altogether? Some things are more clear than others: Verbatim lyrics, for example, are out of the question, as a quick online search could quickly reveal the artist if their lyrics are ever published online. This necessitates a means of anonymizing the lyrics, preferably without redacting them completely. The task resembles something between translation and encryption wherein neither the codes nor the meanings are changed. With enough investigation beyond the material presented here, one could theoretically unearth (re-earth?) my subjects, so it is my task to bury them in a way that is systematic but still plausibly representative. Synonymizing or paraphrasing are close to what is required, though these methods can take many forms: do I try to capture the sentiment of the song even if it requires substituting one metaphor for another (e.g., “I’m feeling blue” becomes “I’m down in the dumps”)? Alternatively, do I try to translate a figurative passage using synonyms (or homophones, or puns) even if the metaphorical meaning is lost (e.g., “I’m feeling blue” becomes “I’m touching teal”)? Is it acceptable, necessary, or even worthwhile to change pronouns? (“I’m feeling blue” becomes “She’s feeling blue”)? Tenses? Pluralization? Spatial-metaphorical reasoning? ("They're under my spell" becomes "You're caught in its web.") How important is it to maintain syllabic patterns or rhythmic phrasing vs., for example, rhyme scheme?

I decide to negotiate between these parameters as needed, with the presumption that discretion of this sort will both aid in anonymizing the performance and in conveying the relative importance of literal, figurative, or rhythmic meaning at any given point. What about the musical notes? A complete transcription could also be problematic because it would permit re-performance of the melody or rhythm and possible recovery of the singer's identity. Two options I consider are to either transcribe the melody and then transform it (perhaps by some uniform but undisclosed set of rules, e.g. shifting the key by a whole step, nudging each note up or down a half step according to a series of coin flips, inverting the melody so that rising phrases fall and falling phrases rise, etc.) or, instead, indicate roughly the up and down movements of the pitch using the Conversation Analysis convention of arrows corresponding to shifts in inflection.

This project of code-switching serves both to obscure individually identifying details while also highlighting a key aspect of the tuning process – namely the blurring between melodic and prosodic modes of expression, including but not limited to "emotional." One appeal of the latter option is that it makes the movement of the singer's voice more commensurable with my voice and the voice of the engineer. Another appeal is that it captures the gist of the melodic movement without providing enough information to recover it in to any identifiable extent. I have already revealed that the engineers in this case are located in Los Angeles, but this is not a problem because the artists they work with are by no means necessarily based (or even recorded) there. To reiterate, decisions about anonymization entail decisions about attribution, identification, creativity, and origination. This methodological question occasions a disassembly of a field recording.

It could be further argued that my presence as an auditor requiring conversational accountability during the tuning process is so unusual and artificial as to limit my claims to be investigating "how voices get tuned" in the first place. The artificiality of my intervention is indeed unavoidable, in that Carl will usually not have anyone looking over his shoulder or asking questions during the tuning process. In this sense, my intervention takes the form of a breaching experiment – my involvement in the act of tuning has the effect of disrupting the performance of habitual actions that may otherwise get taken for granted, and the use of implicit rationales and heuristics that may otherwise go unexamined and undeclared. At the same time, however, this is not entirely a disruptive method because interns do (if only occasionally) get to benefit from demonstrative instruction of this sort during the regular course of studio life. The translational anxieties and practical compromises of this particular ethnographic endeavor provide, however obliquely, a sense of the tension between the performer and the engineer. The engineer tasked with vocal tuning is faced with quite a predicament: changing a vocal performance so as to more adequately present the voice of the performer, striving to perform a certain due diligence with respect to the track without taking liberties. Engineers, like ethnographers, must learn to skirt the hazards of ventriloquism and embarrassment while dealing with the inevitable excesses and inadequacies of someone else's voice.

CHAPTER 6: THE TOPICAL ORGANIZATION OF STUDIO LIFE

Zoom-in far enough on an audio track in Pro-Tools and the squiggles of a waveform will eventually give way to a series of dots representing discrete sample points strung together by imaginary lines. When working in high-end recording, you can expect to find 96,000 of these dots per second. Novice engineers can sometimes get sucked-into this microworld of digital audio, where the sturdy familiarity of a three minute pop song dissolves into 17,000,000 tiny pieces, each of which can be adjusted with the click of a mouse. Good engineers understand the dot world, and know how to move around in it when necessary, but most of the time the work of recording engineering requires zooming-out. Where the previous chapter zoomed-in to the particularities of a recorded voice, this chapter and the next take wider views of studio life. This chapter argues that the recording studio is performed through various acts of place-making. It focuses specifically on practices of listing, mapping, and staging, whereby the studio-as-assemblage is made accountable, and work within that assemblage is made legible.

Placing Voices

The pitch and timing of the voice, and the way that these two qualities concatenate in the domains of rhythm and timbre, are the materials with which studio workers assemble compelling performances. They are also terms drawn from and understood largely in terms of knowledge developed in the fields of psychoacoustics and signal processing. Figures such as Susan Rogers, who migrated from recording engineering to a career in experimental psychoacoustics and music cognition, have highlighted how deeply rooted much of the engineer’s practical knowledge is in experimentally derived understandings of how people hear and experience voices.(Rogers 2013) Figures such as Ethan Winer, a crusader against engineers and equipment manufacturers who claim to operate beyond the bounds of scientifically demonstrated perceptual capacity, have worked to anchor the field of legitimate engineering practices to the realm of scientific knowledge.(Winer 2012)

At the same time, however, the understanding of such basic psychoacoustic phenomena as pitch and timbre are remarkably unsettled. The phenomenon of pitch perception in particular has long been split between so-called “place” and “time” theories. Place theories of pitch perception, also called “tonotopic” theories, conceive of pitch as a local excitation on the body of the listener – typically on the basilar membrane of the ear. Temporal theories of pitch argue instead that pitch perception consists in the discernment of the time interval between events. Though place and time theories imply markedly different perceptual mechanisms, neither is considered sufficient by itself as an explanation for pitch as a perceptual phenomenon. In addition to the partial and contested role of place in the mechanistic understanding of pitch perception, the way that people interpret pitch differences spatially is also interestingly multiple. As acoustic neuroscientist Jan Schnupp explains, topical and temporal experiences of pitch appear to have qualitative differences, such that there are at least two kinds of pitch “height”:

> So both place and rate influence the reported “pitch,” but multidimensional scaling analysis of sound similarity ratings suggests that “place pitch” and “periodicity pitch” may correspond to different “up” directions in “perceptual space” (Chen et al. 2005)… These differences [between harmonic structure of a vowel and the pitch centroid] manifest as quite distinct perceptual qualities, but this important qualitative difference can easily be lost in 2afc experiments based solely on pairwise comparisons. (Schnupp 2014, 2)

As Schnupp argues, experimental set ups in psychoacoustics, such as the two alternative forced-choice (2afc) test can have the effect of collapsing the spatial meaning of pitch to the point that important qualitative differences are lost. Experimental designs do this both by failing to register difference in types of spatial signatures – the thickness of height, for example. Well-controlled listening environments have the side effect of severely restricting the practical context in which the sound is encountered. A thick description of pitch correction would need to account for how local ways of placing and timing sounds *thicken* pitch perception. This chapter offers a thick description of the various ways in which voices are placed (and how places are voiced) in everyday studio life. What follows is an ethnographic elaboration of the concept of tonotopical perception, or the idea that place matters in how voices are produced and perceived in-studio.

Studio as Heterotopia

"A studio is a lot like a pirate ship," Carl likes to say. His precise meaning is something along the lines of its tendency to contain "a bunch of dudes" – mostly male, vulgar, and unkempt, working in close quarters for extended periods of time. The studio-as-ship analogy fits together with a broader set of metaphors of maritime homosociality and conquest within audio engineering and related fields of signal processing. As Tara Rodgers has argued in her extended investigation of sound and metaphor, these typically involve:

> analogies between sound, electricity, and water waves; between fluidity and female corporeality; and between maritime voyage, scientific and colonialist enterprise, and the consolidation of an archetypal white, Western, male subjectivity. Acoustics experimenters and authors aligned the physical properties of sound waves with connotations of fluidity and excess that have been associated with female bodies throughout Western history and philosophy. To analyze and control sound meant to experience the pleasure and danger of unruly waves, and to seek their control from a distanced perspective; both the objectified material of sound, and the subject position of acoustics researcher, were gendered in audio- technical discourse. As well, through themes of maritime voyage and discovery, the experiential navigation and technological control of sound waves were articulated to colonialist paradigms of racial exoticism and conquest. An epistemology of electronic sound was built on the perspective and advancement of a white, Western, male subject, so that the technological control of sound waves became a symbolic containment of gendered and racialized excess. (Rodgers 2011, 55-6)[54]

[54] For more on the role figurative or poetic language in signal processing see (Sterne & Rodgers 2011).

The studio-as-ship analogy is a deeply resonant one. Charles Goodwin, in a study of the construction of space through tool-mediated practices on oceanographic research vessels, employs Michel Foucault's concept of the heterotopia (Foucault 1984), describing it as "a segregated world inhabited by a restricted range of social actors, which contains within it a second space where the phenomena that animate the discourse of a particular scientific discipline are made visible." (Goodwin 1995, 239) The recording studio is similarly heterotopical. It is an exclusive and symbolically weighted place where people go to have their performances captured and prepared for circulation. It has been widely compared to that other heterotopia, the womb. Engineer Eric Sarafin, in his Mixerman books, constantly refers to the studio's control room as "the womb." Musicologist Charles Keil, under the rubric of "womb envy," has compared recording studios to the "flute houses" of the Wogeo people in Papua New Guinea, in which they "initiate boys into the secrets of manhood and punish any woman or child who learns the secrets." (Keil 2014) Engineer Dave Pensado calls his studio "the lair."

Ethnomusicologist and recording engineer Eliot Bates, drawing on Zoe Sofia's concept of "container technologies" has examined the role of studio as "'wombs,' as 'bunkers,' or as 'man-caves.'"(Bates 2012, 2) while also suggesting that "studio-containers (regardless of their gendered readings) … comprise a system of conspicuous constraints on paths of audition, lines of sight, and the nature of the social and musical action that can transpire within." (Ibid, 3) The concept of the heterotopia is useful for understanding recording studios because it draws our attention to how they can simultaneously appear as "other" places, separate from everyday life, while also being the sites of performances that will eventually come to permeate the world outside of them. The tension between the tedious work of tuning a voice and the oversized effect of that voice once it appears in a hit single is one aspect of studio life that a heterotopical framing helps capture.

I found that I consistently experienced this tension as something like a thwarted expectation. Every studio I went to seemed smaller, or somehow more finite than it should be. Walking into Studio A every morning felt eerily similar to the experience of going back to your high school and realizing how small everything is, or seeing a movie star walk down the street and realizing how short he is in real life. It took me a while to not be struck by the fact that each studio had a smell – the tang of warm analog equipment in the control room of Studio B, the sweetened and intensely conditioned hotel air of Seth's upscale studio, competing in one case against the plumes of pot smoke clients were producing, and the stale sedimented incense of studio A's live room. This is a common experience in LA, a city that has been involved in the capture and circulation of itself in sound and light so many times that its sheer realness always comes as a slightly disappointing surprise. One goes into a high-end recording studio expecting either a living music video or a false front; what one finds instead is a workplace, a place where the boundary between the mundane and the extraordinary is routinely produced and managed.

The other-placeness of studio life has much to do both with the fact that studios are largely windowless. It also has to do with the difference between their interior and exterior presentation. Both studios A and B are hidden in plain sight, in bland industrial buildings in downtown Los Angeles. Neither publicly advertise their location, either online or on the building's exterior, which means that clients generally discover the place through word-of-mouth or, less frequently, through correspondence with studio management online. Like most studios, A and B both consist of well-isolated "live rooms" where drums and other acoustic instruments can be performed and recorded, as well as smaller, well-soundproofed isolation booths that can either be used for vocal takes or for housing guitar amplifiers that need to be mic'd individually. The built-in contingency of spatial ordering that characterizes studios means that it is through routine re-performance that the recording setup comes into and out of being. The first of these involves the enumeration and making available of countable things, a set of practices that I bring together under the category of "listing."

Listing

Studio life proceeds in large part through various types of listing. List making is a basic way of performing and giving structure to things, people, events, and the relations between them. Lists contain items, but they also contain interpretations: judgments of similarity and difference, tacit understandings concerning what counts[55] as what, under what conditions it does so, and expectations regarding how accountability of the list, its maker, contents, and audience stand in relation to one another. Two major ways that lists are used in the studio are 1) as ways of organizing action, in the form of the To-Do List or the itinerary, and 2) taking stock, in the form of the inventory. Studio interns are expected to work efficiently and without "getting in the way," which means following a number of unspoken rules governing action within the recording

[55] Socially situated judgments concerning similarity and difference, based on a "rules finitist" reading of Wittgenstein's concept of "family resemblance" have been shown to be fundamental to the production, contestation, and circulation of scientific knowledge (Collins 1992; Palmer et al. 1998; Pinch 1985; Bloor 1989) Ethnomethodologically inflected approaches in STS, operating on the basis of a different reading of Wittgenstein wherein rules are not only finite in their application but in fact shaped by it, have focused more on technoscientific life in terms of the production of socially accountable phenomena. Work on "counting" in particular has shown that ways of counting things are ways of reflexively accomplishing social order.(Martin & Lynch 2009) The lists discussed here serve both as a way of counting (i.e., producing complexly interconnected things as separately countable) and as a way of drawing connections and boundaries between things.

environment. One such rule is that there is always something to be done. Keeping track of these things to be done became a task in itself. At Studio A my tasks ranged widely. Within the first few weeks Carl and I had compiled together a lengthy list of things that needed to be done, reproduced in part here:

> *<u>Shopping/Drop-Off & Retrieval:</u>*
> *Drop-off Wurli to [tech] at [studio]*
> *Get mBox from [engineer]*
> *Pick up Crown DC300 Amp from [studio across town]*
> *Research new speaker to replace broken auratone mono speaker*
> *Order records [i.e., silver, gold, platinum] (for walls)*
> *Sell unused stuff on Ebay*
> *Order MXL V-69 shockmount*
> *Home Depot:*
>> *Cover plate for electrical box*
>> *return light switches*
>> *buy paint: blue & white (get rid of old paint)*
>> *materials for door sound-proofing*
>
> *<u>Cleaning/Organization:</u>*
> *Organize mic closet*
> *Repair and organize broken headphones,*
> *Put on hooks on wall for cables*
> *Organize tech room*
> *Shampoo rugs in tech room*
> *Organize towels*
> *Organize mics*
> *Organize Cables (XLR vs. patch vs. 1/4" vs. gender-benders)*
>> *<u>RECURRING:</u>*
>> *clean, sweep and vacuum*
>> *Bathroom*
>> *Mop*
>
> *<u>Repair/Construction:</u>*
> *Sound-proof door to live room*
> *Replace fiberglass w/ wood in vocal booth (to make sound brighter)*
> *Paint staircase*
> *Build stereo cue cable for control station out*
> *Solder pins & Elco connector to end of 16ch snake so it can show up at patchbay*
> *Test Level-Loc (should pump and breathe)*
> *Compile mic book*
> *Hang-up award records*

Hang curtain in vocal booth
Troubleshoot Shure 55s mic
Mount new TV above console
Get LA-2A fixed (bad gain stage)
Replace batteries on all gear that forgets settings (look for +- tape)
Replace burnt-out lamps on all gear (power conditioner above LA-2A)
Wire Equinox correctly (instead of Central Station for speaker select etc.)
Restring guitars
Troubleshoot & set-up wireless cameras (& wifi)
Troubleshoot left Auratone speaker (not full range)
Fix meter on Obsidian compressor
Wire connectors for dbx 120 (subharmonic synthesizer)
Label patchbay properly
Make 1/4" jack for Tandberg Reel-to-Reel out
    Find out how stereo module works

The overall heterogeneity of the list is significant in that it shows the wide variety of infrastructural care and maintenance work that makes studio life possible. The ways in which tasks are organized under headings is also important as it reveals how the studio as a self-contained space is both conceptually structured and routinely performed. The first heading, "Shopping/Drop-off & Retrieval" encompasses the traffic of things in and out of the studio. In LA this meant a lot of time driving around, getting stuck in freeway traffic, and ultimately learning to take the surface streets. Much of this involved getting equipment from place to place, usually from studio to studio where Carl's colleagues were working. Engineers will work at several different studios over the course of their careers. By working with different engineers from session to session they develop informal networks across which client recommendations, technical questions, and borrowed equipment can be passed.

When I approached Carl about interning with him, his first question was whether or not I had a car. Car ownership and ability to drive is at least as important a qualification for being a studio runner as any recording arts degree, especially in Los Angeles. Having interns shuttle things around town is common practice and serves multiple purposes at once. First, it ensures that expensive and delicate equipment is well taken care of and not entrusted to an anonymous courier. It also helps bolster informal connections between engineers, whose work is often relatively solitary, without taking them away from their DAW and desk. The intern can act as a sort of emissary from one studio to another, checking in on behalf of the engineer and keeping contacts current. Third, it helps bring the intern into the informal network, making them a familiar face within the community of engineers and the personnel of the studios at which they work. Because internships are so brief – rarely more than a couple of months – errands are also ways of setting interns up with places that might take them in once their current gig has elapsed. Learning to get into and out of a studio, where to park, which entrances to take, and how not to get in the way are all important skills for the intern-courier.

The latter two headings, "Cleaning/Organization" and "Repair/Construction" involve the internal organization of the studio. There is a great deal of overlap between these two categories. Repairing broken headphones, for example, fell under the cleaning/organizing heading, while labeling the patchbay – a task one might associate with organization – fell under the repair/construction heading. As a practical matter, the two headings became distinguishable in terms of their frequency and their appropriateness with regard to how the studio was being staged at a given moment. Cleaning and organizing tasks were taken to be persistent concerns – usually things to be done on at least a weekly basis. Repair and construction were lower-frequency tasks. Once you paint the staircase, for example, it should be good for a while. The staging component had to do with the question of whether or not a task is appropriate for when a client is in-studio. Organizing cables may be appropriate during downtime on a tracking session, for example, but replacing fiberglass in the vocal booth was not. Low and high frequency infrastructural tasks need to be balanced so as to make room for work more readily identifiable as “music production.”

Maintenance and creative work come into conflict in surprising ways. The roof of studio B, where guests like to sometimes hang out and enjoy the downtown LA skyline, is lined with thick rubber tiles. Cigarette butts, dirt, bottle caps, and other detritus collect in the spaces between these tiles. One morning, feeling industrious, I set about dragging away and stacking up the tiles in the corner of the roof so I could sweep up these hard-to-reach areas. I am halfway through the task when Andrew, another intern, comes up and informs me that Harold is downstairs trying to figure out why the drum tracks he is recording have this mysterious dragging and flopping sound on them during playback. The tiles, apparently, are not just there to keep people from slipping, but to isolate roof sounds from the sensitive microphones being used two stories below.

Items travel from list to list. The first item “Drop-off Wurli” refers to a client’s Wurlitzer electrical piano that had stopped working during their last session. The Wurli was until recently located under the Repair/Construction heading, as I had spent the better part of two days unsuccessfully attempting to fix it by replacing its fuses with large automotive fuses I found in the studio’s tech room. I ultimately had to take it to Seth, who was working as a tech at a large studio across town. Within a day he had brought the whole thing back to “factory spec,” replacing worn out parts in accordance with the piano’s original technical documentation. When I came to pick it up he was laughing about how some idiot had put car fuses in there. I laughed along, not sure whether he had figured out it was my handiwork he had to fix.

A task has a tendency to multiply itself, and the proliferation of a task into subtasks can only be checked through an act of socially situated interpretation. On a regular basis I would sweep the outdoor entrance to Studio B, which opened out onto a sidewalk next to a loading dock. Dust would blow in from a vacant lot across the street, coating the sidewalk with dirt and debris. The mind of the sweeper is apt to wander, I found, and certain mundane questions come to seem extremely important: At what point, for example, does the studio entrance begin or end? At what point does the dirt end and the pavement begin? Cyrus Mody, building on Mary Douglas' well-known definition of dirt as "matter out of place," has noted the complicated ways in which dirt is involved in technical practices such as materials science. (Mody 2001) Even in the clean room of a materials science lab, dirt is not defined absolutely but relationally, with respect to its role within a practical context. The entrance to studio B, as a boundary between the inside and outside of the studio, is a place where placing and displacing dirt needs to be carried out.

The problem is that the position of the sweeper of an entrance may differ markedly from that of a "user" of an entrance. Thirty minutes of sweeping, and the pavement somehow seems dirtier than it did to begin with. My good intern ethic struggles against a pernicious alliance between my instincts for laziness and existential musing. What is concrete anyway? What is an "entrance?" The fact that I am not being paid to do this, that my labor is being compensated in training and research access, seems to complicate matters further. If I were being paid by the hour, I speculate, the appropriate duration of the task would seem less paradoxical. I could sweep all day and night and it would be, in a sense, worth my while. This is the interesting exchange relationship between money and labor: money can serve as a point for otherwise pointless work. The lack of money makes me look for meaning elsewhere in the act of sweeping. I weigh the ethnographic value of sweeping an entrance. Reflecting on it I recall the advice of Annemarie Mol: if you run into a problem, that is data. Problems-as-data do seem to be piling up.[56]

The connection between sweeping a loading dock and being a recording engineer becomes more clear later that day when I come by Studio A and find Carl "cleaning up" a vocal track for a recording session from earlier that week. For him this task has multiplied, and he has broken it down into a hand-written to-do list of his own, reproduced here:

[SONG TITLE]

Verses AR Ved [checked]
DBL + Triple go thru + Line Up + Tune (even if muted)
CH5 timing + tuning since Big Group (might not need much)
Bridge All 3 Vox Time/Tune

<u>Needs to Look At</u>
[Song from earlier album by Artist]
[Song from earlier album by Artist]
[Song from earlier album by Artist]
(Timing? (.06 VocalComp) Line Up Alt playlists)
NewComp.18(tweaked)
- Check Backsides on 7.4 leave plugins on track - leave note
- [Other Studio] Email post vocals
    LVox separate
    then Appropriate Groups LVo HiVox Harm
    get [echo] chamber prints of all vox for These songs
    [Song titles for current project]

---

[56] I make a list:
1: I am "out here" sweeping when I am supposed to be "in there" learning what engineers do.
2: I am not being paid to do this, which means that:
3: I am actually devaluing labor and depriving career opportunities that might be taken by another intern who has a much more direct ambition to use the skills and connections to become an engineer
4: Even that hypothetical more-deserving intern would also be working for free, thus devaluing labor and continuing to restrict access to "creative" fields such as this one to people who can afford to work for free (with all of the socioeconomic, gendered, and racial biases implied by that restrictive condition).

The song has been captured in a checklist. Carl is reminding himself to check each of the verses, to “line up” and tune the double and triple-tracked vocal lines. He does this to all of the double and triple backing vocal tracks, even if they are “muted” and thus not audible in this version of the mix. He does this so that, if the producer decides that he wants to bring these tracks in they will be pre-corrected and ready to go. He singles out channel 5 in particular for fine-tuning, as well as the three vocal tracks which are used in the bridge. Under the heading “Needs to Look at” he lists three songs from the singer’s previous record. He will use these as “refs” or reference recordings to guide how he edits and polishes the vocals in this song, and as a way of making sure that the “sound” of the voice has continuity with the artist’s earlier work. He includes a reminder to check the timing on a vocal comp (a single recording of a vocal line assembled, or “compiled” from the best elements of multiple takes of the line), which he labels as version .06. The “alt playlists” for this comp, which contain all of the takes from which the comp was assembled, also need to be lined up in case he or anyone else wants to go back in and swap out one part of the comp for another.

It is through this sort of mundane inscription that performances in the vocal booth become itineraries of correction. For the purposes of knowing which parts need timing and tuning-up, a reference sheet like this provides a frame for the close listening required to perceive elements of the song as being in need of correction and carrying out those corrections in an appropriate way. For another comp, *NewComp.18*, he needs to check for differences between how it works on his version of Pro-Tools (HD) as compared with Pro-Tools 7.4, which another engineer on the project is using. In case there are problems, he reminds himself to leave the plugins on the track so they can be dialed-in on the different system – he will include a note explaining this to the other engineer as well. Once he is done with this comp he will email the finished or "post" vocals, divided and grouped according to their role in the song (high vocals, low vocals, harmonies, etc.) He finally reminds himself to get "chamber prints" – versions of the vocal track which have been played into a physical echo chamber at another studio and re-recorded in order to capture the reverberation - for the other songs he will be working on for this project.

This process involves tidying up the components of the track that might have gotten scattered around during the tracking process, making sure vocals are free from recording artifacts, making sure that everything is well labeled and thus legible for further editing and mixing, and making sure that the relevant pieces of the recording are available to the other people working on the project. It also involves comparing elements of the track to "refs" or reference recordings from other artists, so as to make sure everyone is on the same page concerning how the recording is supposed to sound. In the following passage, Carl is cleaning up a vocal track by removing plosives, sibilance, and any other problems at the waveform level. I sit at the DAW screen with him and attempt to follow what is going on. Carl first prepares the clip, raising track levels and using compression to bring out unwanted noises, labeling each change, meanwhile, so that the clip can be brought back to its original settings after the repairs have been made.

C:[57] Sometimes it's nice to um (2.3) just like with tuning have a compressor on (1.5) the vocal so you can hear.
((click))
S:[58] *Going away (.) °↓i:in ↑time° (.) it's*

---

[57] Here I employ Jeffersonian Conversation Analysis notation (Heritage & Atkinson 1984). Parentheticals denote pauses, with (.) denoting a just noticeable pause. Arrows indicate upward or downward inflection. Adjacent brackets and indentation indicate simultaneity of statements. Quietly spoken portions are indicated as in: *°example °*.

O: uhuh
C: And the, uh, Altiverb right on the track that [inaudible]
S: *Going away (.) ↓i:in ↑t-*
↓i:in
O: °Those numbers you put on, the >double and the triple and everything<°
Are those (.)↑decibels?
C: °Y↑eah (.) so you know what you left it at (.)
and when you're done tuning you can change it back°
O: °Oh I see°

Carl finds the jagged waveform of a plosive (what he calls a "sharkfin") and zooms in to smooth it out manually. I first have trouble seeing and hearing the problem and its correction, so he retraces his steps in order to make it more explicit:

C: [Uses the pencil tool to redraw waveform]
J:sss ↑draw in. Whatever you think the waveform should be
S: *↓i:in ↑t-*
O: Haha. Write your ↑name!
C: Yep. See that's what it used to be and then I ((inaudible: disguised?)) it
S: *↓i:in ↑t- ↓i:in [↑t-*
O: [But it still looks like that?
Oh no that's the comparison track
S: *↓i:in ↑t [↓i:in↑t-*
C: [No that's uh, uh it hasn't updated the waveform to the-
O: Oh I see
C: °So you zoom out sometimes it still looks like that.
S: *↓i:in ↑t*
C: It is a little bit of a glitch (.) In the program°

Here Carl has made a distinction between what the program is showing – namely, a non-updated waveform, and what it should be showing if it was working properly – a waveform reflecting the removed "sharkfin."

S: *↓i:in ↑time (.) it's hard to hear your*
↓i:in ↑time (.) it's hard to hear your
it's hard to hear your
it's hard to h-
it's hard to hear yoursel-

[58] Here the looped voice of the pre-recorded singer, *S*, begins to interject. Lyrics have been modified to preserve the artist's anonymity.

C: (look at) that sharkfin there
(3.2)
O: Are those things aud↑ible?
S: *it's hard to hear your [sel-*
C: [Sometimes
O: Yeah (.) It all seems so subtle…
C: Eh esp↑ecially when ↓you compress the ↑shit out of it
O: That stuff pops out
C: Yeah
S: *it's hard to hear yourself*
S: *it's hard to hear [yourself*
O: [Which you will be doing ev↑entually
C: Yeah… in fact we should turn this [compressor] up

In this passage, Carl has projected a setting within which the slight problems with the vocal track will be more noticeable – namely, an instance of high dynamic range compression ("when you compress the shit out of it".) This framing establishes terms under which the distorting sound of the "sharkfin" can be both audible and in need of removal.

S: *it's hard to hear yourself*
it's hard to hear yourself in the
O: [typing notes]
C: Uh…
S: *it's hard to hear yourself*
O: [typing notes]
C: Listen?
O: Sorr…
S:[59] *-n try to*
C: Gonna be right here
S: try to un-
try to un-
O: Oh man (.) [laughs]
You have to listen to it but I hear it now

My note taking has distracted from the careful listening required here. Listening collectively, in this case, involves learning to attend to the shared rhythms of silence and acknowledgement or documentation. Listening correctly is facilitated by use of imagery, as in the case of hearing “clicks” as sounding like being inside of the singer’s mouth:

C: Yeah. And if ya, if it is (.) you hear a lot of em and they’re on a lead vocal and
it just ends up getting… (like freakin’) really mouthy.
Like you’re inside [the singer’s] ↑mouth.
S: Ee (.) try to und
C: See even that it’s like. That’s what (click remove) – sometimes you just cut it out.
O: Yeah [typing]
S: try to understand what a
try to understand what a
try to understand what a taaa
[Sound of a truck driving by outside the studio]
Time
Time
C: Something like that
S: Time
Time
T…
Time
Time. Time. Time
C: But you can’t get rid of her…
S: Time
C: That’s like. A breath or a plosive you could fake if you want. It’s not really part of the verse.

Rachel Prentice has shown how surgeons cultivate forms of embodied action, which allow them to constitute and perceptually inhabit sites of technologically mediated surgical intervention. One surgeon describes “swimming” in the synovial fluid of a patient’s joint while conducting minimally invasive surgery (Prentice 2012, 171) Carl’s description of the mouthy sound as being inside the singer’s mouth orients the sounds on the track in terms of the relationship between listening and singing bodies, providing it with an enfleshed spatial grammar that produces the voice as a site of embodied action. Clicks on a vocal track could be any number of things: digital distortion, glitches in the encoding, latency issues. And, of course, the microphone was never actually *inside* the singer's mouth. Still, by hearing clicks as mouthiness, (or in-mouthness) Carl sets the tuning act in a particular embodied relationship that makes the click legibly tunable. Inhabiting and moving within the spatial grammar of stereo sound is a major part of the perceptual socialization of recording engineers. As he continues declicking and moves to the background vocal track, Carl continues to deploy spatial listening. He explains the waveform phenomenon of phase cancellation ("making it a little phasey") in terms of perceptual distance ("making it sound further away.") We want to be out of the mouth, but not too far from it:

> *S:* *Time.*
> *It's hard to hear*
> *Time*
> *It's hard to hear myself*
> C: The other thing I could do if it's really bugging me. I'm gonna do it on the background vocal. You can go Audiosuite, noise reduction, declicker. And I think it's an easy one.
> *S:* *Nnnnnn ↑Time.*
> C: Ummmm 110. 5.2… there got rid of it.
> *S:* *Time*
> *Nnnnnn ↑Time*
> *Nnnnnn ↑Time*
> C: But every time you do something like this it ends up making it a little phasey or making it sound further away or not as [inaudible… "compari"?] so you don't do it unless you have to. Click. Remove. Click. Remove.
> *S:* *Nnnnnn ↑Time*
> *Nnnnnn ↑Time*
> *It's hard to hear myself*
> C: And we'll do it here. Two clicks removed.
> *S:* *Time*
> *Time… it's*
> *Time*
> *Time*
> O: It's funny when you listen to it you can really hear it
> C: When you listen ↗<u>for</u> it.
> O: For it yeah

For insurance purposes, Studio A needed to inventory every piece of equipment, including serial numbers and dollar value, so I spent the better part of a month on the Borgesian task of taking stock of every single thing in the studio. One fact that becomes painfully apparent when you begin documenting a set of things, in affirmation of Bruno Latour's discussion of the etymology of the word "thing" (Latour & Weibel 2005), is that things are always assemblies of other things. Conversely, an assembly may need to be accounted for as a single thing, and the grouping-together or taking-apart of things into separate line items is always an act of discretion, one that is always undertaken with a particular purpose which needs to be deployed and redeployed if one is to, however improvisationally, develop an accountable and relatively stable set of criteria for categorization. Consider the following, extremely partial, excerpt of the compiled inventory for Studio A:

153	Digidesign	M-Box2	Digital audio interface
154	Tandberg	Model 5 Reel-to-Reel	Reel to Reel tape recorder
155	Apple	Macbook Pro	Laptop Computer
156	Bionaire	BAP1700	Premium Air Purifier
157	LG	LP0910WNR	Portable Air Conditioner
158	Webster	Webcor	Reel to Reel tape recorder
159	LP	Aspire	Conga
160	Weller	PU120T	Power supply for TC201 soldering iron
161	Weller	TC201T	Soldering Iron w/ Stand
162	Swann	2.4Ghz Wireless camera system	Wireless Camera system

The soldering iron itself is legion; the power supply gets its own line item but the stand does not. There is no clear reason that the spool of solder, the multiple bits that screw into the iron's tip, the sponge nested in the stand, should not get their own lines, other than the need to check the endless proliferation of things. Many of the most important pieces of studio property often seem not to have any necessary or sufficient physical presence whatsoever. For example, sound recording cannot proceed without a relatively up-to-date edition of ProTools software, which putatively resides as a barely cohesive monster of bug-fixing patches, updates, and DSP (digital signal processing) plugins, each of which are sustained by the rotations of platter in any one of several hard drives, which may or may not be present during a given session.[60]

This imperfect assemblage cannot even get set into motion, however, without the proprietary iLok key fob – the most formidable obstacle, short of litigation, faced by an unscrupulous engineer's attempt to simply pirate the software. The iLok's small plastic thumb drive, designed for the sole purpose of preventing theft, thus appears as one of the more singular examples of Carl's investment in the main tool of his trade. When not in use, the iLok is kept in a wooden knick knack box next to the monitor. Over the course of the first couple of weeks I repeatedly forget to bring my own iLok, which I would need if I were to do any Pro-Tools work on my own laptop. Over the first month Carl frequently asks me whether I brought my iLok, though two months pass before he actually asks me to use it. The skill of remembering the iLok, I surmise, needs to be mastered before the skill of using the software it unlocks can proceed.

This is not even to mention the millions of other pieces of digital intellectual property, legally accountable and otherwise, continuously passing in and out of the production computer. There is, to begin with, no obvious hierarchy for documentary criteria, let alone a single unambiguous criterion for the status of an individually accountable thing. As an intern, in order to make myself accountable as a member of the studio ensemble, it was my task to make all of the other things accountable. Even the trusty serial number, I find, cannot be relied upon as a

[60] An Audio Engineering Society "Law for Engineers" captures the elusive materiality of computer-based recording in the form of a trinity: "Digital audio does not exist unless it has been backed up 3 separate times on 3 separate drives – Pfaffle's law" Attributed to Engineer Tom Pfaeffle. (Chinn 2014)

universal marker of insurable-thingness; MXR microphones, their friendly customer service representative informs me, have none. Some things, upon sufficient inspection, can be made to turn up number-like inscriptions, but not all of these are legible:

Figure 10: An ambiguously numbered thing. Photo by the author.

The moral is clear enough: things, even well-standardized, human-made, commercially available things, gathered together within a relatively well-contained space, do not present themselves for accounting but must be made accountable through contingent acts of interpretation and material intervention.

Hierarchies are another important form of ordering-by-listing in studio. A hierarchy is both a list of possible positions and a map of how those positions are related. Both studios A and B are simultaneously largely informal, usually minimally populated, and thoroughly hierarchical places. During my first week in Studio A, Carl gave me a photocopied handout he had kept since recording school. It was a collection of article clippings and seminar handouts on the general topic of "being a good intern." As one entry put it, the point of interning was to "develop an understanding of how studios work, how to exist in the hierarchy of the industry" (Hughes 2003)

The spatial distribution and typical orbits of people in the studio are mirrored in a fairly stable industry-wide hierarchy. (Rubel n.d.)

Clients – for whom studio time is booked and paid (either by themselves or by their record labels) typically consist of the performing musician(s) and a producer who, while not necessarily performing, is charged with providing creative guidance to the artist and acting as a go-between between the various domains (technical, artistic, financial, etc.) of the music production process. Antoine Hennion has described the role of the producer primarily in terms of their role as intermediaries between music and market (Hennion 1989). This is still the case, though the role has changed considerably as knowledge and equipment required to record has become more easily accessible. In light of this increased access, and increased demands on saving money, many producers have taken on the role of engineer, and vice versa. The artist and/or producer may choose to bring in their own engineer to run the technical aspects of the session. It is also common for the studio to provide an "in house" engineer, or for the studio to be selected because it has a particular engineer in residence. The studio will typically employ a technician ("tech" for short), who is tasked with keeping everything up and running, an assistant engineer to set things up, break things down, and help the engineer run session, and a runner to retrieve things as needed. At the very bottom of the ladder is the intern, typically unpaid, who fetches food and drink, cleans up, and generally takes care of all the mundane tasks no one else wants to do.

The janitorial work of the intern serves the dual purpose of keeping the studio separate from the outside world, while also utilizing the free labor attracted by the glamor of an exclusive and "creative" profession. As engineer George Sanoff relates (again, quoted in my intern handout packet), "Basically in major studios like these, with multi-million dollar facilities, interns are a way to avoid hiring janitorial and cleaning services. I cleaned the rooms after sessions, I cleaned the bathrooms, answered phones and went on errands." In this way, studio hierarchy is nested within a broader social hierarchy, given the level of privilege required to make working for free possible in the first place. Another packet-contributor suggests that the most important qualities an intern can have are a car and parents paying their bills. Part of the service studios provide their clients is a private atmosphere relatively free from distractions or external observers. Clients, and to a lesser extent studio staff, want to be able to control the way their actions within the studio are ultimately recorded, interpreted, and circulated. Having a participant observer involved is potentially extremely counterproductive in this respect, precisely because it is ambiguous how such a person would fit into the normal hierarchy. This interplay between perception and social order persists in other modes of communication. The broader apparatus of the music industry makes itself known through email correspondence, as engineer Bob describes:

> When I was doing that [project title] stuff I'd be on these email chains with like 30 people. Like, all the managers all the people in charge blah blah blah and the artist and the manager and their people in charge and it's like how do you get all these people to be ok with it? And you literally just have to figure out the social hierarchy of who, who do they think knows what they're talking about? And that's either you, because a lot of times like if you just say it sounds better they'll believe you. And then you can kind of go off that, or you find the person in the organization that they all look to. And then, if he says it's cool then, ok, who cares what anyone else says, I just listen to that guy. Cuz they listen to that guy.

The email thread, for Bob, becomes an extension of the studio apparatus as a play of voices that need to be managed. Ed, a former studio technician working as a programmer for audio encoding standards, makes a similar observation. For him, writing an email is basically the same thing as writing a block of code. Both involve negotiations with a complex system over definitions, interpretations, and orders of operation.

Floor Plans and Session Views

Ways of viewing and moving within space are also important forms of topical organization in the studio. Live rooms and isolation booths generally begin each session as empty spaces. Microphones, stands, cables, headphone monitoring stations, chairs, instruments, are arranged differently on a case by case basis. The control room, however, largely maintains the same arrangement from session to session. Studio B's live room is so large that Harold likes to employ large cylindrical "bass traps," which stand vertically at around six feet, measure about one foot in diameter and are covered with a cream-colored sound-absorbent padding. Up to a dozen or so of these may be positioned as needed around drum kits, amplifiers, string ensembles, or any other sound-making assemblage in the room. These performance spaces are situated peripherally with respect to a central control room, where all of the microphones and instrument inputs that capture the recordings from the live rooms and isolation booths are run into a large single mixing desk or "console", which is in turn routed through additional "outboard" gear such as compressors and equalizers ("EQs") and either a tape machine or a computer running DAW software (typically Avid's ProTools, but occasionally Apple's Logic). The control room, nested within the studio space, thus collects and renders-up for editing and playback the occurrences from the studio's other specialized spaces.

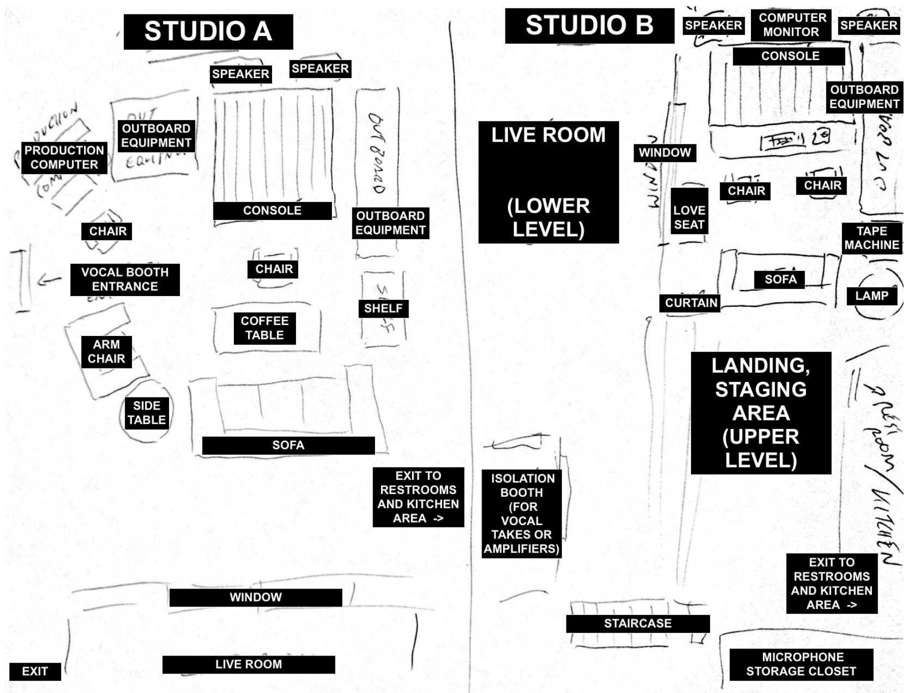


Figure 12: Author's sketch of the partial floor plans for studios A and B, showing position of control room relative to live and isolation rooms.

Like a ship the recording studio is a characteristically self-contained space, with all points leading to a cockpit-like control room. The mixing desk or console, in turn, is the focal point of the control room. Once in the control room, even if you are not directly attending to what is happening on the console or the computer monitor, you are usually sitting facing these things. The engineer is traditionally posted, sitting or standing, at the console and/or computer keyboard and trackball (for some reason, the trackball is preferred over the mouse, at least in LA recording studios.) The assistant is also usually situated close to the console or patch bay, though he or she may frequently leave to check things in the live room. Clients, when not performing, sit in the couch or other, more comfortable seating available slightly further from the mixing console. Entourage members, if present, may also occupy the sofa. Interns, when the other seats are taken, sit or stand at the periphery of the control room. Like assistants, they are liable to be sent to retrieve things in the other rooms – often the kitchen, to fetch coffee. In certain, usually larger studios, interns are not allowed in the control room at all.

For Harold, place-making is important both as an aspect of life within the studio and as a way of thinking about the recordings which take place within it. He compares his approach to a recording project to the design of a house:

> You really kind of have to meditate, and you should do it very early on, and really hear the finished thing. And I think that goes with anything like you know even this thing you're doing now you kind of have to see it completed. And even the details. Like if you're building a house you're like ok I need some 2x4s and some drywall, but also you need to think about what color curtain's in there, or what kind of cool soap dish you're gonna have in your bathroom. All those little details. And the more you envision that stuff and the more it becomes a things where you can walk around and have a 3d tour in your brain of, the easier it's gonna be, and then you commit to that. Not that you're not gonna compromise or you can't take the artist's wishes or respect or anything like that.

At the visual center of the recording process, at least in cases of digital recording, is the DAW display on the computer monitor. Here recorded tracks are depicted as a horizontal stack of color-coded waveforms, running left to right over time. A cursor marks the moment of playback across each track, or "audio region." A given session file is liable to have dozens of individual tracks. The version of Pro-Tools Carl is using does not have a track counting feature, so on some of these shared projects the track numbers are so large as to be literally "countless." Many tracks, which might have outlived their usefulness without being deleted, may lie buried in the nether reaches of the session, muted. In addition to being arranged vertically, tracks can also be routed or "bussed" to one another in order to achieve different effects, or consolidate several tracks into a single one.

Often, when several different engineers are working on a single project, they will share session files by way of physical hard drives. Things become complicated very quickly, and Carl is often irritated to find that a previous user of the session has not labeled their work well enough for him to follow what they have done. Ideally, each person who works on the session will use the same naming conventions to indicate where they are in the course of a particular track or section. Naming conventions vary, however, according to the preferences of the people running the project. It is also common for one engineer to work on a session file using a particular "plugin" or specialized digital signal processing effect. Because these plugins are proprietary, however, it will not be included in the main session file, and it will not sound the same to someone whose system is missing the plugin. Even if the plugins are all available between computers, the type of computer being used, the memory available, and any other vagaries of digital materiality can make things sound different across systems. When things are poorly labeled and certain effects don't travel well, the task becomes somewhat archaeological. A play of effects need to be reverse engineered into decisions by people with particular ideas as to how the thing is supposed to sound.

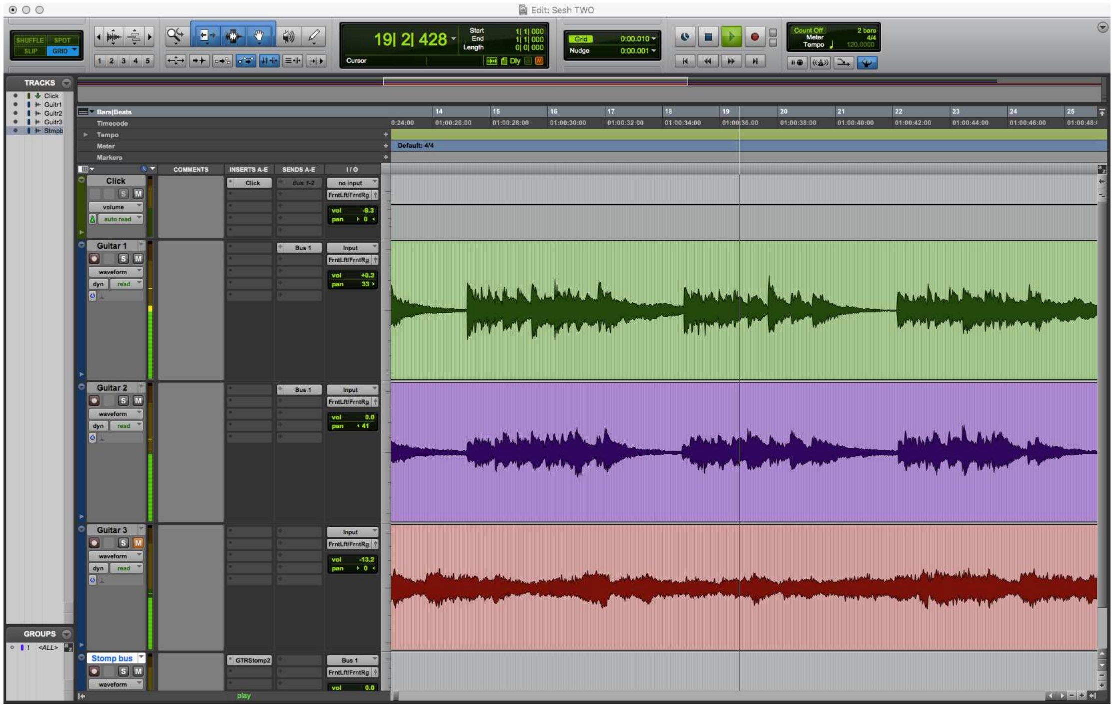


Figure 13: Pro-Tools 10 Session Screenshot by author.

Staging and Gain-Staging

The layout of the studio is further ordered by situated acts of stage management, or "staging." The dramaturgical approach pioneered by Erving Goffman analyzes social phenomena in terms of performance, role, and stage management practices. The theatrical metaphor has proven especially useful in STS work, where scholars have shown the importance of practices of self-presentation and material-semiotic choreography in the production of scientific and technological authority.[61] If scientific and engineering knowledge makes especially frequent and powerful use of the trope of "things (data, experiments, theories, equations, etc.) speaking for themselves," STS scholars have used the concept of performativity to trace the socio-historical

[61] For an STS application of dramaturgical analysis to a case of expert knowledge in a regulatory context, see (Hilgartner 2000)

production of the kinds of speakers who we consider capable of speaking for things that speak for themselves.

Recent work by Dumit and Myers on haptic and embodied aspects of technoscientific performance have challenged performative analyses that presume a self-consciously theatrical social actor. With feminist theorist and historian of science Rebecca Herzig they argue that dramaturgical approaches risk “tak[ing] for granted the bodies of the scientists and the materialities of their experimental apparatuses.” (Myers & Dumit 2011, 243) They instead call for attention to how bodies and materials[62] are not simply the grounds for performance but the products of it as well. Rather than tracing the interplay of active and passive voices (subjects speaking-as or objects spoken-for) they advocate the examination of a “middle voice” the subject of which is “interior to and affected by the action being signified by the verb.”(Ibid 248) Implicit in this critique is a challenge to the metaphor of staging itself, particularly the way that it presumes the binary of front and back-stage and limits ways of thinking about who is or is not “conscious” of the performance as such, for whose benefit the performance is being put on, and whether the actors are being shaped by the performance in anything more than a superficial way.

Staging in the studio is not limited to the deployment of scripted social roles divorced from embodied practice and material difference, however. The studio is an unstable mixture between back and front stage. Managing visibility and invisibility is an important practical element of staging work when engineering a session. As veteran mix engineer Eric “Mixerman” Sarafin illustrates in his memoir (quite popular among engineers, at least in LA) “The Daily

[62] (Pinch 2010)has also called for an increased attention to materiality in dramaturgical analyses in STS, specifically through attention to the “invisible technologies” that can be discovered retrospectively in Erving Goffman’s case studies.

Adventures of Mixerman", the practical production of conditions of visibility and invisibility consists in techniques for setting the scene:

> Willy sat motionless in front of the speakers. I was directly behind him at the couch, but when you've been working at a studio for longer than a day, you figure out that you can watch the expressions of people at the console from the reflection in the glass that separates the room from the Womb. I have a setting on the automatic programmed dimmer selector that allows me the maximum amount of reflection, and I often hit it just before I play someone something. (Sarafin 2002, Day 8)

The simple gesture of hitting the dimmer switch before playback constitutes a densely layered act of staging. At one level, dimming the lights has the effect of framing playback as a special instance of listening, bracketing it from the normal course of experience and presenting the instance of playback as a coherent thing to be evaluated, rather than the contingent assemblage it was before the lights went down. It signals a change in what everyone in the control room is supposed to be doing at that moment, and serves as an invitation to not just listen but to "perform listening." Sarafin brings down the lights in the live room ("the room") where the recording was done, making invisible the instruments, cables, music stands, and other assembled objects which helped constitute the thing listened to. In his later, more practically-oriented book Zen and the Art of Mixing, Sarafin goes into greater detail about the role of lighting as a way to properly stage the moment of playback for the purposes of picturing the mix and monitoring its emotional effect:

> Reducing light creates a more relaxing listening environment and enhances everyone's hearing. It's far easier to hear emotional impact when the lights are low, mostly because there are no visual distractions. The lack of visual stimuli allows your client to visualize the mix as they hear it. Just as this is beneficial to you as the mixer, it's helpful for your clients as well. Shutting off the computer monitor is a must on the first playback, since it can be exceptionally difficult for your client to resist watching their mix rather than

> listening to it.[63] You don't want your client to have fair warning of what's coming next in their own mix (or dismayed by what's been muted); you want them to get the full emotional impact of the work, evaluated from the perspective of the listener. (Sarafin 2010, 152-3)

Carl would use the phrase "scene change" to describe the move Sarafin describes.[64] As the lights dim in the room, the lights in the control room (which Sarafin calls "The Womb") turn the window to the room into a partial mirror, reflecting the control room as a scene where listening is being done. The listeners are able to see themselves at the console, able to witness themselves performing listening. Sarafin stands at the back of the room, where it is too dim for his own image to appear on the glass. From here he is able to witness his clients witnessing themselves performing listening. The fact that he does not appear in this scene is crucial, as it allows him to witness them performing listening as they would without him watching.

Staging also occurs at the level of self-presentation. Interns are advised to "be aware of your body language"(Hughes 2003), pay attention to how the engineer and musicians interact, and "[w]atch the engineer's hand, try to get the gist of what they're doing." (Bingold 2000) Grammy-winning engineer Al Schmitt describes a great assistant as "transparent."(Schmitt, cited in Owsinski 2005, 211) Engineer Trina Shoemaker describes what it feels like when this transparency fails, when one becomes painfully opaque:

---

[63] This rule is also to be found in the production of film sound. I was able to sit in on a sound mixing practicum session at the USC film school near studio B. The instructor repeatedly reminded the students to not look at the Pro-Tools screen while watching their project play back, lest they "not listen right."

[64] For Carl this phrase covers a variety of contexts. While mixing a song, for example, he might accomplish a 'scene change' by dropping the vocal reverb during the bridge or emphasizing a dynamic shift from section to section. Another time, as we walk to a taqueria a few blocks away, the architecture abruptly changes from industrial warehouses to art deco facades – "scene change" notes Carl. The art deco quickly becomes mirrored glass office buildings – "scene change."

> It's a very physical thing… there is a moment where either the back of your neck starts to get very hot because it's not working and you know the artist is behind you and you're nervous and you feel bad and so you don't wanna turn around 'cause they're all looking at you and waiting, or there's that moment where you sit back and think oh this is it. I'm almost done. And I didn't know I was almost done but suddenly it sounds great. (Shoemaker 2008)

Bodily comportment forms an important part of studio life. Just standing around it considered bad form, as is "looking like you're lost." A good intern or assistant should

> know how to sit in a studio and observe without making their presence uncomfortable - basically this means knowing how to take your cues from those around you. Sometimes that means shutting up and being still, but other times it means talking with band members, making the beer run, and being part of the gang. …The biggest thing you learn is how to hang - have a good hang so people aren't having nervous breakdowns and get good takes.[65]

Wardrobe is an element of staging as well. In Studio A it does not seem to matter as much, since it is a smaller and more casual environment. Studio B is different. Andrew, the assistant engineer at Studio B always wears a black t-shirt and shorts. So does Kyle, another Studio B intern who until recently was interning at another studio across town. After spending a day helping them set up, run, and break down a session, it becomes clear to me that I need to start dressing the part. Wearing black acts as a sort of vernacular indicator of being "backstage." It literally makes you seem less in-the-way, while also signaling that you are there to help if needed. Carl always keeps a backpack ready with a clean change of clothes, along with backup cables and hard drives, in case a session runs overnight.

Occasionally staging work occurs at the level of the materiality of studio walls. On a quiet day in studio A, without any sessions booked, Carl and I get to work upgrading the acoustic

[65] From photocopied handout received 2013, original author unclear.

paneling in the vocal booth. Its walls are currently lined with a fiberglass material designed to absorb echoes that might color a vocal take. Carl paces the length of the booth, calling out periodically to determine the more or less reverberant – more "live" or "dead" spaces. Towards the back of the booth he stops and notes that it is far too dead – "try it, it hurts" he says. I stand in his place and feel a sort of dull disorienting pain at the lack of ambience. We spend the rest of the day installing wooden panels to liven up the dead zone.

In addition to material staging practices of this kind, engineers develop a more general habit of attending to music in terms of its production. Many engineers experience this perceptual habit as a mixed blessing. For Barry, who records music for advertisements, hearing the production can be irritating:

> I started writing music and engineering at the same time and I don't think, I don't quite remember when the switch happened, but I haven't listened – I think it might have actually been, I don't think it was this bad until I got to school and out of college maybe, but I haven't been able to hear music the same way since I was 18/19 years old. Um, it's very, uh, I don't know I don't I can't I cannot sometimes I get done from work and I don't wanna hear music. It's difficult for me to passively listen to music. Because I'm not just listening to composition, I'm listening to everything. Like, ground-up, what's that kick drum sample what is that snare drum sample, identifying rhythms, identifying production, identifying weird engineering choices, like blah blah blah. All the time. It's very difficult to kind of turn that off in a really nice way. It's kind of irritating.

The type of hearing Barry describes is a case of what ethnomethodologists have termed a "production account" or an articulation of a "happening" informed by the specific competencies by which it was brought about. (Lynch 1991, 239) Harold Garfinkel drew a distinction between "virtual" and "actual" production accounts, where the former encompasses what "literally" takes place in a situation and the latter concerns what was involved in the staging or production of the situation.(Francis & Hart 1997, 150) A virtual production account of a pop song may, for example, include its genre, feel, or sound. An actual production account, by contrast, would take up the song as a procedural accomplishment, a collection of decisions made in terms of available techniques, materials, and considerations. Where Barry is irritated by his tendency to approach music in terms of a production, engineer Bob, on the other hand, casts his experience in terms of a "love" of production:

> "Well I have a weird listening habit. Some people are lovers of music. I'm a lover of music production. A lot of people will listen to music and love it. Even [Carl] will listen and be like oh this album… this is a great song. He can still appreciate music. Whereas I'll be like – that sounds like crap, why does it sound like that? How did they get that saturation? Is that Decapitator? I wonder how they did that."

Barry's description of hearing a song "ground-up" can be read in two ways, each of which show something about what it means to "hear the production." On the one hand you hear how the song has been "built from the ground up" in the way that a pile of construction materials might be erected into a dwelling. On the other hand, you hear the song "ground-up" as with a mortar and pestle, decomposed so as to be analyzed as a play of substances or a sum of its parts. These two different ways of approaching a song have the shared characteristic of being distinct from how the average listener is expected or supposed to approach it. The consumers of pop music are not generally expected to evaluate songs in terms of their structural soundness or their component parts. They are expected to love and appreciate the "music" as distinct from the production. Bob's self-identification as a "lover of music production" means that he can even love something that "sounds like crap" in terms of its production.

Learning to hear "ground-up" is an act of locational imagination. With the right guidance, you come to hear parts of the sound sitting in places where you would not expect them. Lucille notes, for example, that in the case of hearing an edit involves listening for a particular location in the sound:

> they'd be like ok do you hear, like they'd sit with me and they'd play like acoustic guitar and like ok do you hear this edit here? And I was like, no. I don't hear it. And they're like it's like a whaa sound that's like really low, and it's like a frequency far under the acoustic guitar and we'd just sit there and listen and listen and. Wow oh my god I can totally hear it now. And now I can't not hear it.

Here Lucille is describing being initially guided be a fellow engineer in the form of topological listening required to hear edits even where they might not be expected to appear. They listen together to the same passage as it repeats. Small discrepancies, things that would normally go unnoticed, rise to the surface when a track is looping. Repetition can work to unmoor a recorded segment from the initial context of its production. You do not hear the thing before or after the loop, so you do not perceive it within that frame. In his account of the role of language in developing professional audition among recording engineers, Thomas Porcello has described how engineers must be able to negotiate "spatial metaphors" at various technical and figurative registers. (Porcello 2004, 739) A complaint about a "thin" drum sound, for example, might need to be translated into terms of a spectral distribution wherein there are too many high frequencies and not enough mid or low frequencies. This respatialization may, in turn, lead the engineer to investigate the relative positioning of microphones with respect to the drum kit so as to avoid phase cancellation, or may indicate the need for a "darker" microphone.

We can extend Porcello's insight about the importance of spatial metaphors for collaboration in-studio to include how the mix is itself topically constituted through a variety of perceptual practices. The practice of mixing is often described in terms of locating the various elements of the recording appropriately in the available space of the mix. One key way of doing this is through the division of horizontal space through the use of the stereo image. Because most musical recordings are ultimately mastered to a two channel (left and right speaker) stereo format, the mix engineer is able to position various instrumental tracks along the listener's horizon by controlling how much of a given track is sent to which speaker. This technique is named after the cinematic practice of "panning" wherein the camera is rotated along the horizontal axis. Panning is not the only way to place sounds horizontally, though it is the most common. Another technique draws on the psychoacoustic phenomenon known as the Haas Effect, and involves delivering a sound to one speaker a few milliseconds before sending it to the other one. Whereas panning places sounds in terms of the relative amount of energy given to each speaker, the Haas technique uses subtle differences in time. The latter effect has the advantage of creating a more "durable" stereo image across contexts of playback, as well as a "fuller" sound since both speakers eventually play the same sound at full volume.

Distance and vertical position are less well formalized than stereo space. Pitch shifting or frequency equalization are two ways that engineers place sounds in higher vertical position. This is partly accountable in terms of the standard practice of placing the high frequency tweeters above the low frequency woofers in speaker design. It is also related to the habit of conceptualizing higher sounds as being higher in vertical position. As Bay Area engineer David Gibson suggests in his book *The Art of Mixing*, it also has to do with the phenomenon of lower frequencies traveling better through the floor of the studio and resonating the abdomen while higher frequencies travel through the air and resonate the head. (Gibson 2005) Apparent distance from the listener is usually managed in terms of loudness. The louder something is the closer it is. Distance or closeness can also be achieved through the manipulation of a sound over time, however. Engineer Bob Horn describes using a compressor to make a vocal "sit" closer to the listener. The "release" setting on a compressor controls how quickly the compressor lets the signal level drop after it has been triggered. A "fast release," Horn explains, will move the voice six inches forward.(Pensado 2014a) Distance is also accomplished through time domain effects such as reverb or delay, which extend a part in time and make it sound like it took so long to get to you that it collected reflections along the way.[66]

Making a part "sit" is a common concern among mix engineers, and is often accomplished by "carving out" frequencies through the use of equalizers (EQs) so as to make it fit together with it's neighbor like a jigsaw puzzle piece. While EQs are used to sculpt sounds into particular shapes, they also contain their own topical signatures. One might, for example, speak of an EQ having "beautiful air" which would make it appropriate for orchestral recordings but less appropriate for the claustrophobic and densely-packed mix structure of a death metal song. The scene of a mix leaves its mark on the elements within it. Listening to soloed tracks on a completed mix, you are likely to hear noticeable artifacts that disappear once they are played back with the rest of the mix. A distorted guitar that sounds massive in a rock and roll mix, listened to by itself, will often sound quite hollow because so many frequencies have been "scooped out" to make room for the singer's voice, the cymbals of the drum kit, and the bass guitar and kick drum. Mick Guzauski describes his vocal chain for Daft Punk's Random Access Memories:

> I usually add some top end to make sure they cut through. But I tend to cut around 3-4kHz, very narrowly and depending on the vocals, and use the Dbx 902 de-esser so I can make the vocal brighter without it being sibilant. (Pensado 2013a)

[66] As we shall see in the Appendix on studio shibboleths, the proper spatial interpretation of reverb is a way that engineer's distinguish themselves from their clients.

It is common, as Guzauski shows, for engineers to think and talk about sounds in terms of specific frequency bands. Cutting the "upper midrange" frequencies of 3-4kHz will ensure that the vocal "sits" without competing with instruments such as guitars or synthesizers while boosting the top makes the voice "cut through" the mix. Sounds that cut too much, such as sibilant "ssss" or "shhh" sounds, can be dialed down with the help of dedicated tools known as "de-essers."

Gibson has developed colorful three dimensional representations for how different mix genres should "look" according to these habits of topical listening. While most professional engineers would scoff at the idea of relying on these representations in any serious way, it does convey many of the tacit ways of placing sound that mixers use to put together their recordings. Tony Maserati, for example, makes a habit of planning his mix in terms of scene setting. In his mix for Robin Thicke and Pharrell's hit single Blurred Lines, for example, he designed the mix as a room where a party is taking place: Thicke and Pharrell are in one corner at a party, he explains, the cowbell is in the other corner.(Pensado 2013b)

Concepts of vertically arranged space are not only used to describe the apparent placement of sound within the auditory scene of the listener, but also as ways of describing the limitations on signal level afforded by any given piece of equipment. The amount of signal that can be run through a channel without it distorting, or the difference between "nominal" signal level and the level at which the signal or system may be damaged, is referred to as its "headroom." When the headroom runs out, you have reached the "distortion ceiling." The level of all of the unwanted incidental or background noise within a system is referred to as the "noise floor." The practice of "gain staging" is the management of a signal as it passes through various elements that may, by boosting, reducing, or otherwise changing the signal, change the relationship between headroom and noise floor.

Headroom can differ qualitatively as well as quantitatively. Digital channels, for example, have strictly defined headroom because signals that exceed the threshold are “clipped” by the signal processing algorithm. Analog distortion ceilings are more forgiving, since the distortion is less drastic and often considered more “warm” or “musical.” Carl, for example, will perform the rough mixes of a song “in the box,” meaning within the pro-tools interface, and then send them to his console. Because the console has more headroom, the tracks are able to – so to speak - breathe and stretch out. Carl, meanwhile is able to dial in more precise sounds using the console’s channel controls. Not insignificantly, this “scene change” from sitting in front of a screen to standing in front of a mixing desk, affords a new and reinvigorating perspective on the mix. The mere act of changing posture and removing the pro-tools screen from one’s visual field has a significant effect on how Carl hears and engages with the mix. Where gain staging situates the mix within the limits of the playback system, Carl’s movement to the console re-stages his position as a listener with respect to, among other things, the actual floor and ceiling of the studio, the placement of the speakers with respect to his body, and the signal topology of the physically structured (as opposed to digitally modeled) channel strip.

Standing at the console, Carl is also able to select between a variety of speakers, each of which contain a signature topology in the form of their frequency response characteristics. If a speaker projects the full frequency range evenly, it is described as having a flat frequency response, or simply being “flat.” The industry standard monitoring speaker is the Yamaha NS-10 – both Carl and Harold use it as their primary monitors. The NS-10 is held in high esteem not because it sounds good, but because it is considered unforgiving. Even though its frequency response is not very “flat” the NS-10 has come to be regarded as a sort of impartial arbiter of a mix: if you make it sound good on an NS-10 it will sound good anywhere. When Carl showed me his NS-10s and explained their comparative flatness he conveyed it gesturally by playing a mix through them and holding his palm up to the speaker, fingers pointed upward, as though he could actually feel whether any frequencies in particular were being favored. Studios also make sure to keep a range of lower quality speakers in the chain, in order to give an idea of what things will sound like on a consumer grade speaker system. Many engineers employ the “car test” where they play back a mix over a car stereo while driving around so as to hear what it sounds like through stock automotive speakers and in competition with road and engine noise.[67] Many young engineers like to do at least part of their mixing on laptop speakers or iPhone earbuds, reasoning that this is probably how listeners will actually encounter their work later on.

The bodies of engineers – specifically their hands – are implicated in the topological construction of a mix through the practice of controlling a track’s level in realtime during recording, also known as “riding” the fader. As Al Schmitt recounts, in its earliest incarnation fader riding was also known as “hand limiting” – or manually turning up the level of quiet parts and turning down the level of loud parts so as to produce a more even mix. Hand limiting vocals in this way involves a subtle choreography between the fader technique of the engineer and the microphone technique of the singer. Singers had to learn to come up close to the microphone for quieter notes and pull the mic away during louder sections, and limiting had to take these practices into account in real time:

[67] Engineer Al Schmitt pioneered this practice early on by building small radio transmitters that could broadcast to the radios of cars in the studio parking lot.

> With vocals, the engineers on classic recordings also routinely rode the fader in real time while tracking, either into or out of the tracking compressor — a highly skilled operation now largely unheard of amongst younger engineers, but which can subsequently enable smoother‐sounding compression at the mix stage. ‘When the guy is singing the vocal, I don't just sit there and listen to him,’ explains Mike Clink. ‘I'm feeling the emotions in my fingers, so to speak. When he's down I'm pushing him up so I can hear every single syllable. It's kind of a performance. As you do multiple takes, you actually learn the moves… and I'm following where he's going as far as how much he's projecting it. So by the time I get to the mix, I've done a lot of the fader riding.’ (Anon 2009)

As an embodied practice, fader riding can be likened to the practice of piano tuning as described by David Sudnow in his book Ways of the Hand. Sudnow describes the process of learning to tune a piano by striking two strings and listening for the “beats” between them while adjusting string tension with a tuning wrench. Approaching the task sequentially – that is, striking the strings, listening for the beat, and adjusting the tension – proves unsuccessful as Sudnow is not able to hear the beats. Beats only reveal themselves once the tuner learns “how to constitute and ride on the sound waves of a pair of pulsating strings, with his arm and hand artfully engaged with the hammer and pin.” (Sudnow 1993, 44) In order to make the beats available to the ear as indices of in-tuneness, one has to involve the striking and the tuning hands in a common movement, as an “arm and hammer enterprise.” (Ibid) Just as the tuner’s body is distributed through sense and action across the resonant body of the piano, so are the engineer’s fingers on the fader knob made to both feel the emotion and bring it out through a co-performance with the vocalist. Sarafin also sees fader riding as a way to bring out emotion by placing the vocal in the mix and drawing the listener’s attention: “I often spend an inordinate amount of time riding the vocal in my mix. I want that vocal to push the listener forward, and manipulate the listener’s emotions. In order to achieve this, the placement and balance of that vocal must be entirely deliberate at all times.” (Sarafin 2010, 142) While the fader may be functionally limited to controlling the amplitude of the signal passing through the channel, through skillful handling it becomes a site for articulating the emotions of the listener.

Conclusion

In his Phenomenology of Perception, Maurice Merleau-Ponty described the body in terms of its topical extension, as existing "wherever there is something to be done."(Merleau-Ponty 1962, 291) Perceptual practices such as riding faders, hearing frequency response curves, and making parts of the mix "sit" require specific ways of placing and orienting one's body with respect to the bodies of other people and things within the studio. The possibility of intervening in a sound is what brings together the bodies of engineers and performers within a mix, and what allows them to anticipate and shape the emotional qualities of a recording. The emotion of a mix gets articulated through embodied sense and gesture. Emotional orders of the voice also get projected temporally, into the imagined body of a future listener. The next chapter examines the ways in which this and other temporal relationships are practically articulated in studio work.

CHAPTER 7: THE TEMPORAL ORGANIZATION OF STUDIO LIFE

The studio is not just a domain of topical ordering; it is temporally organized as well. It comes into being through diverse ways of passing, keeping, marking, and making time. The previous chapter has shown how the studio, like many zones of technoscientific production, is heterotopical, meaning that it is a place where multiple spatial settings coexist and make possible the production of heterogeneous representations. Heterotopias, writes Foucault, are "often linked to slices in time—which is to say that they open onto what might be termed, for the sake of symmetry, heterochronies. The heterotopia begins to function at full capacity when men arrive at a sort of absolute break with their traditional time." (Foucault 1984, 6) Some heterochronies – such as those of libraries and museums, are engaged in the accumulation of time.[68] Others are more temporally precarious, oriented not towards eternality and the overcoming of time, but rather a temporality defined by its fleeting quality, or time "in the mode of the festival." (Ibid 7) The studio straddles these chronological modes, being both a place where temporary performances are made to endure over time and also being a place that, for the client, is only occupied for a limited period of time. In this chapter I consider how studio life is organized around particular ways of enacting temporality. I focus specifically on practices of narration, repetition, and material-semiotic practices of "clocking." I close with a discussion of how studio timekeeping has been contested in a case of copyright infringement.

Studio Time and The Proverbial Voice

The idea of temporality as a collective accomplishment and key element of

[68]Foucault does, however, note that there is nothing essential about these labels, and that libraries and museums, up until the end of the seventeenth century, were not universal and public places but private expressions of individual taste.

sociotechnical action is a well-trodden theme in STS. Work in Actor-Network Theory has aimed to trace "how different spaces and different times may be produced *inside the networks* built to mobilise, cumulate and recombine the world."(Latour 1987, 228) In this view, instances of sociotechnical delegation between humans and non-humans appear as ways in which "time is *folded*." (Latour 1992, 231) Lynch has shown how work in a scientific laboratory proceeds in terms of projects and their constitutive local tasks, the proper sequencing of which are made accountable retrospectively over the course of a writing-up process.(Lynch 1985) Social studies of laboratory work, such as *Beamtimes and Lifetimes* (Traweek 1992) and *The Mangle of Practice: Time, Agency, and Science* (Pickering 1993b) have shown how multiple temporalities are woven together in the course of technoscientific life. Historians of science have helped to culturally and historically situate the various ways that we conceptualize and keep time (Galison 2003; Landes 1983) as well as how practices of temporal ordering are deployed metaphorically. (Shapin 1996; Lakoff & Johnson 1980) Dipesh Chakrabarty has used the figure of the "time-knot" as a way of getting at the "plurality of the present" which characterizes the colonial encounter and its play of power-knowledge. (Chakrabarty 2000, 112) Rachel Prentice has shown how the cultivation of craft practices such as those involved in surgical procedures "makes the body into a temporal joint that articulates past and present conditions." (Prentice 2012, 196)

In the 1960s, American Composer La Monte Young began to build his artistic practice around the observation that the question of whether two notes are properly "in tune" with one another is limited not just by the accuracy with which they coincide within the vertical space of a musical staff, but with the duration over which they are sounded together. Two sine waves of slightly differing frequencies will only reveal their difference over time. The degree to which things are in-tune, then, is as much a temporal as it is a spatial question. Young explored this concept most famously through his Dream House installation in Manhattan, which for several years consisted of sustained tones in simple harmonic relation to the 60 cycle hum of the building's alternating current supply.[69] Young's intervention into musical practice opens up questions of how relationships of in-tuneness (taken both in its specifically musical meaning and also in its meaning of social attunement (Schutz 1976)) can be understood as infrastructurally rooted and temporally unfolding accomplishments.

Studio time is a notoriously slippery phenomenon for engineers – a fact well-captured by a set of "Laws for Audio Engineers" compiled by the Audio Engineering Society. These laws playfully combine the genres of "laws of nature," commandments, jokes, and proverbs for studio life. Several of them specifically articulate the protean and procrustean qualities of studio time:

> Hansch's Law: "Work expands to fill the time available for its completion"[70]
> - attributed to mastering engineer Jo Hansch (Chinn 2014)
>
> Westlake's Law: "The first 90 percent of the project takes 90 percent of the time, and the last 10 percent takes the other 90 percent."
> - attributed to the owner of Westlake Audio systems (Chinn 2014)
>
> Marino's Law: "Everything takes longer than you think it will."
> - attributed to mastering engineer George Marino (Chinn 2014)
>
> Blasingame's First Law: "When operating in the vacuum of a studio, time moves faster than anywhere else in the Universe."
> - attributed to engineer Joe Blasingame (Chinn 2014)

Time flies and there is never enough of it. It is often experienced as a function of expectation, (things take as much time as, or more time than, they are given) and location (the "vacuum of a studio.") Harvey Sacks has argued that proverbs are basically "atopical" phenomena, meaning

[69] Young's colleague Terry Riley once argued, in a discussion of just intonation vs. equal temperament, that "Western music is fast because it is not in tune."
[70] Outside of the audio engineering profession this is sometimes known as "Parkinson's Law."

that their sense and use does not derive from any particular empirical instance. They are, instead, self-contained statements, or instances of "single-sentence packing" that – by virtue of their lack of specificity – seem to maintain their validity independently of the domain to which they are applied. (Sacks 1989) Shapin has elaborated Sacks' insight to the concept of "proverbial economies" as scenes of social knowledge production. He notes the rhetorical power invested in the "proverbial voice," the origin of which is thought to be "either lost in the mists of time or… descended from respected ancient authors."(Shapin 2001, 736) Proverbs, in other words, work because they refer both to 'time immemorial' and 'no time in particular.' Beyond the AES laws, proverbs can be found to govern everyday practical action in the studio. Presented with a sloppy or uninspired vocal take, for example, an engineer may comment proverbially on the futility of "polishing a turd" (or, more elegantly if less plausibly, "make a silk purse from a sow's ear.") Even if the vocal take is ultimately salvaged, the proverb remains valid by virtue of its primary atopical domain of application. As the aforementioned laws illustrate, time is both a tool and a topic of proverbial reasoning in-studio. Studio time, in all of its perplexity, is produced and managed through a variety of practices. The first of these we will consider is the practice of narration, whereby studio workers situate their work within stories about their own career arc.

Professional Narratives

The studio hierarchy, in addition to being a stable spatial distribution of types of labor over time, also serves as a narrative structure for the careers of recording professionals. Lessons in the service-custodial ethic, formed during the internship proces, echo throughout the later stages of the career of a recording engineer. This is especially the case where the studio staff is small. As engineer Lucille relates, "cleaning up" is an enduring concern even when you have graduated beyond literally doing the dishes:

> I call myself a janitor sometimes because, not only do I clean up the studio because I'm my boss' only employee so I do all the intern stuff, but, I definitely am, I did this one record that took about a year and it was like, it was a really fun creative record but it was not done by like very, um, great musicians, like the musicianship was low even though the creativity was high up there. And so I felt like I was just, like, trying to like clean things in the session like the entire year of that.

When dues are paid, they are usually paid in time. Part of my 'orientation materials' when I started working in studio A was a Xeroxed collection of anecdotes and commentaries on what people look for when hiring would-be engineers. This had been part of a course packet Carl received while undertaking a postsecondary recording arts degree. One excerpt, taken from an issue of Mix Magazine, illustrates the centrality of housework to the studio economy. Nashville-based engineer Sara Hughes recounts her initiation into a recording career:

> During the summer before my final semester of college, I managed to talk my way into a highly sought-after internship at one of the biggest studios on Nashville's Music Row. I was excited because I was sure that meant I would find hundreds of opportunities to sit in on sessions with major-label artists and top-notch engineers working on cutting-edge gear. Imagine my surprise one day when I was handed a tube of cookie dough, a spatula and instructions on when to bring treats to the in-session clients… I remember standing there, utensil in hand, thinking about how only moments before, I'd been invited by one of the aforementioned clients to sit in on an orchestral tracking session. I protested the assignment to the studio manager, pointing defiantly in the direction of the main control room and asking why one of the other interns couldn't hand the odious task of baking and serving. I pointed out that the whole purpose of internships is to give future engineers the opportunity to learn. And that's when the revelation hit me: She *was* giving me a chance to learn, and the lesson being offered was far more important than anything I could have picked up by watching an engineer's hands hovering over a console.(Hughes 2003)

Placement within the orbits of studio work involves placing oneself in a professional narrative. In Hughes' case, the example of baking cookies is significant in how it inaugurated her own career as a studio worker. Its status as a "dues paying" or induction story is what makes the task of baking and serving significant to her later work.

Interns, in spite of the mostly unskilled quality of their labor, have in recent decades increasingly come to studios with at least some amount of formal training (Porcello 2004). Formal recording arts training is largely oriented towards developing a vocabulary and basic familiarity with the tools used in studio. In interviews, however, it was common to hear that engineers felt that they had gained the vast majority of their most useful skills on-the-job. Until relatively recently, the training of engineers was a largely informal process, based in apprenticeship and tacit, hands-on learning (Horning 2004; Horning 2013). In addition to a source of unpaid labor, the role of studio interns now largely serves as an extended job interview: a way of testing-out the intelligence, reliability and personability of the intern before bringing them on in a more permanent capacity – as a "runner" for example. This gate-keeping function is increasingly important as recording arts programs seem to continue producing greater numbers of would-be engineers than there are jobs available in a professional recording economy that has been contracting roughly since the late 1990s.

Indeed, a major selling point of formal training programs is that they will often work with studios to place their graduates in intern positions. One program, from which several of the engineers interviewed graduated, had a policy of having its graduates call them once they had moved to Los Angeles (or, less frequently, New York or Nashville), upon which they would be given short-notice gigs at nearby studios. Several engineers recounted living out of their cars for extended periods of time, balancing part time jobs with recording school and interning stints. Harold's story is a typical one:

> I was living in my car and working at [fast food restaurant] and going to recording school. And so I had to work from like 6 to 2 at [restaurant], and then drive from 2 to 3. Start school by 3:30 – 4, and then that'd be like from 4 to midnight. And then I'd rehearse and write music at school 'cause you know [recording school] has tons of musician programs so there's just open studio you can go in and rehearse or record or whatever

> you want after hours it's free after midnight. So I'd do that till about 2 or 3 in the morning, and then drive back, sleep in the [restaurant] parking lot, and then work again 6 to 2. And I did that for 6 months. When I finished school, as part of my school I got an internship at [studio], and that was only part of the school's for credit.

Good interns are quick learners, unobtrusive, enthusiastic, and efficient. Bad ones are more trouble than they are worth. All interns, however, are *temporary*. Stints rarely last longer than a couple of months, at which point the touchy ethical and/or legal question of intern employment status and the need for the studio to provide benefits becomes uncomfortably pressing and a new batch of conservatory graduates are lining up for their turn. Interns may come in with a vocabulary but, at least initially, they are not expected to be able to do much more than get coffee and clean dishes. The studio, unlike the recording school, is where vocabulary words and techniques become *storied*, or situated within a temporal account of studio life. Much of this storying is about the movement from mundane to technical work, a temporal distinction that maps onto the contingent arrangements of front and backstage in the studio. For example, some studios do not even allow interns into the control room. You have to work your way up the hierarchy in order to get the opportunity to do anything close to "recording" proper. Some would-be engineers chart their own path, recording clients on the side and learning their chops on their own time. This often involves "stealing time" in the studio when it is not in use, recording friends bands after hours and being careful not to leave a trace. Over time, mundane 'nontechnical' labor shades into technical labor. This also serves the purpose of maintaining the studio as a self-contained space with a particular and exclusive "vibe." There was a fair amount of ambivalence about this system among the engineers I interviewed. At the same time, however, we can also understand it as part of the substantive training of the engineer and the social constitution of the perceptual and affective skills of the engineer.

Repetition as an Infrastructural Temporality

Susan Leigh-Star has pointed to the difficult-to-notice quality of infrastructure as a reason for attending to its social significance (Star 1999). The standards, structures, and practices that support and configure everyday life are so taken-for-granted as to become imperceptible. For Star this also included the question of "*when* is an infrastructure?" (Star 1999; Jackson 2016) In this section I argue that the when of studio infrastructure is located in ways of doing things again. Repetitive labor within the studio can be understood as an infrastructural practice in that it is the performance of a baseline of monotony over and against which creative action can be accomplished. Though the recording studio carries a particular form of cultural mystique, it is as reliant on acts of routine infrastructural maintenance as anywhere else. The medium and product of much of studio work is temporal experience. Again, the studio is heterochronological both in the mode of the archive – collecting and filing away time – and in the mode of the festival – a place rarely and only temporarily visited, often as an escape from mundane temporality.

Try, for example, saying "check" out loud continuously for several minutes, uncertain as to what connections are being made in the other room, who is listening to your voice, and how it sounds to them. It was in this moment that I began to truly understand the injunction given to me by Carl on my first day in-studio: "you can take notes and ask questions, but don't make them repeat themselves." Repetition is serious business in the studio. Making people repeat themselves – either through looping or repeated takes – is a major part of the engineer-client working relationship and must be handled with care. Hearing one's performance in a loop can be something of an alienating process, and managing this effect within the social context of the studio is a key affective and technical skill. Like all studio skills, it depends on the social role of the actor and this role situates the act within an economy of creative and technical labor. In short, the repetitive labor that undergirds accountably "creative" or "artistic" or "original" effort higher up on the studio hierarchy thus falls in large part to the repetitive labor of the intern.

Repetitive labor figures heavily in the process of setting up a session. Once the cables have been deployed and the connections have been made, there is a period of checking to make sure each channel is working properly and connecting the performance in the live room to the board in the control room. Typically the assistant will man the console while the intern or runner will scratch lightly on the screen of the microphone, one by one, in order to provide a noise that will establish a successful connection. If more than a scratch is needed, the intern will be called upon to check the mic through speaking or playing an instrument into it. Pinch has used the example of the mic check as a way of illustrating the practice of "current" testing. Testing, Pinch suggests, can be understood as "projecting" technological action. Technological testing, understood in this way, constitutes the establishment of a similarity relationship across time and space (Pinch 1993, 29) – an instance of testing is significant in that it is supposed to stand for other instances in one or another crucial aspect. Furthermore, the act of repetition as evident in mic checking practice also serves as a staging ground for studio roles, for the sorts of repetition we value and those we do not, and the varieties of creative action available to an individual within the socio-technical assemblage of the studio.

Beyond this mode of repetitive monotony, the engineer-in-training is asked to attend to a wide range of temporalities. Studio time, to a large extent, is of the "hurry up and wait" variety. Corraling musicians and studio personnel is a notoriously difficult task. Once everyone is assembled, time is swallowed in large gulps by the tasks of setting up, mic-checking (or "getting sounds"), and troubleshooting. Once takes have begun, they too require a good deal of patience on the part of those not doing the performing. Musicians take long breaks. Vibe must be established, and vibe takes time. While working in Carl's studio, I was specifically instructed to light incense and set the mood lighting 45 minutes prior to the client's arrival. Sessions at Studio A would often start late in the day and run until the early hours of the morning, and studio staff would need to be able to match the pace (fast or slow) of the client, and help to keep things moving while simultaneously avoiding "rushing" the process. All these practices are directed – in Harold's words – towards "making everything fine" for the client and facilitating the capture of a lively performance.

At the other end of the temporal spectrum, the studio professional must contend with the extremely brief periods of time made available in increasingly higher resolution by the software interface of the DAW. Here previously recorded events can be parsed down to individual excursions of the audio signal. These segments can be set to repeat endlessly (or "looped"), manipulated and scrutinized in any number of ways as they unfold in a virtually endless present. Looped repetition becomes a means of staging for the benefit of the apprentice engineer the variety of audible differences that were not apparent beforehand. The relentless sameness of a loop becomes a staging ground for skilled differentiation. Just as it can threaten to alienate a performer from their own voice, loops permit engineers to take ownership of the performance's hidden artifacts.

"Clocking" as the Material Management of Time

Work in phenomenology and phenomenological sociology has made much of the distinction between "inner" time (or duree) and "outer," objective, or "clock" time. (Husserl 1964; Bergson 1950; Schutz 1976) Schutz, for example, described the mutual tuning-in relationship of musical performance as involving the bringing together of inner and outer time.

Useful as this work is, the implied concept of a singular monolithic "clock time" does short shrift to the rich inner lives of clocks and the relations people have with them. The clock is one of the many black boxes in the studio that need to be opened up. If we take clocks, like any technology, to be products of social negotiation, we need to be able to talk about "clock times" or "clocking" as a material-semiotic social activity that brings together people and technological systems. Clocks are material things. Each clock marks time in its own unique way. Take, for example, what the AES terms "Golden's Law": "A man with one watch knows what time it is. A man with two watches is never sure."[71] What we find here is a proverbial expression of the idea that there are as many clock times as there are clocks. Clocks, moreover, are part of systems that engineers understand as, in some sense, always somewhat out of control. As mastering engineer Bernie Grundman puts it, "Under the most carefully controlled conditions of pressure, temperature, humidity and other variables, the system will perform as it damn well pleases."[72]

A brief digression into an earlier material-temporal order offers a broader perspective on what clocking can mean: Alain Corbin has shown how the regular sounding of the village bell, an early sound technology as well as an early time-keeping technology, helped to create a sense of emotional and territorial identity in 19th century France. (Corbin 1998) Bells also served as important technologies of creating a collective sense of time, as the etymology of clock (from "clocce," or bell) makes evident. The unmarked objective "clock time" that Husserl and Schutz oppose to lived or durational time, has its own material history, as no clock is a purely objective timekeeper. The bells that were used to structure temporal and emotional life in European villages were often named, baptized, and endowed with other human characteristics.(Illich 2000)

[71] Also known as Segall's law. Attributed to mastering engineer John Golden (Chinn 2014)
[72] Grundman's law. Attributed to mastering engineer Bernie Grundman(Chinn 2014)

It is important to remember that even in the studio a clock continues to be something material, which rings out and constitutes a spatial order.

The way in which digital clocks ring out to produce order can be understood as an extension of what Natasha Myers has termed the “materiality of the digital” in the field of protein folding crystallography models. (Myers 2008) Though Myers is referring in large part to how the use of virtual modeling systems can give researchers a deeper material sense of the proteins with which they work, her intervention also invites thinking about how the digital representational system is itself a material arrangement that is always subject to emergence and contingency. Most digital clocks are driven by crystal oscillators, and the material details of the oscillator circuit (including the quality of the crystal used, the type of electricity applied to it, and its relationship to other components in the circuit) can shape the way in which it keeps and signals time.

Often the materiality of clocking involves questions as simple as whether a clock is visibly present at all. The relative lack of visible clocks, combined with the tendency for studios to lack windows, effectively creates a studio temporality that is easily decoupled from that of the “outside world.” As Bob puts it:

> The big thing about studios is you never put a clock in the studio ‘cause then they don’t know how long then how much they owe you, so it’s real hippy-dippy. Like “no worries, this isn’t even a business” and then it’s real awkward afterwards ‘cause it’s like ‘hey about that check’

In fact Studio B did have a small digital clock on the console. Studio A, however, had no easily visible clock at all. Most people in the studio carried timekeepers in their pockets in the form of smart phones, but this meant that checking the time was largely a practice of individual investigation: you could not point to a clock, but you could glance at one in your hand. For an intern, at least, checking the time (especially in the form of glancing at your phone) is taken to be bad form, an indication of impatience and disinterestedness. Gradually over the course of my internship I came to mark time in different, sometimes unusual ways. In the clock-free studio, for instance, the Power Center unit could be relied upon to roughly mark the beginning and end of the normal workday with its voltage intake display. The number indicating available voltage would fall or rise as a nearby office building began or ended its business day, and it would fall to the power supply to compensate for these fluctuations. Much of the studio equipment, older analog pieces especially, would behave differently depending on the voltage available to them.[73]

In spite of the lack of visible timepieces, the studio is in fact filled with clocks. It is just that most of them are dedicated to informing nonhumans. Clock circuits inside of the Analog-Digital converters, analog drum machines, and midi equipment each keep their own time. Some clocks are more reliable than others, and they can drift apart if they are not coordinated regularly. Studio B runs off a pricy Dutch-made master clock, which is designed to have extremely precise and "clean" timing signals. In order for clocks to work properly they have to be put into the proper hierarchy, or as Lucille puts it, you have to identify "which one is the master and get everything else to understand that that's the master." When clocks are not synced properly, especially analog-digital/digital-analog (AD/DA) converter clocks, which are used to translate analog signals (e.g., the voltage fluctuations from a microphone's diaphragm) to digital information (e.g., the quantized and binary-expressible sample values of a digital audio file) sample rates will not be lined up and the signal will degrade. The imaginary mathematical

[73] Some musicians prefer the sound of "voltage-starved" equipment such as dying batteries or outmoded or mismatched power tubes. Certain guitarists have been known to demand that their guitar techs carry nearly-dead nine volt batteries for use in his wah pedals. Others employ voltage attenuators to reliably produce low voltage effects on their pedalboard.

perfection of the digital, so often taken for granted "for all intents and purposes," vanishes when pieces of equipment malfunction or simply fail to get along with one another. As Harold describes, digital and human temporalities can also sometimes collide:

> Like ok the drummer for [band name], we had a drum machine going, and he was playing with the drum machine so it was two tracks, drum machine and live drums. And his kicks were a little bit late, like, every once in a while. I was like dude it's fine. I mean and he wanted me to correct them but he didn't want his drums chopped up. And I was like ok everybody just go on a lunch break, I'm gonna get everything kinda tightened up here kinda nice, and he sat with me 'cause he wanted to see what I was doing, and he's like oh man that sucks I'm off. And you just try to make em feel comfortable. Like dude that's a drum machine from the 80s those clocks are fucking out of control there's no way you'd ever play to it.

In this case it fell to the engineer to negotiate hierarchies of timing, skill, and expressive authority among session participants (human and non-human alike) in order to make the process run smoothly. Harold's move to cast the drummer's playing in terms of temporal errors on the part of the drum machine offers an interesting counterexample to the "ideology of control" (Noble 1978) identified by David Noble in his account of numerically controlled machine tools, which appears in the "variation" parameter of Auto-Tune. By endowing the drum machine with a history (pointing it out as a drum machine from the 80s) and describing its own clocks as "out of control" Harold plays with the ideology of control in the context of drum timing.

The role of the drummer in this particular recording session is also relevant to how time is kept: In this case he is playing a supporting role to a well-known singer-songwriter, for whom the project is named. The songwriter, on his previous releases, had performed most if not all of the instruments, often including the use of a drum machine. The drummer has been brought in to play along with the drum machine – a type of thing that, in the studio's temporal economy, usually has a stronger claim to mechanical objectivity than a human drummer would. The result of this arrangement is that the drummer is expected to both play well with the drum machine and, since he is not the featured artist in this project, to be modest in his own “human” contributions to the performance. The drum machine from the 80s has itself been chosen for this session not because newer drum machines with better clocks were not available, but because it brings a distinct sound, energy, feel, or vibe to the recording. The artist in this case has been using drum machines of this sort for many years, and while it might well have been a practical necessity at the start, it is now a part of his signature nostalgic and eccentric sound. The drummer is stuck, meanwhile, in a blurred domain between human and mechancial repertoires of performative evaluation. Where conventional studio stage management practice calls for Harold to send people off to lunch and fine-tune the drum timing behind the scenes, or during “off time,” in this case the drummer’s perplexing position in the ensemble compels him to stay behind and attend to the editing of his own performance. With the drum tracks in front of him and the drummer anxiously looking over his shoulder, Harold has to finesse the protools track as well as the drummer’s self-presentation.

The idiosyncrasies of older equipment can provide opportunities to re-examine the criteria of what counts as a good mix, an emotionally expressive performance, or an instance of careful and tasteful listening on the part of the engineer. Here the age of the equipment becomes a factor in whether its behavior is considered an instance in need of repair or a positive contribution to the process. For Harold, older equipment encourages close listening and facilitates emotion, precisely because of the vagaries of its operation over a life of service:

H: I think it’s nice because working on old equipment really makes you listen, and have a more emotional quality about it. Because you know, typically it’s not perfect. Like sometime this console has a tendency to be right-heavy

OM: Right heavy? So the right channel?

H: So it can mix, you know, the stereo bus you know? So I'll be mixing and it feels good and I'll be looking at my meters or I'll be printing and thinking fuck that right channel looks a lot hotter than the left channel. But then so I'll ok I'll just double check I'm not doing anything stupid, you know nothing's double bussed, or I didn't knock something up last minute, and then I just really have to kind of listen and make sure it still feels good. And it is right heavy. And a mastering engineer will call me up and say hey I'm looking at your wav files and all your right channels are heavy. And I'll say did you listen to it? Sounds good that way.

As with musicians, equipment acquires experience and gains new abilities with age. Often these abilities take the form of "bugs" that become "features." The Roland TR-808 drum machine sound, for example, was decried early on because it so poorly imitated the sound of an acoustic drum kit. Now it is that very distinctiveness – the way it separates itself from "real drums" – that makes the 808 such a classic instrument. Engineers often speak about "legendary" pieces of equipment, which may be high-maintenance and behave in ways that are not fully accountable, in magical terms. As Seattle-based engineer Randall Dunn puts it, "opinions about gear are really beautiful superstitions."(Kostelnik 2016) These beautiful superstitions are often articulated in ways that position the piece of gear in a particular point in time, or situate it outside of time altogether. Gear might be valued on the basis of its sheer newness, its ability to capture a particular era, or its "timelessness."

Recording software especially tends to consist in the precarious "temporality of the festival," always needing to be patched and updated, constantly shedding and rejecting its history as a way of avoiding obsolescence (or cultivating it in its competitors.) The difficulties of archiving digital audio, which require machines and software that are constantly passing out of currency, are well known. Other elements of the studio are used because they seem to accumulate the past and are used as ways of re-producing the past in sound. Steve Albini, and other engineers who work exclusively in tape are sometimes criticized for being "stuck in the past" or fetishizing the sound of an obsolete format. Albini's stated rationale, however, is actually future-oriented: magnetic tape and the machines that play it do not degrade or go obsolete as do DAW and digital audio formats.

Timing differences might be miniscule, but they are able to accumulate as they run in parallel. Discrepancies in clocking are especially problematic in processes of analog to digital or digital to analog conversion (ADC or DAC, respectively) where the rates at which sounds are sampled or played back need to be closely matched. Discrepancies in clocking can produce a phenomenon known as "jitter." Mastering engineers, who are tasked with producing extremely polished versions of stereo mixes before they are send out for production and distribution, are particularly concerned with the potential effects of jitter. Grammy winning mastering engineer Gavin Lurssen explains how he goes about keeping clocking problems from creating artifacts in his work:

> When we record it back to digital, we have a raw file. That file needs to be scoured for any little tics and clicks because, in today's world, most of what comes in needs cleaning up. Very rarely, if ever, does it happen here that we add tics and pops, because we pay such close attention to our clocking and all that. But in the mix environment, particularly in the new generations of people who are working with plug-ins and all that, all kinds of clocking issues can happen. You're also in an environment where it's so revealing, and they've come from an environment of working on headphones in a bedroom somewhere. We start to hear all this stuff that people haven't heard. The music has to be scoured for that. We have some digital tools that can clean that up. Then we have to make it all faded and nice. We have to then go through sample-rate conversion. If we can, we like to work in high-res, so we master and record a high-res file and then everything else gets made from that. We've just recently gotten comfortable with sample-rate conversion. We never were before, but it's now become a part of our reality. That process has to happen. Any time you do something with a file, it has to be listened to. If you're doing a whole album and you do certain things to it, you have to calculate all the things you're going to do. Every time you do something, it needs to be scoured. It's actually an incredibly time-consuming procedure in production to get that file. Once the mastering is done, once the feel of the sound, the tonal structure has been balanced and vibed out (i.e., mastered), there's a lot of work to be done on that file. It takes a lot of time. (Crane 2016a)

Technical engineer Ted Smith describes his experience of jitter (or lack thereof) as what he calls "touchy-feely" – if people are tapping their toes, for instance, it has to do with all of the timing being perfectly lined up because jitter is not present. (McGowan 2014) Not everyone agrees that jitter is a "real" problem, however. Ethan Winer, a specialist in architectural sound treatments and self-styled crusader against "voodoo" techniques and beliefs in audio engineering, dismisses jitter and related temporal artifacts as being real but practically imperceptible because they occur below the threshold of audibility.(Winer 2014) Tellingly, the debate over the existential status of jitter is not about whether it is a real physical phenomenon (even Winer allows that clocking issues produce jitter effects in the signal), but instead whether it is a phenomenon that is genuinely perceptually available to even the most skilled listener. Winer's debunking argument concerns the discrepancy between the decibel level of jitter-based noise and the decibel level at which humans have been shown to be able to reliably perceive differences between sounds.

In his book *Mastering Audio*, mastering engineer Bob Katz begins his discussion of jitter with the caveat that "in this topsy-turvy digital audio world, sometimes you have to abandon the evidence of your senses and learn a totally new sense, but one that is fortunately based on well-established physical principles"(Katz 2007, 227) He describes the symptoms of jitter as "blurred, unfocused, harsh sound, reduced image stability, loss of depth, ambience, stereo image, soundstage, and space."(Ibid) Interestingly, he notes that digital tuning programs are a rare example of digital signal processing that is not impacted by jitter. He writes:

> With Digital pitch processors such as Auto-Tune(tm) the explanation is a bit confusing, but these are not affected by jitter. Pitch processors are not state machines; due to their randomizing algorithms many of these repitching processors produce a different output from the same piece of music each time they are run. (Ibid 231)

Because they are designed to never run the same way twice, pitch correction tools do not pass along the temporal discrepancies associated with jitter. On the other hand, however, this means that tuning effects need to be "printed" or committed to the track while other effects can be allowed to run in real time every time the session is played back. Carl puts it this way: if there are many plugins running in the background while Auto-Tune is doing its work, it will have a noticeably different outcome than if it was the only plugin running.

Jitter is an example of how the medium through which sonic information is stored imprints itself onto the information itself. Time-based artifacts of audio storage are not unique to digital recording, however. Until the late 1990s, magnetic tape was the go-to medium for mainstream popular music production. Tape can be thought of as a "literal" spatialization of musical time.(Weeks 2002, 363) It is also a important element in what Jonathan Sterne has called sound technology's "compression history" – whereby sonic recording and reproduction is not simply about accuracy of representation but also of increased efficiency of storage and transmission.(Sterne 2012) One of the reasons tape is useful is that it can, for example, be wrapped around a reel and stored indefinitely. Compression, as Sterne argues, always entails artifacts. One major artifact of magnetic tape, as Thomas Porcello has elegantly described in his ethnography of an Austin recording studio, is the "print through" that occurs when magnetic interference causes sounds to leak across layers of the tape reel.(Porcello 1998) The audible effect of print through is to have sounds appearing faintly several seconds before or after they 'actually' happen on tape. Practices such as storing tape "tails out" are deployed in order to minimize the audibility of this sort of material boundary jumping. Jitter is not the only type of temporal bleedthrough in digital audio. MP3 compression codecs, for example, are notorious for having “pre-echo” artifacts, where one is occasionally able to hear echoes of sounds before they occur. (Sterne 2012)

Neither studios A nor B use tape primarily, though both reserve the option to bring in a tape machine in case it suits the client’s needs. Often tape will only be brought in at the end of the mixing process. The entire finished mix may be printed to tape and then re-digitized, in order to lend a touch of “analog warmth.” On the rare occasion that tape is being used as the primary recording medium, new considerations come into play regarding the spatio-temporal boundaries of recorded sound. “Capric Oath” is one of these clients. Their sound draws heavily on early heavy metal bands such as Black Sabbath, so they opt for the studio’s Studer multi-track tape machine over its Pro-Tools rig in order to achieve a “vintage” sound. Studio B’s assistant engineer Andrew calibrates the Studer in anticipation of the band’s arrival, cleaning and demagnetizing the heads and aligning it with the help of a tape of standardized test tones. Sometimes the bias rate on the tape machine needs to be adjusted, so as to avoid what might be described as “analog jitter.” Whereas many bands will opt to record instrumental tracks all once, in order to get a “live” feel and save studio time, Capric Oath decide to put things together track by track. After laying down the basic drum tracks for their first song, the guitarist sets up his amplifier in the downstairs vocal booth.

Andrew mics the amp with a Shure sm57 and an sm7-B, positioned right next to the speaker cone. He runs an instrument cable from the amp’s input jack to a direct inject box, from which an XLR cable connects it to another box in the control room. Here the guitarist sets up his board of effect pedals, hooks up his guitar, and begins tuning. His playing runs through the pedalboard, downstairs back to the iso booth and play through the amp, which will be sent back up to the booth with the help of the microphones. He wears headphones to monitor his performance as well as the previously recorded drum tracks. Whereas it is important in digital recording to avoid sending too high of a signal into the computer, lest it overwhelm the audio converter and cause undesirable digital “clipping,” tape can take louder than intended signals while still sounding good. In fact, “hitting the tape hard” is a time honored way of achieving the distorted sound that the band is going for. When using this technique, however, it is important to keep in mind that the tape consists of actual physical strips of emulsion running in parallel, and that overdriving a single strip can mean that it will bleed over into an adjacent track. Because of this, Andrew is careful to place the guitar tracks far enough away from the drums to avoid any bleed-through. Tape is desirable not just because of its sound, but because of how it structures the session temporally. Tape requires a performer to “commit” to a performance in a way that digital recording does not, because when you re-do a take you need to either record over the last one or start on a new track. Digital recording is described as “non-destructive” because it is possible to undo a take or record multiple takes and put off the decision of which one is best until later. When tracking to tape, the irreversibility of the process is emphasized and time becomes destructive once again.

When tracking vocals for bands like Capric Oath, Harold – a metal vocalist himself – sits quietly at the recording console while the singer is cordoned off in the vocal booth on the other side of the studio. They do a take, and after a quiet pause Harold sets up another take and hits record again. I ask him about that moment afterwards and he explains:

> “Sometimes it’s just that breath that you take that you’re breathing for the other person in there, and it’s just this kind of symbiotic energy thing that’s happening. And I think that flow and kind of being really in tune with their energy helps.

Harold's technique for tuning-in during the tracking process here becomes one of literally conspiring, or breathing-with the vocalist. Being in tune or in sync is equally important once the editing process begins.

Studio life is also structured by what Jackson et al have referred to in the context of scientific research as "collaborative rhythms." (Jackson et al. 2011) which are embedded within local practices and aligned at various institutional levels through planning work. (Steinhardt & Jackson 2014) The work of reconciling plans and rhythms is especially important in the case of major label projects, wherein the recording process entails and responds to the logistical complexities of established promotional and touring cycles. High profile pop projects follow annual "hit cycles" typically comprised of the release of the first single, media blowback, the second single, touring the album, and award shows.[74] Independent acts have smaller promotional networks and less cumbersome touring infrastructure, but are more likely to have to allow for day jobs and limited budgets. Many engineers – in Nashville especially – have an ambivalent relationship with the glut of one-off Holiday season album projects that provide reliable downtime employment but involve recording Christmas standards in late-summer.

Temporal categories, such as era or decade, also serve to index particular sounds or feelings. Engineer Mick Guzauski, for example, describes his mix for Daft Punk's 2013 album Random Access Memories as being the sound of the 70s and the contemporary.(Pensado 2013a) For him this meant the use of certain pieces of equipment and techniques from that era, as well as a general aesthetic orientation. Decades become useful shorthand and mnemonic devices for

[74] See, for example, Tony Maserati's discussion of working on Robin Thicke's "Blurred Lines" (Pensado 2013b)

specific hard-to-articulate sensory and affective qualities. Bob says that whenever someone asks for a "warm" sound he thinks of the 90s. Many engineers make a habit of editing and correcting performances while the performer is away – having dinner, for example. Work carried out on someone's performance right in front of them is potentially difficult. Towards the end of my time at studio A, I ask Carl if he would mind giving me mastering[75] advice on a recording a band of mine I had made back in New York. For the first time since I had arrived, it was my own performance on the table. We had listened together so many times but this time was different because I was also implicated in the recording itself. Carl offered to take a quick stab at mastering the recording – a "fake master" as he called it – and asked if I wanted him to wait until I was not around to do it. I took him up on it.

"Blurred Lines" and The Thickeness of Studio Time

In the previous chapter I discuss the use of reference tracks, or "refs" an aspect of list-making in studio, specifically as a way of translating a recording into an itinerary of correction. Refs can also be understood as objects of temporal organization. Through the use of refs, engineers articulate the sounds of their recordings with past recordings that have already proven successful. This practice – which aesthetically tethers a recording session to strategically chosen genre and artist histories - is largely understood as a way of maintaining perspective and avoiding being "sucked-into" the minutiae of the mix. Refs also offer a way of replicating a desired sound – often that of a particular era, genre, or artist. Working off of well-established

[75] "Mastering" is the process through which a finished mix is fine-tuned, sequenced, and generally polished up before it is ready for the listening public. It involves making all of the separate tracks of an album sound like they fit together. Many mix engineers joke that they try to make their mixes "mastering proof," or good enough that even a mastering engineer cannot mess them up.

refs ensures that the result will be familiar, and that it will "track" with listeners because of that familiarity. Hewing closely to the sounds of previously successful recordings is a strategy for making "safe" mixes that will not raise objections from label representatives or other stakeholders of the recording project. It is a way of testing a mix against one that has been known to "stand the test of time."

From the standpoint of copyright law, however, reliance on refs is an increasingly hazardous proposition. The trial surrounding Robin Thicke and Pharrell Williams' Grammy nominated hit single "Blurred Lines" (2015), which was ordered to pay millions the estate of Marvin Gaye for copyright violations, shows how much a recording's significance depends on the ways in which it is temporally articulated. Looking at how the litigation process retrospectively deconstructed Thicke and Williams' recordings in terms of evidential questions of "what you knew and when you knew it," and the ways in which the artists resisted this deconstruction, shows how the effect of a credibly original musical voice depends on the temporal elaboration of the recording process. The way that the recording process is retrospectively constructed in a legal context helps put into relief the aforementioned ways that engineers and their clients deal with time in situ.

Through the pre-trial taking of depositions, Thicke and Pharrell were made to reconstruct their collaborative production of the album Blurred Lines, particularly its title track. In this process the plaintiff's attorneys worked to establish that Thicke and Williams knowingly borrowed from Gaye in their authorship of their music. Both defendants were extensively questioned, for example, about how young they were when they first heard Marvin Gaye and the songs that they were accused of stealing from. Thicke and Williams had originally shared writing credit on the song, though in court they both testified that Williams had in fact been the sole songwriter. This strategy had to do in part with the plaintiff's reliance on numerous public statements by Thicke that he had deliberately gone into the studio wanting to make something similar to Gaye's "Got To Give It Up." Thicke claimed that none of these statements were made while he was sober, and that he had exaggerated his own role in the song and the role of Gaye in the song's inspiration for the sake of having something to say in interviews. Williams explained that sharing songwriting credit was customary, even though he had in fact written the entire song. Whereas the defendants thus had an interest in distancing Thicke temporally from the reconstruction of the writing process, the plaintiffs worked to place him more squarely within it. At one point, for example, the plaintiffs noted an inconsistency between engineer Andrew Coleman's testimony – in which he claimed that Thicke did not do his first vocal take on "Blurred Lines" until after 10 PM on the day that the song was being recorded – and a time stamp of 6:14 PM on a vocal take from Thicke. (Chelin 2015) By eroding the temporal boundary Coleman set up between the writing of the song and the recording of the vocal parts, the plaintiff's attorneys sought to place Thicke more deeply within the authorship of the song (or, at least, paint the producers of the song as unreliable accountants of their own time in-studio.)

Williams' deposition, meanwhile, proceeds as a contest concerning the ways in which his experience with Gaye's music and the production of his own music can be temporally accounted for. This dynamic can be captured in three specific moments during the questioning process. In the first, the attorney asks whether Williams was influenced by Marvin Gaye. Williams responds "He's an Aries. I respect him." The plaintiff's attorney works to highlight Gaye's precedence with respect to Williams by articulating the question of influence as a matter of whether or not an event had occurred in the past – whether or not Williams "was influenced" by Gaye. Williams responds in the present tense, effectively refusing the question's temporal structure. By referring to Gaye's astrological sign, Williams (also an Aries) counters the questioner's precedential mode of accounting with an horoscopic one, implying that he and Gaye are, in a cyclical sense, coeval with one another. Williams resists the question by responding in terms that do not fit within the questioner's temporal assumptions.

A second example can be found in the way Williams accounts for his own past statements with respect to Gaye's influence. At one point the attorney asks Williams the following:

Q: Did Marvin Gaye's "Got to give it up" ever cross your mind at all at any time while you were creating blurred lines?

A: No

The attorney then reads an excerpt from an interview where Williams says that for the song's percussion parts he was trying to pretend that he was Marvin Gaye recording in a Nashville bluegrass context. The questioner tries to nail this down as an inconsistency in Williams' account. Williams proceeds to maneuver out of it:

Q: You say in this interview "I was trying to pretend that I was Marvin Gaye" so I guess Marvin Gaye did in fact come into your mind when you-

A: You asked me about "Got to Give it Up"

Q: I asked you about "Got to Give it Up"?

A: Yeah you asked me about "Got to Give it Up"

Q: So Marvin Gaye came into your mind when you were creating blu, uh, "Blurred Lines" but-

A: No, when I look <u>back</u>. When I look back.

Q: Do you see here anywhere where you say "when I look back"?

A: No no no I'm telling you. I'm answering you.

Q: You said that "I was trying to pretend that I was Marvin Gaye"-

A: The feeling of the music

Q: Hold on. When you were creating "Blurred Lines" were you trying to pretend that you were Marvin Gaye?

A: At that particular time, no. But as I look back I feel that feeling.

Q: You did say in this interview however that when you created "Blurred Lines" that you were trying to pretend that you were Marvin Gaye. In the interview, correct?

A: And I'm telling you that was, that's not accurate, and I did say that.

Q: You did say it though

A: I did say that. (Williams 2014)

Competing ways of temporally accounting for the song's production can be tracked in the shifting tenses and voices of this exchange. The attorney opens by attributing, in the present tense, Pharrell's previous account of his earlier work on the song: "You say… 'I was trying to pretend I was…'" Williams responds by revising the attorney's account of the previous question, say that it was, in fact, specifically about "Got to Give it Up." The attorney's attempt, in response, to nail down the more general existence of Marvin Gaye within Williams' mind during recording is met with yet another temporal shift, with Williams adding the qualification of "when I look back." The attorney again attempts to bind Williams to his previous statement by asking, in the present tense, whether he sees "when I look back" in the transcript, to which Williams responds by distinguishing his own present position from the previous one. The plaintiff attorneys work to extend Williams' interview statement about Gaye's role to the broader process of the song's production, while Williams works to quarantine that temporal extension of his statement both into the earlier moments of the song's production as well as into his later moments of self-accounting.

A final example from Williams' deposition involves an apparent attempt on the part of the plaintiff's attorney to establish that Williams is not able to read musical notation. Williams' earlier statements concerning the difference between his and Gaye's work had involved references to differences in time signature, types of musical scale employed, and other differences that are apparent in the sheet music. The plaintiff's attorney attempted to discredit these claims of difference by requiring Williams to define terms such as "6/8 time signature" or "pentatonic scale." When Williams refuses to provide definitions, the attorney shows him a piece of musical notation and asks him to identify the written note. Williams repeatedly responds to this request by saying "I'm not comfortable." The attorney responds by asking if Williams means that he is unable to read the note, to which Williams responds again that he is "not comfortable." While the attorney's questioning is oriented towards establishing, as a matter that extends across time and circumstance, Williams' general musical illiteracy, Williams insists on limiting his response to the present situation. By insisting that he is "not comfortable" he implies that his musical authority could in fact be established in a different, more comfortable, situation, and that it is the specificity of the questioning process that keeps him from naming the note. Though the deposition taking process was nominally oriented towards questions of getting at the true details of how "Blurred Lines" was produced in-studio, it is better understood as a game wherein each side works to accomplish a temporal account according to a specific situation within a legal proceeding. Hypothetically speaking, if Thicke and Williams had made the strange decision to produce accounts of the song's production as it happened, regularly reflecting upon and meticulously noting their specific influences and intentions from moment to moment, such an account would likely be markedly different from either the one that the plaintiffs are pursuing or the one that the defendants offer in response. Looking at how accounts of music production work are retrospectively negotiated within an overtly agonistic legal process helps to illustrate exactly how temporality is typically *not* approached in-studio.

The Gaye family's attorneys also retained the services of two musicologists, Harvard professor Ingrid Monson and professional musicological expert witness Judith Finell. They were asked to analyze the similarity between "Blurred Lines" and Gaye's "Got to Give it Up" (1977) as well as the similarity between Thicke's "Love After War" (2011) and Gaye's "After The Dance" (1976). Their analyses, as accomplished interactionally over the course of their testimony under questioning, drew specific connections between Gaye's work and the more recent work of Thicke as a way of establishing a causal connection between Gaye's earlier compositional decisions and those claimed by Thicke and Williams. The plaintiffs structured their case in terms of eight "constellations" of similarity between the compositions, none of which would have been decisive in itself, but which allegedly intersected temporally to an unusual degree. One of the more important constellations consisted of a twelve note vocal passage, which the plaintiffs designate as the "signature phrase" which "Blurred Lines" and "Got to Give it Up" have in common by virtue of four sub-elements[76] Finell noted that the passage occurred six times in Blurred Lines. Monson drew a similar connection between the vocal "hooks" in "Love After War" and "After the Dance," noting that Gaye's hook repeats nine times and takes up 55 percent of the song, while Thicke's hook appears "seven and a quarter times or about 47 percent of the piece."(95-6) By breaking down the compositions in this way, and

[76] These elements are that A: each song begins the passage with repeated notes, albeit not the same ones. The mere fact of repetition is posited as musicologically significant. B: a series of six consecutive notes, of which five are identical in their scale degrees, C: a series of six consecutive notes with identical 8th note rhythms and D: the use of melisma – or the articulation of several notes within a single syllable in both songs. (Exhibit G, 67-70)

constructing structures of comparison that allow similarities to be drawn at multiple levels of musical form, the expert witnesses produce what Goodwin terms "coding schemes" and practices of "highlighting" in the production of their professional stance with respect to the musical works. (Goodwin 1994) Whereas the deposition-taking and temporal interrogation of the song's production in-studio focused on the song as temporally emergent in ways that implicate more or less culpable authors, the musicologists sought to locate similarities in terms of the temporal elaboration of the melodies themselves. The implication is not that Thicke and Williams sought to capture the precise proportion of the hook with respect to the composition, or to identify and reproduce a set of constellations. Rather, the musicologist's professional vision (and professional audition (Porcello 2004)) reorganized the songs temporally as a collection of components which are relevant in how often they repeat, how they rise and fall, and how they are rhythmically distributed.

Digital pitch-time manipulation also played a part in the presentation of the case for copyright infringement. The plaintiffs prepared several "mashups" of the two songs, which involved superimposing the vocal and instrumental elements of each song on the other. In one case, Monson pitch-shifted the accompaniment of "Got to Give it Up" so that it would match the key of Thicke's "Blurred Lines" vocal performance, which was played on top of it. The defense argued that this artificially produced a similarity between the recordings where there was none. The plaintiffs responded by claiming that pitch-shifting as a "standard procedure" in musicological analysis. By situating a change in the pitch of a recording as an analytical rather than a creative practice, and by anchoring it within her expertise as a musicologist, Monson was able to highlight a putative similarity between Blurred Lines and Gaye's composition.[77] Articulating the boundaries of musicology as a knowledge making practice, in this case, became relevant to how record producing as a creative practice is legally sanctioned.

Conclusion

A useful comparison can be made between the question of the appropriate use of studio interventions in music production and questions of legitimate representation in scientific publication. Emma Frow has identified several key tensions within conversations concerning the production and manipulation of images for scientific publication. For example, the increasing availability and ease of use of image manipulation and cleanup software such as Photoshop has troubled the connection between the elegance of a scientific image and the skill of the scientist who produced it. Attempts to "draw a line" between acceptable and unacceptable modes of visual manipulation and presentation, Frow argues, make visible the negotiation over the proper conduct of scientific trust, authority, and objectivity in the digital era. (Frow 2012)

While Frow shows how image making authority gets parceled out in terms of appropriate and inappropriate techniques, the case of music production shows how important temporal articulation is for the construction and maintenance of creative and technical authority. An engineer's perceptual skill, for example, is deeply related to how they understand and perform their own career narrative. The reliability of a studio as an apparatus of representation and circulation of a client's voice depends on repetitive infrastructural work of testing and projection. The management of electrical signals and musical performances involves negotiating multiple

[77] Interestingly, the plaintiffs were not able to make use of the actual recording of "Got to Give it Up" because the copyright claim was based on a piece of sheet music for the song, rather than the recording itself.

material temporalities between diverse human and non-human timekeepers. The peculiarities of studio time become all the more apparent when they are held to account within a separate domain of meaning.

Reconciling these diverse modes of timekeeping and the productive ways in which they conflict means understanding them polyrhythmically. Writing about the relationship between infrastructure, repair, and temporality, Steven Jackson argues that

> Attending to breakdown points us toward the active and ontologically productive nature of ruin, and the irreducible presence of embedded materialities with rhythms and propensities all their own, which can only ever be sometimes, and for a while, slowed, arrested, and aligned. (Jackson 2016, 173)

Reading heterochrony in terms of repair adds nuance to the concept of a sphere of action marked by a “break” with conventional time. In place of a single disjuncture between the studio and the outside world, there appear instead to be a diverse ecology of temporal tears and partial attachments both within studio life and between the studio and the world around it. The leaky syncretism of studio time is simultaneously a continuing project of mending and a basis of creative expression. No wonder the makers of Auto-Tune were so concerned with time management, and the makers of Melodyne so interested in making it disappear.

CONCLUSION

In a widely cited essay on "participatory discrepancies and the power of music" Ethnomusicologist Charles Keil argues that music, "to be personally involving and socially valuable, must be 'out of time' and 'out of tune.'" (Keil 1987, 275) Keil's thesis is almost the precise opposite of the one pre-inscribed in Auto-Tune and other tools explicitly oriented towards projects of pitch correction.[78] While the Auto-Tune model of emotion is that genre-appropriate in-tuneness is a pre-requisite of emotional expression, Keil suggests that genre-appropriate out-of-tuneness is what matters. With this dissertation I have attempted to show how technologies and practices of vocal tuning are articulated between these two idealized poles of in and out-of tuneness. Taken out of context, Keil's and Antares' theses are practically synonymous to the extent that the criteria of the "harmonic fabric and genre of the ensemble" is flexible enough to include conventional degrees of "out of tuneness" into a larger category of appropriate "in-tuneness." In a sense, these polar positions are saying the same thing – that things need to be in time and in tune, but there is no single criterion for what that means outside of a social context. The specificities of how various digital tuning tools are socially constructed and put to use within particular studio contexts help to flesh-out the domain between these two ideal positions. What it means to tune a voice varies depending on the tool being used, and the

---

[78] The relevant passage, which appears in the software's user manual and patent documentation, reads "Voices or instruments are out of tune when their pitch is not sufficiently close to standard pitches expected by the listener, given the harmonic fabric and genre of the ensemble. When voices or instruments are out of tune, the emotional qualities of the performance are lost. Correcting intonation, that is, measuring the actual pitch of a note and changing the measured pitch to a standard, solves this problem and restores the performance." (Antares Audio Technologies 2000a)

construction of tuning tools vary across place and time. Whether and how exactly one is in or out of time depends on how time has been collectively assembled. Being in or out of tune requires the cultivation of a prior tuning-in relationship which establishes exactly what is being tuned in *to*. The tuning-in relationship, and the specific acts of repair and repurposing that is makes possible, is directed towards the production and negotiation of a play of voices.

During my time in-studio I had the nagging feeling that somebody was missing. In cases where there were many people in for a session, I caught myself counting heads to double check that the entire band and studio staff were present. When it was just Carl and me working on a vocal track, I would have to remind myself that there were, in fact, no more or less than two people in the room. I mentioned this to Carl and he suggested that it was the feeling of "missing yourself." That struck me as a good explanation. More than the feeling of one's own absence, however, I felt it as a combination of an anticipated arrival and a recent departure. I cannot speak to the generality of this phenomenon among studio workers, this being the one question that seemed too weird to put to people generally. It is not a topic that gets written about in Tape-Op or The Womb forums. What I can say is that Carl, at that moment, seemed to know what I was talking about better than I did. What I was experiencing, I think, was a partial and processural voice, or a voice in the making. It was the presently absent voice in the Pro-Tools stems during a tuning session, or the residual voice of a band playing together in the live room.

To get at whom exactly this extra and absent person might be, it is helpful to return to the concept of the middle voice. As Myers and Dumit argue, the middle voice is a type of position that shapes action and is shaped by it.(Myers & Dumit 2011) Between the passive and the active voice is a gestural one, which emerges when people and things are involved with one another. It is affectively experienced and affected by experience. In the studio, the middle voice is not just a figure of collaboration, but a deliberate sociotechnical project which bridges affective and perceptual labor. Taking and tuning a recorded voice is a concrete way of dialing-in the middle voice between the various human and nonhuman elements of the studio ensemble. When an engineer gets the feeling that somebody new is about to show up, or that someone is trying to speak but is not yet audible, listening to what that person sounds like or who they might be is just as important as hearing whether or not a particular part is sitting correctly in the mix. That elusive goal of the modern recording engineer – getting emotion out of a singer or a piece of gear – is accomplished at a level of emotional resonance and skilled perception. What makes Harold and Carl so good at their jobs, in part, is a habit of attending to and working to articulate the middle voices of a particular project.

As previous chapters have shown, voice and emotion are elusive objects that nevertheless seem to lurk behind the wide range of techniques and technologies which constitute the field of modern music production. Studio technologies such as Auto-Tune and Melodyne form part of a web of evocative objects. These objects are constructed as specifically evocative through processes of repair, repurposing, and reinscription. They are not simply inscribed but in-voiced. They are imbued with character as well as functionality. Their use is dialogical in nature, both in the sense that they are used to communicate needs between groups (engineers, artists, listeners, etc.) as well as the sense in which their specific application is accountable as a sort of heterogeneous conversation.

Previous ethnographies of the recording engineering have emphasized the importance of mediation and translation in studio life. Hennion has shown how the work of mediation between production and consumption characterizes the role of the record producer. (Hennion 2013) Porcello has shown how communication, or "speaking of sound" is accomplished through metaphorical negotiations among engineers and artists (Porcello 1998; Porcello 2004) Schmidt-Horning has emphasized the importance of the unspoken, or tacit, in the work of the recording engineer. (Horning 2004; Horning 2013) Meintjes has shown how perceptual differences in studio work are negotiated across cultural boundaries. (Meintjes 2003) This dissertation's intended contribution to this body of work is twofold. On the one hand it offers a specific topical focus on the question of how voice and emotion are constituted through sociotechnical practices. On the other hand, it goes beyond the question of how communication is facilitated smoothly, and instead highlights the central role of disruption, partiality, breaching, and repair in the production of new voices that were not there to begin with but arise over time and place. As a heterotopical and heterochronological domain, the studio is as much about accomplishing a break with conventions of time and space as it is about cultivating seamless flows. It is as much about the deliberate production of perplexity as it is about making things come out clearly.

Voice, Emotion, and Algorithm Beyond the Studio

Attending to the practices by which recorded voices are accomplished in pop music production, and the way that the tools of correction are reinterpreted in various ways, may offer insights well beyond the domain of music production. The sociotechnical articulation of emotion and the voice is increasingly prevalent in the field of affective computing, where computer science researchers and software developers are working to facilitate emotional understanding and communication between humans and computers.(Patel & Shrivastav 2011) The same sorts of pitch tracking algorithms used in Auto-Tune and Melodyne are already being deployed in technologies designed to track vocal emotion. Mexico-based startup Emospeech is developing an "Emotion Call Classifier" that can analyze the emotional dynamics of recordings of call center conversations, plotting them on an "emograph" and allowing the user to sort out the most and least emotionally problematic calls. (EmoSpeech 2012) Winifred Poster has shown how the proliferation of "emotion detectors" in India's call center industry have used the waveforms of people's speech to analyze the affect of callers and employees:

> Words themselves are evaluated for emotional content (like "frustrated" and "angry"). Features of the sound like pitch, tone, cadence, and speed can be assessed for more subtle indicators of emotion. Rapid speech or rising tone can signal excitement. Slower speech or moments of silence can indicate distress, discontent, or unwillingness of a consumer, according to a representative of Cisco, to sign up for a health insurance plan… Major call centers GE Capital and Convergys, among the largest in India, use emotion detection to enforce accent neutralization (Mohan 2006) They also use emotion detection to gauge reactions and resistance by customers in the United States and United Kingdom to foreign accents (Hamblen, 2006). (Poster 2011, 882)

BeyondVerbal, based out of Israel, has launched an Application Programming Interface (API) called Emotions Analytics (Anon n.d.) for tracking emotions in vocal intonation. They claim their software has applications in fields of social media, healthcare, and market research. A promotional video for their app "Moodies" features a striking hypothetical scene between two animated characters on a date. The narrator talks us through it: "Finally, she agreed to come up for coffee. You look deep into her eyes and see … Actually, you have no idea what you see. But your mobile app will figure it out for you." (Beyondverbal 2013) Other contenders in vocal emotion analytics include the Empath API released by the Tokyo-based Smartmedical Corp., and the Dutch firms Good Vibrations and Vokaturi, each of which offer software development kits for vocal emotion analysis.

When, in Chapter 4, A-Deuce asked about why his friend's mother thought T-Pain was a medical condition, he might have been on to something. Researchers at MIT have developed open source smartphone applications for detecting major depressive disorder and Parkinson's, using "vocal biomarkers."(Ghosh et al. 2016) Automatic vocal tuning has also be deployed experimentally in the production of emotional experiences. An international collaborative study led by a researcher at the Institut de Recherche et Coordination en Acoustique et Musique (IRCAM) has shown that "covert digital manipulation" of how a person hears their own voice as they are speaking can be used to shape their emotional state. (Aucouturier et al. 2016) The diagnostic potential of vocal intonation tracking, and the therapeutic potential of tuning, are increasingly being investigated across a range of disciplines.

At the same time that technologies of the voice are mobilized to make people's emotional experiences more legible to marketing, therapy, diagnosis, identification, and other modes of analysis, we might also anticipate how communities based on shared vocal or emotional characteristics may also engage in practices of self-fashioning around these technologies. The autism spectrum is one place where technologies of voice and emotion could see important application. The voices of children with autism are often characterized by a "sing song" quality, which may have to do with research that shows autistic children who have difficulty understanding spoken speech are more able to understand the same material when it is sung. (Sharda et al. 2015) People with autism may work to cultivate "neurotypical" speech patterns, teaching themselves to use prosodic cues to convey emotions. Transgender individuals have made use of tools like "Exceptional Voice App" (or EVA) that track their vocal pitch and provide training feedback to facilitate more feminine or masculine modes of speech. (Chen 2015) Audio engineers' ways of extracting and conveying vocal emotion may increasingly seem like a craft practice when presented alongside a broader range of tools that claim to automatically identify or produce emotional effects by way of the voice.

As social groups articulate themselves in relation to, by way of, or against these emerging technologies of the voice, we might expect the concept of the shibboleth to take on important new dimensions. In a situation where vocal emotion is being used to monitor employees or distribute diagnoses, we could imagine the articulation of resistance practices among those who want to have a say in how their voice is taken to speak for them. As the voices of mass media increasingly come under the auspices of pitch tracking and analysis, we might expect to see new ways of speaking that harken back to the new ways of singing that formed around the (mis)use of pitch correction. As technologies of the voice aim to derive emotional information from prosodic information divorced from verbal, historical, or local contexts, we might expect users to develop new ways of pacing and re-placing their voices and the emotions that are assumed to go along with them. Russolo's vision might yet pan out, not through the harnessing of machine noise but through a technological articulation of the voice as a tool of emotional production. As people object to voices that seem to be stripped of emotion, we might expect the essentialist and naturalized views of vocal affect to be displaced by "the most complex and newest sonic emotions, not through a succession of imitative noises reproducing life, but rather through a fantastic association of these varied sounds." (Russolo 1986)

APPENDIX: SHIBBOLETHS IN THE STUDIO

> Then said they unto him, Say now Shibboleth: and he said Sibboleth: for he could not frame to pronounce it right. Then they took him, and slew him
> *Judges 12:6, King James Bible*

> A positive movement of an eyebrow superseeds [sic] any movement of a fader or pot.
> *"Poole's Law," Engineer Mikey Poole* (Chinn 2014)

In his 2004 article "How to Talk About the Body?" Bruno Latour develops the concept of "articulation" by way of an example of the use of "odor kits" in the perfume industry. These kits consist of an array of contrasting scents, the systematic experience of which allows a trainee to "learn to be affected" in new ways and thus become an expert in scent, or a "nose." From here he proposes a new "demarcation" criterion to replace Karl Popper's falsificationism. This new "shibboleth" which he names after philosopher of science Isabelle Stengers and ethnopsychologist Vincianne Despret, goes beyond falsificationism by requiring that the scientist, elements under study, and articulations thereof are all "interested." Drawing on a metaphor of voice-as-agency implied by the concept "articulation," he suggests that "'scientific' means rendering talkative what was until then mute."(Latour 2004, 217) To pass the test of the Stengers-Despret shibboleth, one must perform inquiries in such a way that they provide research subjects a genuine opportunity for recalcitrance, difference, surprise, and re-framing of the questions put to them. Latour's demarcation criterion turns on the question of whether the research results in new articulations (both in the researcher becoming more articulate and the topic of study being articulated in new ways.)

Here I would like to explore a different aspect of articulation by taking up the concept of the "shibboleth" as an empirical topic for STS work, as opposed to an intervention into how science ought to be evaluated. Latour notes that human subjects are especially vulnerable to "uninteresting" inquiries; unlike non-human objects, humans are often so impressed by the research itself that they are unduly "interested" by what are actually relatively "uninteresting" questions:

> when impressed by white coats, humans transmit objectivation obediently: they literally mimic objectivity, that is, they stop 'objecting' to inquiry, in contrast to *bona fide* natural objects which, utterly uninterested by the inquiries, obstinately 'object' to being studied and explode with great equanimity the questions raised by the investigators – not to mention their laboratories! (Ibid)

I would like to focus on the white coat. Taken either literally or as a figure for any number of elements that go into the researcher's performance of scientific authority, the white coat is a shibboleth. If the human subject has indeed made it "more difficult for the social scientists to quickly detect the artifacts of the design" in their experiment (Ibid), then the coat itself is surely one of the most immediately apparent of these artifacts. If we take Latour's proposal seriously, we need to consider the role of the white coat as something that has the remarkable effect of making subjects conform to the expectations of the person wearing it. It marks the boundary between scientist and non-scientist – that is, it is a technology for the local, material performance of "boundary work."(Gieryn 1983) It also enacts a richly layered, asymmetrical, and decidedly *dysfunctional* relationship between the researcher and the subject. In Latour's example, the researchers expect the coat to bolster their scientific authority, though it ultimately has the effect of rendering their findings less applicable to situations where the coat no longer appears. The subject trusts in the researcher's authoritative knowledge and tells them what they think they want to hear, even as the researcher looks to the subject to tell them something they do not already know. The white coat encourages people to let themselves be spoken-for.

As it happens, recording engineers are one professional group that started off wearing white coats but no longer do. Up until the 1960s, EMI studios (later Abbey Road) "resembled a hospital, with orderlies in white coats."(Cleveland 2001, 576) The dress code at IBC studios was "a jacket, collar, and tie for all engineers and assistants. White coats, collars, and ties for the technical department."(Johns 2014, 20) Even the janitors wore brown suits. According to Abbey Road engineer Ken Scott, the coats were eminently practical, since technical engineers had to handle "dirty cables, had to go into damp echo chambers, and they had to wear suit and tie. And they didn't want their suits all dirty."(Eccleston 2014) At the same time, the white coats served as a powerful symbol. For pioneering reggae engineer Graeme Goodall, the epithet "White Coats" referred to the old guard engineers whose by-the-book techniques were inadequate with respect to the emerging musical styles of the 1960s.(Hitchins 2014) The material-semiotic distinctions afforded by the white coat – keeping dirt off one's suit, distinguishing technical engineers from janitors – became liabilities as clients sought a more casual variety of engineer and the distinctions between janitorial and engineering staff began to erode. As explained earlier, engineers are now far more likely to wear black.

Drawing on examples both in and out of the recording studio, this chapter advances the concept of the "shibboleth" as a variety of boundary object which specifically does (or is specifically for doing) boundary work. To rehearse a distinction which will be familiar to STS scholars, "boundary work"(Gieryn 1983) refers to the ways in which communities build up and maintain key social distinctions, for example, the line between science and non-science. "Boundary objects" are Star and Griesemer's term for objects that straddle boundaries between social groups, are "loosely structured" in common use while being "well-structured" in local use, and facilitate "collaboration without consensus" by virtue of this structural asymmetry (Star & Griesemer 1989) Though the "work" and the "object" have very different, and in some ways quite opposite meanings, it is common to see them presented implicitly as “twin” concepts.[79] The term “shibboleth” serves as a complementary concept within STS. I go on to discuss its specific analytical affordances and illustrate them with cases from an ethnography of Los Angeles recording engineers.

A literal shibboleth is a word which, depending on its pronunciation, reveals whether or not the speaker is a member of a particular group. The word has a broader meaning, however, as a way of talking about objects the use of which can be understood as a species of boundary making. A shibboleth is a sort of perverse or *anti* boundary object. Following Mitroff’s analysis of the “counter-norms”(Mitroff 1974) which structure scientific work by standing in contrast to normative structure espoused by Merton, we might also call a shibboleth a “boundary *counter*-object.” Shibboleths have long been the subject of sociological attention (c.f Giddings) but have not been much-used in any central systematic way within the Science and Technology Studies literature. What happens when we use the term as a descriptive category of investigation rather than a programmatic injunction? I argue that it has largely been neglected as a category of empirical investigation and should be attended to as a conceptual tool in its own right. That said, my formulation still retains and builds upon this metatheoretical function: It is a useful tool (indeed, a shibboleth!) for distinguishing the oft-conflated concepts of “boundary object” and “boundary work” in STS.

If a boundary object is something that straddles communities of practice, affording intergroup collaboration through its interpretive flexibility and ability to retain identity across

[79] Boundary objects are, for example, sometimes treated as “components” of boundary work (Tisenkopfs et al. 2015). Alternatively, “boundary work” is sometimes used to refer to work done with or around boundary objects (Nunes et al. 2016)

groups, a shibboleth is the reverse. Shibboleths are things that, by virtue of their being subject to interpretively flexible use, quietly enact and reinforce social boundaries. Where boundary objects are able to survive the passage from one group to another, shibboleths have the tendency to self-destruct along the way. Conversely, this process of self-destruction is productive of a boundary.

Interpretive Brittleness and the Practical-Symbolic Distinction

All socio-technical formations are subject to interpretive flexibility, but shibboleths are characterized by the way that this flexibility is structured with respect to the domains of the symbolic and the practical. From a given perspective, a shibboleth will typically have a practical, or "incidental" set of uses, the social distribution of which serves as the basis for its use as a test of group membership. It is useful, in large part, because its incidental practical functionality obscures its social function to out-group members. In this way it can be distinguished from more formal tools of accreditation such as diplomas or badges, which are primarily used for group identification and have only limited and secondary practical uses.[80] The difference between practical and symbolic is of course a flexible one, not least because even symbolic uses are practical ones. It is also a distinction that changes with one's position. A dollar bill is both a symbol of monetary value and a practical way of getting something out of a vending machine. The practical and the symbolic qualities of the bill are not absolute, but relational and situation-specific. The bill's ability to be exchanged for a candy bar takes on a symbolic quality if one is stranded on an ice floe, at which point the bill's ability to burn and produce heat (a quality one normally would consider symbolic, say in a political protest or gratuitous display of wealth)

[80] Fantastical accounts such as vampire-halting crosses or police badges catching bullets can be understood as exceptions that help prove the rule.

becomes its major practical affordance. The difference is important because it is in the movement and blurring of that the practical/symbolic boundary that shibboleths often operate.

During the 1980s the band Van Halen gained notoriety for including in their 11-page performance contract with the venue (or "tour rider") a request for m&m candies in their dressing room, minus the brown ones. While many assumed this was a case of petty tyranny on the part of overly pampered rock stars, singer David Lee Roth later claimed that the brown m&ms clause was, in fact, a safety precaution:

> The promoters frequently didn't read the contract rider and we would have structural, physical issues because there wasn't the proper electricity, load-bearing stress, et cetera. So, in the middle of a huge contract rider… I had them place a clause that, just out of the middle of nowhere it would say for example "there will be twelve ampere high voltage sockets placed at fifteen foot intervals not to exceed the load-bearing et cetera, and then just out of the middle of nowhere it would be "no brown m&ms in the back stage area or the promoter will forfeit the show at full price." What was the point? If I came backstage, having been one of the architects of this lighting and staging design, and I saw brown m&ms on the catering table, then guaranteed the promoter had not read the contract rider, and we had to do a serious line check. 'Cause frequently we had danger issues, or accidental issues. Well after seeing brown m&ms and ceremoniously and very theatrically destroying the dressing room to try to get the message across, lest we have a disaster like recently happened, the word got around that, hey, take this seriously. Invariably I would show up backstage an hour or two early and there would be two old gals in their 70s, just like cafeteria ladies… and they would be separating the brown m&ms out of the several five pound bags into their own jar and that jar would then be sealed and whisked away…When we played I believe New Mexico… their new building had just installed a brand new basketball rubberized surface… and the promoter had not read all of the physical requirements in the contract, and our super heavy stage sank about six and a half inches into their brand new rubberized floor. Backstage, not knowing this, I walked in and saw brown m&ms, so I trashed the dressing room a little bit. Basic cost, probably about $200 worth of food fight and a torn cushion with plenty of feathers for effect. However, the sinking of the stage into the rubberized floor… did something like $470,000 worth of damage. Cut to the media mulching machine who reported this as David Lee Roth discovered brown m&ms in his dressing room and trashed the dressing room, causing close to half a million dollars worth of damage. (Ganz 2012)

The brown m&ms, in combination with the contract, were used as a shibboleth. In this case, it was the m&m's apparent lack of practical consequence, and the ease with which it could seemingly be ignored as a practical concern, which made it useful as an indicator of whether the promoter had read and fully complied with the contract. In fact it had an important symbolic function, and one that was closely tied to much more clearly consequential practical realities regarding the structural requirements of the production. The brown m&m became a shibboleth by the way it was embedded within the text of the contract, placed like a timebomb on the dressing room table, and enacted as a distinction between a responsible and an irresponsible promoter (as well as that between a dangerous and a safe venue, or a trashed and an undisturbed dressing room.) Shibboleths can come into play where the relation between text and object, often within a contractual context, is placed in an unexpected configuration, which in turn allows for a social distinction to be made. Historian David Mindell, for example, has observed that the owner's manual for a 747 jumbo jet is not designed to provide any insight into the operation of the aircraft but is instead a document designed to limit manufacturer liability by delimiting precisely the modes of use that Boeing are prepared to endorse.(Davis 2015) The manual may at first glance seem to be a way of transferring the intimate knowledge of the inner-workings of the plane, but is in fact a tool for separating, in advance, the manufacturer and the airline, on the basis of a distinction between responsible and irresponsible pilots.

The way shibboleths appear as a form of play between text and object is further illustrated by the example of the moisture sensors that Apple embeds within iPhones and MacBooks. When exposed to moisture, the sensor permanently changes color, thereby voiding the owner's warranty. It is important that the sensor be inconspicuous, lest a savvy user discover it and attempt to "turn" its allegiances through an act of anticipatory waterproofing. The sensor's inconspicuousness, which constitutes its asymmetrical significance between Apple and the user, arises from the fact that the sensor is generally not seen as practically significant for the average user (at least not until they spill water on their keyboard.) The sensor sits quietly at the boundary between Apple and the user, waiting to register the user's irresponsibility. The power relations could, indeed, reverse if a user was able to remove the sensor and reinstall it once they had inevitably soaked their iphone. The sensor, in this case, would remain a shibboleth but with a different interactional outcome.

Just as a shibboleth might come into being by way of a distinction between practical and symbolic registers, it also works to reproduce this distinction and distribute work across it in new ways. In its biblical context, the word *shibboleth*'s practical semantic affordances - as a way to refer to an ear of grain or, perhaps the flow of a stream - is displaced by its use as a symbol of ethnic difference, after which it becomes a practical way of referring to such a symbol rather than its "original" or "literal" meaning. Symbols have practical applications and practices have symbolic meaning. Still, we are able to roughly comprehend the difference between something that is done for a practical reason (wearing boots when going for a hike) and something that is done for a symbolic purpose (wearing a cap and gown at graduation.) Shibboleths trouble this difference: say when a veteran hiker recognizes a novice by their sneakers. The novice uses the sneakers to protect their feet, while the veteran hiker uses them to indicate novice-ness.

Hsu's concept of "coding style" (Hsu 2015) can be taken as an example of where shibboleths come into play: conventions of readability, strategies for structuring algorithms, and other stylistic habits of coding may make little to no difference to the computer that compiles the code – at that register it is technically insignificant. Nonetheless, styles of coding powerfully structure communities of practice in computer programming. They draws boundaries between different communities. They also shape patterns of versioning, modification, and borrowing among software developers. Coding style serves as a shibboleth among app developers. A piece of code might work just as well for one programmer as another, but for one it carries the wrong stylistic markings and is thus not legible, or legible as being the "wrong way to do things."

Closely related to shibboleths are what Mody and Lynch, in the context of laboratory science, call "test objects" (Mody & Lynch 2010). Clear cases of test objects, they write, feature "highly regular and reproducible visible properties that can be used for testing instruments and training novices."(Ibid, 1) One such test object, the 7x7 "reconstruction" of silicon atoms widely used in materials science, served as a "shibboleth, yardstick, and epistemic thing" by which "surface scientists could recognize each other as a member of a body of practitioners requiring their own methods, tools, career paths, and institutions." (Ibid, 5) Here the role of the shibboleth becomes apparent in an area where institutional boundaries are in flux around a particular professional subgroup. Test objects work by telling you more about the person and things that are used to analyze them than they do about themselves. Shibboleths need not simply exist as symbols of group identity, however. The idea of shibboleth as "badge" of group status neglects the most interesting aspect of the shibboleth, namely the way it registers difference through practice.

Partiality, Asymmetry, and Breaking

> *We seek those ruled by partial sight and limited voice-not partiality for its own sake but, rather, for the sake of the connections and unexpected openings situated knowledges make possible… Science becomes the myth, not of what escapes human agency and responsibility in a realm above the fray, but, rather, of accountability and responsibility for translations and solidarities linking the cacophonous visions and visionary voices that characterize the knowledges of the subjugated. A splitting of senses, a confusion of voice and sight, rather than clear and distinct ideas, becomes the metaphor for the ground of the rational. (Haraway 1988, 590)*

Shibboleths are often things that work by breaking. An increased sociological attention to technoscientific shibboleths could offer new ways of critiquing and dismantling social barriers, as well as a way of cultivating and theorizing practices of repair-as-resistance. Thinking with shibboleths invites strategies oriented towards repair: if shibboleths are designed to break, how do people go about repairing them anyway? How does one go about repairing something, like a moisture sensor, that is made to break? Exactly what kind of breadown does such an artifact entail? It might be more precise to say that shibboleths are made to "cleave" in particular ways. A shibboleth breaks not by failing completely but by working *partially* or *asymmetrically*. It is an element of infrastructure that is meant to break quietly and partially and produce and partial visibility (audibility, palpability, etc.) with respect to a particular social relationship. Boundary objects afford "collaboration without consensus"; shibboleths build asymmetrical consensus and partial agreements, which elaborate social barriers and divide up domains of work. Thinking about a shibboleth as a mere password or makeshift badge of group status is not wrong, but it is a thin conceptualization. A thicker account of a shibboleth lets us think about how they exist as objects and the specifics of how they function as social things.

Saying that shibboleths work or break "partially" is to say that they "take part" in social relationships. Boundary objects, by contrast, are impartial – they provide common ground between worlds and are able to elaborate themselves for particular worlds. By contrast, shibboleths are partial both in the sense that they are only "partly" broken, working, visible, or invisible as well as the sense of being partial as having a preference or tendency. Partiality here implies being simultaneously *a part of* an arrangement and *apart from* it. This is an important factor in distinguishing a sociology of technoscientific shibboleths from a sociology of error. A sociology of error falls into the trap of assuming in advance that a social practice or belief is wrong with respect to some fact of the matter lying beyond the domain of social explanation. It thereby neglects to describe and thereby erases the social and practical basis of the position that is taken to be “true.” It is the asymmetry of a shibboleth, its partiality and multiplicity with respect to the multiple groups that encounter it, that make it useful for symmetrical accounts of technoscientific objects. An asymmetrical account of an object would assume that it presents its features as identical to all parties concerned and that only some are able to overcome their social context in order to perceive it as it is. Thinking in terms of shibboleths allows us to retain the basic undecidedness of what counts as working or not working, all the way down to the materiality of the object.

Susan Leigh Star has suggested that one characteristic of infrastructure is that it is transparent except in cases of breakdown. Performing an “infrastructural inversion,” an analytical move aided by attention to cases of breakdown, involves “foregrounding the truly backstage elements of work practice.”(Star 1999) Doing so, moreover, make visible how infrastructures can also serve as barriers to action for certain users. Shibboleths are infrastructural relations that break down partially, and thus become visible asymmetrically. When a philosopher specializing in the American Pragmatists hears a colleague pronounce Charles Sanders Peirce’s surname so that it rhymes with “fierce” instead of “purse,” she knows that the colleague’s familiarity with Peirce must be primarily textual, and to that extent the colleague is not “in the know.” Unbeknownst to the colleague, the details concerning the material and practical makeup of his knowledge have become visible through a breakdown in pronunciation. The fact that pronunciation has little direct bearing on the actual “substance” of the speaker’s knowledge is almost irrelevant. Similar deployments of useful uselessness can be seen in the production of human-machine boundaries. A CAPTCHA[81] is designed to be illegible for optical character recognition but legible to a human. It is designed to cleave across social boundaries, including those which constitute human/non-human relationships. Because it is designed to be needlessly difficult in a very particular way, it works because it reveals an algorithm to be non-human, without the algorithm being able to tell why.

Shibboleths in the Studio

Given any two arbitrarily similar sounds, there will always be at least one audio engineer who will claim to be able to hear the difference. – Hawley's Law of Differences[82]

Because they are performed at, and work to reproduce, social boundaries, shibboleths are most apt to occur in situations where social groups are coming together under a particular set of conditions. These include cases where 1) work and self-identity are strongly related; 2) there is close interaction between workers and clients, such that workers are accountable to the judgment of the client, i.e. where they are obliged to go along or at least humor the client; and 3) there is a highly skilled basis to the work but one that is challenged from the outside, i.e. where there is a need for practitioners to perform their skill for one another as well as for the client. The first two of these conditions come directly from Howard Becker's study of jazz musicians and their audiences. Becker famously analyzed the complex and often strained relationship between the professional dance musicians of midcentury Chicago, who are interested in "blowing jazz," or playing "hip" music, and their audience of "squares" who only want them to play "corn" (i.e.

[81] These are tests of human-ness encountered whenever a website fears being inundated with automated requests. It is an acronym for "Completely Automated Public Turing test to tell Computers and Humans Apart."
[82] Attributed to Nashville-based engineer and educator Scott Hawley (Chinn 2014)

music they find uninteresting.) Becker suggested that this relationship of resentment arose from the fact that the musician is both a "service" worker, and thus held accountable to the preferences of the lay client, and an "artist" who invests a great deal of self value in the quality of their work. The dilemma of either turning "commercial" by catering to squares, or going broke by not doing so, was negotiated in a variety of ways. One saxophone player talks about sneaking jazz lines into the commercial song he was playing, only to be caught out by the bartender:

> The second set, I was playing Sunny Side, I played the melody for one chorus, then I played a little jazz. All of a sudden the boss leaned over the side of the bar and hollered, "I'll kiss your ass if anybody in this place knows what tune you're playing!" And everybody in the place heard him, too. What a big square! What could I do? I didn't say anything, just kept playing. Sure was a drag. (Becker 1951, 141)

Becker describes how some musicians, when faced with a venue without a proper barrier between them and the audience, will improvise physical barriers out of chairs and strategically positioned instruments, so as to avoid having to talk to and take requests from the squares. Were these barriers not also identifiable as non-barrier club furniture, it is less likely that they would have been accepted. Occupational slang proves to be another important resource, as it "readily identifies the man who can use it properly as someone who is not square and as quickly reveals as an outsider the person who uses it incorrectly or not at all."(Ibid, 144) The word "square" for example, can be

> "used as a noun and as an adjective, denoting both a kind of person and a quality of behavior and objects. The term refers to the kind of per- son who is the opposite of all the musician is, or should be, and a way of thinking, feeling, and behaving (with its expression in material objects) which is the opposite of that valued by musicians." (Ibid, 137)

Like Becker's dance band musicians, recording engineers cultivate ways of telling capable from incapable practitioners in the studio. Shibboleths provide one source of "cues" which are used heuristically to determine how skill and authority is structured within the studio scene. Susan Rogers, who engineered several albums for Prince before getting her PhD in cognitive neuroscience at McGill and going on to teach auditory cognition at the Berklee College of Music, explains that while she was able to hear intuitively whether a musician is an expert or not, her more recent work in experimental psychology has shown her the actual "cues in the signal"(Rogers 2013) that were allowing her to draw that distinction in the first place. Engineers take expertise to be something that is felt intuitively but also makes its way into the audio signal. In both of these cases, however, it must be identified with reference to cues.

Cable Wrapping: The Over-Under Method

One example of an in-studio shibboleth can be found in the mundane practice of cable management. Dozens of cables, of various kinds (unbalanced instrument, Tiny Telephone (TT) and balanced XLR microphone cables being the most common varieties) are used to structure the signal path of the session. They connect instruments to amplifiers, microphones to patchbay channels to busses to outboard hardware to headphones and monitor speakers. They are crucial but easily tangled, lost, or damaged. When one cable breaks in the chain it is difficult to identify the problem cable without systematically disassembling and testing each cable one by one – a frustratingly massive time sink, especially for clients that are being billed by the hour.

Accordingly, one of the first and most enduring skills I learned in-studio was the correct way to wrap cables. Being able to quickly and tidily gather and deploy cables is the studio version of "knowing the ropes." After I had received the introductory tour and orientation materials for Studio A, I got to work tidying up the live room from the previous day's session. As I started wrapping an XLR cable tightly around the palm of my hand and the back of my elbow, Carl stopped me – "That'll get you fired wrapping a cable like that!" He went on to demonstrate the preferred technique, which is known as the "over under" method. The uninitiated are liable to, as I did, wrap a cable as one would coil a rope, gathering it in loops or coiling it tightly around the arm. In the over-under method, the trick is to alternate the orientation of the cable as you form loops and bring them together, building a self-opposition into the cable's torsion and the loops it forms. This is a difficult technique to express verbally, primarily because it requires at several points a contingent negotiation between what that particular cable wants to do and what you want it to do. Negotiating a cable appears as a mundane example of what Andrew Pickering called a "dance of agency."(Pickering 1993b)

Because the way cables are used changes the way they behave in the future, wrapping can also be thought of as a form of inscription, or the sedimentation of modes of use into material form. Over time a cable will become trained a particular way. It will retain a memory of the way it was previously stored or used, and will want to return to that shape. Cables that are not well cared-for may turn unwieldy. They will bend wildly, forming bulky and tangled coils as you try to collect them. To wrap correctly you need to recognize the train of the cable and gather it up in such a way that it fits against itself. At the same time – and this is the key move of the over-under method – you must build into the coil a desire for the cable to not twist into previously made loops and form knots.

This trick is accomplished by alternatingly putting a slight twist into the advancing cable with the thumb and forefinger of the gathering hand. You let the loop of the cable form either over or under the beginning of the next loop. When you gather the cable without twisting, the loop will fall under your gathering hand. When you twist it, the loop forms over the top of your hand and you proceed by tucking the gathered line under the loop and into your left hand. The test of a properly wrapped cable comes in the unfurling; you should be able to hold onto one end, toss the bundle across the room, and have it roll out without forming tangles. Wrapping technique is both a means of keeping the studio organized and a powerful shibboleth among recording professionals. Particularly in work environments where people are new to one another, proper cable wrapping technique is a key way of signaling that you know what you're doing. In this way it functions as a technology of trust as well as material organization.(Porter 1996) Good wrapping technique tells others that if something in the chain breaks it's probably not your fault, and that you would be a good person to get in touch with if you need help on a future session.

The practice of the over-under technique trains both the cable and the person wrapping it. It is a process of skilled sensory and expressive movement (Ingold 2011), through which the body of the worker is attuned to the cable's previous habits of movement and storage, and the cable is recomposed according to the embodied knowledge of the worker. In this way, cable wrapping can be understood as an instance of what Prentice has termed the "mutual articulation" of object and subject (Prentice 2005). The body of the engineer is trained along with that of the cable, and through this repeated training they are both made increasingly legible within the studio context and made more sensitive to future movements within it. This technique functions both as a technical shibboleth, demonstrating one's membership within (or outside of) a community of practice, and bears directly on the broader problem of constituting social order within the studio. Over-under techniques may undergo involutionary changes in particular cases, elaborating on the basic form according to local tastes and affording yet another degree of resolution as a marker of skill. As Nashville-based engineer Sara Hughes recounts, a failure to recognize local flavors of the method can cause problems for novice engineers:

> "He [the studio's resident second engineer] and I repeatedly butted heads over the issue of cable wrapping, because he insisted that all cables had to be coiled right-handed, starting at the female end. I didn't agree: I figured over-under was the same regardless of

> which hand or which end started. Eventually, he explained that he had a methodology for tossing out cables that he felt was impacted by my left-handed wrapping. He felt my refusal to comply adversely affected the efficiency of his setup, and that in turn adversely affected his ability to perfectly complete the session engineer's setup. Even though I didn't agree that it mattered either way, I decided to do it his way, if only to keep the peace. His whole demeanor lifted. Instead of wasting time debating the merits of cable-wrapping, he had the time and the inclination to share seconding tips with me." (Hughes 2003)

The over-under technique works as an even more elaborate shibboleth in the case of handedness precisely because its social demarcation function is concealed by – to the extent that it is identified with – its practical function. Hughes did not fully realize what the handedness of cable wrapping meant in that situation, and indeed what practical use it was in terms of the effect that it would have on the demeanor of the studio's second engineer. At one level, Hughes is right that a cable wrapped left-handed and male-out[83] is likely to pass the unfurling test just as well as a right-handed female-out coil. For her, the practical equivalence of these methods masked its symbolic, demarcational difference for the second engineer. Switching to a right-handed female-out technique, thereby "framing to pronounce" the shibboleth correctly, ultimately proved to have practical affordances for Hughes, as it improved her relationship with the second engineer.

Giving Wingmen Fake Faders, Cheating Compressors

Some shibboleths in the studio work precisely because their practical and apparent affordances are unevenly distributed. Harold, for example, describes how surprisingly handy a non-functioning fader knob on the mixing console can be:

[83] The pieces that connect cables and other electronic equipment are described through a gendered metaphor. A "male" plug fits into a "female" receiver. Cables with unusual connective pieces, or that translate from one standard to another (XLR to quarter inch, for example) are sometimes called "gender benders."

> And then also you know, just knowing, giving the wingman[84] a fake fader to have so he feels like he's doing something. Shit like that. That probably never got noticed, but got noticed by a few people. And it kind of made the difference, in the end, because they saw all this background work that was getting done that was making the session a lot smoother regardless.

It is (perhaps deliberately) ambiguous what Harold means when he speculates as to whether anyone "noticed" this technique. On the one hand, he could mean that the wingman noticed and appreciated the extra level of involvement he was apparently being given in the production process. On the other hand, he could be talking about the few people who may have seen through the façade; perhaps the artist noticed and got a kick out of it, perhaps it gave a boost of morale for the intern who had been yelled at by the wingman for getting his lunch order wrong. To describe Harold's fake fader trick as a form of deception would not be incorrect. However, an uncritical use of the word "deception" here is misleading if it is taken to imply that studio life is adequately describable in terms of true and false percepts. It would also distract from the genuine practical affordances of this technique in the broader context of studio life. Even if it has no immediate effect on the electrical signal being sent to the speakers the fake fader has profound affordances when it comes to the redrawing of social distinctions. For the wingman and other entourage members, noticing their colleague getting a say in the mixing process could have a powerful effect on understandings of intra-entourage hierarchy, as well as the entourage's inter-group relation with the studio staff, the recording artist, etc. It may, of course, simultaneously have the opposite effect from the perspective of members of the technical

[84] "Wingman" here refers to an artist-client's friend, guest, entourage member, support staffer, or generic hanger-on. Etymologically it has to do with Air Force lingo – a fellow pilot "at your wing" or who has "got your wing" while flying in formation. Outside of the studio context, a wingman may also refer to a friend who supports your attempts to impress a potential romantic partner in a bar or other social situation.

staff, who are able to witness it as a performance of the engineer's shrewdness as well as a reassurance that they know something that the clients don't know about how to go about recording.

In looking at how boundaries are productively maintained between humans and nonhuman technologies, it is useful to compare Harold's use of fake faders to what he calls the practice of "cheating" equipment. While discussing his approach to compression, for example, he explains that "sometimes I want something darker and I know the compressor will sound darker. So then you have to kind of cheat the compressor into making it darker and having it sound but not compressing the signal." In this example the compressor is being tricked into doing everything except for the thing for which it was primarily designed. It is being repurposed into producing a particular effect. Engineers often take on a trickster role with respect to their clients and equipment, and this role is often performed by way of, and as a way of producing, shibboleths. In the case of compressors, engineers traffic in shibboleths as ways of distinguishing their work from that of a piece of equipment. Broadcast compressors were first developed to replace the work of a broadcaster's hand on the level fader.[85] By working according to precisely set thresholds and according to pre-set response curves, compressors were meant to replicate the work of the board operator. Rather than accept that their domain of skill had been annexed by a machine, however, engineers developed ways of using the compressor to squeeze new sounds out of their recordings. So-called "brick wall" limiters, originally designed to keep dangerously loud signals from damaging ears and speakers, actually facilitated unprecedented loudness levels in recorded music by allowing the engineer to reduce the difference between the loudest and

[85] For a discussion of the role of "hand limiting" in the collaborative production of emotion, see pp 219.

quietest parts of a signal. The tool designed to automate audio engineer labor and protect listeners from especially loud sounds became both an important addition in the engineer's toolkit and the standard-issue weapon for the so-called "loudness wars" of the late 1990s. Engineers developed techniques of listening and recording which made it possible for them to hear and perform compression *as such*. Harold's habit of using compressors to alter tone color rather than dynamic range continues this process by which engineers distinguish themselves from their equipment and their clients through practices that distribute perceptual and social significance unevenly.

Shibboleths like the fake fader enact a sort of "placebo" narrative. A placebo seems to be for one thing (treating an illness) but is in fact being used to draw a number of social distinctions: between people who are actually responding to "real" treatment and those who only think they are responding to a treatment, between effective and ineffective treatments, between real and imagined illness. At a slightly higher level of abstraction, the presence or absence of a placebo within an experimental design is used to demarcate scientific from unscientific biomedical research. Shibboleths can be produced retrospectively through 'placebo effect' tests. As Harold describes, problematic client requests will occasionally be met with appeasement rather than candid explanation of the trouble it raises from the engineer's perspective. Describing strategies for managing the client in-studio, he mentions that "some of it's shady, like, I don't like lying to people, but 'yo dawg take off that compression' I'm like 'ok.' Nope, no compression, check check ah sounds perfect… didn't touch a thing!" By failing to not hear a difference after the request to remove a non-existent effect, the engineer and client fall into a sort of asymmetrical agreement, which has the primary effect of allowing the recording process to proceed. This is a quite common tactic, as engineer Barry relates:

> Sometimes I try to straight up fake people out. You do that a lot as an engineer, too. Like you kind of tell them you're gonna do something and then you don't really do anything and they're like that's great! Y'know. That is a pretty common thing particularly with A: people who don't really know much about music and B: people who don't know anything about engineering. Like we used to do that with bands every once in a while, like someone who was making kind of an absurd request and then you like click a few things and they're like oh that's it – and you're like I didn't do anything.

Sometimes the tables can turn and engineers will catch themselves (or, less frequently, be caught by someone else) working on a piece of gear that is either turned off or not connected to the signal flow. The trope of the fake fader, in such cases, turns against engineers and forces them to respond to the apparent fact of having applied a shibboleth against themselves. Online forums are one place where these stories get confessed. As an Amsterdam-based engineer, for example, relates having "spent many times tweaking in graphic mode, even with the singer and the producer agreeing with the difference they hear, only to realize it was in bypass."(Straus 2009) This experience of "accidental bypass" is common enough that it serves as a professional "in joke" and shared empathetic narrative. There is a popular online meme[86] that features a picture of the character Morpheus from The Matrix films, captioned with a mindboggling revelation. The music production version of the meme poses the question, "what if I told you that compressor you've been tweaking for five minutes is in bypass?" The photo's Facebook comments illustrate how common this experience is:

---

[86] While I first encountered the term "meme" in Richard Dawkins' The Blind Watchmaker, where he uses it as a cultural analog to the biological "gene," in this context "meme" refers specifically to an idiom of internet humor, usually involving the juxtaposition of an image with text which forms a variation on some theme. The Morpheus meme used here, always has the same image (Lawrence Fishburn's character from the Matrix at the moment that he reveals to Keanu Reeve's character that he has been living in a computer simulation) with text beginning with "what if I told you…" typically followed by some mundane revelation with which the audience of the meme may identify and find humor.

[redacted] Lead sound tech at one of the venues I work for, decided to "tune the room". The next gig I covered I noticed he adjusted the crap out of the main's GEQ, but never activated it.... smdh

Upon activating his setting, I discovered what a tin can, in a cave at the bottom of the ocean sounds like.

Like · Reply · 4 · 18 hrs

[redacted] Hahaha, yeah it's frustrating and embarrassing. It comes a time when that issue belongs to the past for everybody but we've all been there

Like · Reply · 1 · Yesterday at 11:50am

[redacted] Been there, and of course not just with compressors!!!

Like · Reply · 8 · Yesterday at 7:11am

[redacted] So. Many. Times.

Like · Reply · 2 · Yesterday at 7:39am

[redacted] There is a certain well known 20yr old digital Yamaha sound board, used as a monitor desk at one of the venues I'm employed. Can't count the number of times we've had 3 sound guys scratching their heads trying to figure out why it wasn't passing signal, only to realize the mute isn't a mute, it's a "channel on" button....

Also the idiocracy of the outboard patch box, has reversed gain structure. Aka -60 = +10, and 20 = -70...smdh

Like · Reply · 14 hrs · Edited

Write a reply...

[redacted] Had my sound card in mono for a whole mix. Convinced myself that all my panning and stereo delays were sounding dope.

Like · Reply · 2 · 22 hrs

[redacted]

Like · Reply · 4 hrs

Write a reply...

[redacted] (Most singers "artists" never know the difference! Lol)

Like · Reply · 1 · 11 hrs

[redacted] Hvor mange gange har jeg ikke siddet og arbejdet på noget med mono knappen i bund?..

[redacted] how many times have i been working on something with mono button in the bottom?..

Automatically Translated

Like · Reply · 1 · Yesterday at 7:43am

[redacted] haha alt for mange gange! ;o

Haha too many times! ;O

Automatically Translated

Like · Reply · Yesterday at 8:25am

Write a reply...

[redacted] I was tweaking channel 11 instead of 10, the other day for about 5mins

Like · Reply · 1 · 22 hrs

[redacted] I've done this. And when you realize you're like "oh fuck...."

Like · Reply · 9 hrs

Figure 15: Facebook comments on Morpheus "accidental bypass" meme
Screenshot by Author

In cases of accidental bypass, engineers find themselves at the other end of a technique they might use on clients. When stories are shared among engineers in this way, however, it becomes a way of strengthening their in-group identity. The in-joke, and the knowing "been there" response that it invites, becomes a sort of positive variation on a shibboleth. Having had the experience of using something when it was turned off becomes something of an inside joke for engineers, as opposed to an "outsiding" event.

Crucially, the distinction between true and false percepts in-studio occurs only in a relational context because it always has a socially performative aspect. For example, Ed, a studio tech, related a story from when he was working as an assistant to a prominent engineer/producer. The producer had spent several minutes fine-tuning a vocal performance, only to discover that the equipment he was using was in bypass. This was embarrassing for the engineer/producer, but worse for the Assistant. "He never trusted me again" Ed explained. Here the assistant, not the engineer, is the one who gets blamed, in part because it is his responsibility to make sure everything is turned on and working properly, but also because the act of skilled perception, the performance of expression and creative taste, is undergirded by a social structure of objective reference. Similarly, Lucille relates:

> I was like trying so hard and I cut this thing up and was like and I realized I was listening to one and was editing the plugin of another. But that I could tell what I was doing was not like making it better – but you definitely fake yourself out sometimes and, um. It's kinda like when if someone really likes you and you're like check out this song and they wanna like this song, they'll probably have a much higher chance of actually liking it, even if it's not as good a song as they might've heard, even if they're not in the right mental space to listen to it.

A failure to hear something that is there (or not hear something that is not there) is a point of social breakdown – specifically, a breakdown of trust and the acts of perception it makes possible. Bateson suggested that having authority means others work very hard to make what you say turn out to be correct (Despret 2004; Bateson 1987). Foucault famously defined the "author function" as the principle of thrift in the proliferation of meaning (Foucault 1979). Authorship, authority, creativity, intention, and artistry proliferate through collective action and it is only through the elaboration and demarcation of social relationships, boundary work which tapes-off the areas where different displays of technical and artistic skill can be said to occur, that these become accountable after the fact. We can think of the studio as a place where voices are socially produced through processes of authorization and articulation.

The Case of the Laptop on the Neve

A shibboleth can also appear as something that remains a transparent tool for one side of an interaction and becomes a point of breakdown for the other side. Bob was working at a particularly prestigious and well-equipped studio when a young male pop star came in for a session:

> He brought a MacBook and sat it on top this Neve 50, uh 80 series. Which is one of the most renowned Neves there is. Like you know Neve 2254s, the 1081s, the 1080s all came from that. And this old 60s API sidecar, with all the original 512s and 525s in it, like, beautiful gear. He sat a USB controller, a MacBook and a mouse on top of it. So we had to build a special case for them to lay on so he could lay them right on top, right on the face of it. Like [laughs] what? But at the same time how cool is that, that this punk ass kid is coming in here with his MacBook and they have this room full of tech nerds making that crap sound great.

For the pop star, the custom-built case on the esteemed "Neve" brand mixing console was a convenient tool, a way of continuing his work without having to worry about his laptop slipping off the console. For the engineers who had to build the thing, haul it out every time he comes over, and maintain the console that sits underneath it and goes under-used, it serves as an enduring reminder of the power and skill disparities between him and them. For the engineers, who loved and cared for the console, the laptop case was an obstacle to their work. Bob and his colleagues pride themselves in "making that crap sound great" – continuing to "love the production" in spite of the hindrances imposed by the actual client. The laptop case facilitates collaboration but it does so by shoring up and clarifying the social boundaries that already marked the interaction.

808 Snares and the N-Word in Waveform

Sounds, both as-heard and as seen in their visual waveform representations, also work as shibboleths in-studio. Sounds that would have previously gone unnoticed are not only made noticeable, but invested with deeper meaning through the instruction of senior studio personnel. Engineer Lucille relates her experience in familiar maritime terms:

> Have you ever heard that theory about like the Native Americans not seeing Columbus coming because they had never seen a ship before? … I feel like it's kind of a little bit like that… you're listening so hard to this acoustic guitar that you don't like realize that there's something two octaves lower that's making a little whaa sound when there's a bad edit, or. … initially it's helpful when someone's like takes the time to show you that stuff, but after a while once you can start like, once you open your ears to not just listening to the acoustic guitar that's in this frequency range and your ears just get accustomed to kind of trying to hear things everywhere, then you're able to learn a lot more yourself. But like that first step of being like just because it's an acoustic guitar that's fucked up doesn't mean that the sound is gonna sound like it's coming from the acoustic guitar, it may sound like its coming from somewhere else. It might sound like it's in the drums or something, but just 'cause everything affects each other.

These moments of perceptual initiation, and the differences they bring to life, need to be socially cultivated. Lucille found she was in some ways able to gain certain skills more effectively by virtue of the fact that male engineers were more willing to show her things they knew rather than keeping them secret. On the other hand, she found that she did not fit the implicit profile of an engineer and, even after engineering a session for several days, would find the clients sometimes surprised to realize that she *was, in fact,* the session engineer. She is able to interact with clients in ways that her male colleagues cannot, but finds certain things men would be able to say are unavailable to her. Less than five percent of audio professionals are women, according to Terri Winston, who is a recording engineer and founder of the Bay Area nonprofit Women's Audio Mission. (Haruch 2010) This disparity in gender representation shapes everyday affective experiences in-studio, which in turn shape habitual perceptual action. For Lucille, particular sounds can seem to index both bodily difference and social distance, in this case the hi-hat and snare samples of an 808 drum machine:

L: I think that like not only are people acting differently in the studio but people are kind of expecting you to. I don't know, that's a really good question but I think that sometimes we definitely hear differently. Like I have a thing with 808 hi-hats and snares, like some reason they're like the most painful things to my ears.

O: But they're everywhere

L: They're everywhere. And I don't mind them when they're blended in with another snare. As long as that particular sound is not the loudest, is not like soloed and really loud, I'm totally fine. But like, like [Producer] can crank that's like it's a stupid example but [Producer] can crank 808 snares – and I can crank other things super loud and he's like what are you doing? And I'm like this sounds great. And he's like, nnnnn I'm in pain. I think sometimes like the differences are really highlighted by like what's really unappealing to me can be appealing to a room full of dudes. And like what's going on, why is everyone just like giving their O-faces on this sound that's like super grating to my ears? But I mean like women have different hearing curves than men.

O: Oh, oh yeah ok…

L: The way women have better peripheral vision but like men have better focused forward vision is like I think there, I don't think anyone's like totally figured it out on the hearing differences, but, um I know we definitely have different curves. And I think sometimes that's like kind of sad because I think a lot of times there's a lot of music on the radio that's like dance music or stuff that is like targeted towards women, but there's not that many women that are like mixing, and there's not that many women that are actually creating it… I have a lot of friends who love pop music but they – even friends that don't know anything about sound, like girlfriends who are like "I like really like this song but there's something annoying about, like, I don't like having it in my headphones" I've

> heard that so many times that they don't like listening to anything on earbuds and it's just like, I wonder if it's all made by men [laughs]

Enjoying or failing to enjoy a particular sound, in Lucille's account, becomes a way of reinforcing and making sense of pre-existing social differences. The ability to listen to snare drum sounds from an Roland 808 drum machine, in her experience, has come to resonate with a distinction between masculine and feminine modes of listening. She articulates these differences, like those between microphones or speakers, in terms of frequency response patterns – or "hearing curves." What might have otherwise been a simple matter of personal preference, in this case, has been projected onto broader social formations within and outside of studio life – namely the pervasive gender disparity among engineers. An existing gender disparity within the studio can get articulated onto a particular sound-object, such that Lucille connects sound qualities with more general and less technically specifiable feelings of difference. It also becomes a resource for her to anticipate and relate to the experiences of the people who will listen to the music she helps produce, giving her a way of claiming expertise when it comes to understanding what, for example, women listeners want to hear.

As with Harold's compressor, the 808 snare was itself first designed to approximate and replace the sound of an actual percussionist striking a snare drum. Rather than eroding the distinction between human drummers and drum machines, however, it found enduring popularity precisely because people learned to hear the difference between it and a "real" snare. Many drummers now include trigger pads in their kits, which allow them to incorporate 808 samples into their live drum performances when they want its distinctive drum machine sound. As a universal drum kit sound, the 808 failed; its durability in music came from its ability to *not* sound like a traditional acoustic drum, and in doing so to carry specific generational, genre-based, and – as Lucille shows - even gendered associations as it passed from one context to another.

Just as the experience of gender difference manifested for Lucille in the form of a snare drum, for Carl, language of racial difference occasionally gets etched visibly into waveforms:

C: It's funny you'll get to recognize certain wavefo— words - by their waveform.
O: Like what? Like in the song or?
C: (ha) yeah
…
O: I'm like curious which words. Is it like from song to song or just within the song?
C: Aaaaaauuh. (there's a) few ones
O: Yeah?
C: Pretty sure I know what n****r looks like
O: [laughs]
C: 'course you know it matters who's saying it
For all the clean edits.

Here Carl is referring to yet another form of "cleanup" work – the removal of foul language (including racial slurs) from tracks in order to make them "clean" or broadcast friendly. Carl, who is white and has engineered many hip-hop projects, has had to erase this word so many times that its image as a waveform has been etched into his way of seeing waveforms. He seems, understandably, ambivalent about having acquired this skill. The word itself is a powerful shibboleth in the modern English-speaking world – the meaning and appropriateness of its use depend strongly on the interplay of, among other things, the cultural identity of the speaker, the particularities of its pronunciation, and its lyrical context. In Hip Hop music it is deployed with diverse and layered meanings – the explication of which I will not attempt here. As an engineer, Carl has come to recognize it as a waveform, something that jumps out as potentially requiring correction if the sound needs a clean edit. It is a particular way of "hearing the production" so as to not presume to evaluate the meaning of the performance itself. As Lucille and Carl demonstrate, engineers orient themselves towards the gendered and racial sensitivities of particular sounds and words, not so much in terms of how they personally experience these differences, but rather as a question of how others, as members of particular social categories might experience them.[87]

Like many practices in recording engineering, the use of shibboleths is scene-specific. Carl and I are looking at a picture of a Memphis studio and he begins explaining how you can tell who is who: "That guy is new in town, but you can tell he plays horn [by his cabbie hat]" "That guy knows a lot because he's got a beard and he's well groomed." Later, he recounts working on a session with a British producer and lead engineer, which meant that certain terms had unexpectedly different connotations. In the UK, for example, an IEC power cord is called a “mains.” When British engineers tells you that you have made a “brave choice” it does not mean he finds you courageous; it means they think you’re crazy.

Shibboleths also appear in the form of habits of perceptual imagery. This is particular apparent in the case of dynamic range compression. A compressor is a piece of recording equipment used to reduce the dynamic range, or difference in amplitude between the loud and quiet portions of audio signals. It works by continuously monitoring the amplitude of a track, raising the level when things get quiet and lowering it when things are too loud. Compressors first came into use as ways of maintaining a uniform volume level for radio broadcasts, but they have become one of the most important forms of signal processing currently used in music production. Any given track may incorporate multiple compressors, selected for their distinctive characteristics and dialed-in for a particular effect. Engineers speak of compressors “pumping”

[87] I owe this particular insight to Michael Lynch.

when they kick into action and attenuate loud signals, and "breathing" as they release their attenuating grip and let the quieter sounds and ambient noise rise back up to audibility.

Compressors are routinely deployed alongside an intensely physical imagery of squeezing. Think of flexing muscles and held breath during a deadlift, relaxing and exhaling once the weight is dropped. References to compression may include such words as "squash," "crush," "hammer," or "brick-wall." Because dynamic range compression makes tracks sound louder, and because louder is often interpreted as "better," the increasing use of compression is sometimes referred to as the "loudness war." Ricky Martin's "Livin' La Vida Loca," was both the first hit single to be produced completely within the Pro-Tools DAW environment and, at that point, a major contender for the loudest piece of recorded music ever produced – the atomic bomb of the loudness wars. Feminist critiques of recording studio culture have often noted the frequency of this violent, "macho" imagery as perpetuating the male dominance of recording engineering.[88]

The Key to the Reverb

Developing shared perceptual habits is an important part of the engineer-client relationship. We can see it especially clearly in a case where it fails to work. Seth, a tech and former mix engineer, relates his experience of a quite common misunderstanding regarding the relationship between the sonic qualities of reverberation and spatial presence.

S: Something I heard a lot the first few times that really threw me, is uh, I want my voice to be right out in front, very reverby. But the thing about reverb is that [reverb] pushes it back. So it can't be out in front and reverby. So you have to learn how to deal with people, to say well it can't really be… both [laughs]

[88] Pink Noise Online Magazine features two recent critiques of sexually exclusionary habits of speech in studio (Farmelo n.d.; Farmelo n.d.).

For his client, it was obvious that a vocal track could be simultaneously "up front" and "reverby." For Seth, however, these were contradictory requests. Through his training as a mix engineer he had developed a particular way of articulating spatial reasoning with the play of potential sonic effects. The singer, meanwhile, had a different, if no less experienced, way of articulating reverb and spatial position with regard to their voice. It is easy enough to recover their likely line of reasoning: when you sing in a non-reverberant room your voice manifests to you as small and muted, barely escaping the confines of your body as it has no room to bounce around and bloom back toward your ears. Singing in a large reverberant hall, by contrast, means hearing your voice fill the room, hearing it go out and return to you a hundred times over. To a singer, reverberation simply means 'more voice,' a louder voice, and a more physically present voice.

Because of the way they have learned to parse out and make spatio-temporal sense of a mix, engineers are more likely to understand reverb as a sort of blurring of the edges of the voice, making it harder to distinguish from other mix elements. To the singer, reverberation is something that extends the voice's reach. To the engineer, it is something that eats away at the contours of the voice, blurring it like a dense fog. Each of these perceptual habits blends the sensory codes of sound, light, spatial movement, and bodily awareness in distinct ways. Because these ways of perceiving are so engrained in habit and previous experience, it is difficult to switch ways of perceiving, even when it would make the job easier. As much as he wanted to give the client what they wanted, Seth simply could not hear reverb without also hearing distance. I ask Seth whether he ever told a client that they were asking for the impossible:

S: Well at first, and then you get yelled at. So you have to figure out what they actually want, and do it. And if you're in the right sync and you know 'cause that's why you might be the best engineer or whatever but you might not be right for the session because if you're not in sync with him on a non-verbal level… him or her, something like that, yeah. So yeah understanding, and that wasn't as big of an issue from an engineering standpoint because you used to have the traditional producer that was kind of like a bridge between the artist and you. Could fill in a blank. So you know as soon as he [the artist] asked that you could look at him [the producer] and go which one should I emphasize? And he'd say well the vocal, put him more up front, back the reverb off a little.

Harold cites another example of a client apparently asking the impossible:

> yo the reverb's in the wrong key" I had a guy do that. Everything was key. Like the key could be anything. The reverb could be the wrong key, or his double could be in the wrong key. Or like just any error was the key of the song. Which was great.

To an engineer, the idea of reverb being in a particular "key" is preposterous; reverb plugins, unlike pitch correction plugins, simply do not have "key" settings. To Harold, the client's use of "key" is catachrestic – a misuse or malapropism. His comment that this use "was great" was meant in terms of it being "hilarious" or absurd. To the client, however, the key of the reverb was apparently a genuine problem. Outside of the engineer's habitual language game, however, it is entirely workable to conceive of reverb, which simulates the ambient effect of a sound in a large resonant room, as being oriented to some key or another. Reverb can be tuned to higher or lower sets of frequencies, which often means certain notes of a scale are emphasized over others. Rather than further pursue the apparent misunderstanding between the client and himself, however, Harold took a more pragmatic "whatever you say" approach.

This example shows how shibboleths take on a catachrestic aspect. Catachresis means more than misuse; it also denotes a mixed metaphor or a tortured figure of speech, as in the phrase "take arms against a sea of troubles." A metaphor entails a carrying-over from one domain of meaning to another. A mixed metaphor means that two metaphorical gestures are colliding in some way. To an engineer, a key is a musical quality with a domain of application limited to the level of a set of pitch values. A song has a key in that it contains a set of notes that bear a certain pitch relationship with one another and with a set of frequencies with respect to a particular temperament and instrumentation. The concept of "musical key" is thus metaphorical in that it is a way of unlocking the logic of pitch relationships within the song, or decoding it as with a map. Reverb, meanwhile, refers to the sound of early reflections of sounds off of surfaces. Carrying the concept of key over to the domain of reverberation is very close to a mixed metaphor. In an industry where the roles of engineer and producer are increasingly being occupied by the same person, it is not clear how future confusions of this sort will be dealt with. There are more than enough interpretive ambiguities when it comes to assembling shared sensory experiences in the studio.

Just as often, however, it is the job of the engineer to perceive things in a particular way so that the client does not *have* to. One way of saying this is that engineers learn to hear the production in the music – the things that average listeners do not hear. They also learn to produce things – to make productive interventions – in ways that don't get in the way of the music itself. If, following Susan Leigh Star's approach (Star 1999), we attend to the infrastructural conditions of these practices, we find that the production of the music *qua* music is accomplished in this act of engineers working to obscure their own activities so as to emphasize those of the performer. The imperceptibility of a good production choice is never absolute and is always socially situated: engineers can hear (or, at least, claim to be able to hear) other engineers' production choices. Indeed, as shown in Chapter 6, many complained that hearing the production got in the way of their ability to hear the music, or at least enjoy it as "music" and not "production".

A production decision will frequently be recognizable as such primarily in terms of the *absence* of something in the signal. Engineers learn to perceive artifacts (or the absence of artifacts) in terms of their own practical experience with producing, removing, and managing artifacts. Skilled perception involves intentional (that is, goal-oriented, always-already-meaningful) coordinated sensori-motor movement, the sum of which calls the recorded object into being. Perceiving the marks of production involves bringing one's knowledge about the possibilities of production into one's assessment of it. In this way it is a form of perceptual knowledge that cannot be reduced to narrative, but is always situated within narrative. "Tuning" practices provide a good case of how engineers pride themselves on the invisibility (inaudibility, properly speaking) of their work.

A good engineer, for example, can digitally tune a vocal performance so as to make it sound as though it has not been tuned at all. For whom, then, is this act of tuning transparent? It is not transparent to the people who pay to have it done (clients: artists, producers, managers, record labels), at least in to the extent that they are deciding whether it happens. It is not invisible to other skilled engineers. They simply assume it has been done, they have done it before, they would do it if they were put into a similar situation. It *is* invisible to the artist in the sense that the artist does not necessarily want to know whether it has been done. It is, for the most part, inaudible to the public: regular listeners are not supposed to know (or, more precisely, care) that something has been fixed imperceptibly. The transparency of this work is often reflected in difficulty in getting credit for performing it. It is common for work of this sort to go uncredited (if not unpaid): Harold and Carl speculated together, only half-joking, that they would probably each have a Latin Grammy on their mantle if they were more diligent about making sure they were credited. Transparency is both an asset and a liability for the engineer, which means that it is best managed in partial, rather than absolute, terms.

ABX Tests

If the facts do not conform to the theory, they must be disposed of. – Perry's Law[89]

One example of a shibboleth that is typically used to constitute distinctions within the recording engineer social group itself is the listening test. One such test, The ABX, is a standard procedure developed at Bell Labs in order to determine whether there is a perceptible difference between any two stimuli.(Munson & Gardner 1950) It is also oft-touted as the ultimate arbiter of sophisticated gear and ears when engineers disagree about recording equipment or techniques. In cases of audio testing, ABX tests begin with two known sound sources, say two different analog-to-digital converters. The test then presents a third unlabeled sound, chosen at random from A and B. If A and B are perceptually indistinguishable, then the guesses will be random and the listener will only be right half of the time. Skeptics of expensive audio equipment like to point to the difficulty that engineers have identifying the difference between their favored gear and its less expensive alternative.

The fiction of the ABX test is that the relationship between the perceiver and the object perceived does not change over time: that the extent of a person's ability to perceive differences and the extent of an object's ability to produce perceptible difference is exhausted in a single setting. As Allen Farmelo writes, "The problem is that these tests assume that because two things

[89] Attributed to mastering engineer Ken Perry (Chinn 2014)

are close enough in a quick test that the difference will also be indistinguishable over long stretches of time. However, this assumption totally misses how it is that we tend to actually experience things in our very real lives." The ABX test is a way of staging similarity and difference both at the perceptual and the social level. That is to say that the ABX test as a shibboleth: for those favoring the "scientific" approach, the blind listening test is the ultimate standard of objectivity in-studio. For those who go by their ears, it indicates a failure to understand the importance of contextual and extended engagement with a sound.

Controversies over AB testing pitch two repertoires of recording engineering against one another: objective-scientific and skilled-artistic. On the one hand engineers pride themselves on providing a solid objective basis of knowledge that supports emotionally compelling performances. In this mode they serve as a proxy for a broader public. On the other hand they claim "big" or "golden" ears, meaning that they can hear things that others cannot. These claims culminate in two characteristic and diametrically opposed narratives. The first is what we might call a "emperor's new clothes" narrative, wherein an engineer claims to hear a difference that is later revealed to be impossible (often by way of the aforementioned "accidental bypass.") The second, opposite narrative might be termed "oracular" or "prophetic" – where only the engineer is able to perceive something that is subsequently revealed to be real. For the oracular engineer, the difference was always there but it is only revealed to everyone else at a later point.

In large part, engineers engage in these sorts of debates because they have a stake in two competing ideological commitments: 1) objectivity and clear-sighted understanding of the limitations of equipment and 2) refined perceptual skill, which can access differences others cannot. This is the basic ideological tension for the recording engineer. Marc Perlman has noted the closely related distinction between audio engineers and audiophiles.(Perlman 2004) Ethan Winer's, who falls decidedly in the "meter reader" camp, has become a sworn nemesis of Eric "Mixerman" Sarafin on this issue. Sarafin drew a distinction between Mix engineers and audiophiles, claiming that pragmatic compromise regarding the sound of equipment characterized the mix engineer's work, while the audiophile's narrow focus on fidelity may lead them to be overly credulous of certain technologies. He claims, for example, to be able to hear the difference between digital and analog summing (the process whereby the many tracks of a mix are combined into two master stereo tracks.) The difference between the camps represented by Winer and Sarafin is very close to the difference between "golden ears" and "meter readers" in Marc Perlman's 2004 study of hi-fi audiophile culture. Whereas Perlman's story was about the people who dedicated themselves to the skilled consumption of musical recordings, however, Winer and Sarafin's colleagues are interested in the skilled production of musical recordings.

According to Perlman, these groups organized around different practices for measuring the quality of a stereo system – meter readers put the most stock in the mechanical objectivity and visual readouts of signal analysis instrumentation. Their standard is one of transparency – the reproduction of a recording with the minimum added to the signal – which could in turn be measured by comparison of input and output signal tracings. Golden ears, by contrast, point to their own perceptual accounts as evidence, arguing that the inability of instruments to measure the difference they perceive is evidence of the inadequacy of the instruments and/or their corresponding theories. While meter readers call for double blind tests, golden ears reject this methodology as being overly artificial by not allowing them to spend the typical amount of time working with a piece of equipment and learning to hear the difference within a particular context.[90]

It was at a 2009 Audio Engineering Society conference panel that Ethan Winer made his first major assault on what he deemed unscientific audio engineering. Winer's position remains that engineers overestimate their ability to hear things, and are able to trick themselves into thinking that they are hearing something they are not. While he admits that many of the techniques he criticizes do have effects, he suggests that they are not perceptible in any meaningful way. Sarafin responded in his online forum:

> Ethan's missing the bigger point. ALL current "bronze age" audio transduction tools are woefully inadequate. IT'S ALL WHIM at this early juncture. Have fun, make music, buy stuff, sell stuff, lose mind. All else is folly… You could even buy "Real Traps" if you wanted to pretend that "flat" exists. Hey... don't laugh... the earth was considered flat at one point…. In fact, mixing is completely a game of compromise. Yet Ethan has proclaimed anyone calling bunk on his claims must be an audiophile. This is just further evidence that Ethan really has no clue what is involved in mixing. You can't possibly be an audiophile AND a mixer. The whole thinking process of an audiophile rejects compromise. This is why audiophiles will spend hundreds, if not thousands of dollars on a wire. Ethan constantly brings this up, as if any decent mixer would buy into that nonsense. We don't. Mixing actually precludes that sort of thinking. There is way too much gear involved in a mix to avoid a compromise. If I have two 1176s available on an analog mix, and I have three tracks that I want to use the 1176s on, I have to make a compromise. This sort of compromise is a way of life for a mixer. (Slipperman 2010)

The distinction between audiophiles and audio engineers is relevant here largely in terms of the fact that Winer began his career as a public skeptic through his writings on "audiophoolery" or the credulity of audiophiles with regard to such dubious and expensive products as "vinyl record demagnetizers" and high-end speaker wire. Audio engineers are apt to distinguish themselves as practitioners and producers of recorded sound, whereas audiophiles are consumers

[90] One example that a golden earist might point to is a 2013 study that found that skilled listeners were able to perceive differences beyond the so-called Fourier uncertainty threshold. It has long been held that there is a fundamental uncertainty relationship between the identification of a note's timing and its frequency composition. (Oppenheim & Magnasco 2013)

and thus more easily duped because they are less engaged in the actual production of the thing they are listening to. We might consider a culinary metaphor: a sophisticated diner may know much about cuisine, but a chef knows more by virtue of having made the meal. These two categories – engineer and audiophile - of course have a great deal of overlap, though they serve as professional repertoires that an engineer may deploy depending on the situation.

Conclusion: Cracks in the Voice

One way in which shibboleths might be thought to "break usefully" is suggested by studio tech Seth. He compares the distinction between the technical and the social with the distinction between a thing that is clearly broken a thing that "isn't quite working right."

S: "the technical end is – to me doing the tech stuff is a lot easier. Because there isn't that grey matter, there isn't that communication stuff. For my outside work it's 'this is broken – fix it.' Stuff that's broken is much easier to fix than stuff that isn't quite working right, 'cause when something's broken it's broken and it's much easier to find. When something's not quite working right you have to figure out which piece isn't quite working right. Much harder to find.

OM: And maybe sometimes difficult to agree that there's a problem in those sorts of situations.

S: Right, whereas that's where the ear still comes in. But yeah its much more technical end, fixing it's much more by the numbers. There's no… fixing other peoples gear that's already been designed and there's a factory spec of how other people expect it to work.

When a social situation is not quite working right, it may well prove more effective, from a given perspective, to go about producing a moment of breakdown. When engineers describe "faking out" their clients, or consider the possibility that their colleague's expensive AD converter will prove undetectable in a listening test, they are looking for breaking points that seem to provide a partially unambiguous answer to an otherwise ambiguous social situation.

Deception in science is rarely applied directly at a peer-to-peer level – this would be considered unethical (if occasionally effective, see: Blondlot's N-Rays, The Sokal Hoax, and

Hilgartner's discussion of on the methodological design of the Sokal Hoax.(Hilgartner 1997)) In scientific practice, deception is usually reserved for subjects: subjects in experimental psychology, laboratory animals, and orchids (Hustak & Myers 2013). Letting actual practitioners believe false things, even temporarily, is considered a dangerous proposition; thus double blind trials are used instead of trials where scientists allow themselves to be systematically misled by their colleagues. The field of recording engineering, by contrast, is one that is interested in the production of impressions which are understood to be subjective, emotional, irrational, etc.; at the same time, it is committed to a set of techniques rooted in fields such as electrical engineering, psychoacoustics, and music theory, all of which are considered relatively objective, ontologically sturdy, and generally not up for negotiation. As such, a recording engineer is very likely to manage the relationship between the technical and the artistic by quietly allowing disagreements between the perspective of the musician-client and the technical arrangement of the studio to proliferate. This is a different sort of "joint" through which perception is articulated in-studio. Such a move is likely to occur in a case where the client is thought to not know better – to not actually know what it is they want (at least in terms of how what they want ought to be achieved.) An engineer, for example, may tape over the meter of a compressor so that the client cannot see the VU meter being "pinned" – indicating that the compressor is being overburdened. An engineer, when asked by the client to turn up this or that effect, may choose to simply raise the playback level.

For Bergson, Husserl, and Schutz, the body was where inner temporality (or duration) and outer (objective or clock-time) temporality intersected. The body, in this sense, is a sort of boundary or fault line. The original shibboleth was a body betraying itself, a voice betraying a speaker, a dialect enlisted by an enemy. As a turning of a body against itself a shibboleth is a kind of torture – a practice Elaine Scarry has described as a way of making and unmaking the world. Scarry notes the key role of the voice as an existential boundary between self and world in the history of Western narrative: for both Oedipus and Lear "the voice becomes a final source of self-extension. only in silence do the edges of the self become coterminous with the edges of the body it will die with." (Scarry 1987, 33) To "speak for" is, first of all, to silence by stopping the proliferation of meaning from the thing spoken for. To repair is, first of all, to break by ceasing to use, by approaching a thing as broken, and by taking apart so as to apply repair. Shibboleths reside in the tension between catechesis - a discursive performance of pious fidelity - and catachresis – a semiotic breaching and heathenistic misuse. Garfinkel's breaching experiments can be read in part as the deliberate application of catachrestic social action in order to disrupt the instrumental transparency of various social methods and, in so doing, describing how they work. In Actor-Network terminology, shibboleths work by delegating acts of social demarcation to objects.

In 2007, the Columbian sculptor Doris Salcedo opened a 548 foot crack in the cement floor of the Tate Modern. The work, which she titled "Shibboleth," began as a tiny fracture beneath a wastebasket, eventually spreading and branching across the length of the gallery. It was meant as a commentary on boundaries, raising questions about "the interaction of sculpture and space, about architecture and the values it enshrines, and about the shaky ideological foundations on which Western notions of modernity are built."(Salcedo 2007) In its first eight months the crack caused 15 injuries. Following the exhibition the crack was filled in, leaving a visible scar. Salcedo had created a scultpure that could not be dismantled and added to a collection, but had instead to be repaired.[91] If Salcedo explored the concept of a shibboleth with a crack concrete, we can explore it further with the phenomenon of cracks in the voice. In speech pathology a voice is said to "break" or "crack" when it transitions between registers. A yodel, for example, involves a deliberate break between the modal and falsetto registers of the voice. Voices break during puberty, as the larynx grows and changes shape. Yodels are the exception that supports the rule that vocal cracks are typically heard as unintentional, often embarrassing things. It is one of the more common ways that a body seems to betray itself. Not coincidentally, it is also taken as an index of emotion. It is perhaps not surprising that the sound of overt auto-tuning, often likened to an electronic yodel, has been heard both as a tool of emotional expression and a marker of technical incompetence.

[91] Salcedo had in fact described much of her earlier sculptural practice as "art as repair" specifically the repair of the past through acts of commemoration and forgetting, as in her 2003 installation "1550 Chairs Stacked Between Two City Buildings."

BIBLIOGRAPHY

ABC, 2002. The World According to Cher. *ABC News: Primetime*. Available at: http://abcnews.go.com/Primetime/story?id=132108&page=1.

Ahl, D., 2009. Auto-tune - how do you hear it ?? *Gearslutz forum*. Available at: https://www.gearslutz.com/board/so-much-gear-so-little-time/379786-auto-tune-how-do-you-hear.html [Accessed April 7, 2009].

Akrich, M., 1992. The De-Scription of Technical Objects. In W. Bijker & J. Law, eds. *Shaping Technology / Building Society*. Cambridge, Mass: MIT Press, pp. 205–224.

Albini, S., 2015a. Mix With The Masters #23: Steve Albini. Available at: https://www.youtube.com/watch?v=TQonl6exhNU.

Albini, S., 2015b. *Steve Albini Analog Mix Session*, USA: Pro Studio Live. Available at: https://prostudiolive.com/archive/steve-albini-analog-mixing-session/88.

Almklov, P.G. & Hepsø, V., 2011. Between and beyond data: How analogue field experience informs the interpretation of remote data sources in petroleum reservoir geology. *Social Studies of Science*, 41(4), pp.539–561.

Alpert, M., 2014. Author Interview with Marco Alpert, VP of Marketing Antares Audio Technologies. April 29.

Anon, BeyondVerbal. Available at: http://www.beyondverbal.com/#home-section6 [Accessed January 1, 2016].

Anon, 2009. Classic Compressors. *Sound on Sound*. Available at: http://www.soundonsound.com/techniques/classic-compressors.

Antares Audio Technologies, 2000a. *ATR-1a Auto-Tune Intonation Processor Manual*, Scotts

Valley, CA: Antares Audio Technologies. Available at: http://www.antarestech.com/download/product-manuals.php.

Antares Audio Technologies, 2001. Auto-Tune 3 Owner's Manual. Available at: http://www.antarestech.com/download/product-manuals.php.

Antares Audio Technologies, 2011. *Auto-Tune 7 Owner's Manual*, Scotts Valley, CA: Antares Audio Technologies. Available at: http://www.antarestech.com/download/product-manuals.php.

Antares Audio Technologies, 2000b. *Auto-Tune Intonation Correcting Plug-In User's Manual*, Los Gatos, CA: Antares Audio Technologies. Available at: http://www.antarestech.com/download/product-manuals.php.

Assar, V., 2009. In Defense, Unsuccessfully, of Auto-Tune. *Tape Op*. Available at: http://tapeop.com/blog/2016/07/20/in_defense_unsuccessfully_of_a/.

Aucouturier, J.-J. et al., 2016. Covert digital manipulation of vocal emotion alter speakers' emotional states in a congruent direction. *Proceedings of the National Academy of Sciences of the United States of America*, 113(4), pp.948–53. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26755584 [Accessed October 10, 2016].

Aumann, P., 2011. The Distinctiveness of a Unifying Science: Cybernetics' Way to West Germany. *IEEE Annals of the History of Computing*, 33(4), pp.17–27.

BarelyHuman11, 2012. *The Autotune Ascension*, USA. Available at: https://www.youtube.com/watch?v=Powi4XNUV-o.

Barthes, R., 1977. *Image-Music-Text: Essays Selected and Translated by Stephen Heath*,

Bates, E., 2012. What Studios Do. *Journal on the Art of Record Production*, (7). Available at: http://arpjournal.com/what-studios-do/.

Bateson, G., 1987. *Steps to an Ecology of Mind*, Northvale, NJ: Jason Aronson.

Baxandall, M., 1988. *Painting and experience in fifteenth century Italy* 2nd ed., Oxford: Oxford University Press. Available at: https://newcatalog.library.cornell.edu/catalog/1517131.

Becker, H.S., 1951. The Professional Dance Musician and His Audience. *The American Journal of Sociology*, 57(2), pp.136–144.

Bergson, H., 1950. *Time and free will*, London: George Allen & Unwin Ltd. Available at: http://philpapers.org/rec/BERTAF-3 [Accessed June 16, 2014].

Beyondverbal, 2013. Access Moodies to Analyze Your Emotions and Attitude in Real Time. Available at: https://www.youtube.com/watch?v=t75uUoY4zdI.

Bhabha, H.K., 1994. *The Location of Culture* 1st ed., New York: Routledge.

Bijker, W.E., Hughes, T.P. & Pinch, T., 2012. *The Social Construction of Technological Systems*, Cambridge, Mass.: MIT Press.

Bingold, B., 2000. An Interns' Confession. *Tape Op*. Available at: http://tapeop.com/interviews/bonus/interns-confession/ [Accessed March 20, 2016].

Bloor, D., 1991. *Knowledge and Social Imagery* 2nd ed., Chicago and London: U of Chicago Press.

Bloor, D., 1989. Wittgenstein on Rule Following: The Old and the New Individualism. *The Polish Sociological Bulletin*, (3–4), pp.27–33.

Boersma, P. & Weenink, D., 2013. Praat: doing phonetics by computer [Computer program].

Branch, C., 2014. T-Pain Doesn't Think Kanye West Uses Auto-Tune "Correctly." *The Huffington Post*. Available at: http://www.huffingtonpost.com/2014/10/28/t-pain-kanye-west_n_6063160.html.

Buchla, D., 2007. Don Buchla Lecture. In *Red Bull Music Academy*. Toronto, ON. Available at:

https://www.youtube.com/watch?v=yfOyAUv-NwA.

Callon, M. & Latour, B., 1981. Unscrewing the big Leviathan : how actors macro-structure reality and how sociologists help them to do so. In K. Knorr-Cetina & A. V. Cicourel, eds. *Advances in Social Theory and Methodology: Toward an integration of micro and macro-sociologies*. Boston: Routledge & Kegan Paul, pp. 277–303.

Cannam, C., Landone, C. & Sandler, M., 2010. Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files. In *Proceedings of the ACM Multimedia 2010 International Conference*. Available at: http://sonicvisualiser.org.

Celemony, 2005. *Melodyne Users Manual*, Munich: Celemony Software GmbH. Available at: http://ww.celemony.com/.

Chakrabarty, D., 2000. *Provincializing Europe: Postcolonial Thought and Historical Difference*, Princeton, N.J. : Princeton University Press.

Chelin, P., 2015. Gayes' lawyer says he has track of Robin's vocals time-stamped 6:14pm. But engineer Coleman says vocals did not begin until after 10pm. *Twitter*. Available at: https://embed.scribblelive.com/Embed/v7.aspx?Id=1108754&Page=0&overlay=false [Accessed March 4, 2016].

Chen, A., 2015. How Transgender Women are Training Their Voices to Sound More Feminine. *Smithsonian*. Available at: http://www.smithsonianmag.com/innovation/how-transgender-women-are-training-their-voices-sound-more-feminine-180957537/.

Chen, H., Ishiharar, Y.C. & Zeng, F.-G., 2005. Pitch discrimination of patterned electric stimulation. *Journal of the Acoustical Society of America*, 118(1), pp.338–345.

Chinn, R., 2014. Laws For Audio Engineers. *PNW Audio Law Review Committee*. Available at: http://www.aes.org/sections/pnw/laws.htm [Accessed July 1, 2016].

Clayton, J., 2016. *Uproot: Travels in 21st-Century Music and Digital Culture* 1st ed., New York: Farrar, Straus and Giroux.

Cleveland, B., 2001. *Joe Meek's bold techniques*, Vallejo, CA: Hal Leonard. Available at: https://newcatalog.library.cornell.edu/catalog/4167194.

Collins, H., 1992. *Changing Order: Replication and Induction in Scientific Practice*, Chicago: University of Chicago Press. Available at: https://books.google.com/books/about/Changing_Order.html?id=gPDHOo2IUbMC&pgis=1 [Accessed February 23, 2016].

Collins, H.M., 2010. *Tacit and explicit knowledge*, Chicago : The University of Chicago Press,.

Collins, H.M., 1975. The Seven Sexes: A Study in the Sociology of a Phenomenon, or the Replication of Experiments in Physics. *Sociology*, 9(2), pp.205–224. Available at: http://soc.sagepub.com/content/9/2/205.abstract%5Cnhttp://soc.sagepub.com/content/9/2/205.short [Accessed June 16, 2014].

Cooper, G., 1975. *Wish You Were Here Songbook*, London: Music Sales.

Coopman, C. et al. eds., 2014. *Representation in Scientific Practice Revisited*, Cambridge, Mass: MIT Press.

Corbin, A., 1998. *Village Bells: The Culture of the Senses in the Nineteenth-Century French Countryside*, New York: Columbia University Press.

Coroneos, T., 2014. Strait's "Cowboy Rides Away" Album Butchered By Auto-Tune. *Saving Country Music*. Available at: http://www.savingcountrymusic.com/george-straits-the-cowboy-rides-away-album-butchered-by-auto-tune/.

Crane, L., 2016a. Gavin Lurssen: Mastering with a Master. *Tape Op*, p.94. Available at: http://tapeop.com/interviews/112/gavin-lurssen/.

Crane, L., 2016b. Intro/TOC #115. *Tape Op*. Available at: http://tapeop.com/columns/toc/115/.

Crane, L., 2000. Scott Fritz of KRCW. *Tape Op*. Available at: http://tapeop.com/interviews/20/scott-fritz-krcw/.

Crane, L., 2013. Welcome to the Service Industry. *Tape Op*. Available at: http://tapeop.com/columns/end-rant/98/.

Crockett, Z., 2015. The Inventor of Auto-Tune. *Priceonomics*. Available at: http://priceonomics.com/the-inventor-of-auto-tune/.

Csordas, T.J., 1990. Embodiment as a Paradigm for Anthropology. *Ethos*, 18(1), pp.5–47.

D, M., 2009. Auto-tune - how do you hear it ?? *Gearslutz Forum*, p.1. Available at: https://www.gearslutz.com/board/so-much-gear-so-little-time/379786-auto-tune-how-do-you-hear.html [Accessed April 6, 2009].

Dasgupta, P., 2015. I am part of an industry that is completely Fake. *Times of India*. Available at: http://timesofindia.indiatimes.com/city/kolkata/I-am-part-of-an-industry-that-is-completely-fake/articleshow/48209393.cms [Accessed September 30, 2015].

Daston, L., 1995. The Moral Economy of Science. *Osiris*, 10, pp.3–24.

Daston, L. & Galison, P., 1992. The Image of Objectivity. *Representations*.

Davis, E., 2015. Expanding Mind - Our Robots, Ourselves: Interview with David Mindell. Available at: http://prn.fm/expanding-mind-our-robots-ourselves-12-10-15/.

Dear, P., 2006. *Intelligibility of Nature*, University of Chicago Press.

Despret, V., 2004. The Body We Care For : Figures of. , 10(September 1904), pp.111–134.

Dickinson, K., 2001. “Believe”? Vocoders, Digitalised Female Identity and Camp. *Popular Music*, 20(3).

Dovey, R., 2009. Death Cab for Cutie takes a stand against Auto-Tune. *Paste*. Available at:

http://www.pastemagazine.com/articles/2009/02/death-cab-takes-stand-against-auto-tune.html [Accessed September 30, 2015].

Downey, G.L., Dumit, J. & Williams, S., 1995. Cyborg anthropology. *Cultural Anthropology*, 10(2).

Dror, O., 1999. The scientific image of emotion: Experience and technologies of inscription. *CONFIGURATIONS*, 7(3), pp.355–401.

Dror, O.E., 2000. Mesmerized: Powers of Mind in Victorian Britain. *The Journal of Interdisciplinary History*, 31(1).

Drordy, K.S., 2015. How to recognize if Auto-Tune has been used in a recording? Is it needed? *Stack Exchange Forum*, p.1. Available at: http://sound.stackexchange.com/questions/26892/how-to-recognize-if-auto-tune-has-been-used-in-a-recording-is-it-needed [Accessed December 24, 2015].

Dunbar-Hester, C., 2010. Listening to Cybernetics : Music, Machines, and Nervous Systems, 1950-1980. *Science, Technology & Human Values*, 35, p.113.

Earwicker, D., 2010. How to recognize if Auto-Tune has been used in a recording? Is it needed? *Stack Exchange Forum*. Available at: http://sound.stackexchange.com/questions/26892/how-to-recognize-if-auto-tune-has-been-used-in-a-recording-is-it-needed [Accessed December 17, 2010].

Eccleston, D., 2014. Inside Abbey Road With Beatles Engineer Ken Scott. *Mojo*. Available at: http://www.mojo4music.com/13343/inside-abbey-road-beatles-engineer-ken-scott/.

EmoSpeech, 2012. EMOSpeech. *Speech emotion Recognition*. Available at: https://emospeech.net.

Fales, C., 2002. The Paradox of Timbre. *Ethnomusicology*, 46(1), p.56. Available at:

http://www.jstor.org/stable/852808?origin=crossref [Accessed September 27, 2016].

Fallforward, R., 2009. Auto-Tune - How do you hear it? *Gearslutz forum*. Available at: https://www.gearslutz.com/board/so-much-gear-so-little-time/379786-auto-tune-how-do-you-hear.html [Accessed April 7, 2009].

Farmelo, A., Show Us Your Racks! – The Sexualization of Recording Gear As Exclusionary Ritual | pink noise. Available at: http://pinknoisemag.com/essays/show-us-your-racks [Accessed March 17, 2016a].

Farmelo, A., The Facebook Rant That Started Pink Noise | pink noise. Available at: http://pinknoisemag.com/letters/the-facebook-rant [Accessed March 17, 2016b].

Foucault, M., 1984. Of Other Spaces: Utopias and Heterotopias J. Misckowiec, ed. *Architecture /Mouvement/ Continuité*, (5), pp.46–49.

Foucault, M., 1979. what is an author? *Paul Rabinow, The Foucault Reader*. Available at: http://publica.fraunhofer.de/izm/mrs.htm [Accessed June 16, 2014].

Fouche, R., 2006. Say It Loud, I'm Black and I'm Proud: African Americans, American Artifactual Culture, and Black Vernacular Technological Creativity. *American Quarterly*, pp.639–661.

Francis, D. & Hart, C., 1997. Narrative Intelligibility and Membership Categorisatoin. In H. Stephen & P. Eglin, eds. *Culture in Action: Studies in Membership Categorization Analysis*. University Press of America, p. 150.

Frere-Jones, S., 2008. The Gerbil's Revenge. *The New Yorker*. Available at: http://www.newyorker.com/magazine/2008/06/09/the-gerbils-revenge.

Frow, E., 2012. Drawing a line: Setting Guidelines for digial image processing in scientific journal articles. *Social Studies of Science*, 42(3), pp.369–392.

Galison, P., 2003. *Einstein's clocks, Poincaré's maps*, New York: W.W. Norton.

Galison, P., 1997. *Image and logic : a material culture of microphysics /*, Chicago : University of Chicago Press,.

Galison, P., 1994. The Ontology of the Enemy: Norbert Wiener and the Cybernetic Vision. *Critical Inquiry*, 21(1), p.228.

Ganz, J., 2012. The Truth About Van Halen and Those Brown M&Ms. *National Public Radio*. Available at: http://www.npr.org/sections/therecord/2012/02/14/146880432/the-truth-about-van-halen-and-those-brown-m-ms.

Gardner, E., 2011. T-Pain Takes on Auto-Tune Maker in New Lawsuit. *Billboard*. Available at: http://www.billboard.com/articles/news/468941/t-pain-takes-on-auto-tune-maker-in-new-lawsuit.

Garfinkel, H., 1986. *Ethnomethodological Studies of Work*, Oxford, UK: Routledge & Kegan Paul.

Garud, R. & Karnoe, P., 2001. *Path Dependence and Creation*, Psychology Press.

Gates, Henry Louis, J., 1988. *The Signifying Monkey: a theory of Afro-American literary criticism*, New York: Oxford University Press.

Ghosh, S.S. et al., 2016. Speaking one's mind: Vocal biomarkers of depression and Parkinson disease. *Journal of the Acoustical Society of America*, 139(4), pp.2193–2193. Available at: http://scitation.aip.org/content/asa/journal/jasa/139/4/10.1121/1.4950530 [Accessed October 10, 2016].

Gibson, D., 2005. *The Art of Mixing* 2nd ed., Boston: Thompson and Artist Pro Publishing.

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

Giddings, F.H., M.A. & Ph.D., 1910. *The Elements of Sociology*, Available at:

https://books.google.com/books?id=w2U0yf-NQkkC&pgis=1 [Accessed March 18, 2016].

Gieryn, T.F., 1983. Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists. *American Sociological Review*, 48(6), pp.781–795.

Gimenez, L., 2009. Auto-tune - How do you hear it ?? *Gearslutz forum*. Available at: https://www.gearslutz.com/board/so-much-gear-so-little-time/379786-auto-tune-how-do-you-hear.html [Accessed April 7, 2009].

Gombrich, E.H., 1977. *Art and Illusion* 5th ed., London: Phaidon.

Goodwin, C., 1994. Professional Vision. *American Anthropologist*, 96(3), pp.606–633. Available at: http://doi.wiley.com/10.1525/aa.1994.96.3.02a00100.

Goodwin, C., 1995. Seeing in Depth. *Social Studies of Science*, 25(2), pp.237–274.

Goodwin, M.H., Cekaite, A. & Goodwin, C., 2012. Emotion as Stance. In M.-L. Sorjonen & A. Perakyla, eds. *Emotion in Interaction*. Oxford: Oxford University Press, pp. 16–41.

Gray, R.M., 2005. The 1974 origins of VoIP. *IEEE Signal Processing Magazine*, 22(4), pp.87–90. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1458295 [Accessed August 25, 2015].

Haase, R., 1969. *Geschichte des harmonikalen Pythagoreismus.*, Wien: Lafite.

Halliburton, OpenWorks® Project Data Management. Available at: https://www.landmark.solutions/OpenWorks.

Haraway, D., 1988. Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, 14(3), p.575. Available at: http://www.jstor.org/stable/3178066?origin=crossref [Accessed August 19, 2015].

Haraway, D., 1999. The Promises of Monsters: A Regenerative Politics for Inappropriate/d

Others. *Politica y Sociedad*, (30), pp.121–163.

Haruch, S., 2010. Women account for less than 5 percent of producers and engineers — but maybe not for long. *Nashville Scene*. Available at: http://www.nashvillescene.com/news/article/13034234/women-account-for-less-than-5-percent-of-producers-and-engineers-but-maybe-not-for-long.

Hayles, N.K., 1999. *How We Became Posthuman*, Chicago: University of Chicago Press.

Heidegger, M., 1996. *Being and Time*, Albany, NY: State University of New York Press.

Hein, E., 2015. Why is there so much Auto-Tune on everything? *The Ethan Hein Blog*. Available at: http://www.ethanhein.com/wp/2015/why-is-there-so-much-auto-tune-on-everything/ [Accessed April 2, 2015].

Henke, C.R., 1999. The Mechanics of Workplace Order: Toward a Sociology of Repair. *Berkeley Journal of Sociology*, 44, pp.55–81.

Hennion, A., 2013. An Intermediary Between Production and Consumption: The Producer of Popular Music. *Science, Technology & Human Values*, 14(4), pp.400–424.

Heritage, J. & Atkinson, J.M. eds., 1984. *Structures of Social Action: studies in conversation analysis*, Cambridge, UK: Cambridge University Press.

Hero, B. & Foulkrod, R., 1999. The lambdoma matrix and harmonic intervals (vol 18, pg 61, 1999). *IEEE Engineering in Medicine and Biology Magazine*, 18(3).

Hero, B. & Foulkrod, R.M., 1999. The Lambdoma matrix and harmonic intervals. *IEEE Engineering in Medicine and Biology Magazine*, 18(2), pp.61–73. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10101670 [Accessed September 28, 2015].

Hildebrand, H.A., 2015. A Message from Dr. Andy. *Antares Audio Technologies*. Available at: http://www.antarestech.com/email/2015Archive/2015_Mar_FlexTune.html.

Hildebrand, H.A., 2009. Auto-Tune: Expert Q&A. *Nova*. Available at: http://www.pbs.org/wgbh/nova/tech/hildebrand-auto-tune.html.

Hildebrand, H.A., 1976. *Nonlinear Filters For Estimating an Insect Population Density*. University of Illinois at Urbana-Champaign.

Hilgartner, S., 2000. *Science on Stage: Expert Advice as Public Drama*, Stanford, CA: Stanford University Press.

Hilgartner, S., 1997. The Sokal affair in context. *Science, Technology & Human Values*, 22(4), pp.506–522.

Hitchins, R., 2014. *Vibe Merchants: The Sound Creators of Jamaican Popular Music*, Burlington, VT: Ashgate.

Hong, A., 2013. iZotope: Alloy 2 Plug-in. *Tape Op*. Available at: http://tapeop.com/reviews/gear/94/alloy-2-plug-in/.

Hong, A., 2008. Safe Sound Audio: Dynamics Toolbox. *Tape Op*. Available at: http://tapeop.com/reviews/gear/64/dynamics-toolbox/.

Hong, A., 2014. The Inaugural Advanced Audio + Application Exchange (A3E) in Boston: A Gear Geek Love Letter. *Tape Op*.

Horning, S.S., 2013. *Chasing Sound*, Baltimore, MD: Johns Hopkins University Press.

Horning, S.S., 2004. Engineering the Performance. *Social Studies of Science*.

Hsu, H., 2015. *The Appsmiths: community, identity, affect and ideology among Cocoa developers from NeXT to iPhone*. Cornell University. Available at: https://newcatalog.library.cornell.edu/catalog/9255386.

Huet, E., 2015. Resurgence in Vinyl Records Means Booming Business - And Growing Pains - For Factories. *Forbes*. Available at:

http://www.forbes.com/sites/ellenhuet/2015/07/08/resurgence-in-vinyl-records-means-booming-business-and-growing-pains-for-factories/#6458e4107419.

Hughes, R., 2016. Pitch Correction Shoot Out. *Pro Tools Expert*. Available at: http://www.pro-tools-expert.com/home-page/2016/7/26/pitch-correction-shoot-out-vote-in-our-listening-test [Accessed July 26, 2016].

Hughes, S.A., 2003. What I Wish I'd Learned in School. *Mix*.

Hughes, T.P., 1983. *Networks of power*, Baltimore, MD: Johns Hopkins University Press.

Hughes, T.P., 1983. *Networks of Power*, Baltimore, MD: Johns Hopkins University Press.

Hui, B.A., 2012. Sound Objects and Sound Products : Standardizing a New Culture of Listening in the First Half of the Twentieth Century. *Culture Unbound*, 4, pp.599–616.

Husserl, E., 1964. *Phenomenology of Internal Time Consciousness*, Bloomington, IN: Indiana University Press.

Hustak, C. & Myers, N., 2013. Involutionary Momentum: Affective Ecologies and the Sciences of Plant/Insect Encounters. *Differences*, 23(3), pp.74–118. Available at: http://differences.dukejournals.org/cgi/doi/10.1215/10407391-1892907 [Accessed June 16, 2014].

Illich, I., 2000. The Loudspeaker on the Tower.

Ingold, T., 2011. *Being Alive*, London and New York: Routledge.

Ingold, T., 2000. *The Perception of the Environment*, New York: Routledge.

Isaacson, B., 2013. You've Never Heard Cell Phone Music Like This Before. *The Huffington Post*. Available at: http://www.huffingtonpost.com/2013/12/16/christopher-kirkley-saharan-cell-phones_n_4383777.html.

Isaksen, J.L., 2011. "Word Up": Hip Hoppers' Smart Pose Outwits Masculine Scripts. In A.

Wannamaker, ed. *Mediated Boyhoods: Boys, Teens, and Young Men in Popular Media and Culture*. New York: Peter Lang Publishing, pp. 145–65.

Jackson, M., Cascio, E. & Porte, J., 2010. Breaking News.

Jackson, S.J. et al., 2011. Collaborative rhythm. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work - CSCW '11*. New York, New York, USA: ACM Press, p. 245. Available at: http://portal.acm.org/citation.cfm?doid=1958824.1958861 [Accessed March 4, 2016].

Jackson, S.J., 2016. Speed, Time, Infrastructure: Temporalities of Breakdown, Maintenance, and Repair. In *Sociology of Speed*. Oxford, UK: Oxford University Press.

Jackson, S.J., Pompe, A. & Krieshok, G., 2012. Repair worlds. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work - CSCW '12*. New York, New York, USA: ACM Press, p. 107. Available at: http://dl.acm.org/citation.cfm?doid=2145204.2145224 [Accessed March 11, 2016].

Jackson, T.J., 2010. Sounds like Jason Malachi to me too... Available at: https://twitter.com/tjjackson/status/1505173436366848.

Jarzabkowski, P. & Pinch, T., 2013. Sociomateriality is "the New Black": accomplishing repurposing, reinscripting and repairing in context. *M@n@gement*, 16(5), pp.579–592.

Johns, G., 2014. *Sound Man*, New York, New York : Blue Rider Press,.

Kadry, H., 2015. The Straight Eight. *Pink Noise*. Available at: https://www.youtube.com/watch?v=ZXUIQzU8M6o.

Karp, H., 2014. The Secret Behind Pop Hits: Swedish Songwriters. *The Wall Street Journal*. Available at: http://www.wsj.com/articles/SB10001424052702303546204579435363582503386

[Accessed October 18, 2015].

Katz, B., 2007. *Mastering Audio: The Art and the Science* 2nd ed., Waltham, Mass: Focal Press.

Katz, M., 2010. *Capturing Sound: How Technology has Changed Music*, Berkeley, Los Angeles, London: University of California Press.

Kealy, E.R., 1979a. From Craft to Art: The Case of Sound Mixers and Popular Music. *Sociology of Work and Occupations*, 6(1), pp.3–29.

Kealy, E.R., 1979b. Solid Gold: The Popular Record Industry. *American Journal of Sociology*, 85(2), pp.453–455.

Keil, C., 1987. Participatory Discrepancies and the Power of Music. *Cultural Anthropology*, 2(3), pp.275–283.

Keil, C., 2014. Womb Envy: Recording Studios, Man Caves and Rehearsal Spaces as Modern Flute Houses. *Pink Noise*. Available at: http://pinknoisemag.com/essays/womb-envy-recording-studios-man-caves-and-rehearsal-spaces-as-modern-flute-houses.

Kelley, F., 2014. Tiny Desk: T-Pain. *National Public Radio*. Available at: http://www.npr.org/event/music/359661053/t-pain-tiny-desk-concert [Accessed October 29, 2014].

Kelly, J.L. & Lochbaum, C.C., 1973. *Speech Synthesis*,

Keltner, D. & Ekman, P., 2015. The Science of “Inside Out.” *New York Times*.

Kepler, J., 1978. *Weltharmonik*, Darmstadt: Wissenschaftliche Buchgesellschaft. Available at: https://newcatalog.library.cornell.edu/catalog/806160.

Kline, R., 2009. Where are the Cyborgs in Cybernetics? *Social Studies of Science*, 39(3), pp.331–362. Available at: http://sss.sagepub.com/cgi/doi/10.1177/0306312708101046 [Accessed June 16, 2014].

Kline, R.R., 1995. Construing "Technology" as "Applied Science": Public Rhetoric of Scientists and Engineers in the United States, 1880-1945. *Isis*, 86(2), pp.194–221. Available at: http://www.journals.uchicago.edu/doi/10.1086/357153 [Accessed September 27, 2016].

Kline, R.R., 2015. *The Cybernetics Moment, or, why we call our age the information age*, Baltimore, MD: Johns Hopkins University Press,.

Kline, R.R., 2004. What is information theory a theory of? *American Society for Information Science and Technology*, pp.13–28.

Knorr-Cetina, K. & Amann, K., 1990. Image dissection in natural scientific inquiry. *Science, Technology & Human Values*, 15(3), pp.259–283. Available at: http://sth.sagepub.com/content/15/3/259.short [Accessed June 16, 2014].

Kostelnik, A., 2016. Randall Dunn: Supporting Superstitions. *Tape Op*. Available at: http://tapeop.com/interviews/114/bonus-randall-dunn/.

KVR, 2013. Peter Neubäcker: A deep appreciation of musical instruments and a fascination with the science of sound. *KVR Audio*. Available at: http://www.kvraudio.com/interviews/peter-neub-cker-a-deep-appreciation-of-musical-instruments-and-a-fascination-with-the-science-of-sound-23118).

Lakoff, G. & Johnson, M., 1980. *Metaphors We Live By*, Chicago: University of Chicago Press.

Landes, D.S., 1983. *Revolution in Time: Clocks and the Making of the Modern World*, Cambridge, Mass: Harvard University Press.

Lapatine, S., 2006. Neko Case Hates Auto Tune, Not Babies. *Stereogum*. Available at: http://www.stereogum.com/2489/neko_case_hates_auto_tune_not_babies/wheres-the-beef/.

Latour, B., 2004. How to talk about the body? The normative dimension of science studies. *Body & society*, 10(2–3), pp.205–229. Available at:

http://bod.sagepub.com/cgi/doi/10.1177/1357034X04042943 [Accessed June 16, 2014].

Latour, B., 2005. *Reassembling the social : an introduction to actor-network-theory*, Oxford: Oxford University Press,.

Latour, B., 1987. *Science in action : how to follow scientists and engineers through society*, Cambridge, Mass: Harvard University Press,.

Latour, B., 1992. Where are The Missing Masses? In W. E. Bijker & J. Law, eds. *Shaping Technology*. Cambridge, Mass: MIT Press, pp. 225–258.

Latour, B. & Weibel, P., 2005. *Making Things Public - Atmospheres of Democracy* B. Latour & P. Weibel, eds., Cambridge, Mass: ZKM Center for Art and Media Karlsruhe.

Law, J., 2012. Technology and Heterogeneous Engineering: the Case of the Portuguese Expansion. In T. Bijker, Wiebe, Hughes, Thomas, Pinch, ed. *The Social Construction of Technological Systems*. Cambridge, Mass. : MIT Press, pp. 105–127.

Lawyer, L.C., Bates, C.C. & Rice, R.B., 2001. *Geophysics in the Affairs of Mankind: A Personalized History of Exploration Geophysics*, Tulsa, OK: Society for Exploration Geophysics.

Lepore, J., 2014. The Disruption Machine, What the Gospel of Innovation Gets Wrong. *The New Yorker*. Available at: http://www.newyorker.com/magazine/2014/06/23/the-disruption-machine.

Libertini, C., 2016. Melodyne 4 Vocal Editing and Harmony Creation - Warren Huart: Produce Like A Pro. Available at: https://www.youtube.com/watch?v=0r3P8l50a-M.

Lining, M., 2015. How Can You Tell if a Voice is Autotuned? *Quora*. Available at: https://www.gearslutz.com/board/so-much-gear-so-little-time/379786-auto-tune-how-do-you-hear.html [Accessed December 10, 2015].

Lynch, M., 1985. *Art and artifact in laboratory science: a study of shop work and shop talk in a research laboratory*, London; Boston: Routledge & Kegan Paul.

Lynch, M., 1991. Laboratory space and the technological complex: An investigation of topical contextures. *Science in Context*, 4(1), pp.51–78.

Lynch, M., 2013. Ontography: Investigating the production of things, deflating ontology. *Social Studies of Science*, 43(3), pp.444–462.

Lynch, M., 1988. The externalized retina: Selection and mathematization in the visual documentation of objects in the life sciences. *Human Studies*, 11(2–3), pp.201–234.

Lynch, M., 1984. "Turning up signs" in Neurobehavioral Diagnosis. *Symbolic Interaction*. Available at: http://onlinelibrary.wiley.com/doi/10.1525/si.1984.7.1.67/abstract [Accessed June 16, 2014].

Malin, B.J., 2011. Electrifying Speeches: Emotional Control and the Technological Aesthetic of the Voice in the Early 20th Century US. *Journal of Social History*, 45(1), pp.1–19. Available at: http://jsh.oxfordjournals.org/cgi/doi/10.1093/jsh/shr001.

Martin, A. & Lynch, M., 2009. Counting Things and People: The Practices and Politics of Counting. *Social Problems*, 56(2), pp.243–266. Available at: http://socpro.oxfordjournals.org/cgi/doi/10.1525/sp.2009.56.2.243 [Accessed August 19, 2015].

Marx, L., 1994. The Idea of "Technology" and Postmodern Pessimism. In Y. Ezrahi, E. Mendelsohn, & H. Segal, eds. *Technology, Pessimism, and Postmodernism*. Dordrecht: Kluwer Academic Publishers, pp. 11–28.

Massumi, B., 1995. The Autonomy of Affect. *Cultural Critique*, (31), pp.83–109.

Massy, S., 2016. *Recording Unhinged*, Milwaukee, WI: Hal Leonard.

Matousek, M., 2015. The Last Laugh: Is T-Pain the most influential popular musician of the past decade? *Student Life: The Independent Newspaper of Washington University in St. Louis*. Available at: http://www.studlife.com/cadenza/music/2015/12/06/the-last-laugh-is-t-pain-the-most-influential-popular-musician-of-the-past-decade/.

Maynard, D.W., 2013. Social Actions , Gestalt Coherence , and Designations of Disability : Lessons from and about Autism. *Social Problems*, 52(4), pp.499–524.

McDuling, J., 2014. "The Problem With Music" has been solved by the internet. *Quartz*. Available at: http://qz.com/202194/steve-albini-the-problem-with-music-has-been-solved-by-the-internet/#/h/64811,1/.

McGowan, P., 2014. Why Jitter Matters: Interview with Ted Smith. Available at: https://www.youtube.com/watch?v=7FyKVZ9D3E0 [Accessed January 24, 2014].

Meintjes, L., 2003. *Sound of Africa!: Making Music Zulu in a South African Studio*, Durham, NC: Duke University Press.

Merleau-Ponty, M., 1962. *Phenomenology of Perception*, New York: Routledge.

Mills, M., 2011. Deafening: Noise and the engineering of communication in the telephone system. *Grey Room*, (43), pp.118–143. Available at: http://www.mitpressjournals.org/doi/pdf/10.1162/GREY_a_00028 [Accessed June 16, 2014].

Mills, M., 2012. Media and Prosthesis: The Vocoder, the Artificial Larynx, and the History of Signal Processing. *Qui Parle: Critical Humanities and Social Sciences*, 21(1), pp.107–149.

MItroff, I., 1974. Norms and Counter-Norms in a Select Group of the Apollo Moon Scientists: A Case Study of the Ambivalence of Scientists. *American Sociological Review*, 39(4), pp.579–595.

Mody, C. & Lynch, M., 2010. Test objects and other epistemic things: a history of a nanoscale objects. *The British Journal for the History of Science*, 43(3), pp.423–458.

Mody, C.C.M., 2001. A Little Dirt Never Hurt Anyone:: Knowledge-Making and Contamination in Materials Science. *Social Studies of Science*, 31(1), pp.7–36. Available at: http://sss.sagepub.com/cgi/doi/10.1177/030631201031001002 [Accessed November 27, 2016].

Mohan, S., 2006. Technology: Call centers that feel your pain. Available at: www.cioinsight.com.

Mol, A., 2002. *The Body Multiple : Ontology in Medical Practice*, Durham, NC: Duke University Press.

Mudry, A. & Mills, M., 2013. The early history of the cochlear implant: a retrospective. *JAMA otolaryngology-- head & neck surgery*, 139(5), pp.446–53. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23681026 [Accessed August 19, 2015].

Mulkay, M.J., 1976. Norms and ideology in science. *Social Science Information*, 15(4–5), pp.637–656.

Munson, W.A. & Gardner, M.B., 1950. Standardizing Auditory Tests. *The Journal of the Acoustical Society of America*, 22(5), p.675. Available at: http://scitation.aip.org/content/asa/journal/jasa/22/5/10.1121/1.1917190 [Accessed November 21, 2016].

Myers, N., 2008. Molecular Embodiments and the Body-work of Modeling in Protein Crystallography. *Social Studies of Science*, 38(2), pp.163–199. Available at: http://sss.sagepub.com/cgi/doi/10.1177/0306312707082969 [Accessed June 16, 2014].

Myers, N. & Dumit, J., 2011. Haptic Creativity and the Mid-embodiments of Experimental Life.

In F. E. Mascia-Lees, ed. *A Companion to the Anthropology of the Body and Embodiment*. Blackwell, pp. 239–261.

Myerson, A.R., 1996. Halliburton Will Acquire Landmark Graphics in Stock Swap. *The New York Times*. Available at: http://www.nytimes.com/1996/07/02/business/halliburton-will-acquire-landmark-graphics-in-stock-swap.html?_r=0.

Najm, F., 2014. *Intro "Stoicville,"* USA. Available at: https://www.youtube.com/watch?v=KaIwNbN4AVA.

Najm, F. & Khaled, K., 2008. Karaoke.

Narcoman, 2009. Auto-tune - how do you hear it ?? *Gearslutz forum*, p.1. Available at: https://www.gearslutz.com/board/so-much-gear-so-little-time/379786-auto-tune-how-do-you-hear.html [Accessed April 7, 2009].

Neubäcker, P., 1994a. Harmonics and The Glass Bead Game. , pp.1–27.

Neubäcker, P., 1994b. What does ' just intonation ' mean ? , pp.1–30.

Neyfakh, 2014. The Sadness of T-Pain. *The New Yorker*. Available at: http://www.newyorker.com/magazine/2008/06/09/the-gerbils-revenge.

Nielson, E., 2010. "Can't C Me": Surveillance and Rap Music. *Journal of Black Studies*, 40(6), pp.1254–1274.

Noble, D.F., 1978. Social Choice in Machine Design: The Case of Automatically Controlled Machine Tools, and a Challenge for Labor. *Politics & Society*, 8(3–4), pp.313–347. Available at: http://pas.sagepub.com/cgi/doi/10.1177/003232927800800302 [Accessed June 16, 2014].

Nunes, F., Rajão, R. & Soares-Filho, B., 2016. Boundary work in climate policy making in Brazil: Reflections from the frontlines of the science-policy interface. *Environmental*

Science & Policy, 59, pp.85–92. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1462901116300338 [Accessed November 16, 2016].

Oppenheim, J.N. & Magnasco, M.O., 2013. Human Time-Frequency Acuity Beats the Fourier Uncertainty Principle. *Physical Review Letters*, 110(4), p.44301. Available at: http://link.aps.org/doi/10.1103/PhysRevLett.110.044301 [Accessed October 5, 2016].

Osumosan, 2009. Auto-tune - how do you hear it ?? *Gearslutz forum*. Available at: https://www.gearslutz.com/board/so-much-gear-so-little-time/379786-auto-tune-how-do-you-hear.html [Accessed April 6, 2009].

Owsinski, B., 2005. *The Recording Engineer's Handbook*, Hal Leonard Corporation. Available at: https://books.google.com/books?id=Emke4izOwVsC&pgis=1 [Accessed March 20, 2016].

Oyola, O., 2011. In Defense of Auto-Tune. *Sounding Out!* Available at: https://soundstudiesblog.com/2011/09/12/in-defense-of-auto-tune/.

Palmer, A. et al., 1998. Peter G. Winch (14 January 1926-27 April 1997). *Social Studies of Science*, 28(1), pp.169–175.

Patel, S. & Shrivastav, R., 2011. A preliminary model of emotional prosody using multidimensional scaling. *Proceedings of the Annual Conference of the International Speech Communication Association*, pp.2957–2960.

Pensado, D., 2014a. Engineer and Producer Bob Horn - Pensado's Place #178. *Pensado's Place*. Available at: https://www.youtube.com/watch?v=2X2c5-UX9Po [Accessed August 28, 2014].

Pensado, D., 2013a. Mix Engineer Mick Guzauski - Pensado's Place #115. Available at:

https://www.youtube.com/watch?v=aKmvfgCILHg.

Pensado, D., 2012. Pensado's Place #67 - Producer/Mixer/Author Eric "Mixerman" Sarafin. *Pensado's Place*. Available at: https://www.youtube.com/watch?v=u0k_pytZIss.

Pensado, D., 2014b. Producer Howard Benson - Pensado's Place #172. Available at: https://www.youtube.com/watch?v=ZXUIQzU8M6o.

Pensado, D., 2013b. Tony Maserati - Pensado's Place #126. Available at: https://www.youtube.com/watch?v=2LUWjr-gpaU.

Perlman, M., 2004. Golden Ears and Meter Readers. *Social Studies of Science*, 34(5), pp.783–807.

Pettman, D., 2011. Pavlov's Podcast: The Acousmatic Voice in the Age of MP3s. *Differences: A Journal of Feminist Cultural Studies*, 6(2–3), pp.140–67.

Pickering, A., 1993a. *The Mangle of Practice*, Chicago: University of Chicago Press.

Pickering, A., 1993b. The mangle of practice: Agency and emergence in the sociology of science. *The American Journal of Sociology*, 99(3).

Piekut, B. & Stanyek, J., 2010. Deadness: Technologies of the Intermundane. *The Drama Review*, 54(1), pp.14–38.

Pinch, T., 1993. "Testing - One, Two, Three...Testing!": Toward a Sociology of Testing. *Science, Technology & Human Values*, 18(1), pp.25–41. Available at: http://sth.sagepub.com/cgi/doi/10.1177/016224399301800103 [Accessed August 19, 2015].

Pinch, T., 2010. The Invisible Technologies of Goffman's Sociology From the Merry-Go-Round to the Internet. *Technology and Culture*, 51(2), pp.409–424. Available at: http://muse.jhu.edu/content/crossref/journals/technology_and_culture/v051/51.2.pinch.html [Accessed August 19, 2015].

Pinch, T., 1985. Theory Testing in Science-The Case of Solar Neutrinos. *Philosophy of the Social Sciences*, 15(2), pp.167–187.

Pinch, T. & Trocco, F., 2002. *Analog Days*, Cambridge, Mass: Harvard University Press.

Pinch, T.J., 2016. “Bring on Sector Two!” The sounds of bent and broken circuits. *Sound Studies*, 2(1), pp.36–51.

Pinch, T.J. & Bijsterveld, K., 2003. “Should One Applaud?” Breaches and Boundaries in the Reception of New Technology in Music. *Technology and Culture*, 44(3), pp.536–559. Available at: http://muse.jhu.edu/content/crossref/journals/technology_and_culture/v044/44.3pinch.html [Accessed August 19, 2015].

Polanyi, M., 1966. *The Tacit Dimension*, Chicago, IL: University of Chicago Press.

Porcello, T., 2004. Speaking of Sound: Language and the Professionalization of Sound-Recording Engineers. *Social Studies of Science*, 34(5), pp.733–758.

Porcello, T., 1998. “Tails out”: Social Phenomenology and the Ethnographic Representation of Technology in Music-Making. *Ethnomusicology*, 42(3). Available at: http://www.jstor.org/stable/852851 [Accessed June 16, 2014].

Porter, T.M., 1996. *Trust in Numbers*, Princeton, NJ: Princeton University Press.

Poster, W.R., 2011. Emotion Detectors, Answering Machines, and E-Unions: Multi-Surveillances in the Global Interactive Service Industry. *American Behavioral Scientist*, 55(7), pp.868–901.

Prentice, R., 2013. *Bodies in Formation : An Ethnography of Anatomy and Surgery Education*, Durham, NC: Duke University Press.

Prentice, R., 2005. The Anatomy of a Surgical Simulation: The Mutual Articulation of Bodies in and through the Machine. *Social Studies of Science*, 35(6), pp.837–866. Available at: http://sss.sagepub.com/cgi/doi/10.1177/0306312705053351 [Accessed August 19, 2015].

Puddin, 2009. Auto-Tune: Tell-tale signs of its use? *Steve Hoffman Music Forums*. Available at: http://forums.stevehoffman.tv/threads/auto-tune-tell-tale-signs-of-its-use.187190/page-3.

Raf, 2009. Auto-Tune: Tell-tale signs of its use? *Steve Hoffman Music Forums*. Available at: http://forums.stevehoffman.tv/threads/auto-tune-tell-tale-signs-of-its-use.187190/ [Accessed July 3, 2009].

Read, M., 2010. How to recognize if Auto-Tune has been used in a recording? Is it needed? *Stack Exchange Forum*. Available at: http://sound.stackexchange.com/questions/26892/how-to-recognize-if-auto-tune-has-been-used-in-a-recording-is-it-needed [Accessed December 9, 2010].

Reuters, 2010. *Teddy Riley talks about producing tracks on Michael Jackson's first posthumous album "Michael,"* UK: Reuters. Available at: http://www.itnsource.com/en/shotlist/RTV/2010/12/13/RTV3180810/?s=teddy riley#.

Rodgers, T., 2011. *Synthesizing sound: metaphor in audio-technical discourse and synthesis history*. McGill University. Available at: http://digitool.library.mcgill.ca/webclient/StreamGate?folder_id=0&dvs=1475015088687~879.

Rogers, S.E., 2013. Berklee Music Cognition Clinic: Susan Rogers.

Rubel, M., The Studio Production Hierarchy. *Blackbird Academy Blog*. Available at: http://blog.theblackbirdacademy.com/the-studio-production-hierarchy.

Russolo, L., 1986. *The Art of Noises*, New York, N.Y: Pendragon Press.

Sacks, H., 1989. Lecture Thirteen: On Proverbs. *Human Studies*, 12(3/4), pp.365–378.

Sacks, H., 1992. *Lectures on Conversation Volumes I & II* G. Jefferson & E. A. Schegloff, eds., Malden, MA: Blackwell Publishing.

Salcedo, D., 2007. Doris Salcedo: Shibboleth. *Tate Modern*. Available at: http://www.tate.org.uk/whats-on/tate-modern/exhibition/unilever-series-doris-salcedo-shibboleth.

Sarafin, E., 2002. *The Daily Adventures of Mixerman* 1st ed., New York, NY: Backbeat Books.

Sarafin, E., 2010. *Zen and the art of mixing*, Milwaukee, WI: Hal Leonard.

Saxena, M., 2013. Sukhwinder Singh: Best singer today is an auto-tune device. *Gulf News Music*. Available at: http://gulfnews.com/life-style/music/sukhwinder-singh-best-singer-today-is-an-auto-tune-device-1.1195829 [Accessed September 30, 2015].

Scaggs, A., 2006. Q&A: Brandon Flowers. *Rolling Stone*.

Scarry, E., 1987. *The Body in Pain : the making and unmaking of the world*, Oxford, UK: Oxford University Press.

Schafer, R.M., 1967. *Ear Cleaning: Notes for An Experimental Music Course*, Toronto, ON: Clark and Cruickshank.

Schegloff, E.A., 2013. Ten Operations in Self-Initiated, Same-Turn Repair. In M. Hayashi & G. Raymond, eds. *Conversational Repair and Human Understanding*. Cambridge, UK: Cambridge Univ Press.

Schegloff, E.A., Jefferson, G. & Sacks, H., 1977. The Preference for Self-Correction in the Organization of Repair in Conversation. *Language*.

Schnupp, J., 2014. Pitch Perception, Models. *Encyclopeida of Computational Neuroscience*.

Schönherr, M., 2011. The Sound of a Stone. Available at: https://www.youtube.com/watch?v=u573PyXo-pY.

Schutz, A., 1976. Making Music Together: A Study in Social Relationship. *Collected papers II*, 4. Available at: http://link.springer.com/chapter/10.1007/978-94-010-1340-6_8 [Accessed June 18, 2014].

Scott, E., 2015. Willow and Jaden Smith have some very deep thoughts on autotune. *Cosmopolitan*. Available at: http://www.cosmopolitan.co.uk/entertainment/news/a33991/willow-jaden-smith-autotune-comments/ [Accessed September 30, 2015].

Scott, P., Richards, E. & Martin, B., 1990. Captives of Controversy: The Myth of the Neutral Social Researcher in Contemporary Scientific Controversies. *Science, Technology & Human Values*, 15(4), pp.474–494.

Seabrook, J., 2012. The Song Machine: The Hitmakers Behind Rihanna. *The New Yorker*. Available at: http://www.newyorker.com/magazine/2012/03/26/the-song-machine [Accessed October 5, 2015].

Seaver, N., 2009. Nothing Natural About Vocal Skill. *Noise For Airports*. Available at: http://noiseforairports.com/post/147647000/nothing-natural-about-vocal-skill [Accessed July 23, 2009].

Shapin, S., 2001. Proverbial Economies: How an Understanding of Some Linguistic and Social Features of Common Sense Can Throw Light on More Prestigious Bodies of Knowledge, Science for Example.

Shapin, S., 1996. *The Scientific Revolution*, U of Chicago P.

Sharda, M. et al., 2015. Fronto-Temporal Connectivity is Preserved During Sung but Not Spoken Word Listening, Across the Autism Spectrum. *Autism Research*, 8(2), pp.174–186. Available at: http://doi.wiley.com/10.1002/aur.1437 [Accessed October 10, 2016].

Shoemaker, T., 2008. Songwriters to Soundmen - Trina Shoemaker - Being a Sound Engineer. Available at: https://www.youtube.com/watch?v=LOsYmgRm6Po.

Sillitoe, S. & Bell, M., 1999. Recording Cher's "Believe." *Sound on Sound*. Available at: http://www.soundonsound.com/sos/feb99/articles/tracks661.htm [Accessed September 3, 2015].

Sjuggerud, S., 2015. Dr. Andy Hildebrand – Auto-Tune Creator and Options Trader. *Inside True Wealth*. Available at: http://insidetruewealth.stansberry.libsynpro.com/ep-6-dr-andy-hildebrand-auto-tune-creator-and-options-investor [Accessed October 1, 2015].

Slipperman, 2010. BIG NEWS!! Ethan Winer gets facepalmed on GS. Shunned by bottom feeders everywhere. Read here! *The Womb Forums*. Available at: http://thewombforums.com/showthread.php?14282-BIG-NEWS!!-Ethan-Winer-gets-facepalmed-on-GS-Shunned-by-bottom-feeders-everywhere-Read-here!/page14 [Accessed April 19, 2010].

Smith, J., 2008. Tearing Speech to Pieces: Voice technologies of the 1940s. *Music, Sound & the Moving Image*, 2(2), pp.183–206. Available at: http://gateway.library.qut.edu.au/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=vth&AN=38213186&site=ehost-live.

Smith, J.O., 2015. Lecture at Faust Programming Workshop, Center for Computer Research in Music and Acoustics, Stanford University.

Sourgens, H., 2007. Prof. Dr. Prof. Dr. med. Hildegard Sourgens Friends & Family- Music & Computers. *sourgens.de*. Available at: https://www.sourgens.de/medical-expertises-documentations/music-computers.php [Accessed October 18, 2015].

Stanfield, G., 2014. Great River: PWM-501 compressor (500-series). *Tape Op*. Available at:

http://tapeop.com/reviews/gear/102/pwm-501-compressor-500-series/.

Star, S.L., 1999. The ethnography of infrastructure. *American Behavioral Scientist*, 43(3), pp.377–391.

Star, S.L., 2010. This is Not a Boundary Object: Reflections on the Origin of a Concept. *Science, Technology, & Human Values*, 35(5), pp.601–617.

Star, S.L. & Griesemer, J.R., 1989. Institutional Ecology, `Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19(3), pp.387–420. Available at: http://sss.sagepub.com/cgi/doi/10.1177/030631289019003001 [Accessed March 4, 2015].

Star, S.L. & Strauss, A., 1999. Layers of Silence , Arenas of Voice : The Ecology of Visible and Invisible Work. *Computer Supported Cooperative Work*, 8(1995), pp.9–30. Available at: http://link.springer.com/10.1023/A:1008651105359 [Accessed August 19, 2015].

Stearns, P.N. & Stearns, C.Z., 1985. Emotionology: clarifying the history of emotions and emotional standards. *The American historical review*, 90(4), pp.813–836.

Steinhardt, S.B. & Jackson, S.J., 2014. Reconciling rhythms. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing - CSCW '14*. New York, New York, USA: ACM Press, pp. 134–145. Available at: http://dl.acm.org/citation.cfm?doid=2531602.2531736 [Accessed April 8, 2016].

Sterne, J., 2003a. Bourdieu, Technique and Technology. *Cultural Studies (London, United Kingdom)*, 17(3–4), pp.367–389.

Sterne, J., 2012. *MP3: The Meaning of a Format*, Durham, NC: Duke University Press.

Sterne, J., 2015. Space within Space: Artificial Reverb and the Detachable Echo. *Grey Room*, 60, pp.110–131. Available at: http://www.mitpressjournals.org/doi/10.1162/GREY_a_00177

[Accessed September 20, 2015].

Sterne, J., 2003b. *The Audible Past*, Durham, NC: Duke University Press.

Sterne, J. & Rodgers, T., 2011. The Poetics of Signal Processing. *Differences*, 22(2–3), pp.31–53. Available at: http://differences.dukejournals.org/cgi/doi/10.1215/10407391-1428834 [Accessed June 16, 2014].

Straus, 2009. Auto-tune - how do you hear it ?? *Gearslutz forum*. Available at: https://www.gearslutz.com/board/so-much-gear-so-little-time/379786-auto-tune-how-do-you-hear.html [Accessed April 7, 2009].

Suchman, L.A., 1985. *Plans and Situated Actions: The problem of human-machine communication*, Palo Alto, CA: Xerox.

Sudnow, D., 1993. *Ways of The Hand*, Cambridge: MIT Press.

Swift, T., 2014. For Taylor Swift, the Future of Music is a Love Story. *The Wall Street Journal*. Available at: http://www.wsj.com/articles/for-taylor-swift-the-future-of-music-is-a-love-story-1404763219.

Thielemann, H., 2010. Untangling Phase and Time in Monophonic Sounds. *Journal of Signal and Information Processing*, 1(1).

Thompson, E., 2002. *The Soundscape of Modernity: Architectural Acoustics and the Culture of Listening in America 1900-1933*, Cambridge, Mass: MIT Press.

Thornton, T., 2013. Tacit Knowledge and Its Antonyms. *Philosophia Scientae*, (17–3), pp.93–106. Available at: http://philosophiascientiae.revues.org/890 [Accessed November 7, 2016].

Tingen, P., 2015. Inside Track: Josh Groban's album Stages. *Sound on Sound*. Available at: http://www.soundonsound.com/techniques/inside-track-josh-grobans-album-stages.

Tisenkopfs, T. et al., 2015. Learning and Innovation in Agriculture and Rural Development: The

Use of the Concepts of Boundary Work and Boundary Objects. *Journal of agricultural education and extension*, 21(1).

Tiwari, S.V., 2016. Practicing with Ustad Nusrat was scary: Rahat Fateh Ali Khan. *Hindustan Times*. Available at: http://www.hindustantimes.com/music/practicing-with-ustad-nusrat-was-scary-rahat-fateh-ali-khan/story-Gsc0Bw27Z0IftJrPKFZXzN.html [Accessed September 30, 2015].

TLMusic, 2009. Auto-Tune: Tell-tale signs of its use? *Steve Hoffman Music Forums*. Available at: http://forums.stevehoffman.tv/threads/auto-tune-tell-tale-signs-of-its-use.187190/ [Accessed July 3, 2009].

Todd, J., 2014. T-Pain Explains "Future Doesn"t Know How to Use Autotune'. *Vlad TV*. Available at: http://www.vladtv.com/article/181080/t-pain-explains-future-doesnt-know-how-to-use-autotune?page=2.

Tompkins, D., 2010. *How to Wreck a Nice Beach: The Vocoder from World War II to Hip Hop*, Brooklyn: Melville House Publishing.

Traweek, S., 1992. *Beamtimes and Lifetimes: The World of High Energy Physicists*, Cambridge, Mass: Harvard University Press.

Turkle, S., 2007. Introduction: The Things That Matter. In Turkle, ed. *Evocative Objects*. Cambridge, Mass: MIT Press, pp. 3–10.

Turnaround, 2009. Auto-Tune: Tell-tale signs of its use? *Steve Hoffman Music Forums*, p.2. Available at: http://forums.stevehoffman.tv/threads/auto-tune-tell-tale-signs-of-its-use.187190/page-2 [Accessed October 4, 2009].

Turner, K.L., 2016. *This is the Sound of Irony: Music, Politics and Popular Culture*, Surrey: Ashgate Publishing Ltd.

Vernier, 2009. Auto-tune - How do you hear it ?? *Gearslutz forum*. Available at: Strait's "Cowboy Rides Away" Album Butchered By Auto-Tune [Accessed April 7, 2009].

Vertesi, J., 2015. *Seeing like a Rover : how robots, teams, and images craft knowledge of Mars*, Chicago: University of Chicago Press.

Voskuhl, A., 2004. Humans, Machines, and Conversations: An Ethnographic Study of the Making of Automatic Speech Recognition Technologies. *Social Studies of Science*, 34(3), pp.393–421. Available at: http://sss.sagepub.com/content/34/3/393.short [Accessed June 16, 2014].

Walden, J., 2015. Antares Auto-Tune 8. *Sound on Sound*. Available at: http://www.soundonsound.com/reviews/antares-auto-tune-8.

Wang, G. & Applebaum, M., 2013. Music & Technology: Ge Wang & Mark Applebaum. Available at: http://www.loc.gov/today/cyberlc/transcripts/2013/131026mus1400.txt.

Weeks, P., 2002. Performative Error-Correction in Music : A Problem for Ethnomethodological Description. *Human Studies*, 25, pp.359–385.

Weeks, P., 2013. Synchrony Lost , Synchrony Regained : The Achievement of Musical Co-Ordination Author (s): Peter Weeks Published by : Springer Stable URL : http://www.jstor.org/stable/20011105 . Your use of the JSTOR archive indicates your acceptance of the Terms & Con. , 19(2), pp.199–228.

Weheliye, A.G., 2002. "Feenin" Posthuman Voices in Contemporary Black Popular Music. *Social Text*, 20(2), pp.21–47.

Wiesman, B.M., 2003. Do you hear Auto-Tune on this song? *Avid Pro Audio Community*. Available at: http://duc.avid.com/showthread.php?t=87402 [Accessed October 28, 2003].

Williams, C., 2012. "We Gave R&B a New Lifeline": How Teddy Riley Invented New Jack

Swing. *The Atlantic*. Available at: http://www.theatlantic.com/entertainment/archive/2012/11/we-gave-r-b-a-new-lifeline-how-teddy-riley-invented-new-jack-swing/265503/.

Williams, P., 2014. Pharrell Williams “Blurred Lines” Testimony Unsealed. Available at: https://www.youtube.com/watch?v=vg4gGnojFGM [Accessed October 24, 2015].

Winer, E., 2014. Artifact Audibility. *Mix*. Available at: http://www.mixonline.com/news/profiles/artifact-audibility/366521.

Winer, E., 2012. *The Audio Expert*, Oxford: Focal Press.

Winner, L., 1993. Upon Opening the Black Box and Finding It Empty: Social Constructivism and the Philosophy of Technology. *Science, Technology & Human Values*, 18(3), pp.362–378.

Woolgar, S., 1990. Configuring the user: the case of usability trials. *Sociological Review*, 38(51), pp.58–99.