

FORMAL AND INFORMAL PEDAGOGIES

Believing in Race, Teaching Race, Hearing Race

The cultural belief that voices are the unmediated expression and evidence of transpersonal categories, such as gender and race, is strong. When discussing this reality, I often invoke the example of Charles Clifford, an African American man who, in 1999, was convicted for selling drugs on the basis that the perceived sound of his voice made him culpable for the crime.¹ Yet Clifford's imprisonment—a result of what his lawyers called “linguistic profiling”—runs against the grain of influential humanities scholarship that has carefully demonstrated that audiovisual markers of race are highly subjective. Moreover linguists have convincingly shown that word choices and pronunciation are tied to speech communities rather than to innate qualities. In my own field of musicology, critics have noted that influences of vocal styles originating in a given community are complex and often extend beyond social circles and across time, as well as musicians' strategic essentialist positioning. Then again, none of these inquiries has systematically mapped the way vocal timbre is entrained and perceived through racialized listening practices, thereby debunking the assumption that voice is an unmediated essence. This is particularly the case in relation to the large body of scholarship that has explored the rich cultural and performance history of diverse forms of African American music.

Although it may sound at first like a complete exaggeration, it is nonetheless true to say that the extensive scholarship on North American and African American musical traditions has never methodically demystified racial suppositions about vocal timbre. This striking problem is deeply embedded in a long

history of practices that involve measuring race. Audile techniques that render African American vocal timbre largely unquestioned today can be traced back to aspects of nineteenth-century scientific racism. The belief that the perceived racial component of vocal timbre could be scientifically measured originates in debates about craniometry, the pseudo-science that calibrated humans by race. My research shows that formal and informal vocal pedagogy and listening practices, which were built upon these very assumptions, understood that, like a resonating chamber, voice “sounded” the cranial dimensions—measurements that were already racialized. Although I am specifically concerned with the cultural-historical formation of one category of vocal timbre, I address the broader concern that researchers in the humanities have no method with which to account adequately for the micropolitics of timbral difference to which voice is still subjected.

Racialized conceptions of vocal timbre persist. But why has vocal timbre resisted analysis when most aspects of the racialized body have been critically treated? I contend that fundamental misconceptions about voice and vocal timbre have prevented careful and critical analysis. Therefore, if we think about this problem simply through questions around race, and avoid examining basic understandings of voice, we will fail to get to the root of how categories, including race, are constructed through vocal timbre. In other words, we will easily note the attitudes about people that are overlaid on top of voice, but we will not be able to identify the distinct building blocks with which racialized timbre is projected, perceived, manifested, and sustained.

However, if we do change underlying conceptions that sound and voice are expressions of essence, we may be able to analyze, and create adequate responses to, the racialization of vocal timbre. Thus, by critically examining a given group’s listening responses and judgments regarding connections between vocal timbre and a given social category, race, or ethnicity, we may begin to deconstruct the deeply held conceptions about sound, voice, and vocal timbre that give rise to racialized judgments. I’m quite aware of the work this entails and of the minute details that need attention, but I am also confident that once we identify our basic misconceptions about sound, voice, and vocal timbre, the process of listening to voices and the critical-analytical tools we use to discourse about voice will be healthier and transferable across fields.

Recall, then, in the introductory chapter, I identified three correctives to broad misconceptions about voice:

- Voice is not innate; it is cultural.
- Voice is not unique; it is collective.
- Voice’s source is not the singer; it’s the listener.

In this chapter I will further investigate these correctives and misconceptions by examining the vocal-pedagogical process as it unfolds during formal and informal lessons (i.e., everyday socialization). By doing so, we can better track the phase in the micropolitics of listening that I term entrainment. Specifically it is the figure of sound, as it pertains to racialized vocal timbral categories, that is entrained. In other words, the figure of sound is entrained into voices and is subsequently used to authenticate the very value system from which it was born. Voices that are heard and even those voices that sound according to racial timbral categories are akin to planted evidence in a criminal investigation: objects or biological traces intended to serve as proof of a fabricated story of difference. While the actual vocal apparatus or vocal sound has not been replaced by another person's vocal apparatus or sound, daily formal and informal voice lessons plant an investment in race into both vocalizers' bodies and listeners' assessments. In short, this chapter outlines the cultural-pedagogical work carried out to maintain the figure of sound.

By identifying the processes involved in decisions around the broad areas of enunciation, articulation, and intonation, we can begin to grasp—and deconstruct—the constructed aspect of timbre. Together these broad vocal processes contribute to oral and vocal tract shape, which in large part indirectly determines timbral characteristics. Vocalizers can also directly influence timbre by, for example, making a concerted effort to sound happy, disinterested, or stern. As choices around enunciation, articulation, and intonation are repeatedly made, they begin to form a pattern of vocalization that feels natural or second-nature to the vocalizer and is recognized as his or her default and consistent vocalization pattern by the people around him or her. Taking this process as a starting point, we may begin to deconstruct any notion that voice is innate and unique. By denaturalizing timbre and placing it on par with word, enunciation, and intonation choices, we see that voice is not innate and that timbre is also formed as a result of vocalization patterns that are repeated hourly and daily and are favored (or not) depending on their result. In other words, vocalizers adopt, keep, switch between, or discontinue timbral patterns.

Voice is not unique. Vocalization and the resulting timbre are as encultured as is self-expression through fashion. Akin to the ways we are habituated to dress, walk, or throw a ball “like a girl,” each vocal engagement is connected to a collective practice that depends on habituated micro-vocal maneuvers. This daily vocalization practice habituates flesh, muscles, and ligaments, leading to an altered vocal apparatus that, in turn, leads to altered sounds. Finally, the source of the voice is not only the vocalizer, because any statement about a voice arises from a listener's assessment. Thus, relevant to our inquiry into race and

vocal timbre, a given listener's attitudes about race will influence which aspects of a voice he or she will notice and how he or she will make meaning from them. It is when such basic assumptions about voice go unquestioned that we also fail to examine our responses to the acousmatic question: *Who is it who is speaking?*

I seek to explain responses to the acousmatic question through the three correctives about voice: voice is not innate, not unique, nor the singer. In identifying these correctives, I am indebted to Foucault, who famously discussed "the body as an object and target of power." The distinguishing aspect of the modern body, he claimed, is that it "is manipulated, shaped, trained," and "obeys, responds, becomes skillful and increases its forces." Not unlike a machine, it is "built, rebuilt, operationalized and modified."² We may productively consider the formation of this modern body through the concept of *body technologies*. Coined by Mauss, this term describes "one of the fundamental moments in history itself: education of the vision, education in walking—ascending, descending, running."³ Jonathan Sterne has added "the education and shaping of audition" to this list, and here I add the phenomenon of vocal timbre.⁴ While we all move, hear, and sing in idiosyncratic ways, body technologies intone these actions so they both project and affirm social structures of recognition—for example, "masculine" strides, "girl-like" throwing, or "upper-class" enunciation—or, in Michael and Linda Hutcheon's formulation, "To train the voice is to train the body."⁵ As Mauss points out, phenomenology assumes culture. I come to my consideration of voice with the assumption that vocal-timbral recognition is a type of encultured behavioral knowledge that manifests dynamics of difference and normativity. Specifically, sonic timbral markers that suggest a person's race or gender hold little meaning or power outside the cultural context within which they are defined.

Thus, naming vocal timbre is a kind of knowledge. That is, within the range of responses to the acousmatic question—the seemingly innocent *Who is this?*—the dynamic of power relations is played out. In response to the question, a statement is offered; therefore, it is in our trust in the question's validity, and in the assumption that it can be answered, that the micropolitics of timbre is carried out. The power dynamic is enacted through the incessant validation of culturally and socially tinged assessments of vocal timbre as knowledge.

Encultured actions create a certain set of behaviors, and these behaviors are set within the play of power. Vocal timbre is an area of body politics that has not yet been thoroughly examined as encultured performance. I wish to look more deeply at this area of human activity in order to denaturalize timbre and to illuminate some of the ways in which timbre, and listening to timbre, are encultured. I also want to examine vocal timbre as a means of considering how

we may intervene in the continuous cycle of vocal-timbral naturalization. Thus this chapter considers formal and informal vocal training in order to understand how racialized vocal timbre is taught and how the physical vocal apparatus is molded and habituated as a result of encultured listening practices.

One way to break away from this cycle, I posit, is to consider a given response to the acousmatic question from the perspective of critical performance practice—that is, to critically examine and experiment by reproducing the vocal practices that created particular timbres, and to test assumptions about essentialized timbre by experimenting with the range of timbres a given voice has the capacity to produce. Such a performative approach can allow us to deconstruct and denaturalize our responses to the acousmatic question. It can also help us identify and name the building blocks from which such naturalized notions are constructed. And by deconstructing notions of voice as innate, singular, and arising solely from within the vocalizer, we can begin to recognize the micro-politics that are carried out through listening. Finally, this allows us to begin to detail the incestuous process of assessing, manifesting, and validating, and thus to denaturalize vocal timbres. In this first chapter of *The Race of Sound*, then, I aim to establish in greater detail how misconceptions about voice affect not only its experience and perception but also the materiality of a voice and its vocalizer. This process, I assert, takes place through both formal and informal vocal instruction.

Sounding Race

As ocean depth is measured by taking soundings (historically using rope; now using sonar), race is sounded, or assessed, through the process of listening. Classical vocal artists undergo intense training, much of which is dedicated to learning to hear their own voices as the experts hear them. A decade of daily practice, weekly (or more) private lessons, monthly or quarterly master classes, sustained participation within the milieu of classical singers and musicians in the form of summer workshops or university or conservatory training, and opera apprenticeship programs constitute the pedagogical structure and business model for this world. The path toward a professional vocal career is an immersive experience and lifestyle. The following discussion draws on specific examples from the world of classical vocal training. It offers a foundation from which to discuss how any feedback given to a vocalizer about vocal usage contributes to his or her subsequent vocal choices and habituation. I seek to examine how general attitudes around sound play out when voices are listened to within the context of deeply held assumptions regarding difference.

The observations are drawn from my sixteen years of intense and direct participant observation of selected classical vocal music communities and training.⁶ While I am still in touch with the classical vocal world, my immersion in the community, including what I refer to as my period of participant observation, took place in Norway and Denmark (1991–99), New York City (1995–99), and southern California (1999–2007). In addition, over a period of a year I conducted thirteen interviews with voice teachers.⁷ In these conversations I asked general questions regarding what constitutes vocal timbre, how vocal timbre is developed, and what kinds of information vocal timbre is able to convey about the singer. When correct singing—in terms of vocal weight and color, both crucial issues in vocal pedagogy—was discussed, issues of race, ethnicity, and vocal timbre arose. In the thirteen interviews I carried out, all but two teachers told me that they can always tell the ethnicity of the singer by his or her vocal timbre. In the following discussion, I will draw on the sentiments expressed in all the interviews. However, two interviewees stood out as crystallizing these sentiments in their statements, and therefore the specific quotes are pulled from those conversations.⁸

Besides my in-depth knowledge of this performance and its associated pedagogical tradition, another reason to consider the relatively exclusive vocal practice of the classical music world is that, in general, teachers and practitioners of this vocal art are some of the most sensitive to—and systematic adopters of—timbral enculturation. Moreover I chose to concentrate on teachers and practitioners of classical music because of the genre's strict adherence to the written score, and of its uniformity in both pronunciation and overall performance practice. That is, for this study, the practice's institutionalization offered a type of baseline. However, I do not take classical vocal practitioners as exceptional or as different from other vocalizers. To the contrary, I take them to be “first adopters,” and indeed, broadly speaking, the practice is only a subset of formal and informal vocal pedagogy. Like a kind of subspecies that shows certain characteristics more strongly than the general population, classical vocal pedagogical practices can be read as an advance warning system of danger. According to my analysis, despite classical voice professionals' extensive vocal education, this subset of experts does not succeed in breaking the cycle of racialized timbre or in laying bare the process of enculturation. To the contrary, many continue to amplify and re-present these beliefs on respected and prestigious stages.

Vocal timbre is both elusive and poorly understood in the performing arts, humanities, medicine, and sciences alike, yet entrenched positions are held

tightly in each area of specialization. Vocal timbre, and what a given timbre signals, is one of those things most people assume they know. For example, most people trust timbre over words if, say, the words “I’m okay” seem to be contradicted timbrally, or if a voice on the radio sounds like it is an older black man or a young white woman. Classical voice teachers hold clearly articulated positions regarding timbre and its meaning. Rather than the unarticulated taxonomy most listen from, voice teachers’ work is to align vocal training with the vocal characters arising from particular cultural moments. Thus classical voice teachers train and pair timbre profiles with conceptual areas such as identity and authenticity. For both the layperson and the voice teacher, timbre is a barometer of one’s inner state and health and is broadly held, continuously assessed, and reliably acted upon. In fact vocal timbre is used as a diagnostic for gendered mental health issues (“She’s hysterical”) and as a diagnostic for truth statements (“She’s lying”). Major plots in both Western literature generally and the Bible specifically hinge on the characters’ ability to judge authenticity vocally.

Voice teachers tend to crystallize these general sentiments into two prevalent concerns around guiding the aesthetic development of vocal timbre: first, the question of what constitutes healthy and natural singing for the student; second, the need to avoid homogenizing students’ voices in favor of allowing each singer’s “true timbre” to emerge. When we discussed the “correctness” of vocal weight and tone color, which are crucial topics in vocal pedagogy, issues of what kinds of information these aspects convey about a singer also arose. Specifically, when fleshed out, conversations that began on topics of “healthy” vocal use and the “authentic” timbre of a given singer’s voice ended by discussing race and ethnicity.

The notion of “correctness” in vocal weight and tone color returns to issues of maintaining healthy, authentic, and beautiful voices. Interestingly, practices that the teachers I interviewed considered “healthy” and “honest” were ultimately correlated with each student’s race and ethnicity.⁹ One way to describe this situation is that, even when it was not referred to explicitly, race was discussed under the cloak of singing “healthily” or “authentically.” Because race has been thoroughly naturalized, what I describe as racialized vocal timbre is conceived by voice teachers as simply a healthy way of singing that promotes a nonhomogenized sound and that allows students to be “themselves.”¹⁰ Voice teachers and students commonly conceive of voices as unrealized or repressed due to any number of causes, from bad vocal habits—often conceptualized as “tensions”—to evidence of underlying physiological or mental issues. In short, a “healthy”-sounding voice is assumed to be a voice freed from blockages, and

thus is assumed to be an unmediated sonorous conduit for the subject. It follows that whatever voice teachers understand as the singer's "inner essence" will be equated with a sound voice and will be listened for during vocal "diagnosis."¹¹

For example, Dorothy, a soprano and professor of voice for seventeen years, told me that she can invariably identify whether a student is, for example, Armenian, Russian, or Korean from the student's vocal timbre, but she frames her classification of students as a concern about *vocal health*: "There are principles of what is healthy, a balanced sound and all of that, and if [voice teachers] observe that rule, then how can they not hear an Armenian sound or Korean sound and cultivate it?"¹² In this statement Dorothy reasons that if the voice is trained along principles designed to promote a vital, balanced sound, it will "naturally" display its inherent ethnicity, thereby conflating race, national identity, and vocal health.¹³

Rather than considering this strategy as a race- or ethnicity-based categorization of voices, Allison, another longtime teacher, views what she calls "ethnic timbre" as the "unique color" and vocal "fingerprint" of the student, yet associated with a racially categorized group. Pedagogy, then, becomes a matter of bringing out the "true sound" of the student's voice—and that true sound happens to be connected to his or her perceived race or ethnicity. Allison regards this pedagogical philosophy as a means of allowing each student to maintain an element of individuality within the highly cultivated and stylized world of classical singing. During the interview process, I frequently heard such statements regarding the *individuality* of a voice, by which my interviewees meant, I believe, the opposite: "an ethnic vocal timbre," determined by *socially* constructed notions of ethnicity. Indeed, an ethic of multiculturalism has permeated vocal pedagogy; Allison goes so far as to criticize ignorant teachers, who have not been exposed to a variety of "ethnic timbres," for "homogenizing" their students' sounds. And most teachers with whom I spoke stressed the importance of being literate readers of "ethnic" vocal timbres.

When we began to discuss what might cause the varied timbres of different ethnicities, Allison explained that the Central and South American timbre is influenced by Latin people's connection to their bodies. In her view, inhabitants of Latin cultures are motivated by bodily drives, while North American inhabitants are moved by cerebral concerns. She explained that singers' connections to their bodies affect their sounds: "The Mexican culture, for example, is, to me, a very visceral culture. It's not a super heady culture. I think we in the United States of America tend to be more cognitive. You know, the whole Puritan ethics where sex is bad and you just disallow that you have anything below your waist. You know, that is a primary drive in people." To clarify her sense of

the connection between body and timbre, I asked Allison whether she believed that some cultures come by that body-voice connection more naturally, so that even if a singer from one of those cultures studied with an American teacher, or a teacher who is not particularly focused on the development of the body-voice connection, his or her voice would still sound the connection that was “in” him or her from the beginning, and thus would differ from the voice of an Anglo-American growing up in the United States. Allison responded:

Yes. I think [Latin Americans] naturally have that connection. . . . They're . . . connected to their bodies . . . and their guts [said with throaty, “gut sound”], and they make music from their hearts. In European repertoire they talk about that “she broke my heart, I will just lay down and die now” [said with a very “proper” voice], and in Hispanic music, the Latino music: “She broke my heart, she ripped it out of my chest and stomped it on the floor!” [nearly screaming]. And that’s how their music sounds. It’s very gut. Americans—we don’t operate on that level, we tend to be a visual or cognitive society.

Allison expressed her claims in compassionate language and avowed a commitment to allowing the “natural” and “individual” voice to remain untouched through intense classical vocal training. Yet several interviewees used these notions of “naturalness” and “individuality” synonymously (if unconsciously) with ethnic, national, or racial difference.

While Allison and Dorothy articulated strong ideas regarding authenticity, the singer’s “real” voice, and ethnicity, all but two of the interviewees touched on these ideas in some way. And, as we will see below, their sentiments are shared by musicologists. While these appraisals are offered as personal assessments, evaluations cannot be assigned as their personal beliefs, but must be seen as examples of broad and pervasive cultural beliefs. Who cannot recall having made such judgments? These teachers carrying out one-on-one sessions with their voice students are doing what the broader sociocultural milieu sanctions by consuming such trained voices. This reinforcement parallels racialization in other corporeal realms, such as skin hue and hair texture. What is today cloaked in concepts of health, authenticity, and self-expression was, only half a century ago, unhesitatingly described as race and racial qualities.

In a 1957 encyclopedia entry titled “Primitive Music,” for example, the musicologist Marius Schneider posits, “Every being has its own sound or its own song, the timbre and rhythm of which embody the mystic substance of the owner.”¹⁴ Alan P. Merriam and Valerie Merriam observe, “Races are held to have special and mystic abilities, and what the anthropologist attributes to

learning and to culture, Schneider attributes to race.¹⁵ In Schneider's own words, some musical characteristics are "bound up with certain racial factors. . . . In fact, the innermost essence of the more intensely specialized types of song cannot be transmitted at all . . . since the dynamic and vocal timbre which is inseparably bound up with it cannot be acquired by learning."¹⁶ According to Schneider, whichever vocal qualities are heard as expressing "certain racial factors" are understood within this listening framework as nonnegotiable expressions. The consequence of such a listening position is that meaning is formed within a rigid and closed cycle.

Historically speaking, my informants' investment in race, as classical vocal pedagogues, is far from anomalous, although their frankness on this question is worth noting.¹⁷ As mentioned earlier, all but two teachers claimed to hear singers' ethnicities in their vocal timbres. That is, teachers' perceptions of students' ethnicities shape their understanding of how the students might develop as singers, and further direct teachers' ears.¹⁸ These beliefs lead mentors to encourage certain timbral features over others, which causes the re-enculturation of racialized vocal timbre. Not only do these specific teachers teach in this way, but vocal pedagogical ideologies as a whole are built on specific ideas of timbre.

In fact the classical vocal pedagogy practiced today in southern California (and elsewhere in the United States) can be traced back to the formation, during the mid-nineteenth century, of what John Potter has called the modern classical voice. For Potter, the formalization of vocal pedagogy grounded in scientific principles marks the transition from the premodern to the modern classical voice.¹⁹ Modern classical vocal pedagogy's advances and its questionable notion of "the natural" were aided in part by findings encouraged and enabled by colonial racial dynamics and research resting on colonial power structures. As Potter observed, the ideologies powering the formation of the modern classical voice are still present in current vocal practices.²⁰ Despite having substantial specific knowledge about voice, voice teachers—like most people—hear race (or health or authenticity) as communicated through essential timbral qualities that are presumed to tell us something unmediated about a person's internal state.²¹

As noted in the introduction, I suggest that these racialized assessments do not flow directly from, nor are they solely enabled by, racial sentiment in a given culture and society. It is not, contrary to the expression, turtles all the way down. Instead, I would argue, this type of listening is propped up by general assumptions about the nature of sound and its ontology (i.e., assumptions

regarding what we can know about sound and its meaning). It is these underlying misconceptions about sound that lead, on the one hand, to misconstrued notions of voice and to unfounded trust in a given response to the acousmatic question, on the other. That is, the type of listening reported above is supported by unexamined assumptions regarding what kinds of meaning sound is capable of communicating.

Thus, while not comparable in effect or ramification, listening assessments ranging from mood, gender, age, health, authenticity, class, ethnicity, and racial tone arise from a general framework that involves the possibility of quantifying and knowing sound. All of these assessments are enabled by the naturalization of sound as a specific knowable entity (e.g., A-sharp) and as a more ambiguous knowable entity (e.g., as authentic or truthful). While, of course, many racialized assessments do arise from an investment in difference, some arise from a less informed place, a place of wanting to “be right”—to “name the right sound,” to adhere to fidelity—an impulse that is nonetheless equally impactful.

In the same way that I argue within this book that a particular voice is not unique, but rather is created and understood within a specific milieu, I suggest that the vocal teachers’ way of listening is not necessarily directly tied to personal racism, sexism, or any other prejudice. What underpins their assessments are the general beliefs that we can identify and know sound and that essential traits are audible in vocal timbre. Once the assumption that sound is knowable is in place, values and beliefs within a given society—around, say, race, ethnicity, gender, age, or class—then also become “knowable” through sound. Concepts such as “correctness,” “health,” and “authenticity” are vacated prior to their use within a cultural and social situation. Hence, because sound in general, and vocal timbre in particular, have no a priori meaning, whichever sentiments, positions, and values a given society deems important fill this vacuum. In other words, the assumption that we can know sound, and that the meaning we infer from it is stable (and indeed essential), allows for the *projection of beliefs about people onto the sound*. Thus denaturalization of concepts about race and timbre depend on a more general denaturalization of basic assumptions and attitudes about sound. In a nutshell, the naturalization of voice leads us to believe that what we decipher through voice exists in that voice, and the naturalization of race leads us to listen for race everywhere. But this destructive cycle cannot be broken on the level of judgments about race and voice, because it does not rest there. Instead we must look to the more general level of the epistemology of sound and the process of listening to voice.

Racialized Timbral Judgments Are Based on Assumptions That We Can Know Sound

Recall that the acousmatic question's assumption that the voice is essential and innate, and that asking and responding to the question can lead to an objective assessment of the vocalizer, is rooted in broader assumptions about sound. That is, a constellation of beliefs in a stable, knowable sound—what I call the figure of sound (FoS).²² Observations about voice, such as “healthy,” “authentic,” “Asian,” and “speaks white,” are really statements built upon the idea of an existing, knowable, sound with a given meaning. We are conditioned to *hear what we listen for* and to assume that what we hear is indisputable, and this conditioning acts much like planted evidence. As I see it, the dominant Western notions of music making and listening are founded on this paradigm of the figure of sound. Listening that is formed and that takes place within this paradigm is listening that knows *only* how to listen for and through difference from a fixed referent. Because the FoS paradigm assumes a fixed referent, it fosters a specific kind of listening wherein the primary goal is to identify difference from that referent. In other words, within this paradigm, making sound and listening are about degrees of fidelity to an imagined a priori sound and our ability to identify that fidelity. For example, we note observations such as the following response to the acousmatic question:

- This is “ma” (as opposed to “pa,” and “ma” is different from “pa”).
- This is B-flat (different from other pitches).
- This is a too-high or “out of tune” G-sharp (it is not faithful to the a priori G-sharp).

Yet the paradigm of the figure of sound does not end with the drive to know and identify a sound such as, say, G-sharp as the second scale degree of the key of F-sharp major. Nor does it end the drive to know and identify sounds on a unit level, such as syllables, words, or pitches. The paradigm of the figure of sound extends into timbre, and such timbral assessments are used to establish basic information around a sound source in response to the acousmatic question:

- This is a flute (different from other instruments, say, a clarinet).

Nonetheless, beyond basic distinctions such as flute versus clarinet, timbre is often bound up with the assessment of value and identity. For example, in the FoS paradigm, listening to human voices can lead to appraisals such as “This is the sound of a woman’s voice,” based on perceived similarities between a given sound and other, specifically female, human voices and their dissimilarities to

male and children's voices.²³ Likewise the observation that someone is "talking white" has at least two layers: the assumption that the speaker is not white and the assumption that the unexpected racialized vocal style is out of place, necessitating attention to the perceived clash of identity and timbre.²⁴ The statement comes out of a comparison with what the listener believes should have been the case. Because Obama is understood to be a black man, his voice is compared to the FoS of "black man." When his voice is not understood as consonant with that FoS, its infidelity to the FoS is called out. In other words, this observation exemplifies assumptions that race is quantifiable and knowable and that race is timbrally conveyed. This is but one example of how the figure of sound is also bound up with the assumed meaning of an identity, which is often derived from values and assumptions related to visual cues.

Listening within the FoS framework effectuates a circular logic. This logic is akin to a self-fulfilling prophecy that sets up a prediction that it directly or indirectly causes to be true. Per Robert Merton's description, the self-fulfilling prophecy is "in the beginning, a *false* definition of the situation evoking a new behavior which makes the original false conception come *true*." He continues, "This specious validity of the self-fulfilling prophecy perpetuates a reign of error. For the prophet will cite the actual course of events as proof that he was right from the very beginning."²⁵ In the case of racialized vocal timbre, the "false definition" is the belief in race as an essential trait, which causes us to fail to attend to the many ways in which timbre is learned and performed, including those we associate with race, ethnicity, or authenticity. We then listen for those phenomena that we believe to exist; we subsequently hear them, and because we hear them, we believe the perceived meaning to be verified.

For example, "black voice" is an observation born from an encultured notion of sound that expects fidelity to a referent and listens for difference. When voices are reduced to fixed sounds and undergo assessment, they cannot help but be heard within binaries or scale degrees of fidelity and difference. Moreover, due to the ways vocal timbre has historically been aligned with and metaphorized as interiority and truth, the stakes and ramifications of such assessment involve more than just sounds. What is measured is a person's degree of fidelity to and difference from a dominant category. I bring two observations to this reading of a timbral micropolitics.²⁶ First, the persistence of the metaphor of vocal timbre as the unmediated sound of selfhood and subjectivity means that a given society's beliefs lie at the core of its citizens' personhood. Second, culturally trained ears assume, and thus perceive, only formalized vocal practices as encultured. Moreover they tend to perceive enculturation only in certain components of the trained voice, while believing others to be "natural." The naturalization of the

untrained voice as an expression of “essential identity” and the naturalization of aspects of the trained voice according to racial categories are both functions of the micropolitics of timbre.

Independent of the “actual” or intended sound, what a listener ultimately hears depends to a large extent on his or her assumptions regarding the ontology of sound. For example, the belief that it is possible to know something firm about a sound and its source deeply affects the meaning the listener will form around that sound. Such belief arises from assumptions that sound can be known, is stable, and can be unequivocally recognized and unambiguously named. Furthermore such belief assumes a deep connection between sound and its apparent signification, an assumption regarding significance that is taken on through enculturation. Considering the statements made by the voice teachers in light of the broader listening framework helps us to see how, when listening through the FoS, we will listen for and, indeed, hear according to categories aligned with values within a given society (e.g., race). However, we do not need direct feedback (in the form of overt praise, recognition, misrecognition, or punishment) from a teacher or other authority figure in order to fall in line.

The FoS is based on the general practice of listening for similarity to and difference from it, and on the perceived implication of the listener in this process. Such listening springs from the assumed connection between a given sound’s source and its apparent meaning. Therefore, while some of the statements from my interviews may be taken as extremely provocative, I chose these because they are helpful in identifying the FoS’s framing of timbral phenomena as personal, innate, and essential rather than as stylistic performance choices. On the level of the FoS, while observations such as “This is a soprano,” “This is a woman’s voice,” “This person is happy,” or “This person is sad” are not driven by the same urgency, they’re based on the same type of listening. There are no technical differences between these seemingly innocuous observations and the types of observations made by the voice teachers.

Believing Race, Practicing Race, Creating Race

One of the many paradoxes related to timbre is that while vocal timbre is understood as essential, classical vocal pedagogy is built upon the very notion that it is possible to construct timbre. For the initiated ear, the classical vocal soundscape can be heard through such blocks of timbral construction. While for most people, classically trained voices might simply sound “classical” or “operatic,” tone quality is further and more specifically refined within subgroups,

directed by aesthetic and pedagogical concerns. A national school of singing implies both a preferred tone quality and the technique that produces that quality. Tone quality and technique function symbiotically on a national and regional scale and result in differing pedagogical schemes and a corresponding shaping of the voice according to a national tone ideal. (Perhaps the most commonly known national schools of singing are the English, French, German, and Italian, but there are also the Nordic and Slavic.)²⁷ We know that the sounds of these various schools are the result of aesthetic preference and of vocal technique designed to accommodate this preference.²⁸ We also know that they are *not* recognized as the unmediated expression of a people, contra nineteenth-century romantic nationalism.²⁹ A national school of singing simply refers to a region's preferred tonal quality (and the vocal technique that engenders it) and does not, of course, necessarily indicate the nationality of the singer. A Norwegian singer may be educated in a conservatory in Germany and thus develop a German tone. A teacher schooled in Italy might teach in Paris, passing on his or her Italian technique and tone ideal.

The phenomenon of national schools of singing is understood as contextually contingent and acquired. The processes involved in forming vocal timbre are formalized and recognized in extreme detail, even if the resulting timbres are understood within the signifying process only by those with knowledge of the cues. While most people have the ability to recognize operatic timbral characteristics in general, not everyone can distinguish between the various national schools of singing. However, for those initiated into operatic timbre, the national schools are quite distinct from one another.

While national schools of singing seem to most people to be only an esoteric detail, for those invested in these distinctions, the preferred national tone that one must perform in order to remain within the group is a serious matter. This is where we can sense the resonance between judgments about what somebody should sound like ("Obama should sound black") and what they might sound like ("Obama talks white"). The French Ministry of Culture, for example, has employed official inspectors to observe regional conservatories of music in order to evaluate their vocal pedagogy. Richard Miller reports that in the decades after World War II, some inspectors were especially adamant that their concept of proper onset be taught in French conservatories.³⁰ The preferred onset among these inspectors was an "attack," a very strong beginning that is created by a powerful inward thrust of the abdomen. This forces the vocal folds to deal with a high level of airflow, and in response the larynx resists the excess airflow by fixing the vocal folds in a single position. The result is a "held" sound that is slightly above pitch, with a pushed and sharp-sounding phonation. This sound

is now characteristic of the French onset and, because the attack sets up a tense position of the vocal folds, of the French line.³¹

It is also important to note that within the geographical area of a single national school there are many different spoken dialects. In some areas these dialects are so different that they are nearly separate languages. In countries such as Switzerland, students at a single conservatory might have four different mother tongues. However, phonation and, as a result, pronunciation differ in song and speech, and singers learn very carefully how to pronounce words when singing, even in their first language. Even singers with different mother tongues or dialects are unified under a single national school or a single teacher's tonal ideal. In summary, the presence of national schools of singing not only exemplifies the malleability of the human voice and the enormous impact that teachers' and institutions' tonal ideals and pedagogical practices have on the ultimate sound of a classical singer's voice, but also shows that we are fully aware of, and aesthetically depend on, the constructedness of vocal timbre in formally trained voices.

Translated to my tripartite diagnostics, this shows that we know voices are not innate, but cultural, and that voices are not intrinsic, but are shaped by a pedagogical collective. However, feeling this intuitively and knowing it intellectually can be challenging. While we know that voice is malleable and construct pedagogies based on this premise, and while timbre is simultaneously naturalized, essentialized, and used to evidence uniqueness, legible FoSs are embedded in flesh through the process of formal and informal pedagogy. Many of these vocal legibility projects are self-directed and so fluidly part of everyday life that their timbral formation processes are not accessible to us. Vocal timbral constraints are not straightforward; they are categories into which we ourselves shape our voices as well as categories that are pushed upon us, both of which are part of a broader sociocultural landscape. Therefore, while we have acquired knowledge in vocal pedagogy and vocal anatomy and could put that knowledge together with how social categories, including vocal timbral categories, are performed and thus embedded into flesh, it has nevertheless been challenging to decipher the constructedness of timbre and to debunk the myth of its authenticity.

Vocal work is about legibility to the listener.³² The listener should here be understood both as an external listener who comes from the same group as the vocalizer and listens for sameness and belonging to that group, and as a listener outside the vocalizer's group who can identify the vocalizer as different from himself or herself and as part of a given group. The listener is also the vocalizer, who goes through the same listening exercise of identifying with and against others. By bringing meaning to the voice, we not only affect it discursively

sively but also, because of voice's physical formation according to its practice, *form* the voice's material existence according to that meaning. Participation in a collective requires vocal legibility. Conforming to notions of legibility shapes the voice. That legibility is not a single monolithic quality that overrides other timbral effects; it is an aggregate of these effects.

But isn't it contradictory that, while a singer is understood by a vocal community to be simply emitting timbral evidence of his or her, say, "ethnicity," that vocal community also has the capacity to recognize that the same singer—at will and with practice—is able to perform across a wide timbral range? It is not contradictory. I understand these different listening outcomes as arising from a split in listening, with both branches emerging from the FoS listening framework. That is because, while within a Western listening context all sounds are heard through the FoS, timbre is believed to fall into two broad categories. Some aspects are understood as essential, while others are understood as acquired, performed, and somewhat open to interpretation. But those aspects considered essential are no different from those that are viewed as acquired. The validity of each rests on the naturalization of the FoS.

Through daily vocal practice, voice in its material presentation as flesh, ligaments, and tissue is encultured according to its constant comparison to the FoS. Because the voice is formed in conjunction with the body, it too broadcasts the social attitudes and values of the trained body. Every interaction, from state educational systems to the informal lessons imparted when a person receives positive or negative feedback on his or her voice, entrains (auto-)listening and vocal behavior. I believe we may take a crucial step toward untangling the "politics of frequency," to use Steven Goodman's apt term, by considering—through a systematic examination of the micropolitics of timbre—how vocalicity, a learned physical behavior, is trained and perceived.³³ Part of the micropolitics is to begin to render all vocal activity as learned physical behavior; through such activism, vocal timbre and any given categories of it that we know can be denaturalized.

Each and every reading carried out after a person's voice is reduced to the FoS contributes profoundly to that person's feeling of place in the world, to his or her attunement to the world and the self, and to his or her subsequent vocal exchanges. In short, informal vocal exchanges are powerful "voice lessons" that invite or discourage particular vocal practices. They are daily vocal performances that in turn are manifested in flesh, and sounded through it. Therefore, in a vocal encounter, the most productive question is not *Who is this?* Instead I propose that we ask questions such as these: On what naturalized assumptions about sound and voice are responses to the acousmatic question based? Through which unexamined assumptions are culturally created categories such

as race and ethnicity upheld? In short, by turning the acousmatic question to the listener, and ultimately to ourselves, we are encouraged to denaturalize vocal timbre by asking *Who am I?* Who hears this? And how was that sound learned, and through which sets of practices, constraints, desires, and structures of power did its so-called meaning become unquestioned?

The micropolitics of listening is not only applied by others to us, or by us to others. The assumption that it is possible to know sound leads to an overarching listening stance—the acousmatic question—through which the casual listener, the teacher, and listeners listening to themselves (i.e., auto-listening) seek fidelity. When it is assumed that it is possible to know sound, the primary tenet in listening is *identification*. The basic tenet of identification is comparison with an “original,” whether an actual sound or the idea of a sound in the mind’s ear. The success of such listening is then dependent on the listener’s ability to distinguish between similarity to or difference from the ideal. In other words, on a basic yet profound level, such listening entrains listening for sameness and difference.

As argued earlier, the sounds we ultimately produce and hear are based on enculturation; they are not merely essential qualities expressed through timbre in an unfiltered manner. To further clarify, it is because of assumptions around the FoS—for example, the assumption that sound can be identified—that we are unlikely to critically examine listening processes and the meanings they produce on a fundamental level. We thus further extend basic assessments regarding a given sound’s sameness and difference from the ideal. The given categories, which offer the basis for listening for sameness and difference, are, of course, culturally dependent.

The sound categories that can be further identified include distinct pitches, adult voices (versus, say, children’s voices), male versus female, “ethnic” versus “nonethnic,” “authentic” versus “inauthentic.” However, because the premise of *listening* is identification, we do not question the likelihood of the a priori existence of the identified categories. Thus, due to a basic belief in something as seemingly innocuous as the possibility of knowing sound, the question of whether it is possible to identify social categories via listening remains unexamined. And when listening within the ontology of the FoS, what is heard is then understood as evidencing essential and nonnegotiable traits.³⁴

What takes place, then, is a curious division between the way we listen to and assess categories understood by a given society as essential, and the way we listen to those aspects of human vocal timbre that the same society understands as performed. Thus our assumptions around the “innateness” communicated and evidenced by vocal timbre arise within a triangulation between the FoS,

a society's belief in certain essential categories, and listening that is entrained to detect sameness and difference. And detection of sameness and difference, in this case, is the detection of essential categories. The catch-22 is that once these essential categories are called out and grounded in FoS listening, there is no room to question them. The very premise is that they are indisputably *true*. It is such assumptions around the possibility of identifying meaning based on listening to and assessing sound in general, and voice in particular, that forms the basis of racial judgment. To combat racism through the mechanism of vocal timbre, we must examine this process in both seemingly unbiased (A-sharp) and biased ("ethnic") views of timbre.

Listening to Listening

Turn the ear to the listener and listen critically to listening, to the "listening ear" in action.³⁵ In carrying out an analysis that is conscious of the fact that any voice is part of the collective voice, and that listening contributes to shaping that voice, we must listen to how we listen. With the knowledge we gain from listening in this way, we can deconstruct the situations that, without such an analytical breakdown, will serve only to reinforce structures of power. By insisting on voice as event, as encultured even before birth, and as collectively projected, we can understand voice as the result of an ongoing pedagogical enterprise. This understanding, in turn, allows us to align vocal practice, listening to voice, and voice scholarship with the critical turn in pedagogy.

What I term formal and informal voice lessons are the sites where the micro-politics of timbre are played out through the fibers of the vocal folds and the habituation of the cricothyroid. These formal and informal voice lessons are the starting points of any meaning-making through voice. They are also where we can assess the process of meaning-making and present and manifest new ranges of meaning. The methodological framework of critical performance practice may already be recognized as informed by educational theorists such as Henry Giroux. As does Giroux, I want to resist the sentiment that a totalizing dominant culture merely imposes itself on students. I am inspired by the ways critical pedagogy understands resistance as enabling transformation and maintains that goals of hope and emancipation should be central to any curriculum.³⁶ My own resistance has taken the form of demonstrating that the practical experience of exploring more of the wide range of timbral potential inherent in each voice offers a perspective on any single timbre. In this way, it is through deconstructing basic assumptions about voice at the outset of listening, and through attending to the process within which listening takes place—through *listening*

to *how we listen*—that the listening framework becomes apparent and that we can grasp the politics of listening.

By shifting our analytical lens from the so-called sound to observing and understanding the process of listening, we may *listen against* the FoS by hearing voices as entrained. With this shift in listening, we demonstrate that every timbral quality or meaning may be interrogated. Even timbral qualities thought to be innate can be deconstructed as reflective of ways of listening that reproduce, or return, the listener's historical, cultural, social, political, moral, ethical, academic, or any other positionality.

Then, once again, we see that there is no a priori sound (FoS) or meaning. Within a given context, there is only the triad: the consortium of sound, meaning, and listener. Moreover, within this consortium the listener is the point of origin for meaning production.³⁷ The point is that by understanding the relationship between the FoS and the way such general assumptions about sound are acted out within a given society's values, we may begin to grasp some of the ways in which listening is always already political. By breaking down the consequences of FoS listening, we can understand its potential power. But, more important, by enumerating the consequences of the FoS, it is once again confirmed to be false.³⁸ That is, since sound is not always already static and knowable, the "identified and its meaning" are listener-derived. And while the "identified and its meaning" are listener-derived, the assessments produced are assumed to be so indisputable that they are used as evidence of everyday observations and their validity is extended to the American court system.

By performing through a range of timbres and witnessing other voices doing the same, and by critically listening to listening, we can create a counterlogic to any given response to the acousmatic question. *Who is this, vocalizing?* It is President Obama, who "sounds white." By focusing our attention on that listening ear, we hear a *listener* who holds the views that Obama is a black man. Ergo, he is not a white man. Black men sound X. White men sound Y. President Obama sounds Y. Ergo, President Obama sounds like a white man. The implication is naturally that it is only worthwhile to make this observation because President Obama does not sound black. By denaturalizing the assessment that gives rise to these responses to the acousmatic question, we can counter the meanings that arise. In this way we can gain critical distance and harness powerful analytical tools by keeping the critical performative aspect of voice front and center. Consider again the 1999 ruling mentioned in the introduction to this chapter. The Kentucky Supreme Court judge decreed that since no one would find it inappropriate for an officer to identify the voice of a woman, "we perceive no reason why a witness could not likewise identify a voice as being that of a

particular race or nationality, so long as the witness is personally familiar with the general characteristics, accents or speech patterns of the race or nationality in question.”³⁹ With this pronouncement the Kentucky Supreme Court ruled that a conviction was appropriately based solely on a police officer’s identification of a suspect whose voice he had heard in an audio transmission. The officer identified the suspect as a black male, testifying that during his thirteen years as a policeman he had had several conversations with black men and therefore was able to identify a black male voice. We here witness an assumption that because a voice is heard as black, it is emitted from a body that is unquestionably black.

Critical performance practice offers a path to examining assumptions such as “what you hear [i.e., blackness] is what he is.”⁴⁰ In other words, while the expression of many commonly held sentiments around race (or any other category important to a given society) is often curtailed, enactments through everyday listening are carried out, reported on, and deemed sound evidence, including in the Supreme Court. Even in cases where racialized politics are not expressed by institutions, like the Supreme Court, adapting the analytical framework of critical performance practice allows for methodologies to denaturalize racialized timbre and hopefully disrupt assumptions that otherwise would slip under the radar.

Adopting the mindset that listening is always already political has the potential to put intense pressure on the positionality of the listener. That is, listeners are not let off the hook, as they otherwise would be, existing under the radar when it comes to understanding timbral meaning. Keeping in mind that listening is always already political, listeners would examine any interpretation or judgment, acknowledge that it is the process of listening and interpreting that willed that particular meaning into being, and interrogate why it was projected onto a particular vocal timbre. In other words, through such a process, listeners would know that any meaning that arises is based on their own meaning derivation.

An examination of meaning would lie not only in the “objective” sound, but also in a meeting between the sound and the listening stance of the listener who derived those meanings. And, most important, such a critical inquiry would find new sites to deconstruct the process of signification.⁴¹ For example, how and why were aspects of vocal timbre, such as health and race, areas of signification that were understood as innate, and to which no interrogative or deconstructive pressure had been applied? For my part, uncovering the overall pedagogy of FoS listening, and understanding that it is involved in each and every act of listening, made the performativity of these areas very apparent. Therefore, while I continue to stress that each teacher (and every one of us

who has made and inadvertently continues to make racial judgments based on voice) is not necessarily doing so because of overt racism, I do not suggest that this erases the violence committed through microaggressions. On the contrary, each and every one of us is still responsible for how we contribute to our collective upholding of these racialized practices. As a community member, my work is, first, to understand how I listen and, second, to gain awareness about what I *produce* through habituated listening practices.

“Questions of form and politics are frequently subsumed in criticism by racial metaphor,” Jennifer Doyle writes. “The mere presence of race as an interpretative factor” often overshadows the “work’s difficulty and the complexity of its relationship to its context.”⁴² In other words, such artists are primarily understood through racial categories. The analogue is to the singers discussed earlier, who are felt to be singing with their “true voice” when they are heard as whichever ethnic or racial category a listener understands them to inhabit. However, as this chapter has shown, form and function create one another. The form produced through loaded racial metaphors (say, *African American opera singer*) limits what and who may be allowed to inhabit and express that form. In a paradigm of naturalized vocal timbre, instead of examining vocalists in all their idiosyncrasies and the communal projects in which they have participated and of which they are a product, listeners seek to find ways to explain the category or form, such as race, that they believe to be true. However, by attending to the performed aspect of vocal timbral production, individual listeners can denaturalize vocal timbres one at a time while helping to effect a broad cultural shift. It is that shift this book hopes to inspire.