



A System for Describing Vocal Timbre in Popular Song

Kate Heidemann

NOTE: The examples for the (text-only) PDF version of this item are available online at:

<http://www.mtosmt.org/issues/mto.16.22.1/mto.16.22.1.heidemann.php>

KEYWORDS: popular song, vocal performance, vocal timbre, timbre, phenomenology, embodied cognition, mimesis, voice physiology, Aretha Franklin

ABSTRACT: This article presents a system for describing perceptions of vocal timbre via reference to four different areas of sensation involved in the sympathetic mirroring of vocal production. This approach draws from phenomenological and ecological approaches to listening and analysis, and is supported by musicological and scientific literature on the acoustic properties and perception of timbre, and the physiology of vocal production. I demonstrate this descriptive method in brief analyses of Aretha Franklin's vocal performance in the openings of her recordings of "Respect" and "(You Make Me Feel Like) A Natural Woman."

Received September 2015

[1.1] This essay presents a perception-based system for describing vocal timbres.⁽¹⁾ Vocal timbres in popular song contribute significantly to both the immediate pleasures and conceptual meanings afforded by this music, but they resist description. Of course, pleasure and meaning also arise through entrainment to a groove, in hearing in the lyrics a story relevant to personal experience, by following the ebb and flow of rhythmic or intervallic dissonance between different layers of a track, or by noticing sensitive instrumental performances or expert mixing by studio engineers. My focus here is on vocal timbre because it is a highly salient parameter of popular song recordings, and is an affecting and intimate component of song performance. Vocal timbre telegraphs the interior state of a moving body, presenting the listener with blueprints for ways of being and feeling. In listening to vocal music, we may involuntarily mirror the actions we imagine the performer undertaking, thereby "catching" the affect of a performance.⁽²⁾ Through this physical rapport, singers' voices can become sources of comfort as well as pleasure, as the feeling of listening to vocal timbre leads to identification with, as well as enjoyment of, a recorded song.

[1.2] The problem of describing the vocal timbres of popular singers is a subset of the problem of describing timbre in general. When it comes to describing timbre in the context of an interpretation motivated by visceral experience, it is difficult to find satisfying words or representations, and misunderstandings abound. For example, I invite you to think about the timbre of Aretha Franklin's voice, as heard on her 1967 recording of "Respect" (**Example 1**).

By listening to the timbre of her voice, which is conceptually distinct but thoroughly intertwined with the other elements of her vocal performance, I have a physical intuition about what it is like to sing like Franklin.⁽³⁾ But how do I describe the

timbre of that voice, in order to explain what it means to me? How might I go about comparing the affective impact of Franklin's vocal timbre in this recording to Otis Redding's vocal timbre in *his* performance of this song? How can we compare our subjective experiences and resulting interpretations of these vocal timbres with one another? Vocal timbre is complex in its acoustic makeup and modes of production. The relationship between its acoustic and physical features, and between these features and listener perception, is still opaque. This is perhaps why analysts have traditionally avoided in-depth study of timbre, preferring to study elements of musical sound that are easier to measure, like pitch structures and rhythm. The increasing variety of timbre to be heard in all types of music means that this avoidance is no longer tenable (if it ever was). In popular music broadly considered, timbre is one of the most active parameters of experimentation, and a primary means of differentiation among artists and styles. Timbre, and vocal timbre especially, therefore poses a delightful yet frustrating analytic challenge—it is a facet of musical experience that cannot be denied any easier than it can be explained.

[1.3] The system for vocal timbre description I offer here draws on phenomenological and ecological understandings of listening in order to address the creation of meaning that arises as part of individual human perception of vocal timbre. It is based on the premise that one's feeling of what it would be like to vocalize in a particular way is a fundamental yet frequently overlooked part of how we conceptualize vocal performance. It is possible to describe this feeling in an organized way by referring to four types of bodily position, activity, or tension involved in vocal production: the movement of vocal folds, the position of the vocal tract, the location of sympathetic vibration, and breath support. These components are crucial in the production of vocal sound, and are part of what motivates our emotional and conceptual responses. They are what we call upon and feel as we consciously or unconsciously simulate a sound to make sense of it. This organization enables us to ground our descriptions of vocal timbre in our felt response to sound, and provides a shared framework for communication. Following an overview of the acoustics, perception, and production of vocal timbre, and an explanation of my descriptive system based on these features, I will demonstrate the application of this system in an analysis of Aretha Franklin's vocal timbres in her landmark recordings "Respect" (1967) and "(You Make Me Feel Like) A Natural Woman" (1968).

Understanding Vocal Timbre Through the Body

[2.1] A satisfying method for description and analysis of pop vocal timbre should meet the following criteria: it should afford detailed descriptions of the sound of vocal timbre, support intersubjective comparison of listening experiences while minimizing miscommunication and confusion, and recognize and reflect the visceral nature of music listening. There is good new research available that is sympathetic to these goals, but more detailed work remains to be done.

[2.2] In his book *Song Means: Analyzing and Interpreting Recorded Popular Song* (2012, 101–103), Allan Moore offers a four-part methodology for analyzing the sound of a singer's voice, in which the listener-analyst accounts for four "positional aspects" of a singer's voice: register, the cavity of the body where the singer's voice appears to resonate, the singer's heard attitude to rhythm, and the singer's heard attitude to pitch. Serge Lacasse's objection to this model, as cited by Moore, is that it ignores finer variations in timbre. Lacasse's (2010) approach to analyzing popular singing adapts Fernando Poyatos' (2002) model of paralinguistics, which distinguishes between and describes both the steady and variable qualities of voice ("primary qualities" and "qualifiers," respectively) as well as special vocal effects that can be used with or without verbal sounds ("differentiators," such as laughter or yawning) and completely non-phonemic vocal sounds ("alternants"). Both of these approaches incorporate the concept of vocal timbre into a broader consideration of vocal performance in more-or-less systematic ways, but take slightly different sets of vocal descriptors as given. I am interested, however, in a system that enables us to investigate why we even choose a certain descriptor in the first place (even if, after determining that, we settle on a term to use again and again).

[2.3] Nina Eidsheim's research on the flexibility of voice and the relationship between voice and race (2008, 2012), S. Alexander Reed's 2005 dissertation on the semiotics of vocal timbre (with case studies of vocal performances by Laurie Anderson and Louis Armstrong), Zachary Wallmark's (2014) work on the affect and meaning of timbre (noisy instrumental timbre in particular), and David Blake's 2012 work on the meaning of timbre in indie music all focus to varying degrees on the embodied nature of timbre perception and understanding. These scholars use methodologies aligned with or directly stemming from Merleau-Ponty's phenomenology (2012), the feminist musicology of Suzanne Cusick (1994), or Arnie Cox's (2011) work on embodied music cognition. The methodological underpinnings of the systematic approach to description and analysis of vocal timbre I present here draw from those same influences and are also indebted to the works of the aforementioned scholars.

[2.4] One of the first tasks of any study of timbre should be to determine the element of timbre most pertinent to the project

at hand. It is possible to describe timbre in terms of its perceptual, acoustic, and physiological characteristics: vocal timbre is a perceptual impression of an acoustic signal generated by a vocal production system (although the method of vocal amplification and recording, or the addition of distortion via external means, may undermine the sense of a timbre originating from a single, organic production system). I am most interested in the perception of vocal timbre. The perceptual realm can be difficult to study, given the variety among different perceivers and differences in ability when it comes to reporting one's perceptions, but it is also the most intriguing realm for those of us interested in achieving a holistic understanding of how sounds elicit different states of arousal and emotion, and come to acquire meaning. The acoustic and physiological elements of vocal timbre directly impact perception, and an understanding of both increases the precision of one's perceptual descriptions.

[2.5] The acoustic components of timbre typically cannot be characterized according to a single unidimensional scale of frequency, amplitude, or duration, although listener perceptions of timbre are impacted by all these elements of sound.⁽⁴⁾ Research that uses multidimensional scaling (MDS) algorithms to discover the acoustic correlates of listeners' perceptions of timbre suggests that the most perceptually salient acoustic components of timbre are a combination of spectral center of gravity (the dispersal and strength of frequencies above the fundamental frequency) and attack time (the nature of the onset of a sound). Many listeners also seem to perceive spectrum fine structure—e.g., the reduction in dB of even harmonics in a clarinet's timbre—as a distinguishing acoustic parameter. The change in dispersal and strength of harmonics over time may contribute to listener perceptions, but this has been found to be less significant than the aforementioned three elements (McAdams 1999; Caclin et al. 2005). When studies that use MDS, even those that use the same MDS algorithms, are compared, a common finding is that different listeners have different strategies and preferences for acoustic parameters when it comes to differentiating between timbres.⁽⁵⁾ Given this difference among listeners, perhaps the best approach for analysts is to be aware of the acoustic parameters in play when we talk about timbre.

[2.6] With this understanding of timbre's perceptually significant acoustic parameters, it is possible to identify elements of vocal sound that are very closely linked to vocal timbre, but that are not timbral in the strictest sense. Perceptions related to language, such as our recognition of phonemes, manner of enunciation, or type of accent are not timbral perceptions strictly defined, although they are integral to an analysis of vocal performance.⁽⁶⁾ Inflections of pitch (our perceptual correlate to fundamental frequency) that are difficult to represent using musical notation—ornaments such as vibrato or glissandi—are not timbral either, but may occur in tandem with timbral manipulations by the singer. The overlap of the acoustic components of timbre with these other perceived elements of vocal performance means that it will rarely prove fruitful to refer *only* to timbre when discussing vocal sound in the context of an interpretative project.

[2.7] My experience in everyday listening activities is that the arrangement, strength, and flux of overtones that make up the acoustic components of timbre become unified in the process of perception, resulting in a general sense of a quality of a sound.⁽⁷⁾ This notion of “quality” stems directly from a sense of the characteristics of a particular sound source, such as the type of material of which it is made, and an approximation of the manner in which that material is vibrating to produce sound or is interacting with other materials around it (Clarke 2005). This understanding is impacted by a listener's previous experiences with a variety of similar sounds and whatever or whoever produced them (experiences shaped by individual and group identity), as well as the listening task at hand. In the case of synthesized or manipulated sounds, or unfamiliar sounds, we might refer to the closest sound source in our memory, or use those memories to construct an imaginary sound source.

[2.8] We often explain our perceptions by simply naming the sound source as our label for timbre, or through a combination of metaphor, onomatopoeia, vocal mimicry, and gesture.⁽⁸⁾ We might use cross-modal metaphors like “bright” or “harsh” to describe these sensory impressions, with varying degrees of agreement between listeners as to the specific meaning of these terms.⁽⁹⁾ One solution to this problem is to enlist techniques and tools that replace or reduce reliance on descriptive terminology, for example in experiments where listeners compare or match timbres through sound synthesis without the use of linguistic description (Kreiman and Sidtis 2011, 22). For music analysts, the use of spectrograms allows a detailed representation of an acoustic signal to stand in for or concretize certain aspects of listener description.⁽¹⁰⁾ For those of us interested in sharing interpretations of vocal performance, however—what a performance *means* to its listeners—the verbal description of sound is necessarily at least a part of what we want to do, regardless of the difficulty posed therein.

[2.9] To discuss the meaning of timbre, one must grapple with the question of how to describe one's perception and interpretation of that phenomenon without an explicit, shared framework of understanding. Developing such a framework benefits from attending to the physical sensations affiliated with music perception and conceptualization, and refining those observations with reference to scientific findings on the relationship between observed modes of voice production, acoustic

components of vocal timbre, and common perceptual descriptors. When it comes to the human voice, most of us who are able to hear are extraordinarily sensitive to how the timbre of a voice reflects the “motional characteristics of its source” (Clarke 2005, 74). Most of us also have a functioning voice production system and have had ample, varied experiences making and perceiving vocal sounds. These resources readily enable feeling and participation when we listen to vocal performance: we have some basic understanding of what singers are doing with their voices, and what we can do to join in. We get a sense of the movements undertaken in the production of a timbre through conscious and unconscious imitation, or mimetic engagement (Cox 2011). Some mimesis is unconscious, and happening at a pre-cognitive level. This simple, automatic type of imitation is involved when we unconsciously copy observed actions that are already part of our repertoire. Much of our mimetic engagement is, however, conscious or at least consciously accessible. This type of imitation may be characterized as complex imitation or observational learning.⁽¹¹⁾

[2.10] We can call on our capacity for mimesis to help clarify the physical basis of our vocal timbre descriptions. We can turn our attention to how our perception of a singer’s vocal timbre invites us to recall or imagine the motions needed to produce a similar timbre, or to actually try out the movements as we attempt to match timbre while singing along. This activity constitutes what Arnie Cox, in his research on musical affect and mimesis, calls “intra-modal or direct-matching mimetic motor imagery” (Cox 2011, 9). Further, as we compare these embodied responses to our broader realms of experience, we can identify how they become incorporated into our existing mental/cognitive database of musical and extramusical connotations, in which a sound becomes conceptually connected to other sounds, imagery, cultural histories, personal memories, and so on.⁽¹²⁾ The physical sensations of imitating—and, by extension, imagining—vocal timbre, and the conceptually related feelings afforded by those sensations, can help us clarify what we hope to communicate when using certain types of description. If we can agree that part of how we make sense of and ascribe meaning to another person’s vocal performance is by reference to our own vocal experience (even as we compare that particular performance to the sounds of other voices), we can make our interpretations of timbre more detailed and intelligible to others by focusing on the specifics of this embodied engagement with sound.

The Four-Part Embodied Comprehension of Vocal Timbre

[3.1] A clearer understanding of vocal physiology allows listeners to precisely indicate which parts of the body they experience as involved in the mimetic comprehension of a particular vocal timbre. In the process of vocalizing, air exits the lungs, propelled by internal pressure created by the constriction of muscles in the abdomen and ribcage (see **Figure 1a**, from Kreiman and Sidtis 2011, 26). The oscillation of the vocal folds (located in the larynx, colloquially termed the voice box) creates changes in pressure of this escaping air, which we hear as sound. This process is referred to as phonation. The supraglottal vocal tract (**Figure 1b**, from Kreiman and Sidtis 2011, 51.) shapes the sound produced by phonation, and this sound can be dramatically varied through the movement of the lips, tongue, jaw, soft palate (or velum), and other muscles surrounding the pharyngeal cavity (the very back of the throat just beyond where the oral and nasal cavities meet).⁽¹³⁾

[3.2] **Table 1** lists the four elements of vocal production that I propose participate in the embodied perception of vocal timbre, some related and commonly used timbral classifications drawn from systems of vocal instruction and speech research, and an incomplete list of related or overlapping components of vocal performance that one might consider in an analysis. There are three primary means by which singers can alter their vocal timbre: by varying the delivery of air from the lungs, changing the stiffness and position of the vocal folds, or adjusting the shape and position of the vocal tract. Additionally, the sensation of sympathetic vibrations in the body is strongly related to the physicality of vocal production.

[3.3] Thinking about vocal timbre in terms of four perceived elements of vocal production provides a group of organizing questions to consider in analyzing vocal timbre: 1) In what manner do the vocal folds seem to be vibrating? 2) What is the apparent positioning of the mouth and throat? 3) Where do sympathetic vibrations occur in the body? and 4) What is the apparent degree of breath support and muscular anchoring required? Many of the common terms we use to characterize vocal timbre (“belt,” for example) encompass movements and degrees of activation in multiple areas of the vocal production system, and can be related to multiple areas of the four-part organizational structure I propose. The advantages of considering the different parts of voice production even though vocal timbre is typically a unified perception afforded by multiple movements in the singer’s body, are threefold. It gives us a place to start when we encounter a vocal timbre that we don’t already know how to categorize, when we want to investigate and problematize a common categorization, or when we are not certain that the descriptive terminology we would like to use will be clear to others.

[3.4] *Question 1: How do the vocal folds seem to be vibrating?*

Thinking about the manner in which the vocal folds are vibrating draws our attention to the sound source. The vocal folds

(small folds of tissue with variously stretchy and squishy layers) are located within the larynx, and modulate the outgoing air. The intrinsic muscles of the larynx (which connect the different cartilages and control their positions relative to one another) create glottal opening for breathing, bring the vocal folds together and stiffen them for phonation, or completely close the glottis to prevent inhalation of material or to fix the ribcage to provide a rigid frame for lifting or pushing. When listening to a singer, through a combination of muscle memory and mimesis, or complex imitation, I can get a sense of the likely degree of tension in their vocal folds, a rough estimation of glottal opening, the physical condition of the folds, and whether those folds are vibrating efficiently. Although I have taken care to verify my own observations about vocal sound against research on the relationship between the vocal apparatus and vocal sound, I must emphasize that it is not necessary to have detailed knowledge of vocal production in order to have a meaningful embodied response to a performer's vocal timbre—it simply helps with the clear description of that response. Changes along these parameters often result in reliably perceptible changes in vocal timbre; speech researchers and voice instructors often characterize this particular set of timbre descriptors as phonation types. A common, cross-disciplinary set of terms to label these types has yet to be determined. I will use those terms that appear with reasonable regularity in contemporary scientific voice research and pedagogical approaches to popular vocal styles.⁽¹⁴⁾

[3.5] Regular (periodic) vibration of vocal folds, with complete and rapid closure of the glottis when the folds meet, is often regarded by vocalists as simply efficient, healthy vibration or referred to as “modal” by clinicians and researchers concerned with pathologies of voice. Most of us who can speak or sing are capable of varying degrees of this type of vocal-fold vibration—otherwise our voice would not sound, or the use of our voice would be painful. Beyond distinctions of undamaged and functional versus injured or non-functioning, however, there is no truly regular or standard human voice, and a wide variety of singing styles are valued in popular music. Describing and evaluating the impact of a vocal timbre is always a matter of comparison, and the listener-analyst's voice is always involved, even when not formally included in a study. I therefore suggest that listeners describing a vocalist's timbre in a music-analytic setting at least consider the sound of their own voices and the habitual vocal tract setting that produces it.

[3.6] Some phonation types are characterized by the sound of air leaking through an incompletely closed glottis. When vocal folds are low in tension, this results in a sound that is frequently described as breathy, and when the folds are more tense, a breathy voice transforms into what might be characterized as a hissing or grainy vocal timbre. A related vocal timbre descriptor, whisper, usually refers to similar vocal sounds produced without the vibration of the vocal folds. My own voice is normally somewhat breathy, and perhaps because of this habitual setting, I also relate this type of vocal timbre to a general feeling of muscular relaxation throughout my body. Breathily phonation is used for varied expressive ends by Astrud Gilberto in “The Girl From Ipanema” (**Example 2**) and by Michael Hutchence (of INXS) in “Need You Tonight” (**Example 3**). The range of expressive possibility presented by breathy vocal timbre—e.g. relaxation, intimacy, or sensuality—illustrates the distance between the basic description of a vocal timbre and the explanation of its impact in the context of a popular song. Clear identification and description using a shared framework is only the first step toward a discussion of the meaning of a particular vocal timbre in a given musical context.

[3.7] Most artists singing in a popular style have some amount of breathiness in their voice that at somewhat louder volumes becomes a subtle hissing sound (sometimes characterized as an increase in “noise”). Since no part of the vocal apparatus operates completely alone, breathiness and extensions of this voice quality can be impacted by the shape of the vocal tract, as well as how the vocal folds are vibrating. For example, I can increase the breathiness of my voice by relaxing the muscles in my neck so that my pharynx collapses a little bit, and perturbs the flow of air after it passes through my vocal folds. The result is still an impression of muscular relaxation, and supports the physical and conceptual relationship between many popular singing styles and everyday speech.

[3.8] Through mimetic engagement, I can get an embodied sense of whether a singer's vocal folds are vibrating in a rough, aperiodic manner, resulting in a variety of vocal timbres that are often described as harsh or hoarse. Otis Redding in “These Arms of Mine” (**Example 4**) seems to tense his vocal folds and then “overblow” them to achieve the harsh timbres that accent his vocal line. This type of vocal effect can also happen with less tense vocal folds and a high rate of air flow, as in screaming: hear, for example Isaac Brock of Modest Mouse in “Bury Me With It” (**Example 5**). Voice instructors usually caution against this effect because it can damage the vocal folds. It is possible that this is how Redding obtained the hoarse quality of his voice heard at softer volumes (**Example 6**). In “These Arms of Mine,” his singing has an exaggerated version of that noisy hiss I hear as an extension of breathy voice, which I associate with the way my own voice sounds whenever my vocal folds are irritated and I'm having difficulty getting phonation started—after a cold, or a long night of shouting over loud music, for example. In general, my embodied sense of these harsh or hoarse timbres is that they either require high

energy and muscular tension in the area of the vocal folds or throughout the body, or are indicative of the repetitive stress of singing in this manner.

[3.9] Another type of irregular vocal fold vibration often encountered in pop singing is creaky voice, otherwise known as vocal fry, or sometimes laryngealization, which is created when the vocal folds open and close abruptly. When I create this sound, my vocal folds are slightly tensed and gently held closed together, and it is easiest to start the sound at a low pitch, i.e. with a low position of my larynx. Donna Summer uses this vocal timbre as part of a titillating simulation of sexual passion in “Love to Love You Baby” (**Example 7**), while Britney Spears has so thoroughly incorporated this timbre into her singing style as to make it a vocal trademark (as in “Oops! . . . I Did It Again,” **Example 8**).

[3.10] A mode of vibration like creak is used in a variety of popular styles to approximate screaming while avoiding excessive strain. Singers might practice carrying the creak sound higher into their vocal tract by constricting the entire pharynx until the “creaky” vibrations can be felt somewhere closer to the vicinity of the soft palate, thereby protecting their vocal folds from excessive wear.⁽¹⁵⁾ A related vocal effect that relies on vibrations higher in the vocal tract is commonly referred to as “growl.” This effect is created through vibrations in the supraglottal part of the larynx (sometimes called the epilarynx, which includes the ventricular or “false” folds, epiglottis, and the folds of the aryepiglottic sphincter—see **Figure 2** for a relevant vocal tract model derived from vocalist and voice researcher Jo Estill’s pedagogical method). Louis Armstrong (**Example 9**, “What a Wonderful World”), Tom Waits, and many vocalists in different heavy metal subgenres have made notable use of growl.⁽¹⁶⁾ The exact source of all these vocal sounds may be difficult for the listener to pinpoint, although I do believe it is possible to sense whether these different types of aperiodic vibrations are occurring higher or lower in the vocal tract. Even if a singer is using special techniques to create vocal distortion, the embodied sense of roughness and a host of associations related to these different modes of distortion are available to the listener. In my own listening experience, however, there is no substitute for the expressive heft of a scream, or vocal distortion produced in an “unhealthy” way, with intense vibrations near the vocal folds. Vocalists who do this seem to be making a terrible sacrifice in the service of musical expression.

[3.11] The vocal break, or sudden switch from one type of phonation to another (also accompanied by a jump in pitch), is often used as a form of virtuosic vocal ornamentation. This is especially cultivated in country-style singing, for example in the “cry break” as described by ethnomusicologist Aaron Fox (Fox 2004, 276), or Jimmie Rodgers’ yodeling in “Blue Yodel No. 1 (T for Texas),” but it can also be heard elsewhere (e.g., Kurt Cobain’s singing in the *MTV Unplugged in New York* version of “Come as You Are”).

[3.12] An analyst may want to consider phonation onset as well. I have already discussed some relevant descriptors: a singer’s onset may be breathy or creaky, for example. When no concomitant sounds are heard at the beginning of a vocal sound, the onset may be described as “smooth” or “coordinated.” In contrast, when the beginning of a vocalization is accompanied by the sound of air suddenly and forcefully escaping through a previously tightly closed glottis, it is referred to as a glottal onset. Although voice instructors often focus on cultivating smooth onsets, each type of onset can be used to a variety of aesthetic ends in popular singing.⁽¹⁷⁾

[3.13] *Question 2: What is the apparent position of the vocal tract?*

Questions 2 and 3 of this descriptive system, about the shape and positioning of the vocal tract above the larynx, and areas of sympathetic vibration, are connected in that they both relate to the embodied sense of what a vocal resonator is doing, but they attend to two different ways we might intuit the position of the vocal tract and thereby assign meaning to vocal timbre. The supralaryngeal vocal tract (the mouth, pharynx, and nasal passages) shapes the buzzing of the vocal folds into the sound we recognize as a human voice. From the perspective of acoustics, this process is called resonance: the vocal tract serves as the resonator for the vocal folds, vibrating well at certain frequencies and less well at other frequencies, thereby also acting as a filter. Those frequencies most effectively “boosted” by the vocal tract are that system’s resonant frequencies, or the formant frequencies.

[3.14] The movement of the mouth (lips, tongue, and jaw) drastically shapes the frequencies of the first three formants in the creation of different vowel sounds, while our perception of general vocal timbre seems to be related to the impact on higher formants created by the length and shape of the entire vocal tract (Sundberg 1987, 101–102). The length of a singer’s vocal tract is determined to some extent by biological factors like sex, age, and inherited genetic traits, but it can be changed by protruding or retracting the lips, or by raising or lowering the larynx (Sundberg 1987, 20–23). Movements of other parts of the vocal tract control its shape. One may enlarge the pharynx by tilting the larynx (or more precisely, the thyroid cartilage) forward and down and by retracting the walls of the throat, in a combination of motions similar to yawning. It is also possible to constrict the pharynx by moving the root of the tongue or the epiglottis back and engaging muscles in motions

similar to those used when swallowing. The velum, or soft palate, may either be raised or lowered. A raised velum prevents exhaled air from escaping through the nasal passages and can result in a speaker's sounding congested, although the effect of a raised or lowered velum during singing has a somewhat different effect.

[3.15] There are a few general changes in the back of the vocal tract that listeners are reliably able to perceive and for which researchers have identified the underlying physiology of production. More research is needed—and is underway—to further refine the perceptual categories of vocal timbre as related to vocal tract position, and to pinpoint the physiological basis and acoustic correlates of these different timbre types.⁽¹⁸⁾ For the time being, I will cover common descriptions of vocal timbre that relate to the length and shape of the vocal tract, beginning with those configurations that are roughly smallest in area and ending with configurations largest in area. Although a variety of terms are found in pedagogical and scientific literature on this topic, I have adopted the frequently used terms devised by former professional vocalist and voice researcher Jo Estill.

[3.16] When I raise my larynx, constrict the back of my throat, and bunch up my tongue, which simultaneously results in a closed or nearly closed velum, I am able to make my entire supralaryngeal tract smaller to create a sound that I characterize as a “baby” voice. Estill refers to this timbre as “oral twang.” Estill’s usage of the onomatopoeic term “twang” is helpful here, because saying this word (especially the “-ang” part) produces this type of vocal tract configuration. She combines this with the term “oral” (somewhat confusingly, since listeners may often interpret this type of vocal production as “nasal” due to the location of sympathetic vibrations) to indicate the high degree of vocal tract constriction. This vocal timbre is a major feature of actor Kristen Chenoweth’s singing and speaking voice, and can also be heard in Dolly Parton’s voice, in slightly exaggerated form as she sings “I cannot compete with you” in “Jolene,” **Example 10**).

[3.17] A similar but slightly less severe constriction of the pharynx is possible by drawing the epiglottis toward the back of the throat and raising the larynx. This is the primary motion of the vocal tract that produces the timbre Estill refers to as “nasal twang,” and other researchers refer to as simply “twang” or sometimes “pharyngeal voice.”⁽¹⁹⁾ Estill suggests imitating a “wicked witch” laugh (“Yeah-heh-heh-heh-heh”) or horse’s neigh to find this sound and movement (McDonald Klimek, Obert, and Steinhauer 2005b, 41). Voice researcher Ingo Titze (2001, 528) refers to the use of twang in producing what he calls a “resonant” voice, since the efficient conversion of energy produced by this vocal tract setting leads to vibrations felt all over the head and neck (which represents an important component of embodied understanding that I will return to later).⁽²⁰⁾ This type of vocal timbre can be heard in a variety of styles, because it is useful in projecting the voice above other loud noises; it is therefore a common component of belt-style singing. It is used in musical theater, gospel, rock (usually coupled with some vocal distortion), and country, and can be regularly heard in the singing of Chaka Khan, Aretha Franklin, and Robert Plant.

[3.18] On the other hand, some singers use a less constricted pharynx for loud, energetic singing, resulting in a more speech-like vocal tract configuration similar to one used in yelling. Many singers, especially while developing a signature vocal style or prior to professional training, use this type of vocal tract setting, resulting in a timbre that sounds very similar to that heard in energetic speech—e.g., Alex Turner of the Arctic Monkeys in “I Bet You Look Good on the Dancefloor” (**Example 11**).

[3.19] By pulling one’s larynx lower, tilting the thyroid cartilage forward, expanding the pharynx, and drawing the velum upwards (but not completely closing off the nasal passages), it is possible to create a vocal sound that listeners often perceive as much darker, and sometimes quieter, than those previously described. Estill refers to this vocal timbre as “sob” since she relates the vocal setting that produces it to “silent, suppressed sobbing” (McDonald Klimek, Obert, and Steinhauer 2005b, 31). In teaching this sound, vocal instructors might encourage a student to breathe deeply through the nose to lower the larynx and expand the sidewalls of the pharynx, or yawn to aid in tilting the thyroid cartilage, opening the pharynx, and raising the velum. This sound can often be heard in operatic singing, and is an important component of the “crooning” vocal style. It is a regular feature of Bing Crosby’s singing (e.g. in “Christmas in Killarney,” **Example 12**), and characterizes Cher’s singing voice as well. It can require extra energy to maintain an expanded pharynx while singing, but this vocal tract position is typically very easy on the vocal folds—this can make listening to and mimicking this vocal timbre feel rather soothing.

[3.20] A few more general movements of the mouth may impact the perception of vocal timbre. In addition to a narrow pharynx and high larynx, belters, especially when reaching for high notes, often adopt a “megaphone” mouth configuration, with the mouth wide open and the corners of the lips retracted (Titze, Worley, and Story 2011). This is how Idina Menzel hits the climactic high pitch at the end of “Let it Go” (**Example 13**). Alternatively, a singer can adopt a close-mouthed style of delivery for a wide variety of expressive ends—consider Kurt Cobain’s terse vocal delivery in the verse of “Smells Like Teen Spirit” (**Example 14**).

[3.21] *Question 3: Where do sympathetic vibrations occur in the body?*

The location of sympathetic vibration during singing is directly connected to the shaping of the vocal sound by the vocal tract, and certain types of sympathetic vibration often coincide with specific vocal tract shapes.⁽²¹⁾ Voices that produce sympathetic vibrations far forward in the nose or in the nasal passages are often referred to as nasal—and as I mentioned previously, frequently result from a narrowed pharynx and high larynx, or “twang” vocal tract setting. Willie Nelson’s vocal timbre is a classic example of nasal voice, from which I get an impression of strong vibrations focused in the nose (e.g. “Can I Sleep in Your Arms,” **Example 15**). There is not necessarily any connection between “nasal” sounds and the position of the velum, which is somewhat surprising because the term “nasal” is sometimes used to describe a speaking voice produced with a completely raised, closed velum (preventing any air from escaping through the nasal passages, causing speakers to sound like they are suffering from nasal congestion). However, in research on the acoustic and physiological differences between operatic and belt voice types, no consistent position of the opening between oral and nasal cavities could be found among voices that reliably sounded nasal to expert listeners (Björkner 2006).

[3.22] Intense sympathetic vibrations felt in the front of the face and throughout the head can be an important physical marker for singers, because this is often a good indicator that they are producing a “resonant” voice that will cut through or carry over other loud sounds; this is also sometimes referred to as “singing in the mask.” The “resonant” sympathetic vibration is closely related to the “twang” vocal tract setup, although it typically involves less constriction of the pharynx and mouth than that used to produce a nasal timbre. Vocalists admired for their powerful, projective voices may experience (and provoke in listeners) these types of vibrations during singing. Examples are Bruce Dickinson of Iron Maiden in “Run to the Hills” (**Example 16**) and Celine Dion.

[3.23] Different sympathetic vibration locations are directly related to vocal register, and result from coordinated changes in phonation and vocal tract shape. The concept of register can be hotly debated, but may be defined for my purposes here as “phonation frequency range[s] in which all tones are perceived and felt kinesthetically as being produced in a similar way and possess similar vocal timbre” (Sundberg 1987, 51). There is a strong relationship in this aspect of vocal sound between fundamental frequency and timbre—higher- and lower-pitched voices are meaningful not just in the realm of pitch perception, but also in how they communicate information about vibrating bodies. Vocal registers are delineated according to perceived changes in phonation: traditionally, the chest register entails vibration of the entire vocal fold mass, while the head or falsetto registers involve the vibration of only the thinned outer layers of the vocal folds. Changes in nomenclature, the overlap of register ranges, and the substantial variation of register boundaries among different singers make clear definition of registers difficult. Mindful of this challenge, I will begin with the top and front of the head and work my way down the body, providing examples of embodied responses to different perceived areas of sympathetic vibration and related terms commonly used for registral timbres in popular singing.

[3.24] Joni Mitchell’s voice quality as she sings high notes (e.g. in “I Had a King,” **Example 17**) is an example of head voice, which I associate with vibrations in my frontal sinuses. A similar effect in male singing, which can also produce sympathetic vibrations in the head and face, is often referred to as falsetto—here demonstrated by Al Green in “Let’s Stay Together” (**Example 18**). Unfortunately, voice scientists, classical vocalists, and popular vocalists use the term “falsetto” somewhat differently. According to voice researchers Kreiman and Sidtis, falsetto “occurs at the upper frequency limits of a speaker’s vocal range, but also reflects a different vibratory mode for the vocal folds. In falsetto, the vocal folds vibrate and come into contact only at the free borders, while the rest of the fold remains relatively still” (Kreiman and Sidtis 2011, 62). Traditionally, falsetto has referred to a style of singing employed only by male singers, that involves breathy phonation, high pitch, and vibrations in the thinned outer layers of the vocal folds (Nair 2003). In recent pop vocal pedagogy and practice, however, the term “falsetto” has been used to refer to this type of vocal production used by both male and female singers (Chandler 2014, 39). Given the shifting usage of the terms “head” and “falsetto,” it is best to clarify the term with the added description of phonation type. In the case of Al Green, I perceive his falsetto voice as including breathy phonation, which in turn affords less intense feelings of sympathetic vibrations in my head and a greater sense of tension near the larynx.

[3.25] A few vocalists are able to sing in an extremely high, “whistle” register. Mariah Carey is well known for her whistle register flourishes, and I find that her singing in this range presents an incredibly intense feeling of focused vibrations in the top of the head and sinuses (e.g. Mariah Carey’s “Bliss,” **Example 19**). This register is still not completely understood from a physiological perspective, but, like falsetto, it seems to involve the vibration of only part of the vocal folds. A dampening of the vocal folds via supraglottal constriction might help to create this effect (Titze 2008).

[3.26] The term “chest voice” is used to describe singers’ lower register, where they probably use a similar larynx height and

phonation type as in speech, or the “sob” type of vocal tract configuration, with the entire mass of the vocal folds vibrating. It is especially relevant in describing the luxuriously low voices of Gladys Knight (in “Help Me Make It Through the Night,” **Example 20**), and Bing Crosby, which produce sympathetic vibrations down the sternum and across the chest. The vibrations associated with this type of vocal production may also be related to the concept of mixed voice, which is often of special interest to female vocalists trying to increase their “belt” range. Vocal coach Kim Chandler describes mixed voice as a “continuation of a ‘chest’-like tone above the main passagio but with more ‘Twang’” (Chandler 2014, 39).

[3.27] *Question 4: What is the apparent degree of breath support and muscular anchoring?*

Although I list breath support last in outlining the vocal timbre description system, it functions as a modifier of all the areas of perceived vocal production already discussed. Singing can require intense, focused muscular exertion, with continuous small muscular adjustments throughout the abdomen and around the ribcage in order to expel air from the lungs at a consistent and efficient rate, and throughout the head and neck to control the position of the vocal tract. I find that timbres produced with high airflow or pressure—such as harsh phonation or belting—also communicate a sense of full-body engagement due to the energy required for what Estill calls the head, neck, and torso “anchoring” that accompanies these types of sounds. Timbres produced with low pressure and airflow, that usually accompany low-amplitude singing—such as breathy phonation, or the sound produced by the sob vocal tract position—thus communicate a more relaxed, quiet kind of body engagement.

Aretha Franklin’s Vocal Timbre in “Respect” and “Natural Woman”

[4.1] Attending to these four parameters of vocal production can aid the analyst in characterizing the multiple and overlapping elements of vocal timbre in a manner that is grounded in sonic detail and physical experience, and can be used to tie the physical experience of music listening to song interpretation. We can use this system to trace the path from perceiving the sound of vocal timbre and experiencing its affect to conceptualizing the meaning of that timbre in the context of a given interpretive question. To demonstrate this system, I therefore pose the following question: What and how does the timbre of Aretha Franklin’s voice contribute to my understanding of the personae expressed in her performances of “Respect” and “(You Make Me Feel Like) A Natural Woman”?

[4.2] I want to first consider the opening “what” of “Respect,” and the similar vocal timbre Franklin uses as she begins several subsequent short phrases (with “what” and “all”; refer again to Example 1). Franklin’s mode of phonation seems to be clear and energetic—I can hear no pronounced roughness or distortion. The most striking timbral element of this moment is how penetrating this “what” feels. Her voice feels almost painfully resonant throughout my head, and I imagine she produces this sound with an energetic, almost smiling “twang” vocal tract setup, and a powerful, yet not tense, physical anchoring and thrust of air. Franklin is singing an E_b5, but with a remarkably different vocal timbre (and presumably different vocal tract configuration) from the one I typically must use to reach that note. I usually need to switch to a breathier, head, or mixed-type vocal production to reach this note, so the way Franklin hits it with her strong, regular manner of phonation is thrilling to imagine. I find it very difficult to keep all these elements—air flow rate, high pitch, and regular phonation—stable at the same time, without unwanted vocal breaks or harshness. Some harsh, aperiodic phonation does seem to occur as she sings “you need.” I take this as an indication that she is keeping the timbre of her voice *just* under control, adding to my positive appraisal of her vocal sound and skill. The strong sympathetic vibrations I perceive as a result of her vocal sound also extends beyond my current embodied experience to recall other, similar experiences: I imagine that the powerful vibrations of Franklin’s voice set everything in her vicinity ringing—that she literally takes control of the space around her, and fills it up with the sound of her voice.

[4.3] The combination of technical difficulty and strength of execution suggested by my attempt at imitation affords a host of associated stances, all of which impart a feeling of physical confidence. Listening to Franklin’s vocal timbre in this moment is like hearing a ringing shout of righteous indignation, or for a more distant association, like watching a world-class athlete perform at the peak of her ability. As a woman listening to Franklin’s performance, an embodied understanding of her vocal timbre is at the core of the thrilling possibility of power and commanding ability made real in her voice. Through my embodied experience of the timbre of her voice, I have the opportunity to try on a mode of self-expression that is powerful and hugely present.

[4.4] Alternatively, what does Franklin’s vocal timbre communicate in the opening of “(You Make Me Feel Like) A Natural Woman” (**Example 21**)?

As she sings “lookin’ out . . . on the morning rain,” Franklin’s phonation sounds a bit breathier than in “Respect,” with a less

energetic vocal tract configuration and lower breath support. Given my knowledge of her voice in other contexts, I believe this breathiness comes from a relaxation throughout her entire vocal tract—from the sound of air passing through incompletely tensed vocal folds and past the slightly inward drop of her pharynx walls. It is perhaps more accurate to characterize her voice as very slightly hoarse, because it sounds like she must be using more air support and muscular anchoring than I would normally associate with a breathy vocal timbre. This seems especially likely when I consider this vocal timbre in the context of all the high-volume, and potentially high-stress, singing she had engaged in prior to this point in her career. This noise suggests an active, strenuous history of vocal use. Franklin’s moderate-sounding breath support and relaxed vocal tract still produce a sound that projects, however, and I still locate sympathetic vibrations in my nasal passages. She retains a hint of that “twang” vocal quality, which I experience and therefore imagine as a vocal delivery with raised and widened tongue, retracted corners of the mouth, and epiglottis drawn toward the rear wall of the pharynx. I believe there is something about her habitual vocal tract setting that consistently preserves this quality and enables a “resonant” vocal timbre even as she sings in a lower-energy manner.

[4.5] I therefore experience the resonant power of Franklin’s voice in “Respect” as still present in the opening of “Natural Woman,” but latent. Franklin does eventually use this more energetic timbre in the first chorus (**Example 22**), supported by a swell in the instrumentation and backing vocals that very effectively expresses the mix of triumph and gratitude expressed by the lyrics (since I associate both states with powerful surges of energy and emotion in my own body). It is mostly the breathiness, the slight hoarseness of her voice, and her less energetic manner that makes me interpret Franklin as initially expressing a resigned weariness. My embodied engagement with the timbre of her voice affords associations of literal wear and fatigue from the strain of living. This feeling is supported by the text (especially her feeling “uninspired” and “tired”), the moderate tempo of the song, and the relatively spare accompaniment. This weariness brings a relatable vulnerability to Franklin’s expression, before she again displays her expert vocal power in the song’s chorus.

[4.6] The timbre of Franklin’s voice gives me a visceral sense of what it is like to be the wistful or powerful personae in these songs.⁽²²⁾ This is the basis of why listening to these recordings is both so pleasurable and uplifting—they express ways of being that I have felt, and also long to feel. These performances express potential ways of being that are now part of my imagination of the world. The range, control, power, and nuance encapsulated in Franklin’s vocal timbre represents a level of artistic self-expression I can aspire to as I sing along.

[4.7] For the purposes of this article, I have kept the interpretive project for which I examined Franklin’s vocal timbre small. I focused on moments of each performance where vocal timbre inhabited the foreground of my listening experience, and let my interpretation be guided by my immediate, individual listening context. The potential applications of this system are, however, broader and more numerous than can be summarized here. Beyond aiding an investigation of how vocal timbre contributes to the conceptualization of a song persona, this method can be used in concert with a variety of analytic tools (including transcription and spectrograms) for any study requiring a comprehensive analysis of vocal performance. The embodied parameters of vocal timbre I discuss here can be graphically visualized according to any number of timbre dimension “scales” the analyst may want to emphasize: modal to breathy phonation, regular to aperiodic vocal fold vibration, constricted to expanded vocal tract area, sympathetic vibration location (ranging from the front of the face to the chest), or degree of tension and anchoring. These dimensions may be used to visually represent changes in elements of vocal timbre over the course of a recording, or combined to create a comparative visual array of a variety of singers’ vocal timbre components.⁽²³⁾ The ability to describe vocal timbre also has applications beyond the analysis of popular music or the types of vocal sound covered in these audio examples. Any analyst interested in methods that respond to the embodied nature of listening will benefit from this explication of vocal timbre, since our voices are very often the touchstone for understanding all types of sound. In a similar vein, this approach helps characterize the affect and meaning of vocal timbres that are dramatically transformed via technological means—we still can and do conceptualize these altered voices in relationship to our own.

[4.8] Using the four guiding questions about vocal fold vibration, positioning of the vocal tract, location of sympathetic vibration, and breath support, we can better attend to what vocal timbre *feels* like as well as what it sounds like. In this way, we will be better equipped to incorporate well-organized, experientially grounded detail into vocal timbre descriptions, and to bridge the gap between quantitative measures of vocal timbre and the language we use to describe our perceptions of it.

Works Cited

- Barthes, Roland. 1977. "The Grain of the Voice." In *Image, Music, Text*, edited and translated by Stephen Heath, 179–89. New York: Hill and Wang.
- Bastian, Robert. 2016. "Introduction to Larynx, Pharynx, and Airway Anatomy." *Bastian Medical Media for Laryngology* video, 14:59. <http://bastianmedicalmedia.com/anatomy-physiology/>
- Björkner, Eva. 2006. "Why so different?: Aspects of Voice Characteristics in Operatic and Musical Theatre Singing." PhD diss., KTH School of Computer Science and Communication, Stockholm, Sweden.
- Blake, David K. 2012. "Timbre as Differentiation in Indie Music." *Music Theory Online* 18 (2).
<http://www.mtosmt.org/issues/mto.12.18.2/mto.12.18.2.blake.php>
- Bregman, Albert. 1990. *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press.
- Brackett, David. 2000. *Interpreting Popular Music*, 2nd edition. University of California Press.
- Buescher, Randy, and Steven Sims. 2011. "The Female Pharyngeal Voice and Theories of Low Vocal Fold Damping." *Journal of Singing* 68 (1): 23–28.
- Caclin, Anne, Stephen McAdams, Bennett K. Smith, and Suzanne Winsberg. 2005. "Acoustic correlates of timbre space dimension: A confirmatory study using synthetic tones." *Journal of the Acoustical Society of America* 188 (1): 471–482.
- Chandler, Kim. 2014. "Teaching Popular Music Styles." In *Teaching Singing in the 21st Century*, edited by Scott D. Harrison and Jessica O'Bryan, 35–51. Springer.
- Clarke, Eric. 2005. *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning*. Oxford University Press.
- Cogan, Robert. 1984. *New Images of Musical Sound*. Harvard University Press.
- Cox, Arnie. 2011. "Embodying Music: Principles of the Mimetic Hypothesis." *Music Theory Online* 17 (2).
<http://www.mtosmt.org/issues/mto.11.17.2/mto.11.17.2.cox.html>
- . 2001. "The Mimetic Hypothesis and Embodied Musical Meaning." *Musicae Scientiae* 5 (2): 195–209.
- Cusick, Suzanne G. 1994. "Feminist Theory, Music Theory, and the Mind/Body Problem." *Perspectives of New Music* 32 (1): 8–27.
- Eidsheim, Nina. 2012. "Voice as Action: Toward a Model for Analyzing the Dynamic Construction of Racialized Voice." *Current Musicology* 93: 9–33, 152.
- . 2008. "Voice as a Technology of Selfhood: Towards an Analysis of Racialized Timbre and Vocal Performance." PhD diss., University of California, San Diego.
- Feld, Steven, Aaron A. Fox, Thomas Porcello and David Samuels. 2005. "Vocal Anthropology: From the Music of Language to the Language of Song." In *A Companion to Linguistic Anthropology*, edited by Alessandro Duranti. Blackwell Reference Online.
http://www.blackwellreference.com/subscriber/tocnode?id=g9781405144308_chunk_g978140514430817
- Fox, Aaron. 2004. *Real Country: Music and Language in Working-Class Culture*. Duke University Press.
- Frith, Simon. 1996. *Performing Rites: On the Value of Popular Music*. Harvard University Press.

- Geertz, Clifford. 1976. "Art as a Cultural System." *MLN* 91/6 *Comparative Literature*: 1473–1499.
- Godøy, Rolfe Inge and Marc Leman, editors. 2010. *Musical Gestures: Sound, Movement, and Meaning*. New York: Routledge.
- Hammarberg, Britta, B. Fritzell, J. Gauffin, J. Sundberg, and L. Wedin. 1980. "Perceptual and acoustic correlates of abnormal voice qualities." *Acta Otolaryngologica* (Stockholm) 90: 441–451.
- Heidemann, Kate. 2014. "Hearing Women's Voices in Popular Song: Analyzing Sound and Identity in Country and Soul." PhD diss., Columbia University.
- Isshiki, N., H. Okamura, M. Tanabe, and M. Morimoto. 1969. "Differential diagnosis of hoarseness." *Folia Phoniatrica* 21: 9–19. DOI:10.1159/000263230
- Johnson, Mark. 2007. *The Meaning of the Body: Aesthetics of Human Understanding*. University of Chicago Press.
- Kreiman, Jody, and Diana Sidtis. 2011. *Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception*. Wiley-Blackwell.
- Kreiman, Jody, Bruce R. Gerratt, and Mika Ito. 2007. "When and why listeners disagree in voice quality assessment tasks." *Journal of the Acoustical Society of America* 122 (4): 2354–2364. DOI: 10.1121/1.2770547
- Kreiman, Jody, Diana Vanlancker-Sidtis and Bruce R. Gerratt. 2004. "Perception of Voice Quality." In *The Handbook of Speech Perception*, edited David B. Pisoni and Robert E. Remez. Blackwell Reference Online. http://www.blackwellreference.com/subscriber/tocnode?id=g9780631229278_chunk_g978063122927817
- Lacasse, Serge. 2010. "The Phonographic Voice: Paralinguistic Features and Phonographic Staging in Popular Music Singing." In *Recorded Music: Performance, Culture and Technology*. Edited by Amanda Bayley, 225–251. Cambridge University Press.
- Laver, John. 2000. "Phonetic Evaluation of Voice Quality." In *Voice Quality Measurement*, edited by Raymond D. Kent and Martin J. Ball, 37–48. Singular Publishing Group.
- . 1980. *The Phonetic Description of Voice Quality*. Cambridge University Press.
- Leman, Marc. 2010. "Music, Gesture, and the Formation of Embodied Meaning." In *Musical Gestures: Sound, Movement, and Meaning*, edited by Rolf Inge Godøy and Marc Leman, 126–153. Routledge.
- Marozeau, Jeremy, and Alain de Cheveigné. 2007. "The effect of fundamental frequency on the brightness dimension of timbre." *Journal of the Acoustical Society of America* 121 (1): 383–387. Doi: 10.1121/1.2384910.
- McAdams, Stephen. 1999. "Perspectives on the Contribution of Timbre to Musical Structure." *Computer Music Journal* 23 (3), Recent Research at IRCAM: 85–102.
- McDonald Klimek, Mary, Kerrie Obert, and Kimberly Steinhauer. 2005a. *Estill Voice Training System Level One: Compulsory Figures for Voice Control*. Estill Voice Training Systems International.
- . 2005b. *Estill Voice Training System Level Two: Figure Combinations for Six Voice Qualities*. Estill Voice Training Systems International.
- Merleau-Ponty, Maurice. 2012. *Phenomenology of Perception*, translated by Donald A. Landes. New York: Routledge. Originally published as *Phénoménologie de la perception* (Éditions Gallimard, 1945).
- Moisik, Scott R. and John H. Esling. 2014. "Modeling the Biomechanical Influence of Epilaryngeal Stricture on the Vocal Folds: A Low-Dimensional Model of Vocal-Ventricular Fold Coupling." *Journal of Speech, Language, and Hearing Research* 57: S687–S704.
- Moisik, Scott R., John H. Esling, and Lise Crevier-Buchman. 2010. "A high-speed laryngoscopic investigation of aryepiglottic trilling." *Journal of the Acoustical Society of America* 127 (3): 1548–1558. Doi: 10.1121/1.3299203
- Molnar-Szakacs, Istvan, and Katie Overy. 2006. "Music and Mirror Neurons: From Motion to 'E'motion." *Social Cognitive and*

Affective Neuroscience 1 (3): 235–41.

Moore, Allan F. 2012. *Song Means: Analyzing and Interpreting Recorded Popular Song*. Ashgate Publishing.

Mufwene, Salikoko S., John R. Rickford, Guy Bailey, and John Baugh. 1998. *African American English: Structure, History, and Use*. Routledge.

Nair, Garyth. 2003. "The Term 'Falsetto': Navigating Through the Semantic Minefield." *Journal of Singing* 60 (1): 53–60.

Overy, Katie, and Istvan Molnar-Szakacs. 2009. "Being Together in Time: Musical Experience and the Mirror Neuron System." *Music Perception* 26 (5). Doi:10.1525/MP.2009.26.5.489

Poyatos, Fernando. 2002. *Nonverbal Communication across Disciplines. Volume 2: Paralanguage, kinesics, silence, personal and environmental interaction*. John Benjamins Publishing Company.

Reed, S. Alexander. 2005. "The Musical Semiotics of Timbre in the Human Voice and Static Takes Love's Body." PhD diss., University of Pittsburgh.

Rings, Steven. 2013. "A Foreign Sound to Your Ear: Bob Dylan Performs 'It's Alright, Ma (I'm Only Bleeding),' 1964–2009." *Music Theory Online* 19 (4).

<http://www.mtosmt.org/issues/mto.13.19.4/mto.13.19.4.rings.html>

Roubeau, Bernard, Nathalie Henrich, and Michèle Castellengo. 2009. "Laryngeal Vibratory Mechanisms: The Notion of Vocal Register Revisited." *Journal of Voice* 23 (4): 425–438.

Sadolin, Cathrine. 2000. *Complete Vocal Technique*. Shout Publishing.

Sundberg, Johan. 1987. *The Science of the Singing Voice*. Northern Illinois University Press.

Sundberg, Johan, and Margareta Thalén. 2010. "What is 'Twang'?" *Journal of Voice* 24 (6): 654–660.

Titze, Ingo R. 2008. "A Hypothesis About Whistle Voice." *Journal of Singing* 64 (4): 473–475.

———. 2001. "Acoustic Interpretation of Resonant Voice." *Journal of Voice* 15 (4): 519–528.

———. 1994. *Principles of Voice Production*. Prentice Hall.

Titze, Ingo R., Albert S. Worley, and Brad H. Story. 2011. "Source-Vocal Tract Interaction in Female Operatic Singing and Theater Belting." *Journal of Singing* 67 (5): 561–572.

Wallmark, Zachary Thomas. 2014. "Appraising Timbre: Embodiment and Affect at the Threshold of Music and Noise." PhD diss., UCLA.

Zak, Albin. 2001. *Poetics of Rock Composition: Multitrack Recording as Compositional Practice*. University of California Press.

Discography

Arctic Monkeys. "I Bet You Look Good on the Dancefloor." *Whatever People Say I Am, That's What I'm Not*. Domino USA CD, 2006.

Louis Armstrong. "What A Wonderful World." *The Definitive Collection*. Hip-O CD, 2006. Original Release: ABC, 1967.

Crosby, Bing, and The Andrews Sisters. "Christmas In Killarney." *A Merry Christmas With Bing Crosby & The Andrews Sisters*. MCA Records MP3, 2000. Original release: *Merry Christmas*. Decca, 1955.

Franklin, Aretha. "Respect." *I Never Loved A Man The Way I Love You*. Rhino MP3, 2007. Original release: Atlantic, 1967.

Franklin, Aretha. "(You Make Me Feel Like) A Natural Woman." *Lady Soul*. Atlantic & Atco Remasters/Atlantic/WEA CD, 1995. Original release: Atlantic, 1968.

- Gilberto, João and Stan Getz, featuring Astrud Gilberto. “The Girl From Ipanema.” *Getz/Gilberto*. Verve MP3, 1997. Original release: Verve, 1964.
- Green, Al. “Let’s Stay Together.” *The Definitive Greatest Hits*. Capitol Records CD, 2007. Original release: *Let’s Stay Together*. Hi Records, 1972.
- INXS. “Need You Tonight.” *Kick*. Rhino Atlantic MP3, 2011. Original release: Atlantic, 1987.
- Knight, Gladys, and the Pips. “Help Me Make It Through the Night.” *If I Were Your Woman + Standing Ovation*. Motown/Universal Island Records Ltd. CD, 2006. Original release: *If I Were Your Woman*. Soul (Motown), 1971.
- Menzel, Idina. “Let It Go.” *Frozen (Deluxe Edition)*. Walt Disney Records MP3, 2013.
- Mitchell, Joni. “I Had A King.” *The Studio Albums 1968–1979*. Warner Music/Reprise Records CD, 2012. Original release: *Song to A Seagull*. Reprise Records, 1968.
- Modest Mouse. “Bury Me With It.” *Good News For People Who Love Bad News*. Epic CD, 2004.
- Nelson, Willie. “Can I Sleep In Your Arms.” *Red Headed Stranger*. Columbia/Legacy CD, 2000. Original release: Columbia, 1975.
- Nirvana. “Smells Like Teen Spirit.” *Nevermind*. Geffen Records CD, 1991.
- Parton, Dolly. “Jolene.” *Jolene*. RCA/Legacy CD, 2007. Original release: RCA Victor, 1974.
- Redding, Otis. “These Arms of Mine.” *The Very Best of Otis Redding, Vol. 1*. Rhino CD, 1992. Original release: *Pain in My Heart*. Atco, 1964.
- Spears, Britney. “Oops! . . . I Did It Again.” *Oops! . . . I Did It Again*. Jive MP3, 2000.
- Summer, Donna. “Love to Love You Baby.” *On the Radio: Greatest Hits Volumes I & II*. Casablanca CD, 1987. Original release: Casablanca, 1979.
-

Footnotes

1. Thanks to Arnie Cox and Steven Nuss for their helpful comments on early drafts of this work. Thanks also to the two anonymous reviewers for their feedback and comments on future avenues for this research.

[Return to text](#)

2. This characterization of listening to music reflects in a different modality Clifford Geertz’s observation that works of visual art “materialize a way of experiencing,” that they “bring a particular cast of mind out into the world of objects, where [people] can look at it” (Geertz 1976, 1478).

[Return to text](#)

3. I am not alone in noting this kind of reaction to voice: well-known reflections on the impact of the voice can be found in Barthes (1977) and Frith (1996).

[Return to text](#)

4. Listener judgments of timbral brightness, for example, can be fairly reliably correlated to the strength of a sound’s high upper partials, but some listeners also characterize high-pitched sine tones as “bright,” suggesting that frequency in general is strongly related to the perceptual correlate of brightness (Cogan 1984, 135; Marozeau and de Cheveigné 2007). The typical perceptual correlate seems to be that whatever the average brightness of a voice or instrument, a higher pitch will result in a brighter tone.

[Return to text](#)

5. For example, in timbre differentiation tasks, some listeners rely more on differences in attack time while others are more sensitive to differences in spectral center of gravity (Caclin et al. 2005).

[Return to text](#)

6. Dialect, accent, and enunciation are components of vocal sound that can be explored in great depth with the aid of articulatory phonetics. These are the elements of voice that most readily yield our impressions of race, ethnicity, class, and region of origin. For example, well-documented dialectal features of African-American English signal the race of some African-American speakers (Mufwene et al. 1998, 204).

[Return to text](#)

7. In general, humans are very good at recognizing different spectral profiles and changes among partials without necessarily being able to describe the acoustic components of timbre. This sensitivity to sound spectra plays an important role in auditory stream segregation and the grouping of sounds (Bregman 1990, 99–100).

[Return to text](#)

8. For example, Thomas Porcello observed all of these strategies as part of an effective system of talking about timbre shared by musicians and engineers working together in a recording studio (Feld et al. 2005).

[Return to text](#)

9. Voice researchers Jody Kreiman, Diana Sidtis, and Bruce Gerratt have summarized a large body of prior research on the perception of speech quality, a term used in the field of speech perception that encompasses timbre as just described, noting such difficulty of agreement. For example, “in studies of pathological voice [quality] Isshiki et al. (1969) found a ‘breathiness’ factor that loaded highly on the scales dry, hard, excited, pointed, cold, choked, rough, cloudy, sharp, poor, and bad, while a ‘breathiness’ factor reported by Hammarberg et al. (1980) corresponded to the scales breathy, wheezing, lack of timbre, moments of aphonia, husky, and not creaky” (Kreiman, Vanlancker-Sidtis, and Gerratt 2004).

[Return to text](#)

10. Spectrograms have been fruitfully employed in analyses of popular singing voices in Cogan (1984), Brackett (2000), and Rings (2013).

[Return to text](#)

11. Scholarly interest in the role of embodiment and mimesis in music cognition (and cognition generally) has increased over the last three decades. For recent examples of such work, see Cox (2001, 2011), Molnar-Szakacs and Overy (2006), and Godoy and Leman (2010).

[Return to text](#)

12. These connotations correspond to Albin Zak’s concept of timbre’s rhetorical form: “the conventional associations that sounds have, which allow them to stand as symbols suggesting dialogues and resonances beyond the boundaries of the track” (Zak 2001, 62). In this article, I focus on the material, embodied grounds of these types of associations, but I purposely avoid interpretive commentary drawing from the rhetorical or associative realm until my final analysis. Rhetorical aspects of timbres are clearer in specific contexts of genre or region, but become diffuse and open-ended in general listening scenarios, especially among listeners with different backgrounds or training. The rhetorical aspects of specific vocal timbres are best approached as a project separate from, yet related to, creating a system for the clear identification of vocal timbre.

[Return to text](#)

13. A very clear, comprehensive introduction to voice physiology is accessible at <http://bastianmedicalmedia.com/anatomy-physiology/>. Dr. Robert Bastian’s video, “Introduction to Larynx, Pharynx, and Airway Anatomy,” includes a narrated tour of the vocal apparatus as well as explanation of video clips from laryngoscopies of speaking and singing.

[Return to text](#)

14. These terms come from Laver (1980, 2000), Kreiman and Sidtis (2011), McDonald et al. (2005b), and Chandler (2014).

[Return to text](#)

15. Pop vocal coach Cathrine Sadolin discusses this and similar techniques for producing a wide variety of vocal effects (although using somewhat different terminology) in her book *Complete Vocal Technique* (2000).

[Return to text](#)

16. For a more in-depth discussion of the effects created by vibrations in the epilarynx, see Moisik et al. (2010) and Moisik and Esling (2014).

[Return to text](#)

17. Research has revealed that the common perceptual scale of brightness used to describe timbre is related in part to the manner of vocal fold vibration: “Voices characterized by strong, quick vocal fold closure have many high-frequency harmonics in addition to the energy at the fundamental frequency, and are often described as ‘bright-sounding.’ Voices generated by gradual or incomplete closure of the vocal folds have most of their energy at or near the fundamental frequency, and sound ‘dull’ or ‘weak’” (Kreiman and Sidtis 2011, 49). “Brightness” is a cross-modal metaphor that roughly equates high energy in sound production with high intensity in light. The concept of brightness, however, is not typically grouped with other phonation descriptors because of its strong correlation to fundamental and overtone frequencies, which are controlled in large part by the length and shape of the vocal tract (which I discuss in the following sections).

[Return to text](#)

18. Some of this work is being conducted at the UCLA Speech Lab under the supervision of Jody Kreiman, and by singing voice researchers Ingo Titze, Johan Sundberg, and their colleagues and students.

[Return to text](#)

19. Some researchers posit differences in vocal posture resulting in the perception of pharyngeal voice as slightly distinct from twang (Buescher and Sims 2011).

[Return to text](#)

20. The acoustic effect of the narrow pharynx in nasal twang is that it typically raises the first formant and/or second formant, leading to an overall effect of brightening that allows for vocal projection with less effort and strain on the vocal folds (Sundberg and Thalén 2010; Titze 2001).

[Return to text](#)

21. These sympathetic vibrations sometimes get characterized as “resonance,” even though they do not correspond to the physical location where vocal resonance actually occurs.

[Return to text](#)

22. This type of vocal timbre description and interpretation can also form the basis of exploring how recorded vocal performances may inform our embodied and perceptually learned understandings of social categories like gender, race, and class. These are the types of interpretive questions I addressed in my dissertation (Heidemann 2014).

[Return to text](#)

23. This potential for multidimensionality points to a significant avenue for further research. One challenge for creating any graphical scheme using these scales lies in choosing the most salient parameters to portray, since the analyst would presumably be creating, at most, a three-dimensional display.

[Return to text](#)

Copyright Statement

Copyright © 2016 by the Society for Music Theory. All rights reserved.

[1] Copyrights for individual items published in *Music Theory Online* (*MTO*) are held by their authors. Items appearing in *MTO* may be saved and stored in electronic or paper form, and may be shared among individuals for purposes of scholarly research or discussion, but may *not* be republished in any form, electronic or print, without prior, written permission from the author(s), and advance notification of the editors of *MTO*.

[2] Any redistributed form of items published in *MTO* must include the following information in a form appropriate to the medium in which the items are to appear:

This item appeared in *Music Theory Online* in [VOLUME #, ISSUE #] on [DAY/MONTH/YEAR]. It was authored by [FULL NAME, EMAIL ADDRESS], with whose written permission it is reprinted here.

[3] Libraries may archive issues of *MTO* in electronic or paper form for public access so long as each issue is stored in its entirety, and no access fee is charged. Exceptions to these requirements must be approved in writing by the editors of *MTO*, who will act in accordance with the decisions of the Society for Music Theory.

This document and all portions thereof are protected by U.S. and international copyright laws. Material contained herein may

be copied and/or distributed for research purposes only.

Prepared by Rebecca Flore, Editorial Assistant