

Speech, Sound, Technology

Johannes Mulder and Theo van Leeuwen

The Oxford Handbook of Sound and Imagination, Volume 1

Edited by Mark Grimshaw-Aagaard, Mads Walther-Hansen, and Martin Knakkergaard

Print Publication Date: Sep 2019 Subject: Music, Sound Studies

Online Publication Date: Aug 2019 DOI: 10.1093/oxfordhb/9780190460167.013.24

Abstract and Keywords

Johannes Mulder and Theo van Leeuwen investigate how the microphone, the amplifier, and the loudspeaker have changed the semiotic potential of the sound of the voice. Based on a social semiotic understanding, the authors discuss the voice from three perspectives: physical, focusing on how bodily experience from speaking and singing informs the understanding of the sound of the voice (covering implications such as pitch, loudness, and alterations of various kinds); social, based on the way human social and cultural experiences, such as accents and different vocal styles, inform the understanding of sound; and, finally, the technologically enhanced voice, showing that, even though the appreciation of this voice rests on prior understanding of biological and social voices, this voice introduces extraexperiential dimensions that slowly assimilate into the culture as a form of disembodied, nonhuman voice.

Keywords: sound production, music technology, voice, singing, sound semiotics

Introduction

LINGUISTS have rarely paid attention to the sound of speech. What mattered in phonology, the linguistic approach to the study of speech sound, was not sound itself, as an expressive mode, but the function of speech sounds in distinguishing words from one another. And phonology was kept separate from phonetics, which studied the production of speech sounds and their acoustic properties but without reference to their meaning and function. At best, phoneticians saw speech sound as a symptom rather than a sign—an index of age, health, energy level, emotional state, and also of regional origin or ethnicity (e.g., Laver 1980). In this chapter, we take a semiotic approach to the sound of speech. As poets have long understood, spoken language simultaneously makes meaning at two levels—the level of word meaning and the level of sound meaning.

It was not only in linguistics, but also in dominant practices of speech and singing that sound was not treated as a mode of expression in its own right. What mattered in public speech was clear, correct diction that would not draw attention to itself. The key institu-

tions of the nation state, the education system and broadcasting, everywhere propagated homogeneous “educated” forms of speech. And what mattered in high-culture styles of singing, in what Lydia Goehr (1992) refers to as the Imaginary Museum of Musical Works in the silent audience tradition of the modern (post-1800) Western European concert hall, was a single ideal of articulatory beauty, in which the voice can be understood as a musical instrument.

This, of course, did not prevent singers differing in the degree to which they could achieve that ideal but nevertheless ensured that performance would be in the service of the composer’s writing and not allow individual, idiosyncratic styles to upstage the composer.

(p. 474) In the course of the twentieth century all this began to change. Movie actors and singers developed their own, immediately recognizable styles of speech and singing and, thanks to the new recording technologies, these became part of language, and part of the language of music, imitated and transformed the world over. No longer was the voice just an index of unchangeable biological and demographic characteristics: it became a source of meaning and a means of expressing identity.

Clearly, this change was facilitated by the affordances of the new media of telephony, broadcasting, and film, though this is not to say that technology played no role in vocal production before the twentieth century. Architects had developed acoustic technologies of amplification for thousands of years (Blessner and Salter 2006, 199) and masks in ancient Greek theater were also used to amplify the voice and emphasize vocal characteristics befitting specific roles—there is some debate about which of these two were most important:

Vocal resonance masks afford additional possibilities for emphasizing specific vocal qualities and can even facilitate special vocal effects. When the voice is able to orient properly in a vocal resonance mask, effects such as the highlighting of overtones can be achieved. Furthermore, the voice is experienced as fuller than otherwise.(Alström 2004, 136)

But overall it was twentieth-century electronic amplification (Mulder 2013; Wijffjes 2014) that made it possible for public speech, whether broadcast or live, to become conversational (Roosevelt’s fireside chats are often mentioned in this context; e.g., Ryfe 1999) and for “crooning” styles of singing, such as those of Rudy Vallee and Bing Crosby, to emerge (e.g., McCracken 1999). Amanda Weidman notices the swift uptake of electronic amplification in Indian classical music¹ as early as the mid-1930s:

It is crucial to note that the voice with “innumerable nuances” is itself a product of the microphone; as the microphone makes these nuances audible in a physical sense, it also makes them available to an aesthetic discourse about music. (2006, 126)

But the microphone, the amplifier, and the loudspeaker were more than facilitators—they also changed the semiotic potential of the sound of the voice. It was soon noted, for instance, that technology can disembody the voice, make it “acousmatic” (Chion 1994) not only in radio broadcasts, cinema voice-overs, telephone conversations, and so on, but also in live performances—the voice of singers in, for instance, Broadway musicals, no longer came (just) directly from the performers but from loudspeakers positioned outside the proscenium. Sometimes this is seen as a loss: “To disembody a voice is to rob it of individuality and this leads directly to the question of vocal identity, to the voice as person” (Frith 1996, 196). The drama of this electro acoustical aberration is emphasized in R. Murray Schafer’s term of choice to indicate the schism, schizophonia (Schafer 1969, 43). But, whether acousmatic or schizophonic, the disembodied voice can also create new resources for meaning making. It can, for instance, represent the voice of God, as already (p. 475) in the biblical story of God speaking to Moses from a burning bush—or as with the voice of the wizard in *The Wizard of Oz*, as discussed by Michel Chion (1994). And it can stimulate the imagination, the creation of mental images, as for instance in the films of Robert Bresson, who wrote, “The ear goes more to the within, the eye to the outer” (1997, 61). In Bresson’s films, acousmatic sound often cues our imagination to fill in the missing parts of the audiovisual narrative. A famous example from his *L’Argent* (1983) is a scene in which Yvon hits the old woman who is hiding him from the police in her face—or so we imagine, as we can only hear the sound and do not see the act of violence. We will return to the issue of imagination in our discussion of the technological enhancement of sounds in the concluding section of this chapter.

In this chapter, we will discuss the semiotic resources unlocked by contemporary audio technology. In doing so we take a social semiotic perspective (cf., Van Leeuwen 2005; Kress 2010), arguing that meaning making is an activity and that our understanding of the meanings of sound *as sound* rests on our experience as biological and social beings. As such, social semiotics expand on Saussurian and Peircean approaches by underlining the social and cultural drivers of signification. Our account will therefore unfold in three steps: an account of the *physical voice*, based on the way our bodily experience as speaking and singing beings informs our understanding of sound; an account of the *social voice*, based on the way our social and cultural experience informs our understanding of sound; and an account of the *technologically enhanced voice*, showing, on the one hand, how our understanding of the technological voice ultimately rests on our understanding of the biological and the social voice but, on the other hand, how technology also introduces extraexperiential dimensions that only slowly assimilate into the culture.

However, even though audio technology does much more than facilitate, it is not the *cause* of semiotic change. The new resources for the expression of identity developed in the context of the growing power of the entertainment industry and of contemporary corporate culture, which do not need national but global semiotic resources. It can be observed that such resources do not always need homogeneity, but thrive on diversity, branding, and customization. That this impacts also on the sound of the voice is demonstrated by the activities of companies like Acapela² that provide synthesized voices for announcements and “product enhancement,” encourage their clients to “use your own ex-

clusive voice, enhance your brand identity and differentiate your services with your own exclusive synthetic voice,” and offer a wide range of “voices with accents, voices of celebrities, voices that surprise you with their naturalness and custom-made voices.”

The Physical Voice

In this section, we argue that our lived experience as hearing, speaking, and singing beings informs the meaning potential of the sound of the voice. This insight is based on the metaphor theory of George Lakoff and Mark Johnson, who argue that we understand metaphors on the basis of our concrete experiences: “No metaphor can ever be comprehended or even adequately represented independently of its experiential basis” (1980, 19). This includes bodily experiences, such as walking upright, or tensing of the voice. When we tense our vocal musculature, our voice becomes higher, sharper, and brighter because, in their tensed state, the walls of the throat cavity dampen the sound less than they would in their relaxed state. The resulting sound not only *is* tense, it also *means* “tense.” Our experience allows us to recognize tension in our own voice and in the voice of others and to know when voices are likely to become tense—when we are nervous, or intimidated, or threatened, for instance. This then allows a metaphorical transfer from the domain of experience to the domain of more abstract ideas and values and identities that have some component of tension in their meaning. This can express a moment of tension, deliberately or otherwise, but voices might also become habitually tense, expressing “tense” aspects of the identity of individuals or groups. Alan Lomax wrote of the high degree of tension in the preferred female singing styles of cultures with severe sexual repression of women:

It is as if one of the assignments of the favored singer is to act out the level of sexual tension which the customs of the society establish as normal. The content of this message may be painful and anxiety-producing, but the effect upon the culture member may be stimulating, erotic and pleasurable since the song reminds of familiar sexual emotions and experiences. (1968, 194)

Tension is not the only resource for vocal meaning. Voices combine a range of vocal features, each one deriving from a particular articulatory gesture. They are high or low *and* soft or loud *and* tense or lax, and so on, and all these features not only are capable of fine gradation but also contribute to the overall meaning, just as the ingredients of a dish all contribute, in their different proportions, to the overall taste of the dish.

Pitch Range

Our experience tells us that our pitch range tends to flatten when we feel flat and listless, or bored, but also when we are speaking in a soft, intimate voice—some of those sensuous Brazilian bossa nova songs have melodies with a very limited pitch range. When we are excited, on the other hand, our energy level and our pitch range increase—songs with

a wide pitch range are the staple fare of patriotic hymns and other songs that seek to energize people into action.

Our experience of pitch range tells us two other things as well: men's voices are, on average, lower than those of women and children, and smaller resonating chambers (e.g., violins) are sympathetic to higher sounds than larger resonating chambers (e.g., double basses). As a result, low voices can be threatening. In eighteenth-century operas from Beethoven's *Fidelio* and onward, the tenor is the hero and the bass or baritone the villain, for example. Similarly, high voices can be used to make ourselves small and harmless, as when talking to children. Pitch range can convey more gendered, heteronormative, meaning potential. Men use the higher regions of their pitch range to (p. 477) assert themselves and to dominate, while women use the lower end of their pitch range for this purpose. But it is hard to be low and loud at the same time, so women face a difficult choice. They must either speak low (which is assertive) and soft (which is intimate)—which can evoke the “dangerous woman” stereotype—or high (“belittling themselves”) and loud (“being assertive”)—which might be considered “shrill.” In either case, the dominant norms of the public, assertive, “masculine” voice are at odds with the dominant norms of the private, intimate, “feminine” voice. Iconic female voices create different female identities in this way. Marilyn Monroe used a high, yet breathy voice, combining feminine vulnerability and seductiveness. Lauren Bacall used a sensuous low voice. In her autobiography, she recalled how the director Howard Hawks conceived of her character in *To Have and Have Not* (1944) as a “masculine approach—insolent. Give as good as she got, no capitulation, no helplessness” (Bacall 1978, 87). To this end he told her to work on her voice, “to practice shouting, keeping my register low” (92).

Loudness

Loudness relates to power and sometimes to ignorance. The more powerful people or institutions are, the more noise they are allowed to make or are allowed to ignore their impact on a soundscape. The loudest noises, R. Murray Schafer wrote, are always “sacred noises”: “Thunder, the voice of God, migrated first to the cathedral, then to the factory and the rock band” (1977, 179). Schafer also recounted how a group of Hare Krishnas in Vancouver were fined for causing excessive noise while, nearby, the noise of a building site exceeded 90 decibels (201) but without the builders being fined. This example illustrates the ambiguous nature of the term “noise,” as we will argue later, but also how noise, outside of the engineering domain, can be evaluated based on its social desirability or lack thereof.

Our experience also tells us that loudness relates to social distance, both literally and figuratively. As explained by Edward Hall (1966, 154), when people interact at “very close range” (3–6 inches) their voices will be an intimate, soft whisper. At “close range” (7–12 inches) they will still whisper, still seek to be heard only by one person. At “near range” (12–20 inches), they will be soft indoors and “full” outdoors, but still low enough not to be overheard by strangers. At “close neutral range” (20–36 inches), people will tend to speak about “personal matters” in a “fuller voice.” At “public distance” (5.5–8

feet), their voices will be “full with slight over-loudness” conveying public information. Beyond that they will no longer speak to one or two people, but to a larger group. In short, each level of loudness connects to a particular kind of social relation.

Rough, Breathy, and Trembling Voices

In rough voices, we can hear other things beside the tone of the voice itself—hoarseness, harshness, rasp, grit. The opposite of the rough voice is the smooth voice, from which all noisiness is eliminated. Much of the effect of roughness comes from the aperiodic (p. 478) vibration of the vocal cords, which causes noise in the spectrum (Laver 1980, 128). As this is more audible in the lower pitches, it is more easily heard in male voices and lower female voices. Perhaps this is why rough voices are especially common in male singing and highly valued in cultural contexts that encourage assertiveness and enterprise (Lomax 1968, 192). Again, the meaning of roughness lies in what it is—rough. Our experience tells us that roughness comes from wear and tear, whether as a result of smoking and drinking, hardship and adversity, or old age. But roughness is valued differently in different contexts. In Western classical music Maria Callas’s rough and sometimes unpolished voice stood out, whereas in African American music, roughness emotes, in the words of Harold Courlander:

Some outstanding blues, gospel and jazz singers have voices that may be described as foggy, hoarse, rough or sandy. Not only is this kind of voice not derogated, it often seems to be valued. Sermons preached in this voice appear to create a special emotional tension.

(Courlander, quoted in Williams-Jones 1975, 377)

A rough/smooth contrast or continuum can be noted in popular music as well: the lived-through voice of Tom Waits can be contrasted with Leonard Cohen’s pleasant croon, the nasal trademark of Marianne Faithfull to Celine Dion’s warm head voice.

In the breathy voice, another sound mixes in with the tone of the voice itself—breath. Its metaphor potential derives from our experience of what can make our voice breathy—exertion of some kind, or excitement, or sexual arousal. It often combines with a soft voice, suggesting intimacy. Advertisers use it to give their message a sensual, erotic appeal. Singers sometimes use it for the same reason, commanding attention by invoking a need to listen more attentively, maybe even requiring listeners to hold their breath.

The meaning potential of vibrato, the trembling voice, derives from what we know makes our voice tremble—emotion. It plays a key role in the musical expression of emotion. Strings, for instance, are particularly good at producing vibrato sounds and hence are universally used to “pull the heartstrings,” to present and represent love and romance. But other emotions, too, can make us tremble, and vibrato, or alternatively tremolo,³ is also used, for instance, in horror movie music, to express fear and uncertainty.

Articulation

Some vowels are “frontal,” articulated with the tongue in the front of the mouth (e.g., the [i] of *heed*), others are articulated in the back of the oral cavity (e.g., the [a] of *hard*). This has often been commented on in relation to sound symbolism. In many languages words meaning “close” use frontal vowels and words meaning “far” back vowels (e.g., *near* and *far* in English, *hier* and *daar* in Dutch, *ici* and *là-bas* in French, *hier* and *dort* in German). But frontality and its opposite can also be overall articulatory settings expressing a quality of being “upfront,” “confronting” or, of “holding back,” “not coming out with it.”

The same can be said for “aperture.” Some vowels are produced with the mouth comparatively closed and the oral cavity therefore comparatively small, others with the (p. 479) mouth comparatively open and the oral cavity therefore larger. The [i] of *heed* and the [u] of *hood* are less open, for instance, than the [a] of *hard*. This too has mostly been discussed in relation to sound symbolism. Words with an “open” [a] have been said to be “heavy, big and round,” for instance, and words with an [i] “small, light and pointed” (Hildum and Brown 1956). In *Star Wars, The Phantom Menace* (2000), the treacherous Viceroy of Naboo, a character with an inscrutable fish-like physiognomy, not only has a vague Chinese accent, but also speaks with a stiff jaw and an almost-closed mouth, using a breathy, hollow-sounding faucalized voice. Such “articulatory” settings are often used by puppeteers and by the actors who create the voices of cartoon characters—in cartoons, vocal expression plays a key role, as dialogue is usually recorded before the characters are animated (Blanc and Bashe 1989, 75–99).

As mentioned earlier, all the parameters we have discussed in this section combine, in different proportions, in actual voices. The source of inspiration for countless Mafiosi characters, the famous and recognizable voice of Marlon Brando in *The Godfather* (1972), for instance, is comparatively high (and we have seen that men tend to use high voices to dominate). It is also hoarse and rough, signaling the Godfather’s harsh and unforgiving side. And it is articulated with a stiff jaw and an almost closed mouth, suggesting an unwillingness to “give” that keeps us guessing as to what he might be keeping from us. Yet it is also soft and breathy, at times almost a whisper, making his menacing presence disturbingly intimate and attractive.

The Social Voice

Our experience of sound is not only physical, it is also social and includes our experience of “others” who speak and sing in other ways—of dialects and foreign accents, for instance, and of the socially conditioned attitudes to the “others” whose dialects and accents they are. In the era of semiotic uniformity, the meaning of dialects and accents had been indexical, signaling where people came from. Often, they had stigmatized speakers. To upper-middle-class city dwellers, country dialects were backward, working class dialects inferior, and foreign accents could be either suspect or funny, or exotic and even erotic, depending on the kinds of connotation Barthes called “myths” and labeled with

terms such as “Italian-ness” to indicate that they condensed everything the *bourgeoisie* “knew” about Italians as well as the value judgments they attached to this (Barthes 1973, 109–159; 1977, 32–37). Much of this still lingers. The story of *Pygmalion* is still understood. Yet much has changed. In the early 1930s, MGM worried about Greta Garbo’s accent, but her voice was to become another key icon in the new semiotics of voices. Soon Hollywood scriptwriters would recommend the use of accents and dialects:

Dialect ... is as revelatory as make up, as picturesque as costume, as characteristic as gestures, as identifiable as physical disabilities, and as dramatically effective as facial expressions.

(Herman 1952, 198)

(p. 480) The author of this quote, Lewis Herman, a scriptwriter who produced several manuals on the subject, recommended the dropped “h” in cockney, the “d” and “t” substitutes for “th” in German, the “y” substitute for “j” in Swedish, and the dropped final “r” in Southern, Eastern, and Negro speech, together with the dropped “g” in “ing” ending words such as “fightin” and “doin.” Such heavily edited dialects and accents would soon become a new semiotic resource, uncoupled from their ties with a specific place or social group, familiar to movie audiences the world over, their meanings as easily understood as those of the clichés of classical Hollywood music scores and used in public speech as well as to color everyday conversation. Even on the BBC, where not long ago only “the Queen’s English” was permitted, a wide variety of accents and dialects can now be heard, and not only in “vox pops.” And, as we saw, companies like Acapela will flavor the synthetic announcements they produce with dialects and accents as requested by the customer.

The vocal styles developed by iconic actors such as Lauren Bacall, Marilyn Monroe, and Marlon Brando and iconic singers such as Ray Charles, Bob Dylan, Nina Simone, Astrid Gilberto, and others, similarly became a semiotic resource. Marlon Brando’s hoarse whisper is now part of the repertoire of many actors, to be deployed whenever a sense of brooding, yet sensuously attractive menace is required. The same applies to singing. Madonna was perhaps the first singer who used quite different voices for different songs (cf., Van Leeuwen 1999, 150–152). In *Like a Virgin*, for instance, she addressed a “you” who saved her from perdition and made her feel “new” and shiny,” “like a virgin touched for the very first time.” This “you” she addresses in a high, feminine voice, which nevertheless also has an edge of a more hardened, tense, and strident sound to signal scars of abuse and betrayal. In *Live to Tell*, on the other hand, Madonna’s voice is less high and tense, lower, more relaxed, more sure of herself, as well as softer, breathier, and warmer. Only later in the song does an occasional vibrato and increased tension add the strength of relived emotion to some of the key moments. In short, as Chion has said of the actor’s voice in the cinema:

The voice is ceasing to be identified with a specific face. It appears much less stable, identified, hence fetishizable. This general realization that the voice is radically other than the body that adopts it (or that it adopts) for the duration of a film

seems to me to be one of the most significant phenomena in the recent development of the cinema, television and audiovisual media in general. (1999, 174)

Voices now begin to be understood, not on the basis of the meaning potential of individual vocal features, but on the basis of set configurations of values of the parameters we discussed in the previous section, not on the basis of bodily experience, but on the basis of familiarity with the movie roles and songs in which they have been used and the public profile of the actors and singers who developed and used them.

Before we move on to discussing the technological voice, two further points should be made. First of all, what we have said of the voice can also be applied to musical instruments. Instrumentalists, like actors and singers, also developed characteristic individual sounds on their instruments. Just as voices can be soft, smooth and well oiled, or rough, (p. 481) raspy, and cracked, so saxophones or trumpets can be soft and mellow, or tense and strident, sound like a hoarse whisper or a foghorn in the mist—and both now include a large repertoire of howls, wails, groans, and other vocalizations that can be captured neither by alphabetic writing nor by musical notation. Experience here includes not only the parameters we have discussed in the previous section but also our experience of tools and materials and of the actions that produce the sounds—whether a gentle caress or a vigorous attack. Schafer has given an excellent example of the broader significance of such physical modes of sound production and their relation to everyday physical action:

The substitution of the punched string piano for the plucked harpsichord typifies the greater aggressiveness of a time in which objects were punched and beaten into existence by means of industrial processes where once they had been stroked, carved and kneaded into shape.

(Schafer 1977, 109)

But here too, genre-specific configurations of values of the articulatory parameters exist, and here too, the individual sounds of instrumentalists such as Miles Davis, Dizzy Gillespie, John Coltrane, and Stan Getz have become part of the toolkit of many musicians.

Secondly, everything we have said of the voice can also be applied to sound effects and ambient sounds. Their meaning not only derives from their source, from what they are the sound *of*, but also from their qualities. The closing of a door, too, can sound loud or soft, tense or lax, rough or smooth, and so on. In the early days of the sound film, the Hungarian film theoretician Béla Balász foresaw a language of sound along those lines:

It is the business of the sound film to reveal for us our acoustic environment, the acoustic landscape in which we live, the speech of things and the intimate whisperings of nature ... the meaning of a floorboard creaking in a deserted room, a bullet whistling past our ear, the deathwatch beetle ticking in old furniture and the first spring tinkling over the stones. (1970, 179–180)

It is only now, in the age of sound design, with its blended sound effects and tracks that blur the distinction between music and ambient sound, that this vision is beginning to be realized.

The Technologically Enhanced Voice

The development of the new semiotics of the voice (and of music and other sounds) was, as we have seen, facilitated by the global reach of the technological media of radio, TV, and cinema—and, today, of the Internet. But this did not leave the material voice unchanged. We have already discussed the materiality of the voice, but the omnipresent loudspeaker also has its own materiality, which impacts, for instance, on the directionality of sound. True, when we listen to a loudspeaker reproducing voices and other sounds, we do not hear a diaphragm moving along the trajectory of an electronic waveform, we (p. 482) hear instead the materiality of the reproduced sound. But loudspeakers, while recreating sounds accurately, create their own directional patterns of acoustic reflections from floors, walls, ceilings, or foliage, and these patterns are usually quite different from those of human voices or musical instruments. It is this transformation that, in addition to the common acousmatic dislocation, explains why an audience can generally discern between a voice or an instrument reproduced by a loudspeaker or produced by a human in the acoustic space. The sound of the reproduced voice, as it reaches our ears, does not remain entirely unchanged, the intimacy of a voice recorded up close with a microphone remains intimate, even though the reproducing loudspeakers might be meters away adjusting the meaning potential of social distances (cf., Mulder 2014, 2015).

Beyond that, audio technology can *enhance* what the body can do with respect to almost all the parameters we have discussed in the preceding sections and can create new sounds, sounds the human vocal apparatus cannot produce unaided, as we will discuss in what follows.

Enhancing Pitch and Pitch Range

Contemporary digital audio technology includes pitch shifters that can change the pitch range, shifting the voice up or down up to an octave, or extending or decreasing its range, and modern digital signal processors can create an instant “robot effect,” whatever the source input. Digital pitch correction technology not only can change pitch but also can be used to tune speech into song. The expressive potential of digital pitch manipulation and pitch shifting was explored in the Cher effect (Frere-Jones 2008; Lacasse 2000) and in Laurie Anderson’s use of vocoders (and harmonizers) “to give herself a ‘masculine’ pitch and thus subverting the ‘body-in-the-voice’” (Frith 1996, 194). In the Flemish stage director Dirk Thange’s reimagination of Schitzler’s classic play *Reigen* (*La Ronde*, 1897) the five male and five female characters were played by one actor and one actress who differentiated their five characters by using exaggerated pitch ranges enhanced by a real-time pitch shifter—a grave, deep voice for the countess and the baron, squeaky high-pitched voices for the young man and the servant, and so on.

Traditional technology has also been used to enhance pitch range. The low, booming voice of the actor Don Lafontaine,⁴ who recorded the voice-overs for more than five thousand movie trailers, was usually enhanced by the “proximity effect” of the (very) close microphone.⁵

Enhancing Loudness

The manipulation of loudness did not have to wait for digital technology and played a crucial role in the development of the new semiotics of the voice. When a voice is recorded, a whisper remains a whisper and a scream stays a scream, no longer identified by the perceived loudness of its reproduction but by additional voice quality parameters. In other words, the loudness parameter is uncoupled from the social distance and power (p. 483) relations that form its meaning potential: a whisper intended for one person only, conveying a message of a highly personal nature can be reproduced in a movie theater at a screaming, high sound level, and the rock singer’s yelling voice can be heard at a comfortable level on one’s headphones. Social distance is no longer locked on to physical distance. It is therefore freed up to express *imagined* social distance, *imagined* relations of intimacy with distant people we will never meet—or alienating distance from the kind of voices that do actually form part of our everyday environment. As voiced by Frith:

The microphone made it possible for singers to make musical sounds—soft sounds, close sounds—that had not really been heard before in terms of public performance (just as the film close up allowed one to see the bodies and faces of strangers in ways one would normally only see loved ones). The microphone allowed us to hear people in ways that normally implied intimacy—the whisper, the caress, the murmur.(1996, 187)

The compression of loudness, a common element of sound in media, does not simply reduce loudness; it brings down a sound’s dynamic range by, usually, simultaneously bringing up the level of the quietest sounds. Typically observed in the radio voice, this results in an emphasis of breaths, swallowing, gulping, salivating, of vocal collateral, as well as a certain hyperactive pronunciation (aiding intelligibility when done appropriately). On the other hand, such dynamic compacting specifically in conjunction with audio editing can overshoot. Frances Dyson, in a chapter discussing the work of the vocal artist Amanda Stewart, underlines the consequence, for example in an edited radio segment, of cutting out and reducing such vocal collateral, the breath in particular:

What is not uttered is multi-present with what is uttered. In the complexity of the human body, sound, meaning, and emotion are actually synthesized and latent. And in fact the breath is one of the openings for that latency to come out, and that’s what gets sealed off more and more with compressed/media sound.(Stewart, quoted in Dyson 2014, 104)

Noise: Enhancing Tension, Roughness, and Breathiness

The German, French, and Dutch languages have a number of different words referring to different types and applications of noise. Additionally, to the (sound) engineering perspective where noise is something that needs to be avoided, if not eliminated (see Wittje 2016), in English, usage of the word “noise” can be more context specific; Torben Sangild considers:

Etymologically, the term “noise” in different Western languages (støj, bruit, Geräusch, larm, etc.) refers to states of aggression, alarm and tension and to powerful sound phenomena in nature such as storm, thunder and the roaring sea. It is worth noting in particular that the word “noise” comes from Greek nausea, referring not only to (p. 484) the roaring sea, but also to seasickness, and that the German Geräusch is derived from rauschen (the sough of the wind), related to Rausch (ecstasy, intoxication), thus pointing towards some of the aesthetic, bodily effects of noise in music.

(2002)

Noise is present in many elements of sound generation (vocal or instrumental) and in strategies to synthesize (familiar or unfamiliar) sounds. Before reaching a steady state, the spectrum of, for instance, pitched instruments displays many random elements. This attack phase of a sound’s development over time (essential for recognizing different timbres) can be understood as the chaotic result of the tension in a vocal cord, a string, or a membrane being triggered by displacement. A whispered sentence can be experienced as a faltering of the voicing, never making it beyond the process of generation and never failing to alert the listener to the required tension.

The noise element is emphasized in synthesized breathy sounds; these are for instance, easily perceivable in the Shakuhachi and other flutes. Although some flutes are known for their serene, almost plain spectra, the tension displayed in the breathy element of flute sounds is often called on to create suspense.

Breathiness can be enhanced by using filtering and reverb and/or added hiss (another type of noise), suggesting the proximity of a whisper. In a “suspense effects” album produced by the BBC Radiophonic Workshop, a soft sighing breath that almost touches listeners with the overwhelming closeness of its presence happens in a track titled *Spectres in the Wind* (Cain 1976). From the original press release for that record, by Richard Robson:

the versatility of the synthesised effects is such that they can often be interpreted in accordance with what the audience imagines is happening or by what action is taking place.

Roughness can be created, or perhaps we should say suggested, approximated, by using distortion, whether digital or otherwise, and whether live or in recording. The singer David Eugene Edwards of the band 16 Horse Power often alternates, during live perfor-

mances, between a “clean” and a “dirty” microphone, making distortion an expressive choice rather than an overall quality of the amplification, and the Bon Iver singer Justin Vernon also often uses a two-microphone setup, one used for amplification as usual, the other feeding a voice processor creating an autotune or related effect from his voice. The expressive power of roughness can also be found in the distorted electric guitar and its history of broken loudspeakers, which made a technologically enhanced timbre a natural player in pop music and beyond. Nor is distortion limited to electric guitars, as can be heard for instance in the distorted harpsichord on the song *Blood Roses* by Tori Amos (1996).

Enhancing Vibrato

Vibrato is created by often somewhat irregular pitch fluctuations in the human voice or, in musical instruments, by pitch fluctuations brought about by finger or lip gestures.

(p. 485) It became particularly popular when recordings began, perhaps to allow listeners to imagine the “aura” of live performances. Tremolo fluctuations may be faster and/or more regular than the human voice can produce—and this regularity somewhat diminishes its emotionality. Tremolo has long been used in attempts to make musical instruments sound like the human voice, for instance the “vox humana” stops on church organs (cf., Van Leeuwen 2009), or to suggest sustain on instruments that do not have the ability to sustain, for instance the mandolin (and also the piano). Sustain is then simulated by fast repetition of the same note and fluctuation by trills. All this can of course also be digitally programmed. Somewhat similar is the chorus effect, in which two closely related but slightly detuned timbres are overlaid, creating a shimmering quality.

Enhancing Articulation and Acoustics

Digital technology can also manipulate the formant structure of the sound, creating sounds that approximate the articulatory settings of vowels such as [i] or [a] and hence the meaning potentials of aperture and closure and of frontality and back-in-the-mouthness. On contemporary digital keyboards and synthesizers, such vowel sounds (“humanizers”) can be added to musical sounds. On Roland digital keyboards, for instance, a preset called *Aah Choir* adds an [a] sound to an organ-like sound, while the *Ooh choir* adds an [u] and a consonantal attack to a thin, reedy instrumental sounds, and the *Jazz Scat* uses key pressure to create different vowel sounds. Such vowel colors can also be added to atmospheric sounds and sound effects. In other words, speech-like articulatory settings can now be blended with instrumental sounds and sound effects, thoroughly blurring the boundaries between speech, music, and other sounds and beginning to realize a vision that had already been conceived of in the age of Descartes (Van Leeuwen 2009). Even before digitization, some instrumentalists had in fact already tried to achieve this kind of effect, for instance Rahsaan Roland Kirk in his flute playing.

Acoustic parameters can also be enhanced, for example through artificial reverberation, which adds new cues to a represented acoustic environment, so that listeners can imagine themselves in a cathedral, even though they might be sitting in a car. In performance

arts' venues, artificial reverberation can be produced by large numbers of little loudspeakers, hidden (or visible) in walls and ceiling surrounding the audience. A dry theater room can be enhanced to sound like, or offer the experience of, a concert hall, opening up the economics of multipurpose entertainment venues (essentially allowing cities to build a theater, concert hall, and opera for the price of one). Such economic benefits often trigger debates of sonic aesthetics; however, more interesting in the context of this chapter is that such technologies, rather than recreating existing familiar or stereotypical acoustic environments, can be used to create imaginary, novel acoustic experiences. In a venue where acoustics are conditioned by loudspeakers and digital signal processing, the audio parameters that inform spatiality can be changed instantaneously, suggesting specific acoustic experiences such as a cathedral, a theater, or a (p. 486) cinema but, more interestingly, everything in between or beyond those outliers. Blesser and Salter (2006) refer to agents working with that technology as aural architects, guided, and limited, by the agent's imagination. Imagine the subtle difference between the acoustics of a pine-tree or a gum-tree forest, or the atmosphere of a magical forest.

All these will still be understood on the basis of our experience of vocal production and the acoustics of spaces. At the same time, we may be aware of subtle differences, of the all-too clean regularity of digital production that is lacking in human speech and music performance. Whereas, unaided, it is, for example, practically impossible for human beings to achieve perfect rhythmic regularity, for computers it is very difficult to even approximate the subtle irregularities of human activities and those of nature. This contrast can, of course, be used for its potential of contrasting the "human" and the "technological." (See Danielsen, volume 2, chapter 29.)

Recreating Technological Experience

Technology can also produce extraexperiential sounds, sounds that fall outside our experience. The disembodied voice of the radio and the telephone was once such an experience, and it took time to learn not to shout through the telephone and, as historians have documented, not to project one's voice on radio as though addressing a crowd (Cardiff 1986), in short, to become familiar with the particular ways in which such technologies reproduce voices.

With the ubiquity of camera smartphones, many people now have everyday experience of electroacoustic voice recording and playback. Both in technological and experiential terms, the relation between audio and video, very well developed and researched since the inception of sound film, is being redefined by this development. The smartphone's microphone is always where the camera is and never as close to the sound source as we have grown accustomed to from almost a century of mediatized (vocal) sound and social distance. Consequentially, clips shot in high definition video are accompanied by the most basic of sound recording with the auditory perspective (Van Leeuwen 1999, 12), for instance, figure, ground, and field, rendered to a single dimension.

Meanwhile, digital technology has made it possible to simulate older technologies such as lo-fi radio, record players, and the telephone. This makes them available as semiotic choices, in addition to the audible digital artifacts of GSM telephony and Voice Over Internet applications such as Skype (including the now rare rhythmic pulse inference of older GSM phones on analog audio systems). But the predigital media no longer form part of any people's everyday experience and are therefore likely to either become arbitrary signifiers of "period" authenticity, or aesthetic effects used for the meaning potential of their acoustic qualities. This has been the case, for instance, with the vocoder, once an instrument to reduce the bandwidth of telephone signals for economic reasons, but already from as early as the late 1940s used for musical purposes (Tompkins 2010).

(p. 487) **Extraexperiential Sound**

There are also sounds that relate neither to our bodily experience nor to everyday technological experience. Their meaning potential is therefore, at least initially, the "nonhuman," whether, in context, this becomes the "divine" or the "alien" (as in "alien beings")—in either case they will evoke menace or mystery, awe or dread, as can be seen in this review in *The Listener* of a 1959 radio play that pioneered the use of electronic sound:

The three-note phrase was something like the slow swell of the sea but also like a groaning dirge from some inhuman voice ... and as affective as an audible equivalent to nightmare.

(quoted in Niebur 2010, 67)

The "nonhuman" sound of course existed before digital sound, for example in sounds that sustained beyond the time where humans can hold their breath. Such sounds were produced by sacred instruments such as the church organ, or the didgeridoo, and they were the staple fare of music that had to express "eternity" or "nature." They often use pitch fluctuations, which human voices are unable to produce—rapid warblings, metallic shimmerings, and so on, to signify the landscapes of outer space or the inhumanity of technology, as on a track called *Alien Tekno*, from the Russian group Cutworks (2012), where many of these "atmospheric" sounds are overlaid on a bass and drum rhythm: shimmering zings, rapidly vibrating electronic whistles, low rumbles, slow organ tremolos, and so on.

Another use of "nonhuman sound" has been in confronting the human and the "nonhuman." The Cher autotune effect produces sounds that cannot naturally be produced by human beings and stretch the limits of our experience, like a character transforming into a werewolf. Yet it also becomes a new medium of expression, perhaps one that is necessary in a world where technology has come to play such a central role in almost all human activities.

"Nonhuman" sounds have also been used to synthesize voices for robots. Robots, such as R2D2, and more recently Wall-E, might be deprived of vocal parameters that are used to

express human emotion, such as variations in pitch and loudness, and use synthesized voices, such as R2D2's whistles, or mechanical-sounding voices, such as that of the protagonist of an animation movie called *The Iron Giant* (Bird 1999), an enormous metal robot who sounds metallic, his size underlined by mechanical clanking, screeching, and scraping.

But “nonhuman” sounds can also be rehumanized, returned to the realm of culture. As Niebur (2010) has documented in a history of the BBC Radiophonic Workshop, today's scores for electronic music often combine rudimentary musical scores with words evoking human experiences—words such as “squeak,” “scream,” “growl,” “buzzy,” “reedy,” “warm,” “flutey,” “swirly,” “grainy,” “clashing,” “clanking,” “throbbing,” “banging,” “rumbling,” “splashing,” “whooshing,” “ticking,” “clicking,” or references to musical instruments such as “a gong-like sound,” “almost like a bowed vibraphone”—and, over the years, (p. 488) “extraexperiential” sounds have been increasingly incorporated in familiar musical structures, just as keyboards were reintegrated with the knobs and dials of technological interfaces. However, the interfaces of contemporary digital audio technology, unlike the scores of the BBC Radiophonic Workshop sound designers, are structured along purely technical parameters. Innovations in semiotic technology are not always driven by semiotic and social needs, and they may therefore take time to find their role in semiotic practices—or be discarded when they fail to do so.

An example of a successful return to culture is Stephen Hawking's famous voice, produced by a specific, early voice synthesis chip from 1988.⁶ Hawking's artificial voice, exemplary in its monotony and other nonhuman speech attributes, again demonstrates how sounds can increase in meaning potential and as such come to life and lose some of their artificiality. Not only do many people recognize that voice as his, Hawking himself chose to keep using that first, very early rudimentary synthesized voice, even though many improvements have been made to voice synthesis since, potentially making his voice more realistic. The makers of a recent film about Hawking were elated when he offered the use of one of his own (spare) original synthesis units.⁷

Conclusion

By way of conclusion, we will discuss some effects of the developments we have outlined in this chapter and then briefly return to the theme of imagination.

First of all, there is a blurring of the boundary between the human and the technological voice and between traditional and digital musical instruments. The vocoder is a good example of this—a human voice or a musical instrument becomes input for a synthesizer that then reconfigures its resonances, resulting in a voice with a strange, extraexperiential metallic ring. The American singer, instrumentalist, and “vocoderist” Casey Benjamin brilliantly uses it to synchronize and “vocoderize” his singing and his playing of a keytar.

Another boundary that is blurred is one between agency and causality, highlighted in the question: who or what produces the sound I hear? Amanda Stewart uses two microphones for her performances, one connected to loudspeakers on the left and one to loudspeakers on the right of the stage:

Using two microphones gave the voice a rich, spatialized plasticity and ambiguity that hadn't been present in my work before. The voice was materialized as a shifting, sculptural entity.

(Neumark et al. 2009, 178–179)

Like the example of vocalists moving between a “dirty” and a “clean” microphone described earlier, this approach gives the performer a level of agency that is commonly left to a producer or a sound (mix) engineer, hidden from view. The performative gesture of switching between microphones and the resulting change in sound brings the technological enhancement in the open and its essence to the fore.

(p. 489) There is also a blurring between vocal and instrumental sound. The talk box, for instance, well-known from recordings by Peter Dinklage and others, has the sound from a guitar or keyboard relayed through a tube into a player's mouth, with the microphone picking up the sound as it is manipulated by the movements of the mouth. As already discussed, we can now “sing” a musical instrument—or “play” a human voice on contemporary keyboards that can add vowel color to instrumental sounds, combined, perhaps, with hiss, suggesting breath, and vibrato, to simulate further qualities of the human voice. By contrast, “space voices” might then be somewhat “nonhuman,” despite also having vowel color, through reverb and “nonhuman,” fast and shallow vibration, reminiscent of the menacing tremolos in horror and science fiction films.

Finally, there has been a blurring of the boundaries between music and sound effects, both in “product enhancement” sounds such as the Microsoft Windows start up tune, where what once would have been a mechanical sound caused by pressing a button or flicking a switch, has now become an ethereal electromusical theme, and in film and television soundtracks, where half-musical, half-electronic sounds might be used as sound effects or background music. Niebur describes how, in a 1959 BBC science fiction TV series, *Quatermass and the Pit*, a mechanic is working inside a ship when a wrench starts moving by itself, accompanied by an electronic sting. As he tries to outrun this psychic attack,

Rapidly moving flute lines connect crunchily dissonant bass chords, and underneath this both rumbling tape feedback loops and bent sine wave attacks follow the worker. Here it is impossible to separate the diegetic from the non-diegetic, the acousmatic from the atmospheric, even if the radiophonic noises are meant to be heard as sound effects. Their commingling with the brass and the flute make such a unified “music” that the result sounds like a transitional evocation of both suspense and “science fiction.” (2010, 78)

In all this, the question of authenticity has often been raised; for instance, in connection with pitch correction. The implicit criterion is that technologically mediated sound must, in some way, be faithful to some original that has not been technologically manipulated. And it is true that synthesized voices, musical instruments, and other sounds still often strive to be “natural.” In our view, speech, music and sound have always been entwined with technology, as can be seen from our earlier example of acoustic design and theatrical masks and from the history of musical instruments (Théberge 1997) whose qualities and affordances have always changed in accordance with the needs and interests of the societies that developed and used them. Today’s audio technology is no exception. As our exploration in this chapter has shown, we hope, it has created a wealth of semiotic resources whose meaning potential is only just beginning to be realized. If social semiotics can make a modest contribution to the unlocking of these resources, it is because it can offer insights in matters of agency and can disambiguate between what is afforded and what is enforced by technological restriction, context, or tradition. Semiotic analysis, in this approach, therefore, not only allows for theoretical reflection, but can, and should, also inform current and future sound practice.

(p. 490) To finally return to the theme of imagination, we have seen that acousmatic sound has inspired the imagination since time immemorial—from biblical stories in which God is a voice rather than a visible presence, to contemporary films in which events that are heard but not seen are strongly and vividly imagined. We have seen that sound’s uncoupling from the lived face-to-face experience can make us imagine the spaces suggested by the acoustic qualities of music and the intimacies and emotions that come with the physical closeness of whispered voices and tender melodies. And we have seen that technology not only enhances these experiences, but even make us imagine experiences that have no counterpart in real life. In all these instances, sound affects us more deeply than the silent image—a voice choking with emotion will move us more than a silent tear; the immersive reverberations of a cathedral awe us more, and the violent sounds of a thunderstorm stir us more than mere pictures. Sound asks us to surrender ourselves, to immerse ourselves in imagined worlds and to place these at the heart of sensation and existence. To quote Bresson once more, sound “goes to the within” (1997, 61).

References

Alström, T. 2004. The Voice in the Mask. *TDR (1988-)* 48 (2): 133–138. doi: 10.2307/4488557.

Amos, T. 1996. Blood Roses. *Boys for Pele*. Atlantic Records.

Bacall, L. 1978. *Lauren Bacall by Myself*. London: Jonathan Cape.

(p. 491) Balázs, B. 1970. *Theory of the Film: Character and Growth of a New Art*. Translated by E. Bone. New York: Dover Publications.

Barthes, R. 1973. *Mythologies*. London: Paladin.

Barthes, R. 1977. *Image-Music-Text*. London: Fontana.

Bird, B. 1999. *The Iron Giant*. Warner Bros.

Blanc, M., and P. Bashe. 1989. *That's Not All Folks*. New York: Warner Books.

Blessner, B., and L.-R. Salter. 2006. *Spaces Speak, Are You Listening? Experiencing Aural Architecture*. Cambridge, MA: MIT Press.

Bresson, R. 1997. *Notes on the Cinematographer*. Los Angeles: Green Integer.

Cain, D. 1976. Spectres in the Wind. In *Out of This World: Suspense and the Supernatural*, edited by BBC Radiophonic Workshop. London: BBC Records.

Cardiff, D. 1986. The Serious and the Popular: Aspects of the Evolution of Style in the Radio Talk 1928–1939. In *Media, Culture and Society: A Critical Reader*, edited by R. Collins, J. Curran, N. Garnham, P. Scannell, P. Schlesinger, and C. Sparks, 228–246. London: Sage.

Chion, M. 1994. *Audio-Vision: Sound on Screen*. Translated by C. Gorbman. New York: Columbia University Press.

Chion, M. 1999. *The Voice in Cinema*. Translated by C. Gorbman. New York: Columbia University Press.

Cutworks. 2012. Alien Tekno. In *Sunchase vs Cutworks*. Germany: Drone Audio.

Dyson, F. 2014. *The Tone of Our Times: Sound, Sense, Economy, and Ecology*. Cambridge, MA: MIT Press.

Frere-Jones, S. 2008. *The Gerbil's Revenge: Auto-Tune Corrects a Singer's Pitch. It Also Distorts—A Grand Tradition in Pop*. The New Yorker. Accessed June 9, 2008.

Frith, S. 1996. *Performing Rites*. Oxford: Oxford University Press.

Goehr, L. 1992. *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music*. Oxford: Clarendon Press.

Hall, E. T. 1966. *The Hidden Dimension*. Garden City, NY: Doubleday.

Herman, L. H. 1952. *A Practical Manual of Screen Playwriting for Theatre and Television Films*. Cleveland and New York: World Publishing.

Hildum, D. C., and R. W. Brown. 1956. Verbal Reinforcement and Interviewer Bias. *Journal of Abnormal and Social Psychology* 53 (1): 108.

Kress, G. 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. London: Routledge.

Lacasse, S. 2000. Listen to My Voice: The Evocative Power of Vocal Staging in Recorded Rock Music and Other Forms of Vocal Expression. PhD thesis, University of Liverpool.

Lakoff, G., and M. Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.

Laver, J. 1980. *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.

Lomax, A. 1968. *Folk Song Style and Culture*. New Brunswick NJ: Transaction Books.

McCracken, A. 1999. God's Gift to Us Girls: Crooning, Gender, and the Re-Creation of American Popular Song, 1928–1933. *American Music* 17 (4): 365–395.

Mulder, J. 2013. Making Things Louder: Amplified Music and Multimodality. PhD thesis, FASS, University of Technology Sydney.

Mulder, J. 2014. The Iconic Microphone. *Academic Quarter* 10 (Spring 2015): 97–106.

Mulder, J. 2015. Social Distance and the Arena Concert. In *The Arena Concert: Music, Media and Mass Entertainment*, edited by K. Fairclough-Isaacs, B. Halligan, R. Edgar, and N. Spelman, 269–278. London: Bloomsbury Academic.

(p. 492) Neumark, N., R. Gibson, and T. Van Leeuwen. 2009. *Voice: Vocal Aesthetics in Digital Arts and Media*. Cambridge, MA: MIT Press.

Niebur, L. 2010. *Special Sound: The Creation and Legacy of the BBC Radiophonic Workshop*. Oxford: Oxford University Press.

Potter, J., and N. Sorrell. 2012. *A History of Singing*. Cambridge: Cambridge University Press.

Ryfe, D. M. 1999. Franklin Roosevelt and the Fireside Chats. *Journal of Communication* 49 (4): 80–103. doi:10.1111/j.1460-2466.1999.tb02818.x.

Sangild, T. 2002. The Aesthetics of Noise. *Datanom*. <http://www.ubu.com/papers/noise.html>. Accessed January 30, 2016.

Schafer, R. M. 1969. *The New Soundscape: A Handbook for the Modern Music Teacher*. Ontario: Berandol Music.

Schafer, R. M. 1977. *The Tuning of the World*. New York and Toronto: Knopf.

Théberge, P. 1997. *Any Sound You Can Imagine: Making Music/Consuming Technology*. Hanover, NH: Wesleyan University Press.

Tompkins, Dave. 2010. *How to Wreck a Nice Beach: The Vocoder from World War II to Hip-Hop, The Machine Speaks*. Brooklyn, NY: Melville House.

Van Leeuwen, T. 1999. *Speech, Music, Sound*. Houndmills, Basingstoke, and Hampshire: Macmillan.

Van Leeuwen, T. 2005. *Introducing Social Semiotics*. New York: Routledge.

Van Leeuwen, T. 2009. Vox Humana: The Instrumental Representation of the Human Voice. In *Voice—Vocal Aesthetics in Digital Arts and Media*, edited by N. Neumark, R. Gibson, and T. Van Leeuwen, 5–15. Cambridge, MA: MIT Press.

Weidman, A. J. 2006. *Singing the Classical, Voicing the Modern: The Postcolonial Politics of Music in South India*. Durham, NC, and London: Duke University Press.

Wijffes, H. 2014. Spellbinding and Crooning: Sound Amplification, Radio, and Political Rhetoric in International Comparative Perspective, 1900–1945. *Technology and Culture* 55 (1): 148–185.

Williams-Jones, P. 1975. Afro-American Gospel Music: A Crystallization of the Black Aesthetic. *Ethnomusicology* 19 (3): 373–385.

Wittje, R. 2016. Concepts and Significance of Noise in Acoustics: Before and after the Great War. *Perspectives on Science* 24 (1): 7–28.

Notes:

(1.) Neil Sorrell (Potter and Sorrell 2012, 181) points out that the use of amplification with microphones, amplifiers, and loudspeakers is very common in Indian classical music.

(2.) Acapela “vocalize[s] your content with authentic and original voices that express meaning and intent.” On the web <http://www.acapela-group.com>. Accessed January 30, 2016.

(3.) Although in popular uses the terms can be confused, from a technical point of view, vibrato is a continuum of pitch fluctuations, whereas tremolo is the iteration or alternation of notes alternatively the—abrupt or continuous—alternation of amplitude. However, in singing, tremolo is often used to indicate an excessive vibrato. Furthermore, vibrato is often used erroneously as a synonym for tremolo, for instance on some guitar amplifiers and electric pianos.

(4.) This common use of the close mike voice was pointed out to us by the first batch of students enrolled in the Speech, Music, Sound subject that runs at the University of Technology, Sydney, where both authors taught.

(5.) When directional microphones are used short distances from the sound source, the low frequencies are emphasized. This is known as the proximity effect, and can be applied to make a voice sound warmer or give it more gravitas.

(6.) Hawking gives a description of his voice system on his website <http://www.hawking.org.uk/the-computer.html>. Accessed January 30, 2016.

(7.) Related in a blog entry by R. Hiscott: "The Theory of Everything": Stephen Hawking's Story on the Big Screen. *American Academy of Neurology*, Last Modified November 6, 2014, <http://journals.lww.com/neurologynow/blog/breakingnews/pages/post.aspx?PostID=213>. Accessed January 30, 2016.

Johannes Mulder

Lecturer in Sound, Murdoch University

Theo van Leeuwen

Professor of Language and Communication, University of Southern Denmark