

A note from the author:

On page 136, when describing Nina Sun Eidsheim's discussion of racism in operatic listening, I quote a phrase from an article by Eidsheim in which she describes this tendency as "acousmatic blackness," or "the perceived presence of the black body in a voice that otherwise meets all the standards of a professional classical voice." I failed to acknowledge that the term "acousmatic blackness," as cited by Eidsheim, was originally coined by artist Mendi Obadike to refer to the perception of "blackness" in sound even in the absence of its source: Mendi Obadike, "Low Fidelity: Stereotyped Blackness in the Field of Sound" (Ph.D dissertation, Duke University, 2005), 135-77. I apologize for this oversight, which regrettably bears out my own argument about racial erasure in the scene of listening.

I share below a screen grab from a Facebook post by Mendi and Keith Obadike dated October 28, 2022, in which Keith Obadike elaborates on Mendi Obadike's use of the term. The text reads:

"Acousmatic sound is a sound for which the source is not visualized. In the late 1990s, Mendi Obadike coined the term **"acousmatic blackness"** to describe the reading of a sound as 'black' when the source of the sound is not visible. Think of hearing off-screen music in a film scene, or listening to a voice over a public address system. During this period we began speaking about the pleasures and pains of 'acousmatic blackness' as we watched films, listened to transcendent hip-hop tracks, made sound art, and created artworks for the internet. Mendi explored this notion of reading race through sound when it is separated from the image in her dissertation *Low Fidelity: Stereotyped Blackness in the Field of Sound* (2005). This term, 'acousmatic blackness', has since been taken up by musicologists, cultural critics, film scholars, artists, and many others. The concept (of a sound with representational possibilities and an incorporeal source) has been an important area of investigation in our work from the beginning of our practice. — K.O."

Acousmatic sound is a sound for which the source is not visualized. In the late 1990s, Mendi Obadike coined the term **"acousmatic blackness"** to describe the reading of a sound as "black" when the source of the sound is not visible. Think of hearing off-screen music in a film scene, or listening to a voice over a public address system. During this period we began speaking about the pleasures and pains of "acousmatic blackness" as we watched films, listened to transcendent hip-hop tracks, made sound art, and created artworks for the internet. Mendi explored this notion of reading race through sound when it is separated from the image in her dissertation *Low Fidelity: Stereotyped Blackness in the Field of Sound* (2005). This term, **"acousmatic blackness"**, has since been taken up by musicologists, cultural critics, film scholars, artists, and many others. The concept (of a sound with representational possibilities and an incorporeal source) has been an important area of investigation in our work from the beginning of our practice. — K.O.

For information about Mendi + Keith Obadike's works see their website blacksoundart.com.

SOUND OBJECTS

EDITORS

James A. Steintrager and Rey Chow

DUKE UNIVERSITY PRESS

Durham and London

2019

© 2019 Duke University Press
 All rights reserved
 Printed in the United States of America
 on acid-free paper ∞
 Designed by Matthew Tauch
 Typeset in Scala Pro and Scala Sans Pro
 by Westchester Publishing Services

Library of Congress Cataloging-in-Publication Data
 Names: Steintrager, James A., [date] editor. | Chow, Rey, [date] editor.
 Title: Sound objects / James A. Steintrager and Rey Chow, editors.
 Description: Durham : Duke University Press, 2018. | Includes bibliographical references and index.
 Identifiers: LCCN 2018021935 (print)
 LCCN 2018030091 (ebook)
 ISBN 9781478002536 (ebook)
 ISBN 9781478001096 (hardcover : alk. paper)
 ISBN 9781478001454 (pbk. : alk. paper)
 Subjects: LCSH: Sound—Social aspects. | Auditory perception. | Sound (Philosophy)
 Classification: LCC QC225.7 (ebook) | LCC QC225.7 .s68 2018 (print) | DDC 534.01—dc23
 LC record available at <https://lccn.loc.gov/2018021935>

Cover art: Paul Klee, *Fuge in Rot* [*Fugue in Red*], 1921, 69. Watercolor and pencil on paper on cardboard, 24.4 cm × 31.5 cm. Private collection, Switzerland, on extended loan to the Zentrum Paul Klee, Bern.

Contents

VII ACKNOWLEDGMENTS

- 1 Sound Objects: An Introduction
James A. Steintrager, with Rey Chow

I GENEALOGIES

- 23 1 Reflections on the Sound Object and Reduced Listening
Michel Chion
 33 2 Pierre Schaeffer and the (Recorded) Sound Source
John Dack
 53 3 The Fluctuating Sound Object
Brian Kane

II AURAL REIFICATION, SONIC COMMODIFICATION

- 73 4 Listening with Adorno, Again: Nonobjective Objectivity
 and the Possibility of Critique
James A. Steintrager
 94 5 Spectral Objects: On the Fetish Character of Music
 Technologies
Jonathan Sterne

III ACOUSMATIC COMPLICATIONS

- 113 6 Listening after "Acousmaticity": Notes on a
Transdisciplinary Problematic
Rey Chow
- 130 7 The Skin of the Voice: Acousmatic Illusions,
Ventriloquial Listening
Pooja Rangan

IV SOUND OBJECTS AND NONHUMAN RELATIONS

- 151 8 The Acoustic Object: Sound and the Legal Imagination
Veit Erlmann
- 167 9 The Alluring Objecthood of the Heartbeat
Jairo Moreno and Gavin Steingo
- 185 10 On Nonhuman Sound—Sound as Relation
Georgina Born

V MEMORY TRACES

- 211 11 The Sound of Arche-Cinema
John Mowitt
- 228 12 Listening to the Sirens
Michael Bull
- 246 13 Entities Inertias Faint Beings: Drawing as Sounding
David Toop

265 BIBLIOGRAPHY

281 CONTRIBUTORS

285 INDEX

Acknowledgments

This volume is the third installment of an ongoing project and intellectual collaboration that began with a special double issue of *differences: A Journal of Feminist Cultural Studies* that appeared in 2011 and that we titled "The Sense of Sound." We started more with a sense of curiosity and interest than expertise and certainly without realizing that sound studies was rapidly congealing into a field—if, thankfully, not quite a discipline. Several of the contributors to that initial foray return here: Michel Chion, Veit Erlmann, John Mowitt, and Jonathan Sterne. We thank them for sticking with us and continuing in multifarious and creative ways to deepen our explorations of the sonic field. The second installment was James A. Steintrager's translation of and critical introduction to Michel Chion's *Sound: An Acoulogical Treatise*, which appeared in 2016 with Duke University Press. The author's profound engagement with the legacy of Pierre Schaeffer and his notion of the "sound object" helped shape the path we have chosen for this collection.

Along the way, we have benefited enormously from conversations, criticism, and debates, both in and out of academic settings. Special thanks go to Julie Napolin, Dominic Pettman, and Pooja Rangan for organizing and bringing us to the Sonic Shadows symposium at the Eugene Lang College of the New School in the spring of 2015 (JS and RC); to Luis Carcamo-Huechante for the Future of Sound Studies symposium at the National Humanities Center in the spring of 2014 (RC); to Jacqueline Waeber for the Study Day: Voices and Noises workshop at the Franklin Humanities Center, Duke University, held

The Skin of the Voice

Acousmatic Illusions, Ventriloquial Listening

The Voice behind the Face

The turning point of Julie Dash's acclaimed independent short film *Illusions* (1982) takes place in a fictional Hollywood studio's sound booth.¹ Dash's film is shot in black-and-white cel-luloid and edited in the classical film style of 1942, the year in which it is set. The film's statuesque protagonist Mignon Dupree, the studio's sole female producer's assistant—a light-skinned African American woman who is passing for white—is working overtime to oversee the postproduction of a film. The sound technicians inform Mignon and her boss that they have a problem: the sound operator lost the sync while filming one of the musical numbers, a jazzy love song, and as a result the picture and music tracks do not match. Upon viewing the rushes, Mignon remarks, "Leila Grant [the star] looks like she's chewing marbles while somebody else sings." Since the usual solution, to reshoot the scene with Grant following the song, is not an option, the senior sound engineer devises an unorthodox solution: he asks Esther Jeter, the young black female singer who provided the original backup vocals, to return to the studio and sing along with the actress's moving lips.

When the recording begins, all eyes, including Esther's—she has been instructed to "watch the screen"—are locked on the actress on-screen as she sashays around a bedroom in a satin gown and feather boa, mouthing the words to the song that Esther sings (Ella Fitzgerald's *Starlit Hour*, in another act of doubling/dubbing). The disparate acts joined together to create the illusion of screen unity, in both Dash's film and the film within Dash's film, are captured in two symmetrical frontal medium shots of Esther and Leila Grant. Both women move their lips in tune with Fitzgerald's song, and while Esther's face is at first knit in concentration, she gradually begins to mimic Grant's expressive facial gestures. By the end of the scene, Esther no longer seems to be looking directly at the screen. Captured in profile, she sings with ecstatic abandon, having fully inhabited the imaginary role of star.

Mignon is not looking at the screen, either. When the camera pans across the inhabitants of the sound booth shortly before the end of the recording session, the two white sound men are captivated by Grant's image on the screen, but Mignon looks the other way at Esther, her brow furrowed in an inscrutable expression. We only learn later, during an exchange between Mignon and Esther, who has recognized Mignon's ethnic heritage, that this moment "behind the scenes" has been a moment of unmasking for Mignon. Esther tells Mignon that when she goes to the cinema and hears her own voice coming out of an actress's mouth, she shuts her eyes and imagines that she is on the screen in a satin gown. Mignon recognizes the pathos in Esther's cheerful comment: the act of witnessing the seamless concealment of Esther's voice by Leila Grant's face reminds Mignon of how her own racial concealment perpetuates the veiling of black women's labor by the cinematic apparatus. Mignon's disillusionment paves the way, as Judylyn Ryan notes, to a counter-hegemonic manifesto, "when Mignon, ventriloquizing Dash, proclaims her new determination to 'use the power of the motion picture.'"² Her goal? To work against the Hollywood system to tell "real" stories—to "represent real characters that people can identify with."

Although *Illusions* is a fiction film, its aims are anti-illusory. Dash un-masks the façade of Hollywood-style fiction, showing that the romantic ideals of whiteness that unfold on-screen are held in place by the invisible realities of black labor. Intriguingly, image and sound play the roles of villain and protagonist in the story of Mignon's enlightenment, which invites the reading that Hollywood's subordination of the sonic to the visual lies

at the heart of its injustices. Such, for instance, is the reading of *Illusions* offered by the feminist film scholar Patricia Mellencamp, who likens the concealment of Hollywood's racialized politics of labor to the camouflaging of sound in the process of synchronization. She argues, "The seamless union between image and sound, face and voice, with voice subservient to face (as black women are in film to white women) . . . paradoxically seals the dominance of face, of the visible."³ In her disavowal of Hollywood-style fiction to pursue "real" stories about "real" people, Mignon is an avatar of Dash, who has devoted much of her career to producing historical dramas, biopics, and documentaries about the African American experience. Dash's creative trajectory is understandable, since the documentary film genre seeks not only to represent real social actors but also to reverse the audiovisual hierarchy of fiction film, foregrounding the spoken word over the image and, by association, reality over illusion.

The reality principle and democratic impulse associated with liberating minoritized voices from veiling or distortion are a central concern of this chapter. Using the vocal conventions of documentary as an illustrative example, I suggest that the realist pursuit of vocal equality, which restricts the relationship between veiling and power to a visual register, also restricts our understanding of the complex ways in which veiling and unveiling operate in relation to sound and listening. To specify the nature of these auditory illusions, I introduce a term that emphasizes the racialized and gendered perceptual frames that mediate the production and reception of vocal sounds: "the skin of the voice." Using this term, which I evolve in conversation with recent feminist and postcolonialist scholarship on the political economies of sound, I examine how cinematic applications of "acousmatic listening" free from the distortions of vision, as advocated by the composer Pierre Schaeffer, can end up concealing the invisible logics that separate voices into those that are forcibly embodied as an objectified surface and those that are disembodied as a protective disguise. I propose that, to confront and deflect these discriminatory perceptual habits, it is necessary to cultivate an attunement to the seam between the embodied origins of voices and the illusory, surrogate bodies that voices conjure into existence. To this end, I offer a reading of an experimental documentary film in which the artist Mounira Al Solh develops an audiovisual idiom for expressing the ventriloquial basis of vocal sounds as well as the political significance of a *ventriloquial* mode of listening.

Illusions invites the feminist reading that Mignon must *look away* from the illusion of Leila Grant's face to *hear* Esther's voice without distortion. The scenario of this film rehearses a familiar refrain of sound studies, which is that the perceptual and cultural privilege accorded to the image in modern Western art forms foils and even defiles the study of sound on its own terms. Cinema plays on this dynamic when it attributes the sounds we hear to causes the film makes us believe in: the attribution of a black woman's singing voice to the screen image of a white woman in Dash's *Illusions* is just one instance of the ideologically loaded possibilities of this perceptual sleight of hand. In this regard, Dash's film offers a variation on the film sound scholar Rick Altman's argument that the study of Hollywood film must begin by looking past the image—a mere ventriloquist's dummy—to its true sonic source.⁴ But to what extent are this argument, and the remedy it proposes (looking past the image to hear sound objectively), helpful in apprehending the "reality" of Esther's voice? I propose a different line of thinking that foregrounds the ventriloquial basis of vocal sounds, as well as the perceptual biases that predetermine the success of the vocalic bodies that voices conjure, habituating to them as a screen, or outing them as a skin. At stake in this approach is an apprehension of the discriminatory habits of listening to vocal sounds, as well as the limitations of acousmatic listening as a means of undoing these habits.

The notion that the veiling or distortion of hearing by vision can be countered only by another act of veiling—averting the eyes—is at the basis of Schaeffer's theory of the sound object. An engineer and composer by training, as well as the founder of *musique concrète*, Schaeffer proposed that acousmatic sound, defined as a sound that is heard without its causes being seen, could enable a non-preconceived, "reduced" mode of listening. Schaeffer's interest in the technique of reduced listening, also known as acousmatic listening, was technical rather than ideological. Michel Chion explains: "Reduced listening takes the sound—verbal, played on an instrument, noises, or whatever—as itself the object to be observed instead of as a vehicle for something else."⁵ He continues: "Schaeffer thought that the acousmatic situation could encourage reduced listening, in that it provokes one to separate oneself from causes or effects in favor of consciously attending to sonic textures, masses, and velocities."⁶ In a recent book on acousmatic sound, Brian Kane paraphrases Schaeffer's theory as follows: "A sound object only truly emerges when a sound no longer functions for

another as a medium, but rather is perceived *as such*.⁷ In other words, Schaeffer believed that he could cultivate an “objective” attunement to sound, including vocal sounds, by denaturalizing or defamiliarizing the distorting lens that visual context habitually imposes on sound and, furthermore, by “fixing” the sound in the manner of an object. As the editors of this anthology note, a phonographic stylus stuck in a closed groove, permitting the listener to study the same sound over and over under acousmatic circumstances, was the paradigmatic realization of Schaeffer’s concept.

It is interesting to note that the originary sound object that inspired Schaeffer’s technique involved a ventriloquial illusion whose effects were not only technical (in fact, the illusion prefigures technical mediation) but also thoroughly ideological. Chion observes that the term “acousmatic” derives from “the name assigned to a Pythagorean sect whose followers would listen to their Master speak *behind a curtain*, as the story goes, so that the sight of the speaker wouldn’t distract them from the message.”⁸ According to this legend, the uninitiated disciples, known as *acousmatiques*, were allowed to look upon their master as full members of the sect only after spending five years in silence listening to him speak from behind the curtain. The purpose of this technique was to ensure obedience by *disciplining* the disciples to hear the subjective content of the master’s voice as objective truth. Writing about the same Pythagorean illusion, Mladen Dolar suggests that the master’s technique can be seen as a symptomatic response to a logocentric auditory culture. Dolar argues that in the Western philosophical tradition, which is bound by the imperatives of logocentrism, voices are obliged to shed their corporeal encumbrances (these include the embodied traits of accent, timbre, and tone, as well as involuntary utterances such as laughter, coughing, sighs, and so on) to convey their proximity to *logos*, or divine reason. The ritual of listening to their master lecture from behind a veil may therefore have served to train the disciples to hear his voice as a divine utterance—a “vanishing mediator” whose corporeal content (*phone*) evaporated in the act of utterance (*logos*).⁹

The question then arises: would the lifting of the curtain not spoil the illusion of mastery, revealing the master to be a mere mortal? Dolar’s answer is that the five-year ritual of acousmatic listening would have habituated the disciples to the *sound* of their master’s voice, permitting it to function as a virtual screen or veil even after the actual screen concealing his body had been lifted.¹⁰ Counter to the technique it inspired, the Pythagorean ritual of acousmatic listening actually would have *prevented* the listeners from “fixing” and regarding the traits of their master’s voice in the manner

of an object. If we take the master’s voice as the originary sound object, then, the sound of his voice “as such” is not the outcome of disillusioned listening liberated from the distortions of vision. Rather, this sound is *itself* an auditory illusion, curtain, or ideological veil—one that conceals the particularity of its embodied source and subject-position in order to evoke the authority and aura of a divine presence. To put it somewhat differently, the “objective” effect of the master’s voice is the result of a ventriloquial illusion that summons what Steven Connor calls an imaginary “vocalic body.” The vocalic body, writes Connor, “is the idea—which can take the form of dream, fantasy, ideal, theological doctrine, or hallucination—of a surrogate or secondary body, a projection of a new way of having or being a body, formed and sustained out of the autonomous operations of the voice.”¹¹ In summoning an imaginary body of metaphysical proportions, the master’s voice functions as a manner of invisibility cloak that conceals its actual origins, using the power of sound to countermand the evidence of sight.

The idea of the vocalic body as invisibility cloak in this originary scene of acousmatic listening vividly illustrates the often imperceptible but thoroughly ideological practices of perceptual disciplining that cloak the subjective, embodied origins of *certain* idealized voices, framing their traits as divine or disembodied. But if the Pythagorean master’s voice points to how white, male voices come to function as a screen despite their visible corporeality, then *Illusions* dramatizes the opposite: the equally imperceptible disciplinary practices that forcibly relocate voices whose traits depart from this norm in a racialized and gendered body. Esther Jeter also attempts a ventriloquial illusion in that her singing voice conjures an idealized vocalic body that obscures attention from her own black body—a dynamic that cinema only visualizes after the fact by attributing her voice to the face of a white woman as the epitome of an idealized and objectified femininity. However, the film is haunted by the failure of this illusion: the prospect of a (racialized) body whose skin threatens to assert its vocal presence, “outing” or disacousmatizing the body even in its visual absence.

In a revealing historical anecdote, Jennifer Lynn Stoeber notes that in the late nineteenth century, white opera reviewers often advocated “blind listening” as a technique for listening to black female opera singers, suggesting that closing one’s eyes would enable white listeners to judge black vocal performers more objectively, without being skewed by their visible blackness. However, this technique seldom worked. To the contrary, Stoeber observes that white listeners persisted in perceiving the voices of black female opera singers as both hypersexualized and technically in-

ferior to those of their white counterparts.¹² The musicologist Nina Sun Eidsheim has emphasized that there is no technical basis for the type of perceptual bias described by Stoever. Her own research on vocal morphology in the context of American operatic singing concludes that there are no more similarities in timbre within a so-called racial group than there are differences among groups. Eidsheim attributes the discriminatory tendency in operatic listening to “*acousmatic blackness*,” or “*the perceived presence of the black body in a voice that otherwise meets all the standards of a professional classical voice*.”¹³ By this, she means that even under acousmatic circumstances, the absent, visibly “other” bodies of black singers were conjured up as a perceptual phantom projected by listeners onto their vocal timbre.

Esther’s voice calls up the racialized perceptual biases that mediate both the production and the reception of vocal sounds—biases that acousmatic listening cannot necessarily remedy. These disciplinary perceptual processes interweave the racial gaze and its aural counterpart, or what Stoever calls “the listening ear,” in a collusive relationship.¹⁴ As a matter of fact, Dash’s use of Ella Fitzgerald’s voice to represent a black female singer whose voice passes as that of a white woman is a wishful revision of American film and music history. In the 1930s and ’40s, it was not unusual for black vocalists to sing specialty numbers for dramatic films, but the film historian Marsha Siefert reminds us that these numbers were often cut when the films were distributed in the American South for fear that their “natural voices and singing styles might also mark their ethnicity and therefore might limit the market for the film and sound track sales.”¹⁵ In fact, it was far more common for songs in film musicals featuring black casts to be dubbed by white, classically trained singers: a famous example is that of Marilyn Horne, a white, classically trained mezzo-soprano opera singer who was hired to sing the lead female vocals in Otto Preminger’s film *Carmen Jones* (1954), dubbing for the African American actress Dorothy Dandridge, who was herself an accomplished musical performer. This technique both drew on and reinforced a racialized hierarchy of musical styles, between the mannered and cultivated style of opera singers and the so-called spontaneous or natural style of jazz and pop singing.¹⁶ Dubbing was thus used in Hollywood films as a means of audibly “whitewashing” the visual presence of black bodies, producing an effect that the film music historian Jeff Smith calls “a kind of phantasmic body that registers visually as black but sounds ‘white’ in terms of the material qualities of its ‘voice.’”¹⁷

The contrasting attunements of the acousmatic listeners described by Chion and Dolar, on one hand, and by Stoever, Eidsheim, and Smith, on the other, illustrate an insidious and ideologically fraught perceptual tendency that I sum up as follows: idealized voices are heard as a *screen* that resists objectification even when their bodies are visible, whereas minoritized voices are circumscribed in advance as an objectified *skin*—even when they are acousmatic. My reference to the voice as skin adapts and extends Rey Chow’s use of the term “skin tones” to describe how accented voices are not only heard but subject to a type of visual scrutiny that probes and reads them as a racialized visual surface. Chow points out how the conjoined visual and audial connotations of “skin tones” are activated in a postmodern version of Esther Jeter’s vocal makeover: the situation of South Asian and East Asian offshore call center workers who must imitate the vocal mannerisms, styles, and accents of their American customers to be understood by them.¹⁸ I would add that we can better apprehend the complexly political operations of sound and listening if we attend to the ways in which sound—and specifically, the voice—functions as a protective veil or skin, as well as the moments in which that auditory illusion fails and is exposed, denuded. Indeed, the restriction of “veiling” to a visual register to arrive at a more discerning mode of listening—the technique to which Schaeffer subscribes—can have the unexpected effect of concealing the ways in which listening stands in for and conjoins with vision as a means of discrimination, or a way of separating subjects from objects in the field of sound.

In the next section, I briefly turn to scholarly debates surrounding the voice in documentary. While this may seem somewhat of a departure from my discussion so far, the purpose of my turn to documentary is to consider the genre’s vocal conventions—which documentary scholars typically frame as a realist corrective to fictional illusions—as another instance of how anti-illusory cinematic forms can end up concealing the discriminatory perceptual frames of voicing and listening, even when they aim to achieve the opposite. My main contention is that the liberatory impulse driving innovations in documentary’s vocal modes maintain an unproductive focus on the fundamental viscosity of the acousmatic veil. Feminist criticisms of one of documentary’s most derided sonic inventions, “Voice of God” narration, which takes inspiration from the Pythagorean technique of lecturing from behind a veil, offer a case in point. These critiques, and their proposed vocal alternatives, exemplify how realist attempts to achieve

vocal parity using the disillusioning techniques of disacousmatization or acousmatization fail to account for the different degrees and types of auditory scrutiny to which idealized and minoritized voices are subject, as well as the ventriloquial illusions to which they have recourse. This analysis of documentary lays the foundation for my reading, in the final section of this chapter, of a ventriloquial mode of cinematic looking-listening that makes perceptual adjustments to offset the discriminatory frames of vocal production and reception.

Documentary (Dis)illusions

Documentary's difference from fiction film is frequently articulated in terms of its reliance on sound rather than the image. The speaking voice is regarded as a reality principle that brings sobriety and grounding to the flights of fancy that the image might otherwise encourage. Bill Nichols captures this sentiment in *Representing Reality* (1991), a classic of documentary studies, when he writes:

Documentary film often builds itself around the spoken word. Works from Frank Capra's *Why We Fight* series (1942–45) on the reasons for United States involvement in World War II to Ken Burns's *The Civil War* (1990) would be subject to endless interpretation if we had nothing but their extraordinarily diverse and historically intriguing images to guide us. Commentary points us toward the light, the truth . . . Fiction attends to unconscious desires and latent meanings. It operates where the id lives. Documentary, on the other hand, attends to social issues of which we are consciously aware. It operates where the reality-attentive ego and superego live.¹⁹

Even though Nichols has been criticized for associating progressive ideologies with aesthetic features, and himself goes on to deconstruct and complicate this polarization of reality/documentary versus illusion/fiction, documentary scholars continue to associate *voice*—that is, the spoken word—with the pursuit of democracy, disillusionment, and truth.²⁰ The proposed alternatives to documentary's most vilified device, Voice of God narration, evidence the enduring power of these associations. Critics of this device attribute its effects of distortion to a visual illusion and have accordingly sought alternatives to this convention that visualize the sources of idealized speaking voices. At the same time, feminist documentary

scholars have recuperated the veil of acousmatic anonymity as a means of drawing attention to the voices of minorities free from the distortion of the image. I propose that these parallel attempts to unveil the "truth" of sound conceal a larger concern: that the capacity for illusion and deception rests not only with the image but also with sound.

Voice of God narration is the name given to the distinctive type of disembodied voice-over commentary associated with "classical" documentary films of the 1930s and '40s. This type of commentary is usually described in terms of its prototypical features, exemplified by the vocal commentary for the *March of Time* newsreel series—for example, "a white, male, middle-class and anonymous voice," according to Stella Bruzzi; "detached, authoritarian, male," according to Jeffrey Youdelman; and "disembodied . . . stentorian, aggressive, assuming a power to speak the truth of the filmic text, to hold captive through verbal caption what the spectator sees" and "omniscient, omnipresent," according to Charles Wolfe.²¹ Since these descriptions note the absence of the cause or source of Voice of God commentary in their emphases on anonymity, detachment, and disembodiment, this mode of documentary narration can be called an example of what Chion, adapting Schaeffer, calls an "acousmatic voice."²² Indeed, the origins of the word "acousmatic" in the myth of the Pythagorean teacher lecturing from behind a veil are strongly evocative of documentary's etymological root in the Latin *docere*, in its emphasis on pedagogy and the transmission of knowledge by "telling" unencumbered by the distractions of "showing."

The convention of Voice of God narration exemplifies how contemporary audiovisual forms adapt and mutate the ancient Pythagorean technique to disguise the particular, subjective source of certain privileged voices (namely, those of white, middle-class, educated men) in an authoritarian cloak of universality. Documentary scholars have championed two types of democratic alternatives to Voice of God narration. The first type includes vocal conventions such as the recorded conversation, dialogue, and the "talking head" interview that synchronize voices to their "real" or originating bodies as a means of unmasking their location. This practice was pioneered by practitioners of direct cinema and *cinéma vérité* in the 1950s and '60s who rejected voice-over narration in favor of visualizing verbal events. Although these conventions have been critiqued for actively promoting the illusion of authenticity and immediacy, they remain in widespread use in contemporary documentary practices. What is more, they have garnered critical appreciation for invoking the particularity of embodied speech, or, to quote Jeffrey Ruoff, "the material texture and richness of unrehearsed

speech, the grain of the voice."²³ These values are positively opposed to the negative values of objectivity and univocality associated with the idealized male acousmatic voice in Voice of God narration.

The second alternative consists of voice-overs that, while disembodied, are nonetheless believed to foreground the particularity of speakers' embodiment in their tone, gender, ethnic identity, and mode of address.²⁴ Bruzzi, for instance, praises films that employ women's voices in subjective, experimental ways (e.g., first-person narration) that undercut the emphasis of Voice of God narration on authority, detachment, and omniscience by embodying the virtues of indefiniteness, idiosyncrasy, and personal exploration.²⁵ In this regard, Bruzzi is one among several documentary scholars who has extended Kaja Silverman's advocacy of the acousmatic voice as a feminist documentary device. Silverman celebrates the work of feminist filmmakers who borrow the veil of acousmatic anonymity as a means of claiming the authority of speaking subject for women's voices. She argues that the presence of a female acousmatic voice defamiliarizes the gendered articulation of power typical of classical Hollywood cinema, in which acousmatic authority is reserved for male voices, whereas female voices are "pinned" to bodies and thereby subject to the distortion of "to-be-looked-at-ness."²⁶

We can thus categorize the democratizing tactics adopted by critics of Voice of God narration as (1) the attempt to disacousmatize voices, especially idealized voices, by revealing/unveiling their bodies; and, reciprocally, (2) the attempt to recuperate the veil of acousmatic anonymity for minoritized voices. Both tactics interpret the acousmatic veil in literal terms as a *visual* screen. Thus, the powerful effects of offscreen commentary are claimed as a way for minoritized voices to be *heard*, free from the visual distortion of being turned into an image. Through an inversion of the same logic, the incontrovertible effects of Voice of God narration are countered by visualizing the sources of authoritarian or "expert" voices. The goal, as Bruzzi puts it, is to problematize claims to an unproblematic or universal truth. "Narration," she writes, "could therefore be seen as a mechanism deployed to *mask* the realization that this mode of representation, and indeed its inherent belief in a consistent and unproblematic truth, are perpetually on the verge of collapse, that commentary, far from being a sign of omniscience and control, is the hysterical barrier erected against the specter of ambivalence and uncertainty."²⁷

What is salient in Bruzzi's analysis of narration is that she cautions against too literal an interpretation of the acousmatic mask or veil of classical

documentary narration, describing it instead as a structure of disavowal. Her analysis unintentionally sheds light on the ventriloquial illusions involved not just in Voice of God narration but in *all forms* of documentary narration, including the talking head or subjective, first-person commentary. However, when we take into account the centuries of auditory habituation that connect the Pythagorean *acousmatiques* and documentary listeners, we can also see how these illusions might have discrepant degrees of success. Following Bruzzi's logic, we may venture this proposition: idealized voices may continue to exert their powers even *after* their sources have been exposed or disacousmatized. The embodied traits of a mid-Atlantic, middle-class, educated, deep male voice can thus be made to seem weightless, almost immaterial, summoning a metaphysical vocalic body that supplants the evidence of the eyes. Conversely, voices that depart from this zero degree of speech may be disacousmatized *despite* being disembodied; the audible evidence of their race, gender, class, and education can envelop and weigh them down, summoning their bodies in absentia even when their sources are withheld from view. Under these circumstances, the realist documentary discourse of sound's distortion by the image can itself become a veil or "hysterical barrier," to repurpose Bruzzi's term. The specter that is warded off is twofold: that the speaking voice is as capable of illusory flights of fancy as the image, and that discriminatory habits of (looking and) listening determine which of these illusions pass as reality.

The "real" source of Esther Jeter's voice in *Illusions* is an interesting limit case that reveals the contradictions of the critical consensus I have described. In a politically correct sense, the synchronization of Esther's voice to Leila Grant's face is a falsehood—a visual distortion that demands disillusioning corrections of the kind that documentary filmmakers have pursued. But in a more perverse sense, the "false" synchronization is true to the auditory illusion that Esther is attempting (and failing) to produce: the illusion of a voice that *sounds* white, and which seeks to deflect the visible evidence of her own black body. Esther's voice ostensibly emerges from her body, but it conjures up what Connor calls an imaginary vocalic body that may "contradict, compete with, replace, or even reshape the actual, visible body of the speaker."²⁸ This dematerializing vocalic body promises an auditory screen of relative anonymity, allowing Esther to elude the experience of being heard and objectified as a skin. Neither a corrective disacousmatization nor a redemptive acousmatic solution would allow us to properly apprehend the enigmatic embodiment of this vocal performance or the perceptual biases that it acknowledges and seeks to circumvent.

In the final section of this chapter, I turn to *Paris without a Sea*, a short experimental documentary film by the Lebanese artist Mounira Al Solh, as a counterpoint to Dash's *Illusions*.²⁹ I would like to hold this film alongside the predominantly visual interpretations of the Pythagorean veil that have inspired the realist turn to acousmatic voicing and listening among feminist film critics and documentary scholars. Al Solh neither proposes averting our gaze to grasp minoritized voices objectively nor recuperates the cloak of visual anonymity for these voices. Instead, she brings a concrete audiovisual expression to Connor's concept of the vocalic body as a type of auditory veil ("a surrogate or secondary body, a projection of a new way of having or being a body, formed and sustained out of the autonomous operations of the voice"). What is more, she finds an unusual way to subvert the discriminatory perceptual habits that virtually unveil and denude some voices as a means of keeping them in their place while permitting others to function as a dematerializing screen body that consolidates their privileged social position. The absurd ventriloquial performance at the center of Al Solh's film uses the simple but effective technique of reversal to school us in a ventriloquial mode of listening—one that confronts the ways in which the disciplining of the ear and eye hold larger social hierarchies in place.

Ventriloquial Displacements

Paris without a Sea is provocative to think with alongside Dash's *Illusions* because of the irreverent way in which Al Solh connects the two chimerical registers of the voice's skin—objectified surface and performative cover-up—with the broader symbolic economies of voicing in documentary and fiction. Al Solh is Lebanese and works between Beirut and Amsterdam, as well as between the idioms of documentary and performance art. Migration, translation, appropriation, and miscommunication are central themes in her work. Trained as a painter, Al Solh incorporates a variety of media in her cinematic language, including drawing, performance, puppetry, and role-playing. In her hands, all of these media become ways to navigate the world as an Arab woman whose movements and artistic expressions are further restricted by the realities of conflict and its aftermath.

In *Paris without a Sea*, Al Solh comes up with an ingenious solution for occupying the world vocally in ways that her own body does not allow: she steals the voices of Beirut men, as though to proclaim that a voice always speaks from elsewhere than one's body. This short film is part of a multi-

media project that revolves around a group of Beirut men, including Al Solh's father, who swim daily in the Mediterranean, no matter the weather. Al Solh speaks wistfully of the ability of the men to sun themselves shirtless and to strip off their clothing and jump into the water. She compares their daily escape to her own mobility as an art student in Amsterdam, but with one important difference. She writes, "It is as if they emigrate every day to somewhere else through the water, without really leaving their country. They are out of place, yet still immersed in it."³⁰ Al Solh attempts to access the eminently embodied yet out-of-body experience of this masculine public sphere by conducting interviews in Arabic and French with the men right after they emerge from the water. Her interview questions turn the men's love for swimming into an avenue for a dialogue on more intimate topics, such as courtship, religion, romance, and gender. Caught off-guard (and on camera) by Al Solh's polite but playful questions, the men answer in disarming ways that nonetheless reveal their male privilege. One young man admits, for instance, that his mother washes the laundry that he discards after his daily swims. Another, asked by Al Solh how he would respond if his girlfriend—by his own poetic admission, the "love of my life, the veins in my body"—asked him to give up the sea and move to the mountains, answers, "No way! The sea is more important. . . . Go to the mountains by yourself. . . . If she tries to deprive me of the sea?! The sea is part of my eyes!"

Al Solh finds an unusual way to imagine what it might be like if their gender roles were reversed. She dubs the men's answers in her own voice, replicating their intonations and their laughter, their bravado and their embarrassment. This simple technique is both comical and subversive: Al Solh remains offscreen so that her soft, feminine voice not only issues the questions but also answers them, appearing to emanate incongruously from the naked, sunburned torsos of the men, many of whom are clad only in swim trunks and goggles. English subtitles, rendered in two different colors, assist Anglophone audiences in telling the questions from the answers, but since both are voiced by Al Solh, the effect is that of a ventriloquist's performance in which Al Solh projects her voice from offscreen onto the men's onscreen bodies. The performance is humorous because it is unconvincing. Although Al Solh syncs her speech as perfectly to the men's moving lips as Esther Jeter does to Leila Grant's in *Illusions*, the effect is absurd, not naturalistic. Al Solh superimposes two vocal conventions associated with documentary realism—the talking head interview and the acousmatic female voice—by projecting the latter over the former. In the

process, she also undoes the reality effects of these seemingly naturalistic conventions, inviting us to engage with them both as advanced Pythagorean illusions. This has an estranging effect on our habits of looking and listening. Where we would ordinarily hear documentary voices as indexes of bodies, Al Solh urges us to listen for how voices call surrogate vocalic bodies into being and how the reality effects of these vocalic bodies are differentially mediated in every cultural context by the twinned operations of a racialized, gendered gaze and its aural counterpart, the listening ear. The realism of documentary's vocal conventions, she suggests, can leave these perceptual frames undisturbed.

The unexpected reversal of hearing Al Solh's voice emerge from the men's torsos brings a concrete audiovisual expression to the perceptual adjustments required to work against these mediating forms. Al Solh effectively uses the cinematic play of sound and image to claim the protective cover of acousmatic displacement for her voice in the very same movement in which she (doubly) denudes the men as an objectified skin, transforming them from speaking subjects into a ventriloquist's dummies. By the end of the video, it becomes clear that Al Solh is speaking from a place, and calling into being an embodiment, that is unavailable to her *except* as a vocal performance. She ends her video by once again ventriloquizing the second of the men mentioned earlier. Asked whether he can imagine a city without a sea, he declares, "I'd die! . . . There's no city without a sea. It's just that maybe the sea is far . . . you might need to travel thousands of kilometers to get to it." This heartfelt statement becomes a poetic metaphor for Al Solh's voice, in which it is delivered. Speaking through her proxy, Al Solh tells us that using her voice as anything but a protective skin that is both immersed in her body and out of place would be like living in Beirut without a sea. To partake daily of that performance, that displacement, she insists, is as important to her as swimming in the sea is to the Beiruti men.

What makes this humorous spectacle sobering is that Al Solh asks us to understand her absurd ventriloquial performance as a necessary and even natural precondition of speaking and being heard in a logocentric, Eurocentric, and patriarchal society. Her film produces a Brechtian alienation or estrangement effect similar to that which Chow argues is achieved when actors memorize and reproduce lines of dialogue in a foreign language whose grammatical meaning is unknown to them: the actor's body is displaced as the de facto "origin" of the voice and instead appears as a ventriloquist's dummy that is animated by a force that comes from without

and not within.³¹ Al Solh's ventriloquial performance invites a mode of listening that is attentive to the ways in which voices displace their origins, as well as the hierarchizing social forces that animate the desire for this displacement and shape the phantasmic forms that these desires take. Indeed, Al Solh invites us to actively practice a *ventriloquial* mode of listening as a necessary and even natural mode of listening to vocal sounds. Ventriloquial listening brings a political dimension to the technical ritual of acousmatic listening that is especially pertinent in the context of vocal sounds. Acousmatic listening, as advocated by Schaeffer, attunes the listener to the traits of a voice, independent of its causes. Ventriloquial listening displaces the listener's attention from the visible causes of a voice (its embodied origins), attuning them instead to the invisible racialized and gendered perceptual frames—those forces that "come from without and not within"—that shape the production and reception of vocal sounds.

To listen ventriloquially is to behold the bifurcation of a voice's origins and its surrogate forms as one might behold performances of ventriloquists projecting their voices onto dummies. A ventriloquial listener becomes perceptive to the *seam* between the embodied origins of voices and the surrogate bodies that voices conjure into existence. They take note of when the illusion succeeds, magically animating the dummy, and when the illusion fails, reasserting the thingness or the matter of the dummy body. Ventriloquial listening asserts the ideological work involved on the part of both the performer and the audience of these vocal and perceptual illusions. It trains listeners to notice and work against the training of their senses by coercive social forces that operate disparately to "out" the origins of some voices, unraveling their auditory illusions and grounding their vocal traits in embodied matter, or that conversely transform the material traits of other voices into a vanishing mediator, seamlessly allowing the illusion to persist.

Conclusion

"Every emission of the voice is by its very essence *ventriloquism*," writes Dolar.³² While Dolar points to the inherently acousmatic, veiled origins of every voice, Stoeve, Eidsheim, and Chow point to the equally veiled perceptual frames that mediate the prospects of these ventriloquial illusions through their analyses of "the listening ear," "acousmatic blackness," and

"skin tones." Listening, as Eidsheim notes, "does not connote passive reception of information and is not a neutral activity. Rather, in listening we participate in social processes both embedded in and producing cultural forms."³³ I have foregrounded the perceptual frames that mediate the production and reception of acousmatic voices by introducing the interrelated concepts of "the skin of the voice" and "ventriloquial listening." As a conceptual frame, the skin of the voice frames the voice as both an auditory phenomenon and a visual surface that can be read, profiled, and objectified and pinpoints how hearing is complicit in *and* complexly intertwined with vision in distinguishing between idealized and "other" voices.³⁴ I have argued, using this term, that idealized voices are perceived as an acousmatic screen—that is, as detached and displaced from their source—even when they emerge from a visible body, while minoritized voices are heard and objectified as a skin, even when their bodies are invisible. This term not only moves us beyond the persistent dualisms between sound and image, and *logos* and *phone*, that frame discussions of acousmatic listening, but also points to the ideologically inscribed perceptual processes that interweave the habitual operations of looking and listening.

Reconceptualizing the voice as skin also suggests that the voice can displace itself from its embodied source and relocate to an "other" imaginary place even without the assistance of a visual barrier. While Esther Jeter uses a surrogate white body to claim the protective veil of acousmatic displacement that her own voice has been unfairly denied, Mounira Al Solh's surreal projection of her own voice onto surrogate male bodies behaves as a supplement that paradoxically strips the veil from voices that enjoy its protection. I have read these films together in an effort to point out how Al Solh develops an audiovisual idiom that invites us to listen ventriloquially to *every* voice as an acousmatic performance that estranges and disguises the body from which it emanates. The closing image in Al Solh's film, of her father, towel slung over his shoulder, squinting into and adjusting a pair of binoculars, is a visual echo of Mignon Dupree's piercing but unfathomable look at Esther in the recording booth. It is a fitting reminder of the varying levels of scrutiny that different voices—whether speaking or singing, gendered or accented, embodied or disembodied—must "pass" for their illusions to succeed. Ventriloquial listening can function as a type of perceptual adjustment that attends to and actively works against these differential frames of scrutiny so that we may indeed grasp every voice as an acousmatic performance emerging from and contending with asymmetrical social conditions.

Notes

- 1 Julie Dash, dir., *Illusions*, film, Women Make Movies, New York, 1983.
- 2 Ryan, "Outing the Black Feminist Filmmaker," 1322.
- 3 Mellencamp, *A Fine Romance*, 235.
- 4 Altman, "Moving Lips."
- 5 Chion, *Audio-Vision*, 29.
- 6 Chion, *Audio-Vision*, 32.
- 7 Kane, *Sound Unseen*, 25.
- 8 Chion, *The Voice in Cinema*, 19, 1915. Chion attributes the anecdote regarding the *acousmatiques* to Denis Diderot and Jean le Rond d'Alembert's *Encyclopédie* from 1751.
- 9 Dolar, *A Voice and Nothing More*, 15.
- 10 See Dolar, *A Voice and Nothing More*, 67–68. Dolar speculates that, like the eponymous character in *The Wizard of Oz*, the Pythagorean master is capable of exercising his divine powers over his listeners only as long as the source of his voice remains hidden. But while the wizard is revealed to be a "pitiable old man" once the veil is lifted and his voice is disacousmatized, the master's five-year acousmatic ritual wards off the castrating effects of exposure. Thus, the master's voice becomes a "phonic phallus" that allows the fantasy of his power to persist.
- 11 Connor, *Dumbstruck*, 35.
- 12 Stoever, *The Sonic Color Line*, 114–15.
- 13 Eidsheim, "Marian Anderson and 'Sonic Blackness' in American Opera," 647. Eidsheim writes that reviewers of American opera from the late nineteenth century to the mid-twentieth century have persisted in attributing a "distinctly black" timbre to the voices of black female opera singers, a tendency that may have led to the typecasting of these singers in roles such as maid, slave girl, or gypsy.
- 14 Stoever, *The Sonic Color Line*, 79.
- 15 Siefert, "Image/Music/Voice," 60.
- 16 Smith, "Black Faces, White Voices," 37.
- 17 Smith, "Black Faces, White Voices," 37.
- 18 See Chow, *Not Like a Native Speaker*, 7–9. Chow reads the failed attempts of these workers to disguise the *sound* of their unseen but audible embodiment as a form of vocal disfigurement akin to the botched effects of skin-whitening treatments. Both resort, she argues, to self-defacement in order to rectify a mode of self-expression that is falsely deemed defective.
- 19 Nichols, *Representing Reality*, 4.
- 20 See, e.g., Bruzzi, *New Documentary*. Since the publication of the first edition of her book, Bruzzi has been a vocal critic of Nichols's chronology of the various modes of documentary, which, she argues, produces false dichotomies between ideologically regressive and progressive documentary approaches based on films that share formal features.

- 21 Bruzzi, *New Documentary*, 57–58; Youdelman, "Narration, Invention, and History," 9; Wolfe, "Historicizing the 'Voice of God,'" 151.
- 22 Whether such a voice can be described as an *acousmètre*—or, indeed, whether this concept is applicable to documentary—is an open question. Chion's analysis of the *acousmètre*, defined as a disembodied cinematic voice that activates a tension between onscreen and offscreen space in the form of an "acousmatic zone," is restricted to fiction films, and he disavows the capacity of documentary voices to activate an acousmatic zone; see Chion, *The Voice in Cinema*, 22. The coordinates and implications of an acousmatic zone in documentary could profitably be discussed in alignment with Eidsheim's account of "acousmatic blackness," and I hope to address them on another occasion.
- 23 Ruoff, "Conventions of Sound in Documentary," 30. See also the analysis of these conventions in contemporary biographical and indie-rock documentaries in Sexton, "Excavating Authenticity."
- 24 See, e.g., Wolfe, "Historicizing the 'Voice of God'"; Youdelman, "Narration, Invention, and History"; Bruzzi, *New Documentary*.
- 25 Bruzzi, *New Documentary*, 64–65.
- 26 Silverman, *The Acoustic Mirror*, 165.
- 27 Bruzzi, *New Documentary*, 59, emphasis added.
- 28 Connor, *Dumbstruck*, 36.
- 29 Mounira Al Solh, dir., *Paris without a Sea*, video, <https://vimeo.com/40821360> (password-protected).
- 30 Mounira Al Solh, "The Many Metamorphoses of Mounira Al Solh," interview by Nat Muller, *Ibraaz*, June 27, 2013, accessed February 1, 2016, <http://www.ibraaz.org/interviews/82>.
- 31 See Chow, "After the Passage of the Beast," 42.
- 32 Dolar, *A Voice and Nothing More*, 70.
- 33 Eidsheim, "Marian Anderson and 'Sonic Blackness' in American Opera," 665.
- 34 For this chapter, I am bracketing the haptic dimensions of voicing and listening but I discuss them at greater length in my introduction to a special issue of the journal *Discourse* on "Documentary Audibilities" that I coedited with Genevieve Yue: see Rangan, "Audibilities."

IV

Sound Objects and Nonhuman Relations