Vocal Projections
Voices in Documentary
Edited by
Annabelle Honess Roe and Maria Pramaggiore
Contents

List of Illustrations vii
Acknowledgements viii

1 Introduction: Documentary’s Vocal Projections
Annabelle Honess Roe and Maria Pramaggiore 1

Section One Voices, Bodies and Power
1 Playing God: Film Stars as Documentary Narrators
Annabelle Honess Roe 11
2 Auditing the Call Centre Voice: Accented Speech and
Listening in Sonali Gulati’s Nalini by Day, Nancy by
Night (2005) Pooja Rangan 29
3 The Fundamental Lie: Lip Sync, Dubbing, Ventriloquism
and the Othering of Voice in Documentary Media
Patrik Sjöberg 45

Section Two Voices beyond Language
4 Hearing Voices in Singer-Songwriter Music Documentaries
Maria Pramaggiore 63
5 ‘The Voice’ and Sound in Indian Documentary Film: Listening
to Surabhi Sharma’s Bidesia in Bambai (2013) Deborah Matzner 81
6 Hum, Buzz, Gurggle: Ecological Soundscapes in Poetic Ecodocs
Shilyh Warren 101
7 Literally Documenting Her Voice: Valie Export’s Vocal-Visual
Experiments with Screen Media Helen Hughes 119

Section Three Gender, Sexuality and Species
8 The Voice of Mockumentary Sarah Kessler 135
9 Fabulous Thinking, Chimeric Voice: Kathy High’s Lily Does
List of Illustrations

3.1 Adie Russell as Merv Griffin. *Never Had More Trouble* (Adie Russell 2007). Courtesy Adie Russell  55
3.3 Adie Russell as Andy Kaufmann. *Time Can Do So Much* (Adie Russell 2017). Courtesy Adie Russell  57
5.1 Kallu on stage. Still from *Bidesia in Bambai* (Surabhi Sharma 2013). Courtesy Surabhi Sharma  88
5.2 Yadav imitating the lonely village wife. Still from *Bidesia in Bambai* (Surabhi Sharma 2013). Courtesy Surabhi Sharma  89
5.3 Yadav on the Chhath Puja stage. Still from *Bidesia in Bambai* (Surabhi Sharma 2013). Courtesy Surabhi Sharma  90
5.4 Kalpana Patowary recording in her studio. Still from *Bidesia in Bambai* (Surabhi Sharma 2013). Courtesy Surabhi Sharma  94
6.1 A toxic lake in *Petropolis* (Peter Mettler 2010) shatters the idyllic scenery of the opening of the film. Greenpeace/Eamon Mac Mahon  107
6.2 At the centre of *Silent Springs* (Erin Espeile 2011) a montage of cellular material in microscopic detail. Courtesy Erin Espeile  107
6.3 Traversing the surface of the Arctic in *Invisible* (Roz Mortimer 2006) is an enigmatic feminine figure in a bizarre red dress. Courtesy Roz Mortimer  114
8.1 Gloria (Sofia Vergara) entertains the family with her un-closeted ventriloquist's dummy, Uncle Grumpy in 'Yad Sale' (dir. Gail Mancuso). Season 4, Episode 6, *Modern Family* (ABC, created by Christopher Lloyd and Steven Levitan) 31 October 2012. Screenshot CityTV, Toronto, Canada. Rogers Media  148
Auditing the Call Centre Voice: Accented Speech and Listening in Sonali Gulati’s *Nalini by Day, Nancy by Night* (2005)

Pooja Rangan

What can we learn from call centres about the place of voice in documentary? Although these two sites – the offshore call centre and documentary’s vocal conventions – may seem at first to have little to do with each other, I will suggest that their conjunction tells us much about documentary as a site of auditory interpellation. My analysis focuses on the accented speech of non-native English speakers and the techniques of making these voices intelligible, both in the context of the accent neutralization programmes that call centre trainees undergo and in terms of the use of documentary conventions to neutralize and render transparent, or conversely to mark and stigmatize the affective disturbances of accented speech. Call centre documentaries like *Nalini by Day, Nancy by Night* (dir. Sonali Gulati, 2005), in which I ground my analysis, offer a rich vehicle for this inquiry, since they dramatize the tensions inherent in and the potential of documentary as an ‘accented’ genre of cinema (Nafcy 2001) at both a thematic and a formal level. The internal contradictions of this film prompt me to reframe its textual point of view, or what documentary scholars might call its ‘voice’, as an *audit*: a mode of speaking and organizing that both produces and is produced by unspoken norms of perceptual discipline or attention. I ask: What disappears or is neutralized at the level of the speaking voice and the listening ear when we sense such a textual voice in documentary, and what can the ‘placelessness’ of accented speech in the age of the call centre tell us about the stakes of this disappearing act?

My inquiry into the structuring absences of accented speech and listening pivots on a reassessment of ‘voice’ as the prevailing metaphor for a textual point
of view in documentary. Following John Mowitt (2015) I call this reassessment an ‘audit’ of the place of voice in documentary: a form of review that acknowledges the persistent logocentrism of voice as a metaphor, and the ways in which this metaphor informs the way documentary films listen for, organize and distribute the materiality of vocal sounds, as well as how they invite us to listen to the world.

From voice to audit

In English, as Mowitt notes, there is no satisfactory equivalent to ‘point of view’ in the auditory domain; a ‘point of audition’ is not entirely satisfying or coherent since ‘point’ means little in the sonic register (2015: 4). Mowitt therefore proposes the term ‘audit’ in an effort to develop an analogue of ‘gaze’ in the auditory or sonic domain: one that would roughly designate ‘that which exceeds and conditions hearing and organizes the field of the audible’ (5–6). Mowitt describes audit as a ‘hearing’, or a mode of perception that has a primordial tie with aesthetics or what Jacques Rancière calls ‘the distribution of the sensible’ (Rancière 2006: 12). Mowitt further elaborates that this ‘hearing’ refers not so much to a ‘perception or sensible event’ but to a ‘fold where perception turns back on itself, traversing the faculty of hearing with the angle, the posture of listening’ (5). This complicated statement requires a bit of parsing: listening and hearing are often conflated, but hearing is generally considered a more passive mode of auditory reception, whereas listening is considered to involve ‘making an effort to hear something’ (Rice 2015: 99). Mowitt’s definition of the audit as the fold where the posture of listening traverses the faculty of hearing can thus be understood as an effort to argue that (a) aesthetic forms introduce thoroughly ideological, if unconscious, perceptual distinctions, hierarchies or distributions into the conceptual domain of sound, and (b) these aesthetic forms, and the perceptual attunements that they cultivate in their audience, are both produced by and productive of the discursive conditions of audibility in any given conjuncture.

Documentary is one of these aesthetic forms. When we approach it from Mowitt’s perspective, the popular metaphor of the ‘voice of documentary’ can be understood not just as a passive description of a film’s textual point of view but as an articulation of how documentary’s textual conventions are both produced by and actively productive of the audit. In the field of documentary, voice, rather than point of view or gaze, is the prevailing metaphor for a documentary film’s particular world view or social perspective. The popularity of the voice metaphor owes in large part to an early and influential essay by Bill Nichols titled ‘The Voice of Documentary’, in which Nichols defines ‘the voice of documentary’ as ‘that which conveys to us a sense of a text’s social point of view, of how it is speaking to us and how it is organizing the materials it is presenting to us’ (1983: 18). But even as he associates voice with speech, Nichols hastens to add: ‘Voice’ is not restricted to any one code or feature such as dialogue or spoken commentary. Voice is perhaps akin to that intangible, moiré-like pattern formed by the unique interaction of all a film’s codes, and it applies to all modes of documentary’ (18).

I have written elsewhere about the widespread and largely uncritical adoption of Nichols’s metaphor – which stands simultaneously for a film’s perspective, its mode of address and its textual organization, even as it invokes the varied narrative possibilities and political stakes of ‘giving voice’ to actual speaking subjects – by documentary scholars and practitioners over the past thirty years (Rangan 2017). Rather than reiterate these arguments, in which I redefine voice as an ‘audibility’, extending Gilles Deleuze’s analysis of ‘visibilty’ in Michel Foucault’s writings (Deleuze 1988: 52), I would now like to offer some additional reflections on voice. I am interested in how voice, as a sonic metaphor for point of view, has defined documentary realities, by giving rise to conventions that not only construct perceptual conditions of hearing but also ‘highlight some features of reality and hide others’ (Novak and Sakakeeny 2015: 1). I find it noteworthy that Nichols’s metaphorical use of ‘voice’ simultaneously signals and disavows the documentary genre’s literal emphasis on speech and innovation of vocal conventions (such as voice-over commentary, dialogue and the interview) as its distinguishing textual feature. Nichols’s mention of the term ‘codes’ to refer to these conventions (indicative of the linguistic and post-structuralist bent of film studies in the 1980s)1 refuses any naïve empiricism regarding documentary’s social ethic of ‘giving voice’ to actual speaking subjects even as his use of voice to invoke that ‘essential albeit invisible something’ that coheres a film together seems symptomatic of the enduring Western metaphysics of presence that attaches to the voice in all of its conventional forms (Chow 2013: 40).

The same contradiction characterizes Nichols’s recent return to his voice metaphor in Speaking Truths with Film (2016). Nichols proposes here that one of the major historical shifts in the voice of documentary since the genre’s inception, reflective of technological as well as social changes, is the shift away from a voice that aims to speak (often from a unilateral and anonymous perspective, as characterized by intertitles in early silent films or the convention of ‘Voice of God’ commentary) to one that ‘listens with compassion’ to ‘histories
from below' (82; my emphasis). The contemporary voice of documentary conveys an ‘embodied perspective’ to the viewer (82) that foregrounds elements of affect and emotional engagement (84), including inflections, gestures and behaviour (75). Intriguingly, even as Nichols associates documentary’s evolving horizons with the potential for an encounter with an embodied diversity of affects and gestures, he again employs a vocal metaphor to describe that which these sensations ultimately convey: an imperative, interpellative address. He argues: ‘Unexpectedly, someone calls out: “Hey, you!” [...] to be addressed by a film – to sense that a film seeks to engage and speak to us about the world we share – functions as a hallmark of documentary film’ (74). To elaborate, a documentary film’s ‘Hey, you!’ may be explicit, even didactic (the filmmaker speaks to us directly or through a proxy commentator) or it may be implicit or tentative (the editing or the choice of interview subjects conjures a voice or offers a proposal about the world). But Nichols is arguing that despite these differences in who is speaking and how we perceive their address, what is common to all documentary films is the sense we get of an interpellative ‘Hey, you!’ that seeks to ‘engage and speak to us about the world we share’, soliciting a form of attention that listens for a voice.

This is an intriguing formulation because it suggests that what arrests our attention in documentary, what makes us turn around and pay attention in response to its ‘Hey, you!’ is the implied presence of verbal address, the sense of a voice speaking to us, even as our sense of that voice is increasingly fleshed out, embodied and relational, as it were. I would like to offer a restatement of this paradox: the interpellative force of verbal address – whether actual or implied – makes a textual documentary voice palpable and demands a listening or a channelling of attention towards this voice, even as the vehicle or medium that imbues it with force disappears under it. When it is restated in this way, the disappearing materiality of voice in documentary resonates with what Miaden Dolar argues is the fate of voice in the Western metaphysical tradition, and that is to disappear. ‘The voice’, Dolar writes, ‘is the instrument, the vehicle, the medium, and the meaning is the goal. This gives rise to a spontaneous opposition where voice appears as materiality as opposed to the ideality of meaning. The ideality of meaning can emerge only through the materiality of the means, but the means does not seem to contribute to meaning’ (2006: 15). Voice, in this tradition, is conflated with meaning, and specifically, with logos: the referential content or message of speech. The embodied dimensions of voicing, including the phonic qualities of the voice in speech (accent, intonations, timbre, affectations) and non-linguistic, gestural or pre-phonetic utterances (for instance: sighs, echolalia, affectations, laughter, hiccups, stammers) are considered potential obstacles to voice, and thus to the sovereignty of the subject (Weidman 2015: 233) – except in those cases where social and political norms imbue this matter with the capacity to disappear, marking it as non-matter or as a neutral, seemingly immaterial norm (20).

Could we not think of the metaphysical discourse of voice as an audit, or, that which exceeds and conditions hearing, and organizes the field of the audible in documentary, in the form of the voice metaphor? Once we do so, other questions emerge: as a sonic metaphor for point of view, how has voice (as the residue of a logocentric tradition) given rise to aesthetic and rhetorical conventions of voicing in documentary that in turn amplify some features of reality and muffle others? What can we learn from these vocal conventions about what counts as a voice and the world such a voice exerts us to share, and conversely, about the sonic and gestural matter that counts but does not count? Or, to put it in terms of the distinction between hearing and listening: What ideologies about what counts as a voice do we passively receive and tacitly legitimate when we actively attend to a textual voice in documentary and respond to its hail? What invisible and imperceptible material supports – or conventional arrangements of embodied sounds and gestures – induce us to ‘prick up our ears’ and sense such a textual voice or perspective?

Sonali Gulati’s Naïmi by Day, Nancy by Night brings concrete detail to these abstract questions. Gulati’s film is representative of a subgenre of recent call centre documentaries that focus on accent as an embodied material support that imbues speaking voices with ideological meaning even as it is obligated to disappear. At a thematic level, this film pinpoints the ‘neutral accent’ in which call centre agents are trained as an important site in which the social, lingual and geopolitical norms of the audit are consolidated under conditions of neoliberal global capital even as these norms go unheard. Although in principle and ethos, Gulati works against the prevailing norms of the audit, the vocal hierarchies and distinctions that result from her attempt to standardize accented speech and render it transparent illustrate its powerful yet imperceptible influence. I turn now to a discussion of these issues, with the following proposal: the voice of a documentary film (‘that which conveys to us a sense of a text’s social point of view, of how it is speaking to us and how it is organizing the materials it is presenting to us’, as Nichols puts it) becomes discernible as an audit (i.e. as a means not only of speaking but of listening to the world and organizing an attentional economy).
only when there are embodied obstacles to voicing present, whether explicitly or implicitly, as a disturbance or an interference.

The auditory patdown

Gulati offers the following anecdote to explain why she was instigated to make her film: having grown accustomed to hearing her name mangled by American telemarketers (‘Sonali Gelato’, ‘Somalia Gelatin’), she was caught off guard when she received a telephone call from an agent selling Visa Platinum credit cards who pronounced her name ‘as perfectly as people back home’, even though the agent introduced herself as ‘Nancy Smith’. Gulati would learn, by engaging her in conversation, that her given name was Nalini and that she was calling Gulati in Philadelphia from a call centre in New Delhi, where Gulati grew up.

Gulati’s is one of several non-fiction films produced in the early 2000s that investigate the vocal norms in which agents in Indian call centres are trained, including Diverted to Delhi (dir. Greg Stitt, 2002), Bombay Calling (dir. Samir Mallal and Ben Addelman, 2006) and John and Jane (dir. Ashim Ahluwalia, 2005). These films share a fascination with the ‘neutral accent’ of call centre agents as a type of mask that disguises and displaces their class origins, and thus as a symptom of the homogenizing and reterritorializing forces of neoliberal, ‘Anglobalizing’ capital. Each of these films goes behind the scenes of the call centre industry in India to show us the face behind the voice, so to speak. Thus, the democratizing impulse of ‘giving a voice’ (or ‘listening with empathy’ to ‘histories from below’, to borrow Nichols’s language) takes the form, in these films, of ‘giving an image’ to the faceless, nameless call centre agents who communicate using pseudonyms and pseudo-voices. In their cultural, economic, textual and spectatorial contexts, these films closely resemble the exilic and diasporic modes of filmmaking that Hamid Naficy dubs ‘accented cinema’ (2001): they are produced by diasporic filmmakers who have been displaced from their places of origin, under interstitial and transnational contexts of production, and at a textual level they invite a spectatorial engagement that is attentive to linguistic heterogeneity and the necessity of translation. These shared textual and contextual features prompt me to refer to these call centre films as a documentary subgenre even though they vary widely in production context and idiom. For reasons of space, I have chosen to isolate and focus on Gulati’s film as exemplary of this documentary subgenre.

Linguistically speaking, accent is a fuzzy term that is colloquially used to reference a specific ‘way of speaking’ (Lippi Green 1997: 42-44). Shilpa Dave, author of Indian Accents, explains: ‘An accent involves verbal intonations that stress particular syllables so the manner of speaking is just as significant as what is being said … Accent not only includes tonal qualities but also involves word choice, arrangement of words, and cultural expressions that are rooted in national (and regional) expressions of identity’ (2013: 2). Every native speaker of English has an accent demarcated in geographic and social terms, but ‘when a native speaker of a language other than English acquires English, accent is used to refer to the breakthrough influence of native language phonology into the target language’ (Lippi Green 1997: 42-43). The notion of an accent is thus inherently comparative, appearing only in comparison to ‘normal’ or standard speech; empirically there is no such thing as an unaccented voice or a neutral accent. India has a high concentration of English speakers – a legacy of its colonial past – which has made it an attractive location for the business process operations (BPO) of multinational corporations. However, the phonological influence of regional dialects, or the ‘Indian accent’, presents a hindrance, one that the call industry attempts to minimize by stressing a ‘neutral accent’ that is at once ‘global’ and ‘standard’ English (Maitra 2017: 7).

In its inception in the late 1990s, the call centre industry in India operated under a ‘cloak of secrecy’ because American and multinational corporations feared customer backlash for outsourcing their BPO (Carillo Rowe et al. 2013: 9). Aimee Carillo Rowe, Sheena Malhotra and Kimberlee Pérez note: ‘Agents were instructed to act as if they were stationed in the United States; they were required to take pseudonyms, speak with American accents, and deny their geographical and social locations as Indians’ (Carillo Rowe et al. 2013: 9). Although the mandate of secrecy has since fallen, the practice of masking call centre operations still permeates the Indian call centre industry. This mandate, intended to minimize aural and cultural dissonance for native-English-speaking customers in the United States, the UK, Canada or Australia takes the form of rigorous voice training, accent neutralization and cultural competency programmes that can last anywhere from two to eight weeks (10-11). Accent neutralization has marked class implications. In a recent ethnography of call centres in New Delhi, A. Aneesh points out that many call centre trainees tend to be middle- and lower-middle-class migrants from smaller Indian towns who are trained to perform a globalized class identity by adopting Anglicized pseudonyms, ‘switching off’ local and regional linguistic habits and imitating common
features of English speech that persist, to degrees, in the four aforementioned dominant accents (2015: 6–8, 62; Maitra 2017). Accent neutralization aims to scrub off the marks of an accent’s place of origin so that the hearer cannot place it (Aneesh 2015: 4); thus, the neutral accent does not allude to a pre-existing reality so much as it performatively calls into being a certain placelessness, an auditory non-place (59; also see Augé 1995).

Gulati’s anecdote about Nalini/Nancy offers two insights regarding the audial interpellations of the neutral accent. First, this anecdote attests to the capacity for misdirection at the heart of interpellation. Gulati fails to recognize herself in the content of the call, that is, in the call of capitalism, and does not act accordingly. But rather than simply rejecting the call by handing up (the usual response of unwilling customers), Gulati concentrates on the vocal medium and not the message being conveyed by the telemarketer. Nalini/Nancy’s Indian accent makes Gulati suddenly aware of the material support and embodiment of the agent’s voice as an obstinate residue or remainder of accent neutralization (Dolar 2006: 20). Here, another powerful interpellation works in place of the failed one: the ‘correct’ pronunciation of Gulati’s name by Nalini/Nancy triggers an auditory sensation of linguistic and geographic (un)belonging for the filmmaker, hailing her as a New Delhi-ite, even though Gulati had fashioned a new class identity for herself as an expatriate student in the United States. In that instance, Gulati is constituted as a member of an ‘Anglobalized’ and increasingly placeless linguistic community as a response to a call – an imaginary community that is constituted in and through the auditory interpellation of accented speech.

Gulati’s anecdote also points to a second register of the audit/audition that the practice of accent neutralization attempts to anticipate, mediate and obviate, however unsuccessfully, by minimizing the traces of the body in the voice. I am referring to the ‘sound hermeneutic’ (Altman 1992: 252) activated by an acoustical voice, especially one that is accented or somehow marked in auditory terms as ‘other’. We might think of this hermeneutical maneuver as an ‘auditory patdown’ or frisking in which the listener endeavours to visualize the embodied source or location of an unseen voice based on the audible evidence of its ‘skin tones’ (Chow 2014: 7–9) – evidence that only becomes palpable because of its departure from a perceptual norm that is transparent and thus inaudible. It is precisely this racialized presence of the body in the voice that the call centre industry aims to obfuscate or ‘mask’ by training agents in a ‘neutral accent’ that simultaneously masks the accented nature of listening by coding the listening ear of its ideal auditor as a ‘colour neutral’ norm. As a listener, Gulati departs from this colour-neutral norm, and this triggers a minor response to the aforementioned sound hermeneutic: Gulati’s detection of Nalini’s accent serves as a point of narrative intrigue that structures her quest to ‘unmask’ Nalini’s placeless, acousmatic phone voice by localizing it in a ‘real’ face, body and story – that is, to give her back her voice by restoring her identity.

How should we think of this attempt to give voice to – or to lend an ear to – socially constructed sonic identities associated with voicelessness? Do call centre documentaries like Gulati’s resist the audit, by unmasking the ‘neutral accent’ of call centre agents as well as the myth of neutral listening as twinned illusions or obfuscations? Or do the Anglobalizing mediations of the call centre haunt their own excavations of phonic and linguistic heterogeneity, leading them to mask their own accentedness – their own audit – even as they pursue the goal of auditory disillusionment? Gulati’s film brings into focus what I see as a central site of tension within as well as the potential of documentary: that between the destabilizing diversity of vocal soundings opened up by the documentary encounter and the lingering metaphysical imperative that filters, hierarchizes and neutralizes this diversity in the name of ‘giving voice’. This tension, or dialectic, is not new. As Trinh-T. Minh-ha and Brenda Longfellow once argued with respect to ethnographic documentary, the modernizing drive of ‘giving voice’ is at once a symptom and vehicle of the objectification of the world through linguistic colonization. Longfellow notes that ethnographic films from 1930 to 1970 tended to reiterate the political binarisms of colonialism at an aesthetic level, as a ‘separation of space and voice […] where voice is ascribed to the white European investigator and image to the bodies of colonial others’ (2004: 338). This binary resulted in other discriminations or exclusions; to quote Trinh: ‘Language as voice and music – grain, tone, inflections, pauses, silences, repetitions – goes underground. Instead people from remote parts of the world are made accessible through dubbing/subtitling, transformed into English-speaking elements and brought into conformity with a definite mentality’ (1991: 60).

The call centre documentary illustrates the continuation and mutation of these dynamics in the era of neoliberal, Anglobalizing capital: an era in which voice, and not image, has become a vehicle of the ongoing objectification and fragmentation of the (post)colonial other. The phenomenon of accent is enabling in this context: it introduces a spectrum and hierarchy of cultural values in relation to the ‘how’ of voicing (grain, tone, inflections etc.) and its linguistic ‘contents’ where Trinh sees a dualistic gridlock. One scene in particular
from Gulati's film dramatizes, in both content and form, the internalization of such a spectrum as a hierarchy — one in which a 'neutral accent' however chimerical, is deemed indispensable for social advancement. My reading of this scene channels Gulati's own attention to the contradictions between the film's message (its voice) and its vocal medium (its accent), as well as the economies of attention that emerge from their cleave. I offer the following proposal: we can only discern the audit of a documentary film in its accentedness by paying attention to the neutralizing masks and accentuating marks that shape our sense of its textual voice.

'Going nowhere': Masking and marking

The scene takes place midway through Gulati's film and features an excerpt from a failed job interview. We learn at the outset of the scene, from Gulati's voice-over commentary, that it is the candidate's fifth attempt at procuring a job as a call centre agent, and we learn at its conclusion that he has failed the interview yet again, because of his thick North Indian accent. The interviewer, a young Indian woman who speaks English fluently, but whose accent is more difficult to place, remarks to Gulati, who is behind the camera, that the candidate's pronunciation is 'nowhere'.

I would like to draw the reader's attention to two observations regarding this scene that have to do with the audiovisual representation of accented speech — or, to restate this chapter's conceptual refrain, the manner in which Gulati's film audits the job interview (itself an audition of sorts). The first, which will be readily apparent, concerns the way the voice of the candidate, a non-native English speaker, is heard and judged by himself and his interviewer as lacking. The second may be less apparent, even imperceptible: I am intrigued by how this structure of self-consciousness manifests in Gulati's film in the form of textual signs or conventions (voice-over commentary, subtitles, captioning) that produce a hierarchy of accented voices based on their proximity to an unspoken yet palpable norm. To allow close attention to these audiovisual choices in relation to the word choices of the three subjects, I have transcribed the interview scene in detail. The interview takes place in English and is captioned in English, except for Gulati's off-screen voice-over, which is not captioned or subtitled, and the interviewer's final statement, which is subtitled rather than captioned:

Gulati (off-screen voice-over heard over street scenes of New Delhi shot at night — no captions or subtitles): It hasn't taken me long to realize that call centre jobs are indeed the most sought after jobs in urban middle-class India today. I've learned that there are three million people scrambling to land a prestigious job as a telemarketer. I even met with a structural engineer who's desperately trying to get a job at a call centre. This is his fifth attempt at a job interview.

Candidate (framed in medium shot — captioned throughout): Spectramind, I was up in fourth round, fourth round I was out.

Interviewer (off-screen — captioned throughout): What went wrong? What do you think must have gone wrong in Spectramind?

Candidate: I had a heavy regional accent in my English.

Interviewer (off-screen): What kind of a regional accent are you talking about?

Candidate: North Indian accent ... Then I tried copying National Geographic's channel. Whatever they spoke, I tried it. I bought a tape recorder. I spoke, I heard what my voice looks like. I tried to improve on it. Image cuts to medium profile shot of interviewer. I have a friend. She's working at EXL. She helps me a lot in it. Image cuts back to candidate.

Interviewer (off-screen): Okay ... Which was the last movie that you saw?

Candidate: Last movie was Face/Off. Just 2 days before I saw it. I had seen it before, in between I saw it.

Interviewer (off-screen): Two days before where?

Candidate: On the TV, yesterday night I think it was ...

Interviewer (off-screen): Can you describe it?

Candidate: Okay.

Interviewer (off-screen): Narrate the movie to me.

Candidate: In that movie there are two main characters, Nicholas Cage and John Travolta. Nicholas Ca ... sorry John Travolta, he's a cop, and this Nicholas Cage, he's a criminal. What he does ... he makes arrangements with his friends, captures the other fellow, takes him to the hospital, they change their faces. All the other assistants ...

Interviewer (interrupts from off-screen): Change their faces?

Candidate: Yes.

Interviewer (off-screen): How do they do that?

Candidate: Some medical technology. He feeds them and changes it.

Interviewer (off-screen): Mm hmm.

Candidate: Then what happens after that, all his assistants, he kills them. So now there are only two people who know, Nicholas Cage and John Travolta, which is the real one!
Interviewer (off-screen): Right, thank you.
Interviewer (gets up, leaves the room): Welcome. Image cuts back to interviewer.

Interviewer (to an unseen interlocutor, possibly Gulati – subtitled as 'When he spoke, his tongue used to hit his teeth. He had a lisp and he was stammering in between'): To ... it was like, he had a lisp also, jab woh bolta tha to his tongue used to hit his teeth. She demonstrates: 'duh-duh' ... jab main ... voh ... lisp aa jaata hai na ... Then he was stammering also beech beech mein. Then he said, 'I don't want to be a follower, I want to be a leader,' To ... pronunciation was, like, nowhere. Next candidate.

This seemingly unremarkable scene from Gulati's film stands out to me because, in her attempt to listen empathetically to the candidate's predicament, Gulati also visually reinforces his voice as 'lacking': his accented speech requires captioning in English to ensure its intelligibility. In the judgement of his interviewer, the candidate's pronunciation is 'nowhere'. The implication is that the young man is a poor candidate for a call centre job because he has a 'nowhere' voice, a voice whose audible geographic specificity portends limited class mobility for the speaker. If he wants to 'go somewhere' in life, that is, if he desires economic and geographic mobility, however illusory or virtual that mobility might be, he must be able to scrub, mask or disguise his identity: not his face, as one might surmise from his rather ingenious plot summary of the film Face/Off, but the audible vocal markers that 'place' his voice in a location associated with underdevelopment. But this candidate's larynx, teeth, tongue and neural wiring are so stubbornly moulded by his inhabitation of a regional Indian dialect that his accented English utterances tether him to a local identity that he cannot escape or efface.

I am also intrigued by the presence in this scene of a formal hierarchy of audiovisual conventions that positions the three speakers (Gulati, the female call centre agent and the candidate) at various rungs along a class system of accents. Gulati's expository first-person voice-over, which is in direct and seemingly unmediated conversation with the audience, is at the apex of the hierarchy. Gulati's speech is not captioned, subtitled or anchored in an image, even though it is not unaccented. Her American-influenced lexical stresses and drawling vocal delivery intimate her expatriate status as a film student in Philadelphia (a piece of information Gulati volunteers early on in the film) even as its melodic up-down intonations, and rounded vowel sounds, evidence of a British English-medium education, respectively hint at her South Asian nationality and middle to upper-middle-class status.

In contrast, every word uttered by the job candidate is captioned, lassoing our gaze to the content of his words lest we be distracted by his 'erroneous' syntax and pronunciation (he rolls his Ds and Rs and strings words together), which indicate the audible influence of his native language (my guess is Punjabi) on his manner of speaking English. Visually, these captions resemble subtitles, which were first used in a documentary context in the 1970s by the ethnographic filmmakers Timothy Asch, John Marshall, and David and Judith MacDougall. As an alternative to dubbed voice-over for translating indigenous dialogue, subtitles offered a way of retaining the auditory qualities of the original speech, giving access thereby to the intellectual and emotional lives of subjects, and ceding the authority and interpretive monopoly of a voice-over to the synchronous voice of the ethnographic subject (MacDougall 1998: 7, 116). But the visualization of the job candidate's accented English as captions/subtitles has an effect contrary to Gulati's democratic aim of restoring this subject's sovereignty by 'giving him an image' – that is, by showing us the 'real' names and faces behind the anonymized call centre voice. As David MacDougall notes, subtitles can establish what is written as more definitive than what is said, producing the effect of a dialogue between the filmmaker-translator and the audience 'about' non-Western people even when the latter are seen and heard speaking in their own voices. The visualization of the oral as subtitles, he argues, can make audiences visually word-oriented and dependent, inducing a sense of passivity that detracts from a more active engagement with the sonorous grain and non-verbal expressive range of the speaker (1998: 174–175).

In Gulati's film, too, the candidate is transformed into an object/image in a dialogue between Gulati and the interviewer, even as he is metaphorically 'given a voice'. The interviewer's voice is less easy to place, and this may well be because she has been trained in the 'neutral call centre accent' that mimics a mixed set of features of British, American, Australian and Canadian English and is thus ironically 'placeless', even as she criticizes the candidate's voice as a 'nowhere' voice. Her words are captioned as well, but the motivation for this decision could be technical, since she is off-screen and un-miced for much of the exchange, and we cannot see her moving lips. When she finally appears on screen, the interviewer's words are subtitled rather than captioned since she switches between Hindi and English when addressing Gulati, relaxing the contrived norms of the interview and acknowledging Gulati's bilingualism as well as their membership in a shared linguistic class that does not include the job candidate: a class capable of 'masking' its origins. Even though we lose some of the content
of what the interviewer says in translation, the suggestion is that her words do not require captioning, since the sound of her voice, unlike that of the candidate, approaches the aspirational linguistic norm of a neutral, placeless accent that can 'go anywhere.' That Gulati has achieved this global standard of cosmopolitanism by travelling to and living in the United States (a status that fewer call centre agents, if any, do; see A. Aneesh 2015: 64) is confirmed at the end of her trip to New Delhi, when she is offered a job as a voice and accent trainer at a call centre.

In conclusion, how should we think of these textual signs or audiovisual conventions: Gulati's off-screen commentary, the subtitled speech of the interviewer and the captioned speech of the candidate? Are these perfectly ordinary and unremarkable conventions – conventions whose ordinariness signifies the arrival of documentary as an 'accented cinema'? Or are they techniques of accent neutralization? Are they a cloak or disguise that simultaneously mask the accented basis of some voices and mark or accentuate the 'otherness' of others, thus enabling an unspoken yet palpable norm – a textual voice – to emerge? To me they are both. Their unremarkable and conventional quality might even be evidence of the successful camouflage of the accented relation between this textual voice and the listening ear that it hails ('Hey, you!'). But like the call centre voice, perhaps these conventions are an imperfect disguise – an unsuccessful attempt to neutralize the film's accent – that points to the capacity for error and misdirection built into every interpellation. Here I'm reminded of Gulati's accidental misinterpellation by Nancy Smith, by the 'correct' pronunciation of Gulati's name by Nalini/Nancy, which calls her attention to the medium of the voice rather than the meaning being conveyed. Just as Gulati fails to recognize herself in the call of capitalism but, recognizes herself as a member of an increasingly 'placeless' linguistic community that is both nowhere and everywhere, I find myself similarly misinterpellated by the neutralizing masks and accentuating textual marks of Gulati's documentary, which call my attention not to the film's voice but to its auditor. I'd like to think that confronting these conditions of hearing and the economies of attention that they sustain, or conversely, fail to contain, could replace 'voice' as a metaphor for how documentaries hail our attention.

Acknowledgements

I am immensely grateful to my colleague Amelie Hastie for her astute and eye-opening comments on multiple drafts of this chapter. I also thank audiences at

Notes

1 I thank Michael Renov for this observation.
2 In contrast to the other films, which are fairly conventional documentaries shot on video, Ahluwalia's film is shot on 35mm and includes documentary material originally shot on video that was subsequently re-enacted and re-shot on film, producing an observational effect that is estranging and uncanny, much like the call centre voice itself. This interesting formal choice deserves a separate and longer discussion, but I will suggest here that the more conventional treatments of this voice, of which Gulati's film is exemplary, can be equally estranging. I hope to discuss these other examples of the call centre documentary subgenre in another venue.
3 The caption notes: EXL (call centre).

References

Gerillo Rowe, A., S. Malhotra and K. Pérez (2013), Answer the Call: Virtual Migration in Indian Call Centers, Minneapolis, MN: University of Minnesota Press.
The Fundamental Lie: Lip Sync, Dubbing, Ventriloquism and the Othering of Voice in Documentary Media

Patrik Sjöberg

This chapter will address the complex relationship between the represented body and face of a subject and the vocal expressions attached to them in documentary media. I will do this by examining instances in documentary where one voice has been substituted for another, or rather, where a voice has been attached to a body to which it doesn't belong. The discussion borrows from current research on the voice, the sonic arts, dubbing, ventriloquism, media history and lip sync. Although several works are mentioned in the discussion, this chapter focuses on a series of short films, *Covers* (2006 – ongoing), by artist and filmmaker Adie Russell. In her series, Russell re-enacts interviews and conversations with famous people from the past through her individual performance and lip sync in front of a camera. Russell’s films lend themselves to a discussion of audiovisual representations of embodied speech acts and what Rick Altman (1980: 6) once referred to as ‘the fundamental lie’ at the heart of sound film; that is, ‘the implication that sound is produced by the image when in fact it remains independent of it’.

In a brief description of her works on her website, Russell accentuates their ‘temporal spaces’. These spaces get superimposed in the films, Russell suggests, because ‘the historical moment of the audio recording’ shares the space, both with ‘the moment inhabited by myself and the viewer in the present’. Russell is well aware of the complexities that arise from such a seemingly simple trick of voice displacement. What is at stake, and what my discussion here elaborates on, includes questions of identity and performance; the encounter with archival audio; mediations of speech acts and the idea of an unstable audiovisual speaking