### Vocal Alterities: Voice-Over, Voice-Off and the Cultural Addressee

**Esther Peeren**

#### The Limit of Dialogism

Mikhail Bakhtin distinguishes dialogism as a privileged form of intersubjectivity that institutes a productive relation with alterity. As a concept, dialogism refers, first of all, to the involuntary relations of responsiveness between utterances. In this guise, it was taken up by Julia Kristeva and developed into intertextuality. But dialogism also appears as a strategy for establishing productive intersubjective relations between individuals, social groups, and cultures. The aim of dialogism in this sense is to produce what Bakhtin calls “active responsive understanding” (1986a: 69), founded on sustained difference and exteriority. Exceeding dialogue proper (as a back-and-forth discussion between two or more speakers) and opposing itself to dialectics (as a teleological move towards synthesis), dialogism denotes an active, answerable interaction with alterity that implies neither negation nor assimilation.

Less a figure of harmonious agreement than one of sustained difference, dialogism functions, in the words of Paul de Man, as “a principle of radical otherness” that aims to “sustain and think through the radical exteriority and heterogeneity of one voice with regard to any other” (109). Dialogism marks the preservation of vocal alterity, but such preservation implies an attitude of ethical responsibility that cannot be taken for granted. In *Gender Trouble*, Judith Butler points to the cultural and historical specificity of the notion of dialogue and argues that “the power relations that condition and limit dialogic possibilities need first to be interrogated” in order not to lapse into a Habermasian model that assumes equality and shared presuppositions and goals (20).

I want to theorize this intersection between dialogism and social power relations by combining Bakhtin’s concept of the *superaddressee* and V.N. Voloshinov’s related
notion of the potential addressee into an agency I will call the cultural addressee. This agency relates the voice and ear of the individual speaker / listener to the collective voice and ear of a particular community, circumscribing this community’s attitude towards radical alterity. Tied to the voice or tongue (tongue-tied, as it were), the cultural addressee signifies the utterance’s orientation towards a potential understanding that functions as a precondition for its being spoken and for its being heard. I will approach the cultural addressee through an examination of the audio-visual rendering of vocal alterity in two popular cultural objects. First, there is the use of voice-over in the American television series Sex and the City. The series revolves around the love-lives of four single women in Manhattan and each episode employs voice-over to have Carrie, its central protagonist, pose a question ostensibly part of a newspaper column she writes, but whose articulation and addressivity exceed Carrie and her readers. Second, there is Michael Apted’s 1994 film Nell, which stars Jodie Foster as a so-called “wild woman” found in the woods speaking an “unknown” language. Foster’s initial appearance as voice-off – an intra-diegetic voice without a visible speaker – emphasizes the threat her character’s radical alterity poses to the town community. Both voice-over and voice-off prompt a splitting of the voice’s addressivity into the embodied ear of the actual listener and an altogether less substantive auditory agency.

Voice-Over
In Sex and the City, when Carrie’s voice-over formulates her episode-framing questions, she is usually alone in her apartment, writing on her laptop computer. No direct, personal addressee is available. Moreover, posing the questions in voice-over instantly detaches the words – spatially and temporally – from her enunciating presence. Carrie’s voice-over cannot be straightforwardly categorized in film-theoretical terms. Unquestionably interpretative but at the same time closely linked to Carrie’s body and thoughts, the voice-over impossibly appears as both interior monologue and commentary, destabilizing the distinction between intra-diegetic and extra-diegetic filmic space. It is, furthermore, both synchronous and asynchronous: although it speaks in the past tense, it synchronizes visually with the question as Carrie types it into her computer. Finally, Carrie’s voice-over undermines the traditional gendering of the authoritative voice-over as male, non-embodied and non-diegeticized. Her relation of authorship to the newspaper column of which the question is part points to a possible recuperation of the female voice-over as diegeticized, embodied, and feminized, yet still authoritative.²

At the same time, the precarious positioning of Carrie’s voice-over across different levels of enunciation emphasizes how even an authoritative voice-over only ever poses as the point of discursive origin: aligned with the cinematic apparatus, but never fully equal to it. Carrie’s voice-over, which is both contained and containing, enunciated and enunciating, destabilizes the voice-over’s claim to diegetic authority by pointing
to the inevitable alterity of the utterance to its speaker. This loss of authority extends to the question’s addressee. In several episodes, after the voice-over poses the question, the camera pans out of Carrie’s apartment window, moving upwards as it zooms out, as if taking the question to a higher level, to a more fundamental addressee than the readers of her column. Dislodged from mouth and ear, the question is deferred: but onto what?

The French psychoanalyst Jean Laplanche has suggested that every intersubjective message invariably involves an enigmatic addressee, a future “nameless crowd, addressees of the message in a bottle” looming behind the personal addressee (224). In the Sex and the City episode “The Drought” Carrie asks: “How often is normal?” first addressing herself, then her girlfriends, then the whole of New York. Over a shot of a crowded street – a powerful visualization of Laplanche’s nameless crowd – we hear Carrie’s voice-over: “There are 1.3 million single men in New York. 1.8 million single women. And of these more than 3 million people, about 12 think they’re having enough sex. How often is normal?” Here, the personal addressees are separated from the enigmatic addressee.

What Laplanche indicates, however, is that the relationship is one of simultaneity: the intersubjective message at the same time addresses the nameless crowd. In fact, the individual addressee functions almost as an excuse for addressing the enigmatic addressee, as an alibi for releasing a message in a bottle. Read this way, Carrie’s original question as to whether she is having enough sex with her boyfriend is already a questioning of the crowd, of its collective normativity. Significantly, this crowd is not abstract or placed in an undefined future, but concrete, situated, and contemporary, invoking less the metaphor of a message in a bottle than that of an advertising billboard or, as the episode suggests, a computer screen.

At “How often is normal,” the camera cuts to an extreme close-up of Carrie’s computer screen, where we see the question appear letter by letter. The question’s materialization on the shot-filling blue screen dissociates it from both author and concrete addressees. As writing, it evokes Derrida’s notion of iterability, a structural potential for repetition with difference that enables the letter to break with “the collectivity of presences organizing the moment of its inscription” (1988: 9). As digital or virtual writing, it recalls the acceleration of iterability Derrida associates with the new techno-media, where present time is divided from itself, place is freed from territorial rootedness, and where the singular logic of the event becomes ever-more immediately intertwined with the repetitive logic of the machine (2002: 210). In other words, virtuality signals the accelerated, enhanced interference of alterity in identity, of absence in presence.

The close-up of the computer screen, with its unstable pixels, suggests that it is not Carrie who writes, but a more enigmatic author, visualized by the provocatively blinking cursor, which always seems on the brink of writing (or erasing) of its own accord. And again, the same applies to the addressee. Reproducing digital text is
more straightforward than reproducing handwriting or typewriting, so that it can address itself again and again. Digital text renders tangible our lack of control over what we write and whom it addresses.\textsuperscript{4}

As digital writing, Carrie’s query transforms the temporality of Laplanche’s enigmatic addressee. The computer, associated with e-mail and instant messaging, is an appropriate metaphor for the way our words address themselves not so much to a nameless future public, but to a nameless crowd that intervenes at or even before the moment of enunciation. Thus, whereas Laplanche conceptualizes the actual addressee as an alibi for the enigmatic addressee, who stands behind him in another, future time and place, “The Drought” suggests a reversal: Carrie can only address the individual addressee (her boyfriend, whom she does not dare ask whether they are having enough sex) through the implicit addressees of crowd and screen. The enigmatic address becomes the condition for the intersubjective address; it has to sanction the question before it can reach the other. As in Derrida’s spectral hauntology, the absolute alterity of the specter has to be addressed before we can address concrete others: “Such an address is not only already possible, but [. . .] it will have at all time conditioned, as such, address in general” (1994: 12). The spectral address precedes the intersubjective address.

Yet Derrida’s spectral addressee, as an ascetically messianic and featureless “alterity that cannot be anticipated” cannot be relied upon for a response (1994: 65). As with Laplanche’s enigmatic addressee, the speaker is passive in relation to it. Neither Laplanche nor Derrida can account for the active urgency of Carrie’s questions, which expect to be heard and answered in the present. Far from leaving a message in a bottle or attending an undefined entity with no guarantee of reply, her acute questioning of crowd and screen summons a more responsive and socially defined agency, which I will delineate in my combined reading of Bakhtin’s superaddressee and Voloshinov’s potential addressee.

**From Superaddressee to Cultural Addressee**

Bakhtin describes the superaddressee as follows:

\textit{In addition to this addressee (the second party), the author of the utterance, with a greater or lesser awareness, presupposes a higher superaddressee (third), whose absolutely just responsive understanding is presumed, either in some metaphysical distance or in distant historical time (the loophole addressee).} (1986b: 126)

The superaddressee assumes an active relation to the word, imparting an exo-topic validation to the subject’s speech no longer linked to individual others, but to a higher instance that stands above concrete intersubjectivity, making it possible.

My understanding of Bakhtin’s superaddressee differs from that of other critics, who have argued that it signifies an ideal, abstract or transcendental instance of deferred understanding that is never concreticized. Craig Brandist views the superaddressee as...
a godlike instance, an “eternally deferred supreme judge who views the social world from without” (171). Ken Hirschkop emphasizes its messianic dimension, defining the superaddressee as “the one who somehow, beyond our fate in actual history, redeems our words by understanding them correctly” (397). These interpretations are supported by Bakhtin’s undeniable religiosity, but they bar the superaddressee from functioning as the addressee of Carrie’s acute questionings of the distinct present of New York.

Bakhtin’s sparse remarks do harbor an opening towards an interpretation of the superaddressee as a situated, socially specific agency. At one point, he presents it as taking on different names in different eras and cultures:

> In various ages and with various understandings of the world, this superaddressee and his ideally true responsive understanding assume various ideological expressions (God, absolute truth, the court of dispassionate human conscience, the people, the court of history, science, and so forth). (1986b: 126)

This turns the superaddressee from a universal ideal into a socially, historically and ideologically variable construct, located in the social world rather than beyond it. The superaddressee, in outlining the norms of intelligibility, provides the utterance with a defined direction and although this direction may at times be expressed as God, it is not necessarily divine. Unlike Laplanche’s enigmatic addressee or Derrida’s specter, the superaddressee does not remain indeterminate, but is required to be filled by the appropriate ideological expression before the subject can speak with the confidence that she will be heard and understood.

Carrie’s acute questioning of crowd and screen invokes not a maximally distant superaddressee, but an active force of linguistic regulation rooted in her community. Such a force is recognizable in a reading of Bakhtin that construes the superaddressee as the culturally shaped horizon of speech. But it is perhaps more immediately perceptible in V.N. Voloshinov’s work.

In “Discourse in Life and Discourse in Art,” Voloshinov delineates two social aspects of the utterance’s addressivity: the third participant and choral support. The third participant is defined as the personified object of the utterance, characterized as “nameless” and “verging on apostrophe” (103). When taken to its extreme, this third participant “becomes the source of a mythological image, the incantation, the prayer, as was the case in the earliest stages of culture” (103). Here, we encounter once more the featureless, messianic agency of address. However, the second social aspect of the utterance’s addressivity ties it to a concrete community and its specific notion of intelligibility. Voloshinov uses the term choral support to denote the utterance’s relation to “the assumed community of values belonging to the social milieu wherein the discourse figures” (103). If choral support is lacking, “the voice falters” and the utterance becomes unintelligible (103). Choral support, therefore, ties the utterance’s addressivity to the discursive norms of a specific, situated community.
In *Marxism and the Philosophy of Language*, Voloshinov further argues that each utterance addresses itself to a presupposed potential addressee, who takes the form of “a normal representative of the social group to which the speaker belongs” (85). This addressee is emphatically not abstract or universal, because, according to Voloshinov, “even though we sometimes have pretensions to experiencing and saying things urbi et orbi, actually, of course, we envision this ‘world at large’ through the prism of the concrete social milieu surrounding us” (85). Each person has a more or less stable social audience that limits the scope of the potential addressee and makes it specific.

There is no ambiguity here: Voloshinov’s potential addressee and choral support do not stand above the speech community, but constitute it, cementing it together. Members of a speech community do not imagine themselves understood by a transcendent entity, but by those around them: “I give myself verbal shape from another’s point of view, ultimately, from the point of view of the community to which I belong” (86). Accordingly, Carrie strives to shape her utterances to the norms of the Manhattan crowd in order to be understood by its individual members. Voloshinov, earlier and more unequivocally than Bakhtin, outlines the function of what I will call the cultural addressee as the conditional limit of meaningful speech, as that which gives us to be spoken and allows us to speak in our specific cultural contexts.

What Voloshinov does not adequately explain, however, is that the cultural addressee’s guarantee of internal understanding is predicated on exclusion and coercion. The film *Nell*, through its use of voice-off, suggests that the realization of dialogism, of taking responsibility for radical alterity, depends not just on a basic ability or willingness to speak and listen, but on the particular vocal inclusions and exclusions structured by the cultural addressee, which circumscribe ability and willingness and, consequently, may close both mouth and ear to alterity.

**Voice-Off**

Sound without a visible on-screen source was originally theorized by Michel Chion as the acousmêre. Voice-off relates specifically to the human voice divorced from a visibly speaking mouth. In contrast to voice-over, it speaks not *over* the image, but from its margins, Chion associates voice-off with (masculine) authority, panopticism and omniscience (24) and Kaja Silverman invests it with a “threat of absence” (48). Although its source is usually located, there is always a chance it will remain missing.

Voice-off approximates Derrida’s specter as a present-absent trace of marginality, intangibility and lack of body. Like the voice-off, which we cannot answer because we do not know its origin, the specter is an all-seeing agency that does not allow us to return its look: “it is someone who watches me or concerns me without any possible reciprocity” (Derrida and Stiegler 121). Although Derrida considers the voice the only graspable element of the specter – “we must fall back on its voice. The one who says
‘I am thy Father’s spirit’ can only be taken at his word” (1994: 7) – voice-off renders the voice itself spectral, to be heard but not to be touched (sensuous non-sensuous).

Significantly, Derrida’s specter not only makes use of the visor effect – the ability to see without being seen – but also of an accompanying ability to speak and hear voices, to become a “spectral rumor” whose resonance invades everything (1994: 135). Voice-off is such a spectral rumor, an auditory (dis)incarnation whose embodiment paradoxically weakens its authority, like the specter which “disappears in its very embodiment” (Derrida 1994: 6). Both voice-off and specter threaten a proliferation of meaning, compelling a desire to pin down the errant voice / image to a living or dead body. This desire drives Nell’s narrative.

The film opens with a high-angle shot of a forest, over which we hear a mumbling or singing sound that may or may not be a human voice. Later, the process of embodying the sound is initiated through a series of extreme close-ups showing hands and fingers grooming an old woman. Associated with the fingers and the hands, the disembodied sound becomes a voice-off, manifestly human and female. Full embodiment occurs when the town doctor is taken to a forest cabin to find a dead old woman and her daughter Nell, the “wild woman” who owns the voice-off.

However, because of Nell’s incongruous “wild” body and voice, the embodiment of the voice-off does not remove the threat of absence. Flesh and blood cannot exorcise Nell’s radical alterity, suggesting that the recovery of the body does not always “tame” the “wild” voice, make materiality (into) matter, or conjure away the specter. If the voice remains incongruous even or especially in its body, any move toward an origin is foreclosed. The metaphor of the voice as having “body” is lost to the arbitrary contiguity of metonymy: the voice is linked to the body but the two do not fully explain each other.

The town community fears the indeterminacy of Nell’s voice and body and seeks to control it by making Nell’s voice all body, all form without content, invoking what Silverman sees as classical cinema’s “identification of the female voice with an intractable materiality, and its consequent alienation from meaning” (61). Conceiving Nell’s speech as a purely physical, instinctual expression of her body reduces voice to matter and restricts the “migratory potential” expressed by voice-off (Silverman 84). This strategy of containment signals the dependency of sustained alterity on the other’s recognition of the utterance as an utterance. Nell (the product of a rape committed in the community) marks a radical alterity within the community’s identity, but as long as her words are not received as signs they effectively mean nothing. Nell’s voice is linked to her body, but it is not anchored in the communal body. It does not receive choral support. To keep the community safe, her utterances either have to be coded as meaningless or made to conform to the community’s understanding of intelligibility, its cultural addressee. Either way, Nell’s alterity is neutralized.

The community engages in a two-pronged strategy of, on the one hand, translating Nell’s utterances into “proper” English, and, on the other, interpreting her speech – and
her deliberate silence when incarcerated in a psychiatric hospital – as sound / quietude instead of word / silence. The distinction is Bakhtin’s:

Quietude and silence (the absence of the word). The pause and the beginning of the word. The disturbance of quietude by sound is mechanical and physiological (as a condition of perception); the disturbance of silence by the word is personalistic and intelligible: it is an entirely different world. (1986c: 133)

The voice-off in Nell’s opening scene places itself on the dividing line between sound and word, making it uncertain whether what is broken is quietude or silence. This ambiguity is maintained for most of the film, until we learn that she speaks not an “unknown” language, but merely a mangled form of English. Nell’s oscillation between word and sound shows up the relativity of the distinction and identifies the border between silence / quietude and sound / word as a socio-cultural construct, collectively established and patrolled.

This compromises the political weight of silence. As Peter Hitchcock notes, “there is no single strategy for positioning oneself with respect to silence, partly because that place is beyond voluntarism and volition (formed therefore by more than this or that individual consciousness or praxis)” (97). Meaningful silence is not simply someone who does not speak, but someone who is recognized by others as withholding words rather than mere sounds.

By defining Nell’s utterances as sounds rather than words, the community disavows her as a subject and assuages the dangerous implications of a form of speech that straddles both categories. This is why the embodiment of Nell’s voice-off strengthened the threat of absence: coming from nowhere, the voice-off could be dismissed as mere sound. With the voice-off linked to Nell’s lips, however, it becomes an unintelligible word far more threatening than an unintelligible sound. This threat prompts the community’s effort to assimilate Nell and subject her to the shared cultural addressee, which intervenes not so much after words are spoken, but before anyone can speak, before sound becomes word and “it” becomes “I.”

Nell’s problem is not that she was ever without a cultural addressee, but that her particular expression of it was shared only with her mother and a twin sister who died as a child. Both are not only dead, but repudiated by the community’s fantasy of Nell as a pre-subjective “wild child.” Not recognized as members of Nell’s first speech community, mother and sister appear as non-voices, non-specters, inaudible and visible only in the material form of death (as corpse and skeleton). Without them, Nell’s cultural addressee loses its intersubjective dimension and can no longer guarantee subjectivity and understanding to her voice. In Voloshinov’s terms, Nell is reduced to a pure “experience,” a social rootlessness that “tends toward extermination” (1986: 88). In order to make sense of – and in – her new social surroundings, Nell has to re-model her voice to the image of the town’s cultural addressee, to its “we-experience.” This becomes particularly urgent when it becomes clear that this “we-experience”
has a low level of differentiation and is unwilling to incorporate her alien voice except on condition of its full assimilation.

_Nell_ may be read as an allegory of immigration, where one community is left behind for another. The immigrant haunts her new community with her inalterable otherness, creating a fear in the face of which the community is required to take an ethical position. This position is only truly dialogic if the immigrant’s otherness and the otherness of her cultural addressee are welcomed and respected without rejection and without turning the other into another self. Derrida writes:

_One should not rush to make of the clandestine immigrant an illegal alien or, what always risks coming down to the same thing, to domesticate him. To neutralize him through naturalization. To assimilate him so as to stop frightening oneself (making oneself fear) with him. He is not part of the family, but one should not send him back, once again, him too, to the border._ (1994: 174)

_Nell’s_ use of voice-off demonstrates the extent of the cultural addressee’s normative power, which may preclude dialogic interaction with the radically other. Such interaction is possible only under a cultural addressee that values difference over sameness, outsidersness over internalization, understanding over recognition, and transformation over assimilation. For dialogic interaction to occur, each party needs to take responsibility for understanding and respecting the other party’s cultural addressee, without attempting to subsume it under its own. In Derrida’s words, “respect for the alterity of the other dictates respect for the ghost” (2002: 123). We have to take responsibility for (our cultural definition of) absolute alterity before we can take responsibility for the alterity of specific others.
Endnotes

1. See Kristeva’s “Word, Dialogue and Novel” and “The Bounded Text.”

2. For film-theoretical accounts of the voice-over see Doane; Chion; Silverman.

3. George Landow calls the blinking cursor “a moving intrusive image of the reader’s presence in the text” (44). I regard it more as a sign of the presence in the text of the computer as “a terrible and tireless writing machine that is now relayed, in this floating sea of characters, by the apparently liquid element of computer screens” (Derrida 2002b: 123). The cursor’s oscillation between visibility and invisibility also evokes the Derridean specter as “a trace that marks the presence with its absence in advance” (2002a: 117).

4. On digital writing, see Landow; Bolter; Poster.

5. This is not to deny that God is the incarnation of the superaddressee privileged by Bakhtin. In “Author and Hero in Aesthetic Activity,” he writes: “Outside God, outside the bounds of trust in absolute otherness, self-consciousness and selfutterance are impossible, and they are impossible not because they would be senseless practically, but because trust in God is an immanent constitutive moment of pure self-consciousness and self-expression” (144). See Hirschkop for a detailed discussion of Bakhtin’s religiosity.

6. Its capacity to take different forms distinguishes the superaddressee from Habermas’ conceptualization of the ideal speech situation, where communication is guaranteed by ethical rules presumed to be universal and by a bracketing of relations of domination. T. Gregory Garvey articulates the difference as follows: “The superaddressee marks a form of undistorted communication, but a form that is contextualized, rather than universalised, as it is in Habermas’s model” (384).

Works Cited


