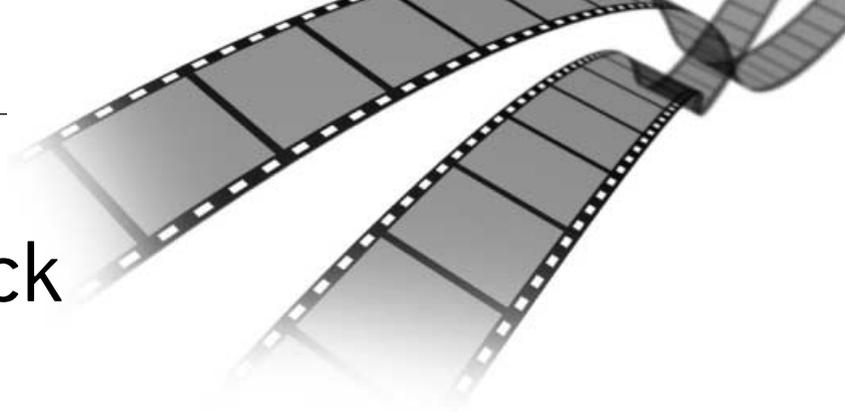


---



# Soundtrack in Mind

Edward Branigan

**Abstract:** Contemporary film theory is noted for its sturm und drang, though in the case of the soundtrack, incompatible attitudes and methods are found mostly below the surface where theoretical presuppositions are ruled by unpredictable melodic contours and accents. This article provides a comprehensive overview of philosophical issues concerning audition. It aims to orient a diverse array of sound theories in relation to a set of core issues involving perceptual processing, language, and mind. The article sounds out various cognitive frameworks, where each type of frame projects a favored description and explanation of sonic phenomena. It argues that what is heard in a sound depends on how one listens, and with what purpose.

**Keywords:** audiovisual media, cognitive film theory, framing, language, perceptual processing, reference, representation, reproduction, voice

## Talking about Sound

*Why, in the case of the ear, is there withdrawal and turning inward, a making resonant, but, in the case of the eye, there is manifestation and display, a making evident? ... Shouldn't truth "itself," as transitivity and incessant transition of a continual coming and going, be listened to rather than seen? (Nancy 2007: 3–4)*

How do film theorists talk about sound? What categories and considerations make up a particular language for explaining the soundedness of film? What "truth" is sound said to reach? The following account examines the choices that have been made in constructing various theories of the soundtrack by focusing on narrative fiction films. The aesthetics of specific types of sound—dialogue, music, noise, silence—as well as the historical development of sound theories and methods of sonic analysis fall outside the scope of this article (on methods see, e.g., Bordwell and Thompson 2008; Chion 1994; Kracauer 1997). Instead the article concentrates on the sorts of frameworks that allow propositions to be formulated about the nature and effects of what is being heard. Is there a ground for sound? What is the status of the duet between sensation and sense in film? Perhaps surprising, the article shows that attempts to the-

orize the soundtrack inevitably lead to theoretical claims about the nature and effects of imagery. As a result, in talking about sound, the medium itself becomes an object of conjecture *sub silentio*.

### **Sound Thinking**

Jonathan Rosenbaum (1978) expressed the anger that had built up over the perceived marginalization and non-theorization of sound within the reigning film theories of the 1970s, which were premised on psychoanalysis, semiotics, structuralism, Marxism, and feminism. Rosenbaum called for an entirely new vocabulary for talking about sound. He advocated a medley of cures premised on an “erotics” arising from “unconscious and collective impulses” that would take full advantage of special cases in film of “sound thinking” and “sound bullying” (1978: 39, 40). The basic question was where to begin in creating new ways of talking about sound. What would be the founding principles and how should a theory be developed? Should one begin with physics prior to metaphysics?

### **The Physical Frame**

Physically, audible sound comes from an object whose mechanical vibrations cause vibratory waves in a medium, as in air, water, flesh, or bone. Sound rubs against and within us. Sonic vibrations are of a very low frequency and their movement is slow ( $10^4$  cycles per second moving through air at 1/5 mile per second). The situation is quite different for visible light, which is an electromagnetic wave, requires no medium, and is lightning quick ( $10^{16}$  cycles per second at 186,000 miles per second).

The physical asymmetry between sound and light strongly influences our experience of audiovisual media. These physical differences suggest that sound has an entirely different relationship to our bodies than does light; that is, sound and light are differently instantiated in our daily life and language. The assumption here is that what is being lived—through our experiences in an environment, gestalt perception, handling of objects, bodily movements, plans, and problem solving—acquires special significance for our thinking. Notably when a mixture of embodied ideas must be projected into abstracta to engage a new something or some thing anew (Branigan 2006; M. Johnson 1987; Lakoff and Johnson 1999; Varela et al. 1991), such as within an aesthetic context or when devising a theoretical language about an aesthetic context.

The large difference between operating frequencies means that sound is more flexible, bending around corners, while the straight lines of light lay down a solid grid that (seemingly) makes space isotropic and offers a high degree of visual resolution where edges, colors, and distances do not usually blur. Some contemporary forms of analytic philosophy spring from these ordinary ways in which light interacts with human vision. The analysis is thoroughly Cartesian in spirit (Branigan 2006: xv passim). Its goal is to achieve a

**The physical asymmetry between sound and light strongly influences our experience of audiovisual media.**

“clear and distinct” mental conception to match the crispness and stability of visible objects: line of sight, in sight, insight. By comparison, sonic phenomena are much less “clear and distinct,” less precise for modeling a given ontology. Hence we usually speak of the cause and source of sound: someone who is walking in a boot on a floor causes a creaking noise, or else we speak of a dispositional property of the floor to make a creaking noise under certain conditions. In contrast, we speak about the effect and look of light: the book has a red cover. Thus, from an experiential standpoint, a theoretical language about sound should begin in a different place than a language about light.

Philosophers have seized on the fact that sound seems to be transitory—coming from an event that is happening somewhere at the very moment of its perception, while light seems to create permanent, intrinsic properties that are essential—being at a definite point, possessed by and defining a material thing or substance continuously (the book is red even when the lights are out). It is through an “emanation” of light, Kant says, which is unlike the motion of sound, that “a point in space is determined for the object” ([1798] 1996: vii, 156). Moreover, though we may make sounds, our bodies are unable to create light. Thus, light seems relatively external, objective, and disembodied whereas sound is within us and personal. For Aristotle color is the type of predicate that is simply “present *in*” its material object although “never predicable *of* anything” (*Organon*, “*Categoriae*”: 1<sup>a</sup> 27–28; emphases added). Light seems to determine the logic of things while sound is variable and contingent. The rhetorics of philosophy and film theory prefer metaphors based on the (apparent) stability and certainty of illumination, where “light” may be cast on a topic, rather than based on the unpredictability of sound.

### **Material Heterogeneity**

The mutability of sound as opposed to light, together with the many material, physical elements in an actual environment that contribute to the exact quality of an aural event in a medium (e.g., locale acoustics, microphone placements, character movements, limitations and distortions of technology, re-mixing by sound designers, the fashioning of synthetic sounds), has convinced Rick Altman that sound undergoes a fundamental, ontological shift when it is recorded for later presentation in a film.

*Recorded sound substitutes an apparent sound event for the original phenomenon. Revealing its mandate to represent sound events rather than to reproduce them, recorded sound creates an illusion of presence while constituting a new version of the sound events that actually transpired. (Altman 1992c: 29).*

It is true that the psychological literature documents many fascinating aural illusions and constancies. Do these departures from normal perception, how-

---

ever, warrant a far-reaching skepticism about recorded sound that can become the mainstay of an aesthetic theory (Allen 1997)?

Notice that Altman maps the duality representation/reproduction onto copy/original, illusion/actuality, and presence (in a movie theater)/absence. This amounts to thinking about filmic sound (and, separately, imagery) as a *surface* phenomenon where an ideological “exploitation of culture” (1992c: 30) occurs through “infinite possibilities for creation and obfuscation” (31). Beneath a film’s agitated, arbitrary, and seductive surface, somewhere in the *depths*, resides a (perhaps inaccessible) true core or origin, “the original phenomenon” (29) that has been lost. A forefather of this approach to sound was Walter Benjamin (1969), who posited an “aura” surrounding an “original,” and who discovered in mechanical reproduction a number of new features after the “aura” had been stripped away and replaced through the copying process, including an underlying sonic and optical “unconscious” (Ryder 2007). Psychoanalysis, with its theory of human consciousness related to a deeper unconscious heterogeneity, and semiotics, with its theory of signs that relate an ostensible presence to an absence, provide two ready techniques with which to elaborate the gap between surface simulacrum and missing original.

For this tradition, which focuses on the unattainable original, “sound in film is itself a kind of echo, re-presented and reproduced, never actually ‘there’” (Lawrence 1991: 3). Accordingly, the aesthetics of such theories tend to emphasize a split within reality, which fosters such conditions as deception, masking, loss, emptiness, alienation, ambivalence, nostalgia, forgetting, false consciousness, psychic excess, liminal states, interpretations against the grain, and sudden revelations in chance details that permit a return of the (social, political, or psychic) repressed, all of which shroud or reroute the death drive (Doane 2002). A kind of 1970s “suture” theory (Branigan 2006: 133–45) is applied to the soundtrack so that certain sonic techniques are said to work in conjunction with image editing in order to conceal discontinuities in the material of the medium as well as to obscure the heterogeneity and otherness of the world, thus creating an illusory surface of homogeneity and deceptive truth on the screen, a truly false thereness (Doane 1985a, 1985b). A “sound bridge,” for instance, becomes a device for crossing over/concealing an important (unseen, unheard) gap lying somewhere beneath or between, rather than being, say, a link in a network, in a “flat” heterarchy. The key question for this tradition of thinking about sound remains, “what is beneath and obscured, what has been lost?”

In addition, there are many flirtations with a so-called apparatus theory (Baudry 1986a, 1986b) and with subject-positioning theory of the 1970s. Alan Williams states that in watching and listening to a film, “*signs*” and “pseudo-perceptions” are being substituted for “a physical situation,” forcing us to “accept the [electronic and photochemical] machine as organism, and its

---

'attitudes' as our own" (1980: 58; see also Altman 1992d and Gorbman 1987). These sorts of skeptical ideas underpin explanations of how a patriarchal apparatus confers on women an artificial lack through widespread sonic "construction" of gender "positions" in film texts (Carroll 1996a; Greene 2009; Lawrence 1991; Silverman 1988). Also important has been the notion of the "chora" (derived from Plato's *Timaeus*), which is regarded as a pre-symbolic soundscape of tense ambivalence in (not only) the womb (Doane 1985b; Flinn 1992; Kristeva 1984). The latter formulation is perhaps an attempt to avoid Freud's topographical model of the psyche (surface/depth) in favor of fitting the supposed dynamism of acoustical phenomena to Freud's dynamic and economic models of the mind. But is sound necessarily and always "dynamic"? Perhaps not.

### ***Nonidentity (of Copy/Original) Versus Identity Theories***

The theories of Altman (1992), Williams (1980), and Thomas Levin (1984) that postulate the fundamental nonidentity of copied and original sound—and hence the loss of at least some important part of an original—are built on a pervasive skepticism about percepts and perception, and are part of a class of sound (and presumably image!) theories that occasionally appeal directly to Plato's cave (Altman 1992a; Baudry 1986a, 1986b). James Lastra (2000: 123–53) has expertly delineated the intricacies of these acoustic theories as well as the details of an opposing class of "identity" theories.

Christian Metz adopts an "identity" position when he claims that *auditory aspects, providing that the recording is well done, undergo no appreciable loss in relation to the corresponding sound in the real world: in principle, nothing distinguishes a gun shot heard in a film from a gun shot heard on the street.*" (Metz 1980a: 29; see also Balázs 1970; Baudry 1986b; Cavell 1979; Mast 1977)

For identity theorists, this conclusion is strengthened by the fact that a gunshot heard in a film may have been produced in innumerable ways without the use of a gun. Indeed, an audience may fail to recognize a real gunshot that has been recorded because it does not resemble a movie gunshot, but may recognize rain and surf even though the sounds have been produced by turning a cylinder filled with peas and nails at different speeds. In another example, the crunching noise of a character walking in snow may best be duplicated by a sound designer's fingers strolling through cornstarch (Carello et al. 2005; Eidsvik 2005; Jarrett 2000; Mancini 1985: 365).

What is important for identity theorists is that a particular filmic sound must be believed by an auditor to be typical of such sounds and sufficiently familiar so as to elicit an intended meaning or response. The original acoustic situation is not relevant. "To understand' a perceptual event is not to describe

it exhaustively in all its aspects but to be able to classify and categorize it” (Metz 1980b: 65). A liberal identity theory would allow sound to be an index, icon, symbol, or a combination of these depending on convenience and use (Plantinga 1997). Identity theorists argue that sound plays to a variety of insti-

***Identity theorists argue that sound plays to a variety of institutional norms, narrative conventions, textual functions, and a perceiver’s expectations, and not to the standard of a unique, profilmic event . . .***

tutional norms, narrative conventions, textual functions, and a perceiver’s expectations, and not to the standard of a unique, profilmic event, which has originals in great abundance. Realism, for an identity theorist, is a complicated issue that is rarely as simple as engaging and recording a documentary event. The reason is that human memory, judgment, and imagination are assumed to be complex processes enmeshed in significant behavioral patterns and customs (Black 2001; Branigan forthcoming

a; Lefebvre 1999). Especially important to identity theorists is the ability to fictionalize and to entertain hypotheses, even to initiate fictions within a fiction (Branigan 2002; Smith 1995b), all of which may lead to true lies that, in fact, teach us about the real world.

Where, then, should the authority of sonorities and imagery be located? Identity theories stress that human behavior occurs within a social setting, a “form of life,” that gives rise to language-games (Vescio 2001; Wittgenstein 1967), including film theory language-games (Branigan 2006 and forthcoming b). The orientation is epistemological and historical, based on “contingent universals” and “mild realism,” because what counts as “being familiar with a thing” is relative to the present life of a community (Bordwell 2008; Dennett 1991; Dutton 2006). The basic principle is that meaning and critical thought cannot be fixed by reference to sensations alone or by skepticism about sensations. One can discern within identity theories the beginning of a shift away from details of technology (sound’s material track or simply the sound track) and medium specificity (Costello 2008) toward an investigation through media archaeology of how we are disposed to hear within a cultural register (soundtrack); from measuring a sound event toward its inscription by a discursive context; from objective properties to the alchemy of fascination (Kahn 1999; Sarkar 1997; Siegel 2007; Sterne 2003); from the ontological in of “what’s *in* sound (that can’t be copied)?” to the cognitive “what can sound *do*?” (Branigan 2009).

By contrast, nonidentity theorists tend to celebrate—or to mourn the loss of—an essential groundedness of material reality/heterogeneity. Echoes of a Bazinian theory appear whereby sound and image are believed to place a real perceiver where he or she would have been if he or she could have been present at a scene to hear and see (Branigan 2006: 76–80). The filmic techniques selected as appropriate to this task are modeled on the faculties of sensory perception, not on the production of higher order meanings, mental abstrac-

tions, and heuristic inventions. With respect to the latter, a first-generation cognitive theory tends to favor some sort of invisible, ideal, imaginary, or possible witness present at a scene (Branigan 2006: 112, 308 n. 26). For nonidentity theorists, such a “witness” is merely trickery calculated to evade material reality.

André Bazin worries about the ways in which aspects of imagery may be made, in effect, “nonidentical” to the world through excesses of filmmaking. Bazin observes that a film image is a

*fragment of concrete reality in itself multiple and full of ambiguity, whose meaning emerges only after the fact, thanks to other imposed facts between which the mind establishes certain relationships. . . . Each image [is] just a fragment of reality existing before any meanings. (1971: 37)*

As an ontological matter, “Every image is to be seen as an object and every object as an image” (Bazin 1967a: 15–16; but see the subtle reinterpretation of Morgan 2006). Likewise, sound for Bazin is designed to translate “the physiology of existence” into “the concrete integral” in a way exactly parallel to the image, yet independent (Bazin 1967b: 131, 133, 139). Although some nonidentity theorists despair of recovering from film anything more than a slight trace of an original reality or else a delicate symptomatic incongruity that points toward the missing or repressed, the standard of measure for all nonidentity theorists remains that which once was.

### **When Is Sound?**

*If we consider for a moment the lengthy process of conception-investment-production-distribution-exhibition-reception, we recognize that the completed film constitutes the only step in the progression representing apparent unity. Until completion, the film is characterized by the multiplicity of its conceptors; after distribution, the film is characterized by the multiplicity of its receptors. (Altman 1992a: 4)*

Sound exists in many places and times. Each of Altman’s steps above may be further subdivided into a series of connected components. For example, a labor market yields a screenwriter, who dictates that an actual physical cause (e.g., a prop) be employed to become a fictional cause (e.g., a prop-as-gun) of the vibration in a diegesis that disturbs air for an actor playing a character. That character, in turn, drives a recording apparatus that eventuates in someone in an audience listening to, and then hearing (conceptualizing, remembering, associating, emoting, fantasizing, visualizing) a particular sound in a discursive genre, which satisfies audience demands. The question then arises: which is the sound that once (upon a time) was?

---

Where should one draw a frame around these many events and conditions in order to theorize the nature of sound? How large or small a frame? Non-identity theories tend to draw a rigid frame around technology and industrial production, which is understood to leave an indelible mark. However, there may be important theoretical limitations to the idea of such a frame as a *container* for empirical data (Quine 1980b). Not every effect in a film refers to its cause; nor is an effect necessarily explained by its cause; nor is our understanding of an effect necessarily dependent on knowing its actual cause. Perhaps instead the question should be which sound at which time is taken to be *the* sound at the end (and is the end only to be found at the conclusion of an action, scene, or story)? That is, if a sound is more than a given pitch, timbre, loudness, attack, transition, tempo, accent, beat in a physical frame, and/or is tied to a film frame on screen, then which set of criteria and assumptions on which occasions is relevant when choosing to describe the sound event for a given purpose? The claim here is that “specifics” and “film specificity” exist only when a frame and a goal have been selected to elaborate relevant properties. There is no sound in itself, no authenticity that is self-validating or self-evident apart from choosing a relevant frame for one’s talk about sound.

### ***Framing Audition***

Donald Davidson (2001) asserts that there are three kinds of propositional knowledge that exist within subjective, intersubjective, and objective frames of reference (see also Hernadi 1995; Quine 1980a). One might even add a fourth frame—the interobjective—based on irreducible emergent properties within complex dynamical systems as described by, for instance, chaos theory (Abbott 2008; Cilliers 1998) and manifested in, for instance, massively multiplayer online role-playing games or the collective responses to a popular film. Another approach, inspired by Wittgenstein’s notion of family resemblance, expands these four into fifteen descriptive frames or appearance types (Braniġan 2006: 97–149; cf. Sibley 2001b, 2001c). David Black (1987) also deploys multiple frames to create an erudite form of apparatus theory.

What does a sound sound like in different frames? In fact, how does any device of film function, and become distinctive, within each of these frames? Lastra (2000: 147–150) invokes Derrida’s well-known essay, “Signature Event Context,” in working to dis/entangle nonidentity-identity theories. It would seem that whatever approach is adopted for contextualizing a film experience, the aim should be to multiply and weigh, rather than to collapse, the number of language(-game)s in frames that may become pertinent when describing a sound. Recall Metz’s example of a gunshot heard in a film produced without the use of a gun. Is such an “imitative sound” that is produced artificially somehow “less trustworthy or ‘authentic’ than any other type of

---

sound? Or does it simply remind us, again, of the shifting, tragic, comically unbreachable gaps between different ways of knowing?” (Walsh and Kubick 2006: 18).

### ***Is Sound an Adjective or Noun?***

It may be true that in thinking and hearing within a strictly physical frame—the site of Davidson’s (2001) “objective knowledge”—that sound collapses into an adjective. Metz says that sound is an adjective while vision is a noun. When we see a physical object, he says, its identification is complete and all that could be added would be merely adjectival, as in a “tall, reading” lamp. By contrast, when we hear a sound the identification remains incomplete. A “whistling” sound still needs to be specified: the whistling of what thing? from where? because of what action? Thus, for Metz, sounds function as adjectives/verbs that point to and characterize substances that are visual and properly named by nouns/gerunds.

Metz traces this distinction between vision and audition to the subject-predicate structure of Indo-European languages and to the distinctions among primary, secondary, and tertiary qualities made by the Western philosophical tradition. Philosophy and language are seen as instances of social and cultural phenomena closely joined to our everyday perception of the world. What we take to be “intelligible,” and what is seen and heard, depends on how we have learned to classify and categorize as well as, presumably, what social actions and goals are encouraged or permitted under specific descriptions/prescriptions of the world (Metz 1980a: 25–28). Metz’s comments illustrate that a theory of language use may be fundamental to both identity and nonidentity theories. The problem is to decide which analytical language or linguistic theory is most appropriate.

Kendall Walton formulates the adjectival status of sound in another way by declaring that “vision is frequently more effective than hearing as a means of identifying particulars, as a source of *de re* rather than mere *de dicto* knowledge” (1988: 352). The word “frequently” here is well-chosen because Walton’s claim depends on ordinary situations, or rather, on what has been taken to be prototypical for a situation when assessing the status of visual and aural experience. But, one may ask, are aesthetic situations always ordinary? And doesn’t the choice of a prototype merely act as a frame, hence both enabling and limiting the resulting descriptions? Furthermore, may different frames for an experience be chosen with different purposes in mind? And do the qualities that are discovered in music, noise, and dialogue derive from the same prototypical situation or must there be different prototypes for each?

Here are a few sonic situations, any one of which qualifies as prototypical for hearing and thus available for judging and measuring some or even all of the sounds in a film:

---

- conversation (Devereaux 1987; Kozloff 2000; Young 2001);
- report;
- lecture;
- commentary (e.g., voice-over; Smoodin 1983; Wolfe 1997);
- moral lesson;
- interview (e.g., apparent questions from off-camera as in a Godard film; or, fictional narration in a question and answer format, contest, or mystery; Carroll 1996b);
- theatrical performance;
- oral storytelling;
- reading aloud (Williams 1985);
- free association;
- music (e.g., melody, rhythm, and sonic pattern; Bordwell 1980; Mitry 1997);
- ambience (naturalism; found sound);
- a perennial favorite, the sound of the human voice (Appelbaum 1990; Chion 1999; Ihde 2007; Karpf 2007; Mitry 2000; Smith 2007b).

In short, what do we expect from ourselves when we expect something from an object that sounds? Against which norm is the sound to be measured? What do we hear when we hear it against different norms (habits, practices)? Audiences of the past may have had in mind norms provided by vaudeville, music hall, opera, radio, phonograph, telephone, public address, and live commentary spoken by a lecturer speaking for a silent film. Thus there may be distinct models or modes of/for listening. When competing mental prototypes and soundscapes are overlaid on historical practices and developing technologies, an innovative method of research is established that Altman terms “crisis historiography” (2004; on “acoustic ecology,” see Hilmes 2008; Lastra 2008).

### **Sound in Mind**

Auditory adjectiveness appears in a great many guises in theories of film (Nasta 1991) and its status, according to Walton, as “mere *de dicto* knowledge” explains why it is commonly held that a film image may easily demonstrate that a voice-over narration is unreliable (*equivocal*) while the converse is not true (Fleishman 1992; Kozloff 1988). Mitry asserts that “*Visible* complexity does not have an *acoustic* equivalent” because a film image is capable of reflecting the “duality of *being* and *seeming*,” while sound reflects only the latter (1997: 244). But is this the only way to see sound? Perhaps the human mind is not so single-minded.

Rudolf Arnheim discovers two possibilities for sound to become noun-like. In one approach he places his faith in the brute expressivity of sound to over-

---

whelm and defeat our natural (i.e., non-aesthetic) inclination to create out of a sonic event explanatory mental images by means of “the inner eye” that would otherwise be “missing” (Arnheim 1986: 136, 137; Cardinal 2007). To be aesthetic for Arnheim, sound must avoid the inner eye. In theorizing the medium of radio in 1936, he stressed that “the elemental force lies in the sound, which affects everyone more directly than the meaning of the word, and all radio art must make this fact its starting-point” (1986: 28). Arnheim mentions “intensity, pitch, interval, rhythm and tempi” as ways to defeat “the inner eye,” and offers the following illustration:

*The direct expressive power of a hammered-out rhythm and a soft blurred sound, a major and a minor chord, a fast and a slow pace, a sudden or a gradual rise and fall in pitch, a loud or a soft tone—these are the most elemental and the most important creative means for every form of acoustic art, for music just as much as for the arts of speech and sound! (1986: 29, 30).*

In effect, Arnheim bypasses a nonidentity theory to postulate a liberating power in the sensuous qualities of the sounds themselves, which he calls the “entirely unexplored” art of “blind hearing” (1986: 145, 147, 226). Arnheim’s theory applied to noise would perhaps yield *musique concrète*. His theory aims to put an auditor in direct touch with objects without sighting them. Related concepts are Michel Chion’s “reduced listening” (1994: 29–33) and Roland Barthes’s (1977) “grain of the voice.” Husserl is one source for these concepts (phenomenological bracketing and reduction); aesthetic goals may include the experience of liminal states, the sublime, excess, and surplus—a kind of textual overtone or undertone—and *jouissance*.

Arnheim (1986) explores a second path to explain a noun-like soundedness. He devotes several chapters to showing how a sense of direction, distance, movement, perspective, and resonance can be heard by a listener, allowing an “emancipation from the body” (chaps. 3, 4, 7: 15, 120, 148). He suggests that a listener may glide, jump, or wander through layers of an imagined deep space that has been created and developed acoustically through such techniques as juxtaposition, reduction, transposition, discordant superimposition, and dispersion. (Will the listener thereby become an acoustic *flâneur*?) Arnheim also imagines movements through acoustic, temporal densities (chap. 5). In addition, one might include the audio technique that creates multiple and simultaneous rhythmic and semantic centers in a scene, as in Robert Altman films of the 1970s (Chion 2009: 121–23). Film scholars would do well to pay greater attention to the literature on radio aesthetics; exceptions to this deafness are Altman (1994) and Vertov’s “radio-eye” (Fischer 1985).

We may think of Arnheim’s two approaches to freeing sound from sight as focused primarily either on involuntary, bottom-up, stimulus-driven percep-

***We may think of Arnheim’s two approaches to freeing sound from sight as focused primarily either on involuntary, bottom-up, stimulus-driven perceptual skills . . . or else on top-down, goal-oriented, inferential skills . . .***

tual skills (“blind hearing”) or else on top-down, goal-oriented, inferential skills operating apart from online bodily sensations (“emancipation from the body”). It would seem that these two broad kinds of mental activity involve different relationships to “reality”: a bottom-up “deductive realism” versus a top-down “inductive realism” (Black 1987), or an abductive realism (the logic of schemata), where we may label a feeling, opinion, or view as “sound” or “unsound.” Tellingly, this

illustrates that in a top-down approach, a view is being submitted to sonic judgment, rather than an image giving the lie to words.

There would also seem to be two different kinds of time associated with Arnheim’s two approaches. Bottom-up processing moves continuously and irreversibly with the “river of time” as does the ribbon of celluloid through a projector, stamped with an optical or magnetic sound stripe. We perceive the “present” of this directed time to last about two-and-a-half seconds. Chion (2009) proposes various sonic types, including nondiscontinuity, *Shining* effect, temporal linearization, temporal splitting, temporalization, and vectorization. By contrast, top-down processing treats time as “spatial” and (temporarily) “frozen,” such that time can be manipulated by moving forward and backward, cycling, stretching and compressing, as is typical of judgment heuristics, schematic thought, dream processes, and various types of memory (e.g., the phonological rehearsal loop and the 150-second cycle of working memory). For example, we conceive of time spatially as a quantity or commodity subject to manipulation (Lakoff and Johnson 1999) when we say: “You may *take* your time, but for me time is running *out* because I *lost some* last week and now time is *short*.” Here time is circumscribed by a speaker’s purpose or unspoken drive.

The concepts top-down and bottom-up are themselves a bit schematic and susceptible to criticism on the basis that they only reinstate a high-low (surface-depth) way of thinking about thinking. To be persuasive, the concepts would need to be elaborated into a more complete cognitive architecture, which has become the goal of the cognitive sciences (Wilson and Keil 1999).

### ***Sound in Time***

When not standing in the “river of time,” a person is guided by various top-down schemata under which novel, even non-adjacent, juxtapositions may emerge and be heard to create a significant regularity or periodicity that is not dependent on simple duration, as revealed in, for instance, Deleuze’s crystalline time, Riffaterre’s hypogram, and McGann and Samuels’s deformative diagramming. Simple illustrations are found in sonic forms of foreshadowing,

---

closure, overlap, and metaphor. Thus if there exist two types of cognitive processing and two types of time, it may be that sound can escape its physical adjectival status from time to time. Consequently, a range of fresh aesthetic possibilities opens for the soundtrack (Branigan 1997a: 108–110).

Another way to pose the question of whether sound can escape its adjectival status is to ask whether there is an aural equivalent to the time of an image freeze frame. There are four different answers to this question. No, because sound, like the river of time, is defined by its duration and cannot stop moving (Chion 1994: 44; Levin 1984: 62). Yes, even though sound and image are separate, both offer equivalent channels for “information,” so that, for example, a single musical note may be heard as being “stretched” (“long” held, “frozen” in time) through the device of a sound loop (W. Johnson 1985, 1989). In this spirit Arnheim asserts that “rhythmic monotony is the absolute zero of all motion” (1997: 48). Sometimes, in certain cases, one or the other (Chion 1994: 10). And, finally, the answer may be simultaneously “yes” and “no” in the case of a sonic rhyme or motif appearing in different places in a film.

It does not help to be frustrated over the question, “But what is sound really?” Problems about the nature of sound are not empirical questions, but theoretical ones involving the grammar of words like “flow,” “stream,” “drift,” and “freeze,” themselves subspecies of polymorphic words like “motion” and “change.” Specific conclusions about sound would seem to depend on the framework chosen to justify a preferred set of descriptions; whether sound is being heard through folk knowledge about “the physical frame”—an embodied physics—or heard in a political context or heard when attentive and mindful of a matrix of aesthetic and narrative meanings. Thus, as Wittgenstein (1967: §§ 515, 671) says, a rose in the dark may be seen either as quite black or as red. In similar fashion, the sounds of silence may differ when heard bottom-up (an emptiness) and top-down (a pause, an interruption, an interval). In the former, silence can only be “broken” or “shattered”; in the latter, silence functions within an ensemble and may even acquire a color, “golden.”

### ***Sound in Language***

Sound and language are both temporal phenomena. They may, however, share an even closer bond. We use language(-games) to express our thoughts and feelings about film to others and, most important, to ourselves through inner speech, as in the gist of semantic memory and the phonological rehearsal loop of working memory. That is, we generate meanings through language in order to describe the meanings we find and hear in film and hear within ourselves. Language, like sound, is an action and its use depends on our goals in acting. From this perspective, it would seem that when a person’s goals are tied closely to the identification and location of the attributes and actions of a depicted physical object, then listening/comprehending proceeds

largely in a bottom-up manner. That is, we detect quickly and automatically such attributes as color and shape, motion and loudness.

By contrast, when a perceiver's goals are tied only to a name that is being used to refer to a depicted object through one or more categorical concepts that may fit the object in a variety of ways—and are filtered and transformed through mental schemata—then he or she is attending top-down to an audiovisual display. This nominal use of language may summon an unusual or non-tangible quality of an object; or point toward a second object that shares some literal or figurative quality; or remind us of another object through some chain of association or logical relationship (e.g., contradiction). These top-down aspects of an object are indirect and relatively more abstract than bottom-up aspects of an object. What is at stake with nominal, acoustic reference is knowledge of one or another world through (contingent) cognitive dispositions, rather than knowledge of a fixed world (Goodman 1978).

In general, these two ways of thinking and speaking about filmic experience—through what may be called existential reference (physical, attributive) as opposed to nominal reference—would seem to closely correspond to Keith Donnellan's (1966) synthesis of Russell and Strawson in the terminological pair "attributive" versus "referential" uses of definite verbal descriptions (see also Branigan 1997a). Another possibility would be to examine the intricacies of "attributive" versus "predicative" uses of adjectives (Sibley 2001a). In applying the logic of these sorts of methods, the notion of "language" would need to be expanded from operations and relations among words to the use of labels and descriptions that are employed to pick out, or apply to, an object or class of objects. Labels may be verbal predicates or else nonlinguistic elements drawn from a myriad of systems, such as pictures, photographs, diagrams, gestures, color samples, and, presumably, sound samples (Goodman 1976).

One may conclude that once sound is defined through a particular theory of language and placed within the processes of meaning and language use, sound then acquires the potential to escape the constraints imposed by its narrowly construed physicality in order to function and interact more broadly with many facets of mind.

### **Sounding an Image**

*Sound is, when it comes to its localization, in a constantly unstable state with respect to the image. Either it is included, or it includes, or, in a third possibility, it "roams the surface." One could say that sound in the cinema is "that which seeks its place." And this seeking process is played out for each sound in a specific way—only rarely do all the sounds heard at the same time participate in the same emplacement. (Chion 2009: 485)*

---

A number of practical issues of filmmaking bear on the relationship of sounds to one another—sonic figure and ground—as well as bear on the fluid relationship of sounds to an image, e.g., the relation of sound to color and graphics or sound to camera movement.

### ***The Film Frame***

Should the microphone be placed near the position of the camera and pointed in the same direction as the lens, creating a sense of “sound perspective” that is coordinated with the image (i.e., submitted to the objectivity of space)? Or should the microphone be placed near the most important source of sound at that moment in the scene so that the dialogue is heard clearly even in the background of a shot or while a character is turning away from the camera, creating a “narrativized perspective” or “psychological realism” (Altman 1992d; Doane 1985a; Lastra 2000)? This amounts to asking whether realism should be perceptual or psychological. A subset of the latter is “point-of-audition” sound, which parallels the character subjectivity of a point-of-view shot. Indeed parallel and independent are apt descriptions of sound when considering issues of filmic point of view, narration, and story structure because all of the functions proposed by narrative theories (Branigan 2006: 25–63, 116–117) may be conveyed equally by employing sound or image.

Much debate has centered on whether sound must vigorously declare its complete independence from the framed image through any one of several types of asynchronous relationships with the image (Anderson 1996; Arnheim 1997: 40–42, 47–51; Eisenstein et al. 1988; Thompson 1980; Wollen 2003). Pudovkin affirms that “only by . . . counterpoint can primitive naturalism be surpassed” (2006: 207; see also Pudovkin 1970). Nonidentity theories tend to favor types of asynchrony because of a worry that otherwise an “inaudible” sound style will be created to simply serve the “invisible” editing of Hollywood’s classical and supposedly illusionist narrative.

The issue of asynchrony emerges clearly in the relationship of music and image. Music that has been deftly coordinated with motions on the screen is known as “Mickey Mousing.” If it is believed that synchrony makes sound a slave to the image and promotes a false sense of unity (or at least reduces the complexity of a film in favor of mindless spectacle), then the term Mickey Mousing—along with the handy connotations of its name—is employed to stand for all cases of synchrony. Rick Altman (1980) manages to reverse this prejudice by contending that it is imagery that is redundant in classical narrative film; what is viewed as moving on the screen is actually an elaborate act of “ventriloquism.”

The debate about the effects of asynchrony/synchrony intersects with a set of topics involving speech that is being heard with an accompanying image of moving lips. This debate affects the perceived status of dialogue, voice,

---

voice-under, voice-over, and the status of character in film (Naremore 1988; Smith 1995a).

### **Sound Unseen**

A film frame displays things, but also excludes other things; here, too, there are opportunities for sound to play a role along with the unseen. Noël Burch (1973) has identified six areas of “fluctuating or non-fluctuating,” off-screen space by envisioning a camera inside a consecutive series of spatial cubes. Additional areas of unseen space may be brought into play through permutations of non-diegetic places and times (e.g., with voice over a space). Sound issuing from off screen is said to have unique and peculiar powers (Balázs 1970: 209–216), perhaps because it appears to be disembodied and apart from us. (A major exception, however, is location ambience, which usually furthers a sense of continuity [Eidsvik 2005: 73–74 on Walter Murch’s psychological-aesthetic rules].) Chion has defined one type of voice-off, which he calls “acousmatic” sound in the person of an “acousmètre,” who communicates through the agency of an “acousmachine” (Chion 1994, 1999). Think, for instance, of the thundering voice of the Great Oz in *The Wizard of Oz* hiding behind an apparatus of curtains, grimacing masks, and smoke; recall also the unseen Dr. Mabuse in *The Testament of Dr. Mabuse*, and the mother’s voice in *Psycho*. Nevertheless, the balance of knowledge may be neatly reversed from Chion’s acousmatic sound when sound plays to an off-screen eavesdropper (Weis 1999). Here power resides with the quiet eavesdropper, heard in the silence of off-screen space.

Sounds either from, or for, an unseen person may act to destabilize a text with anxiety, uncanny moods, and self-consciousness. Nevertheless, off-screen ambient noise or music does not normally produce these aesthetic effects. And the invidious effects attributed by feminist theory to a woman’s voice-off do not always occur (Doane 1985b; Lawrence 1991; Silverman 1988; see also Smith 2007a). It would seem, therefore, that the issue is not just the visibility or not of a sound’s source, but of the importance of what is being heard, that is, how sound cues or redirects our expectations in relation to a narrative context. In other words, how the unsightly sound is being narrativized. Again, what seems to determine the character of a sound is derived from a top-down narrative context that defines the predicaments of a character.

In the same spirit of separating sound off from an image, Robert Bresson advises filmmakers, “When a sound can replace an image, cut the image or neutralize it. The ear goes more towards the within, the eye towards the outer” (1997: 61). The potent effects of isolating and concentrating sound as recommended by Bresson may be due partly to the fact that the amygdaloid complex, a site of emotions in the brain, has rich interconnections with the nearby auditory system, which is more lateralized than vision. Also it is known

---

that waves of inaudible infrasound cause internal vibratory feelings that may be experienced in a deeply emotional way. Thus there may well be a physiological basis to the power of “unseen” sound, though it may not lie in mere formal considerations like the placement of frame lines.

A striking exception to the fluid movement and malleable aesthetic of unseen sounds moving across spaces and levels of narration (e.g., an oscillation between diegetic and non-diegetic narrations) is diegetic “object noise,” which is distinct from “colored” noises, e.g., the pink noise of a waterfall and white, random noise. Nondiegetic noise is exceedingly rare in film. The reason is that noise seems either extraneous (the word “noise” comes from the Latin, *nausea*) or else affixed firmly to a concrete, diegetic object (cf. Kracauer 1997: 124-125). It has been argued that non-speech sounds are treated differently than speech by our precategorical acoustic memory. A person’s voice, too, contains noise components that may be “amplified” in a variety of aesthetic ways in order to block the intelligibility of a character, potentially decentering the image of a speaker in favor of realities other than the words uttered (Barthes 1977; Chion 1992) and transforming a subject into a brute object.

Noise, music, and language are processed in different areas of the brain. Perhaps, then, it is a mistake to begin with the idea that noise, music, and dialogue in film are “sounds,” because one may continue to look for commonalities among the three when differences are more important. What is seen and unseen about a sound may be relative, relational, and a matter of degree. This issue cannot be settled by resorting to distinctions based on formal or technical, bottom-up features of a film, such as space marked by a frame line or time moving smoothly and irreversibly, like celluloid through a projector.

### ***(Non-)Images of Sound***

A final matter to contemplate about the relationship of sounds to an image is whether sound and image are identical in some fundamental way. Various answers may be given within various frameworks. Here is an identity theorist arguing for the nonidentity of sound and image:

*[In cinema] one does not hear an image of the sounds but the sounds themselves. Even if the procedures for recording the sounds and playing them back deform them, they are reproduced and not copied. Only their source of emission may partake of illusion; their reality cannot. Hence, no doubt one of the basic reasons for the privileged status of voice in idealist philosophy and in religion: voice does not lend itself to games of illusion, or confusion, between the real and its figurativity (because voice cannot be represented figuratively) to which sight seems particularly liable. Music and singing differ qualitatively from painting in their relation to reality. (Baudry 1986b: 304–305)*

For Baudry the divide between sound and image would seem to rest on a belief that sound is causal and literal (indexical) while the image reaches our imagination by being copied, imitated, altered, and refigured (iconic). Lastra has suggested (in private correspondence) another interpretation: that Baudry may be attributing the divide between sound and image to the fact that a represented sound is still a sound while a represented action is only an image of the action. This difference between sound and image would seem to also stem from Baudry's use of Freudian methodology, which insists that all memories and unconscious thoughts, even the most abstract, are displaced and refigured toward pictorial substitutes, making dream-work strictly visual not sonic; that is, it is images, not sounds, that give voice to psychic events. Baudry's fear of imagery and imitation in 1975 may be imagined to be greater today when computer-generated imagery and screen software written to "emulate" the look of other software are widespread.

Chion, too, asserts that "Sounds and images are intrinsically too different to be commensurable" (2009: 498). One difference Chion finds is that sounds have no frame (470-471) whereas "the frame preexists the appearance of the image and survives its extinction" (469). But is this true? Or, rather, in which frames of reference—under which descriptions—do sound and image "frames" acquire the peculiar status of never existing even abstractly in one case and never not existing abstractly in the other? What theory is presupposed?

Conversely, Noël Burch remarks, "It seems that the essential nature of the relationship between sound and image is due not to the difference between them, but rather to the similarity between them" (1973: 91). For Burch, all parameters of film are subject to plastic manipulation and available equally to articulate space and time (see Branigan 1997b). Burch follows Sergei Eisenstein, who wrote a great deal about sound and for whom sound was an independent montage unit equivalent in all respects to the movements of actors and light, to shifts of a spectator's attention, and to the appearance of new emotions and ideas (see Bordwell 2005). For Eisenstein the common denominator of filmmaking and audience reaction was dialectical movement. All film elements were regarded as forces capable of producing movement. What is fundamental for Eisenstein and Burch is the flow of perceptual forms spurred by the sound-image interaction. The focus is on a filmic form that creates a series of interchangeable, middle-level relationships that are not based on the physicality of sound and image, the perceptible qualities of material, technical specifications, or on the personal, expressive potential of the medium.

### ***Sound within Space: Sound Visualized***

Another reason Chion gives for the nonidentity of sound and image is that sound can have no space (no room) of its own. For sound "the temporal dimension seems to predominate, and the spatial dimension not to exist at all"

---

(Chion 1994: 44). Admittedly, sound may carry acoustical information about the space it is moving through—what Altman calls “spatial signature” (1992c: 24). Sound may also be displaced within a diegetic space—what Chion (1994) calls “spatial magnetization,” where, for example, a spectator perceives the sound of footsteps as following a character, who is moving across and off-screen, no matter the actual location of a theater’s loudspeakers. That is, sound may be read perfectly into the image without regard to its physical source in the theater. Nonetheless, “sounds have no images,” and are not imaged or visible, maintains Balázs (1970: 216).

But why shouldn’t sound be able to acquire a “spatial form” in ways other than a cocktail party effect? Is it possible that a sound could reach and define a space, and be visualized spatially, because of its connection to temporal duration? Consider the following fact that is often overlooked by theorists concerning the intimacy of sound and time. Dialogue in film clearly operates to imprint celluloid with at least one sort of Bazinian real time captured intact from the world, because a spectator would instantly notice the smallest deviation during playback from the normal time of speech or any disruption in a character’s visible gait on screen. Does this suggest that sounds in some cases may be “frozen” into the simultaneity of an image and then deployed like a color field as evidence for actual happenings in space? One must still ask, however, in which theoretical frame sound is able to testify to the image of a person heard/seen speaking and walking. Perhaps sound that has no image and no frame exists only within a technological frame of a non-identity theory. In other words, it exists only as coming from a sound track, not being on a soundtrack.

***It turns out that in some theoretical frames sound is not excluded from being imaged and having a spatial dimension.***

It turns out that in some theoretical frames sound is not excluded from being imaged and having a spatial dimension. Melinda Szaloky (2002) has demonstrated that some silent images nevertheless provoke a mental hearing because they embody sound. A silent film is not truly silent because it is obliged to provide a rationale justifying how sounds can be heard. A spectator does not presume characters to be deaf or the world mute. For example, a silent image of someone playing a piano, or perhaps merely the image of a piano by itself, may bring about an appropriate mental sounding. Furthermore, when reading a dialogue intertitle in a silent film, a spectator may hear his or her own inner voice or else the “imagined voice” of a character. There are still other, more sophisticated illustrations of “visualized sounds” (Szaloky 2002). This phenomenon of visualized sound was not unknown to early film theorists, but it is remarkably unappreciated by modern writers despite the fact that there are visible objects in sound films that are not part of the sound’s (physical) track, though clearly sounding to a spectator. Visualized sounds are

more pervasive and fundamental than Chion's notion of "phantom" or "negative" sound (2009: 483).

Sound and image may also be made interchangeable through synesthesia—through an exchange of sound and color (Hanssen 2006)—an issue of great interest to Eisenstein. "The melodic blending [of visual elements] with sound is probably achieved most distinctly through nuances of light which are inseparable from nuances of color" (Eisenstein 1987: 394).

### ***Inaudible Sound Images***

If sound may be imaged, must it always sound, or may it be inaudible? As indicated earlier, one might argue that sounds and images are equally productive of a film's narration; any narrative function may be realized by either sound or image. There are many narrations in film, or levels of narration, that assume indirect forms; that is forms that are implicit, circumstantial, presumed, ambiguous, suppressed, or repressed (Branigan 1992). In this sense, inaudible narrations exist on the soundtrack, because even a self-conscious voice-over cannot acknowledge all of its functions and sources, and thus may contain important *sotto voce* elements. However, controversies among theorists have ensued over the need to postulate the existence of certain narrations in a text that are said to be both powerful and not immediately perceptible. This raises the problem of what can, or should, be said by a critic or spectator about such inaudible and invisible facets of a text as the implied author, "non-narrated" objective knowledge, silent-invisible observers, free indirect discourse, inner speech, heteroglossia (Bakhtin 1981), and filmic "enunciation" generally (Buckland 1995).

One answer is to reconceive each level of narration as a specific bracketing of knowledge that is designed to indicate the scope of relevant quantifiers, connectives, predicates, and variables of the propositions constructed by a person. Levels of narration, then, would appear as a set of patterns in a text that warrants particular sets of descriptions offered by a critic or spectator about how and when a particular kind of propositional knowledge may be acquired. Inaudible sound images in narration are then given voice through the responses of critics and spectators as well as through a person's readiness to speak about a film experience.

Especially labile is the concept of "voice," mentioned by Baudry (1986b), which may assume numerous inaudible and unspeakable forms in a narratology (Banfield 1982; Branigan 1992; Gibson 2001; Plantinga 1997), sometimes also associated with an unspoken attitude, for example a film's judgmental tone (Bordwell 1985: 61). An association with point of view is perhaps inevitable because voice connects to belief and intention ("actions speak louder than words"), and easily projects onto numerous wider social and cultural activities and contexts vital to an individual's life and goals. Thus a film may be

---

said to give voice to a marginalized class or an idea whose time is arriving. In effect, the notion of “visualized sound” has now moved still deeper into imagery to reflect the occasions in which texts are spoken about, recounted, interpreted, and valued (Black 1997). The reason is that there are many ways in which texts may be said to be efficacious for a person. Some of these ways are not directly seen or heard on the screen, but are understood sonically.

Inaudible voice returns us to the problem of “sound in mind,” discussed in conjunction with Arnheim’s theory. Consider, for example, the following question (inspired by Balázs 1970: 217): How long might a vivid or enigmatic sound or image—that is, an imaged sound or sounded image—remain available in mind to interact with subsequent material appearing onscreen or through a loudspeaker? It may happen that a dramatic impression from a film lingers in mind like an echo or shadow to become a faint double whose shape may continue to shift in engaging ongoing experiences. “When you listen to Mozart, the silence that follows is still Mozart” ([http://en.wikiquote.org/wiki/Sacha\\_Guitry](http://en.wikiquote.org/wiki/Sacha_Guitry)). Such an inquiry would turn from analyzing a sound track to hearing a soundtrack—from studying a sequence of physical sounds, local syntax, or industrial practices to assimilating scattered responses into a network of memories.

### **A Conclusion Made Explicit**

*Sounds and speech are not “chemically pure” art-products as tones of music are to a certain extent; they are products of nature and reality. From this it follows that they are not strictly definable. Of course the artist moulds them . . . by stylising them with the help of those musical means, but there always remains . . . a vestige of the untamed and the incalculable. (Arnheim 1986: 33–34)*

Arnheim’s wise conclusion holds as well for the photographic image, for the nature of the audiovisual medium as a whole. The properties selected as important and grouped to form a definition of the film medium may not be fixed, but instead responsive to competing new forms of media as well as open to our present memories, activities, and embodiment. There is no certain way to talk about sound without selecting a framework that limits discussion according to the properties and goals one has in mind for talking about and using sound to assign values to a form of life.

### **Acknowledgments**

The author thanks Irina Burns, Ira Konigsberg, James Lastra, Paisley Livingston, Carl Plantinga, Melinda Szaloky, and Charles Wolfe.

---

**Edward Branigan** is professor in the Department of Film and Media Studies at the University of California, Santa Barbara. He is the author of *Point of View in the Cinema: A Theory of Narration and Subjectivity in Classical Film* (1984); *Narrative Comprehension and Film* (1992); and *Projecting a Camera: Language-Games in Film Theory* (2006).

## References

- Abbott, H. Porter. 2008. "Narrative and Emergent Behavior." *Poetics Today* 29 (2): 227–244.
- Allen, Richard. 1997. "Looking at Motion Pictures." Pp. 76–94 in *Film Theory and Philosophy*, eds. Richard Allen and Murray Smith. New York: Oxford University Press.
- Altman, Rick. 1980. "Moving Lips: Cinema as Ventriloquism." *Yale French Studies* 60: 67–79.
- . 1992a. "General Introduction: Cinema as Event." Pp. 1–14 in *Sound Theory/Sound Practice*, ed. Rick Altman. New York: Routledge.
- . 1992b. "Introduction: Four and a Half Film Fallacies." Pp. 35–45, in Altman, *Sound Theory/Sound Practice*.
- . 1992c. "The Material Heterogeneity of Recorded Sound." Pp. 15–31 in Altman, *Sound Theory/Sound Practice*.
- . 1992d. "Sound Space." Pp. 46–64, in Altman, *Sound Theory/Sound Practice*.
- . 1994. "Deep-Focus Sound: *Citizen Kane* and the Radio Aesthetic." *Quarterly Review of Film and Video* 15 (3): 1–33.
- . 2004. "Cinema Sound at the Crossroads: A Century of Identity Crises." Pp. 13–46 in *New Perspectives in Sound Studies*, eds. Dominique Nasta and Didier Huvelle. Brussels: P.I.E.-Peter Lang.
- Anderson, Joseph D. 1996. *The Reality of Illusion: An Ecological Approach to Cognitive Film Theory*. Carbondale: Southern Illinois University Press.
- Appelbaum, David. 1990. *Voice*. Albany: State University of New York Press.
- Aristotle. 1941. *Organon*: "Categoriae." Pp. 3–37 in *The Basic Works of Aristotle*, ed. Richard McKeon. New York: Random House.
- Arnheim, Rudolf. [1936] 1986. *Radio*. Trans. Margaret Ludwig and Herbert Read. Salem, NH: Ayer.
- . [1928–1934] 1997. "Sound Film." Pp. 29–51 in *Film Essays and Criticism*. Trans. Brenda Benthien. Madison: University of Wisconsin Press.
- Bakhtin, M. M. [1934–1935] 1981. "Discourse in the Novel." Pp. 259–422 in *The Dialogic Imagination: Four Essays*. Ed. Michael Holquist. Trans. Caryl Emerson and Michael Holquist. Austin: University of Texas Press.
- Balázs, Béla. [1945] 1970. *Theory of the Film: Character and Growth of a New Art*. Trans. Edith Bone. New York: Dover Publications.
- Banfield, Ann. 1982. *Unspeakable Sentences: Narration and Representation in the Language of Fiction*. London: Routledge & Kegan Paul.
- Barthes, Roland. [1972] 1977. "The Grain of the Voice." Pp. 179–189 in *Image-Music-Text*. Select. and trans. Stephen Heath. New York: Hill and Wang.
- Baudry, Jean-Louis. [1970] 1986a. "Ideological Effects of the Basic Cinematographic Apparatus." Pp. 286–298 in *Narrative, Apparatus, Ideology: A Film Theory Reader*, ed. Philip Rosen. New York: Columbia University Press.

- . [1975] 1986b. "The Apparatus: Metapsychological Approaches to the Impression of Reality in the Cinema." Pp. 299–318 in Rosen, *Narrative, Apparatus, Ideology*.
- Bazin, André. [1945] 1967a. "The Ontology of the Photographic Image." Pp. 9–16 in *What Is Cinema?* Vol. I. Select. and trans. Hugh Gray. Berkeley: University of California Press.
- . [1951] 1967b. "Le Journal d'un Curé de Campagne and the Stylistics of Robert Bresson." Pp. 125–143 in Bazin, *What Is Cinema?* Vol. I.
- . [1948] 1971. "An Aesthetic of Reality: Neorealism." Pp. 16–40 in *What Is Cinema?* Vol. II. Select. and trans. Hugh Gray. Berkeley: University of California Press.
- Benjamin, Walter. [1936] 1969. "The Work of Art in the Age of Mechanical Reproduction." Pp. 217–251 in *Illuminations: Essays and Reflections*. Ed. Hannah Arendt. Trans. Harry Zohn. New York: Schocken Books.
- Black, David Alan. 1987. "Cinematic Realism and the Phonographic Analogy." *Cinema Journal* 26 (2): 39–50.
- . 1997. "Synopsis: A Theory of Symbolic Representation." *Yale Journal of Criticism* 10 (2): 423–436; corrected illustration in 11 (1): 313.
- . 2001. "Homo Confabulans: A Study in Film, Narrative, and Compensation." *Literature and Psychology* 47 (3): 25–37.
- Bordwell, David. 1980. "The Musical Analogy." *Yale French Studies* 60: 141–156.
- . 1985. *Narration in the Fiction Film*. Madison: University of Wisconsin Press.
- . 2005. *The Cinema of Eisenstein*. New York: Routledge.
- . 2008. "Convention, Construction, and Cinematic Vision." Pp. 57–82 in *Poetics of Cinema*. New York: Routledge.
- Bordwell, David, and Kristin Thompson. 2010. *Film Art: An Introduction*. 9th ed. New York: McGraw-Hill.
- Branigan, Edward. 1992. *Narrative Comprehension and Film*. New York: Routledge.
- . 1997a. "Sound, Epistemology, Film." Pp. 95–125 in *Film Theory and Philosophy*, eds. Richard Allen and Murray Smith. New York: Oxford University Press.
- . 1997b. "To Zero and Beyond: Noël Burch's Theory of Film Practice." Pp. 149–168 in *Defining Cinema*, ed. Peter Lehman. New Brunswick, NJ: Rutgers University Press.
- . 2002. "Nearly True: Forking Plots, Forking Interpretations—A Response to David Bordwell's 'Film Futures.'" *SubStance* 97: 105–114.
- . 2006. *Projecting a Camera: Language-Games in Film Theory*. New York: Routledge.
- . 2009. "Of Prepositions: Lost and Found." *Velvet Light Trap* 64: 95–98.
- . Forthcoming a. "If-then-else: Memory and the Path Not Taken." In *Interactive Frictions: A Decade of Discourse on Digital Culture*, eds. Marsha Kinder, Tara McPherson, and Alison Trope. Berkeley: University of California Press.
- . Forthcoming b. "Teaching Film Theory." In *Teaching Film*, eds. Lucy Fischer and Patrice Petro. Options for Teaching series. New York: Modern Language Association Publications.
- Bresson, Robert. [1975] 1997. *Notes on the Cinematographer*. Trans. Jonathan Griffin. København: Green Integer Books.
- Buckland, Warren, ed. 1995. *The Film Spectator: From Sign to Mind*. Amsterdam: Amsterdam University Press.
- Burch, Noël. [1967] 1973. *Theory of Film Practice*. Trans. Helen R. Lane. New York: Praeger. Reprinted in 1981 with some omissions by Princeton University Press.
- Cardinal, Serge. 2007. "Radiophonic Performance and Abstract Machines: Recasting Arnheim's Art of Sound." *Liminalities* 3 (3). <http://liminalities.net/3-3/cardinal.htm> (accessed 1 December 2009).

- Carello, Claudia, Jeffrey B. Wagman, and Michael T. Turvey. 2005. "Acoustic Specification of Object Properties." Pp. 79–104 in *Moving Image Theory: Ecological Considerations*, eds. Joseph D. Anderson and Barbara Fisher Anderson. Carbondale: Southern Illinois University Press.
- Carroll, Noël. 1996a. "Cracks in the Acoustic Mirror." Pp. 336–342 in *Theorizing the Moving Image*. Cambridge: Cambridge University Press.
- . 1996b. "The Power of Movies." Pp. 78–93, in Carroll, *Theorizing the Moving Image*.
- Cavell, Stanley. 1979. *The World Viewed: Reflections on the Ontology of Film*. Enlarged ed. Cambridge, MA: Harvard University Press.
- Chion, Michel. 1992. "Wasted Words." Pp. 104–110 in Altman, *Sound Theory/Sound Practice*.
- . [1990] 1994. *Audio-Vision: Sound on Screen*. Ed. and trans. Claudia Gorbman. New York: Columbia University Press.
- . [1982] 1999. *The Voice in Cinema*. Ed. and trans. Claudia Gorbman. New York: Columbia University Press.
- . [2003] 2009. *Film, A Sound Art*. Trans. Claudia Gorbman. New York: Columbia University Press.
- Cilliers, Paul. 1998. *Complexity and Postmodernism: Understanding Complex Systems*. New York: Routledge.
- Costello, Diarmuid. 2008. "On the Very Idea of a 'Specific' Medium: Michael Fried and Stanley Cavell on Painting and Photography as Arts." *Critical Inquiry* 34 (2): 274–312.
- Davidson, Donald. 2001. *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press.
- Dennett, Daniel C. 1991. "Real Patterns." *Journal of Philosophy* 88 (1): 27–51.
- Devereaux, Mary. 1987. "In Defense of Talking Film." *Persistence of Vision* 5: 17–27.
- Doane, Mary Ann. [1980] 1985a. "Ideology and the Practice of Sound Editing and Mixing." Pp. 54–62 in *Film Sound: Theory and Practice*, eds. Elisabeth Weis and John Belton. New York: Columbia University Press.
- . [1980] 1985b. "The Voice in the Cinema: The Articulation of Body and Space." Pp. 162–176 in Weis and Belton, *Film Sound*.
- . 2002. "The Object of Theory." Pp. 80–89 in *Rites of Realism: Essays on Corporeal Cinema*, ed. Ivone Margulies. Durham, NC: Duke University Press.
- Donnellan, Keith S. 1966. "Reference and Definite Descriptions." *Philosophical Review* 75 (3): 281–304.
- Dutton, Denis. 2006. "A Naturalist Definition of Art." *Journal of Aesthetics and Art Criticism* 64 (3): 367–377.
- Eidsvik, Charles. 2005. "Background Tracks in Recent Cinema." Pp. 70–78 in Anderson and Fisher Anderson, *Moving Image Theory: Ecological Considerations*.
- Eisenstein, Sergei. [1939–1947] 1987. "Nonindifferent Nature." Pp. 216–396 in *Nonindifferent Nature*. Trans. Herbert Marshall. Cambridge: Cambridge University Press.
- Eisenstein, Sergei, Vsevolod Pudovkin, and Grigori Alexandrov. [1928] 1988. "Statement on Sound." Pp. 113–114 in *S.M. Eisenstein: Selected Works*. Vol. I. Ed. and trans. Richard Taylor. Bloomington: Indiana University Press.
- Fischer, Lucy. [1977] 1985. "Enthusiasm: From Kino-Eye to Radio-Eye." Pp. 247–264 in Weis and Belton, *Film Sound*.
- Fleishman, Avrom. 1992. *Narrated Films: Storytelling Situations in Cinema History*. Baltimore: Johns Hopkins University Press.
- Flinn, Caryl. 1992. *Strains of Utopia: Gender, Nostalgia, and Hollywood Film Music*. Princeton, NJ: Princeton University Press.
- Gibson, Andrew. 2001. "'And the Wind Wheezing Through That Organ Once in a While': Voice, Narrative, Film." *New Literary History* 32 (3): 639–657.

- Goodman, Nelson. 1976. *Languages of Art: An Approach to a Theory of Symbols*. 2nd ed. Indianapolis: Hackett.
- . 1978. *Ways of Worldmaking*. Indianapolis: Hackett.
- Gorbman, Claudia. 1987. *Unheard Melodies: Narrative Film Music*. Bloomington: Indiana University Press.
- Greene, Liz. 2009. "Speaking, Singing, Screaming: Controlling the Female Voice in American Cinema." *Soundtrack* 2 (1): 63–76.
- Hanssen, Eirik Frisvold. 2006. *Early Discourses on Colour and Cinema: Origins, Functions, Meanings*. Stockholm, Sweden: Almqvist & Wiksell International. Also available at <http://su.diva-portal.org/smash/record.jsf?pid=diva2:189727> (accessed 1 December 2009).
- Hernadi, Paul. 1995. *Cultural Transactions: Nature, Self, Society*. Ithaca, NY: Cornell University Press.
- Hilmes, Michele, ed. 2008. "In Focus: Sound Studies." *Cinema Journal* 48 (1): 115–146.
- Ihde, Don. 2007. *Listening and Voice: Phenomenologies of Sound*. 2nd ed. Albany: State University of New York Press.
- Jarrett, Michael. 2000. "Sound Doctrine: An Interview with Walter Murch." *Film Quarterly* 53 (3): 2–11.
- Johnson, Mark. 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.
- Johnson, William. 1985. "The Liberation of Echo: A New Hearing for Film Sound." *Film Quarterly* 38 (4): 2–12.
- . 1989. "Sound and Image: A Further Hearing." *Film Quarterly* 43 (1): 24–35.
- Kahn, Douglas. 1999. *Noise, Water, Meat: A History of Sound in the Arts*. Cambridge: MIT Press.
- Kant, Immanuel. [1798] 1996. *Anthropology from a Pragmatic Point of View*. Trans. Victor Lyle Dowdell. Carbondale: Southern Illinois University Press.
- Karpp, Anne. 2007. *The Human Voice: How This Extraordinary Instrument Reveals Essential Clues About Who We Are*. London: Bloomsbury.
- Kozloff, Sarah. 1988. *Invisible Storytellers: Voice-Over Narration in American Fiction Film*. Berkeley: University of California Press.
- . 2000. *Overhearing Film Dialogue*. Berkeley: University of California Press.
- Kracauer, Siegfried. [1960] 1997. *Theory of Film: The Redemption of Physical Reality*. Princeton, NJ: Princeton University Press.
- Kristeva, Julia. [1974] 1984. *Revolution in Poetic Language*. Trans. Margaret Waller. New York: Columbia University Press.
- Lakoff, George, and Mark Johnson. 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.
- Lastra, James. 2000. *Sound Technology and the American Cinema: Perception, Representation, Modernity*. New York: Columbia University Press.
- . 2008. "Film and the Wagnerian Aspiration: Thoughts on Sound Design and the History of the Senses." Pp. 123–138 in *Lowering the Boom: Critical Studies in Film Sound*, eds. Jay Beck and Tony Grajeda. Chicago: University of Illinois Press.
- Lawrence, Amy. 1991. *Echo and Narcissus: Women's Voices in Classical Hollywood Cinema*. Berkeley: University of California Press.
- Lefebvre, Martin. 1999. "On Memory and Imagination in the Cinema." *New Literary History* 30 (2): 481–498.
- Levin, Tom. 1984. "The Acoustic Dimension: Notes on Cinema Sound." *Screen* 25 (3): 55–68.

- Mancini, Marc. 1985. "The Sound Designer." Pp. 361–368 in Weis and Belton, *Film Sound*.
- Mast, Gerald. 1977. *Film/Cinema/Movie: A Theory of Experience*. New York: Harper and Row.
- Metz, Christian. [1977] 1980a. "Aural Objects." *Yale French Studies* 60: 24–32.
- . [1977] 1980b. "The Perceived and the Named." *Studies in Visual Communication* 6 (3): 56–68.
- Mitry, Jean. [1963] 1997. *The Aesthetics and Psychology of the Cinema*. Trans. Christopher King. Bloomington: Indiana University Press.
- . [1987] 2000. *Semiotics and the Analysis of Film*. Trans. Christopher King. Bloomington: Indiana University Press.
- Morgan, Daniel. 2006. "Rethinking Bazin: Ontology and Realist Aesthetics." *Critical Inquiry* 32 (3): 443–481.
- Nancy, Jean-Luc. [2002] 2007. *Listening*. Trans. Charlotte Mandell. New York: Fordham University Press.
- Naremore, James. 1988. *Acting in the Cinema*. Berkeley: University of California Press.
- Nasta, Dominique. 1991. *Meaning in Film: Relevant Structures in Soundtrack and Narrative*. Berne: Peter Lang.
- Plantinga, Carl R. 1997. *Rhetoric and Representation in Nonfiction Film*. Cambridge: Cambridge University Press.
- Pudovkin, V[sevolod] I. [1925] 1970. *Film Technique and Film Acting*. Ed. and trans. Ivor Montagu. New York: Grove Press.
- . [1934] 2006. "Asynchronism as a Principle of Sound Film." Pp. 199–207 in *Selected Essays*. Ed. Richard Taylor. Trans. Richard Taylor and Evgeni Filippov. New York: Seagull Books.
- Quine, Willard Van Orman. [1948] 1980a. "On What There Is." Pp. 1–19 in *From a Logical Point of View: Nine Logico-Philosophical Essays*. 2nd ed. Cambridge, MA: Harvard University Press.
- . [1951] 1980b. "Two Dogmas of Empiricism." Pp. 20–46 in *From a Logical Point of View*.
- Rosenbaum, Jonathan. 1978. "Sound Thinking." *Film Comment* 14 (5): 38–41; corrections in 14 (6): 79.
- Ryder, Robert G. 2007. "Walter Benjamin's Shell-Shock: Sounding the Acoustical Unconscious." *New Review of Film and Television Studies* 5 (2): 135–155.
- Sarkar, Bhaskar. 1997. "Sound Bites: Fragments on Cinema, Sound, and Subjectivity." *Spectator* 17 (2): 22–35.
- Sibley, Frank. [1959] 2001a. "Adjectives, Predicative and Attributive." Pp. 154–175 in *Approach to Aesthetics: Collected Papers on Philosophical Aesthetics*, eds. John Benson, Betty Redfern, and Jeremy Roxbee Cox. Oxford: Clarendon Press.
- . 2001b. "Aesthetics and the Looks of Things." Pp. 24–32 in Benson et al, *Approach to Aesthetics*.
- . 2001c. "Why the Mona Lisa May Not Be a Painting." Pp. 256–272 in Benson et al, *Approach to Aesthetics*.
- Siegel, Greg. 2007. "Babbage's Apparatus: Toward an Archaeology of the Black Box." *Grey Room* 28: 30–55.
- Silverman, Kaja. 1988. *The Acoustic Mirror: The Female Voice in Psychoanalysis and Cinema*. Bloomington: Indiana University Press.
- Smith, Murray. 1995a. *Engaging Characters: Fiction, Emotion, and the Cinema*. Oxford: Clarendon Press.
- . 1995b. "Film Spectatorship and the Institution of Fiction." *Journal of Aesthetics and Art Criticism* 53 (2): 113–127.

- Smith, Susan. 2007a. "Singing Outside the Frame: The Female Voice-Off in *Gigi*." *CineAction* 72: 32–41.
- . 2007b. "Voices in Film." Pp. 159–238 in *Close-Up 02*, eds. John Gibbs and Douglas Pye. London: Wallflower. 238.
- Smoodin, Eric. 1983. "The Image and the Voice in the Film with Spoken Narration." *Quarterly Review of Film Studies* 8 (4): 19–32.
- Sterne, Jonathan. 2003. *The Audible Past: Cultural Origins of Sound Reproduction*. Durham, NC: Duke University Press.
- Szaloky, Melinda. 2002. "Sounding Images in Silent Film: Visual Acoustics in Murnau's *Sunrise*." *Cinema Journal* 41 (2): 109–131.
- Thompson, Kristin. 1980. "Early Sound Counterpoint." *Yale French Studies* 60: 115–140.
- Varela, Francisco J., Evan Thompson, and Eleanor Rosch. 1991. *The Embodied Mind: Cognitive Science and the Human Experience*. Cambridge: MIT Press.
- Vescio, Bryan. 2001. "Reading in the Dark: Cognitivism, Film Theory, and Radical Interpretation." *Style* 35: 572–591.
- Walsh, Anne, and Chris Kubick. 2006. "Inventory/F Is for Foley." *Cabinet* 23: 18–19.
- Walton, Kendall. 1988. "What Is Abstract about the Art of Music?" *Journal of Aesthetics and Art Criticism* 46 (3): 351–364.
- Weis, Elisabeth. 1999. "Eavesdropping: An Aural Analogue of Voyeurism?" Pp. 79–107 in *Cinesonic: The World of Sound in Film*, ed. Philip Brophy. North Ryde: Australian Film Television & Radio School.
- Williams, Alan. 1980. "Is Sound Recording Like a Language?" *Yale French Studies* 60: 51–66.
- . 1985. "Godard's Use of Sound." Pp. 332–345 in Weis and Belton, *Film Sound*.
- Wilson, Robert A., and Frank C. Keil, eds. 1999. *The MIT Encyclopedia of the Cognitive Sciences*. Cambridge: MIT Press.
- Wittgenstein, Ludwig. [1953] 1967. *Philosophical Investigations*. Trans. G. E. M. Anscombe. 3rd ed. New York: Macmillan.
- Wolfe, Charles. 1997. "Historicising the 'Voice of God': *The Place of Vocal Narration in Classical Documentary*." *Film History* 9 (2): 149–167.
- Wollen, Peter. 2003. "24 Mismatches of Sound and Image." Pp. 221–230 in *Soundscape: The School of Sound Lectures, 1998–2001*, eds. Larry Sider, Diane Freeman, and Jerry Sider. London: Wallflower Press. A slightly different version appeared in 2002 as "Mismatches (& Acousmètres)." Pp. 242–254 in Wollen, *Paris Hollywood: Writings on Film*. London: Verso.
- Young, Kay. 2001. *Ordinary Pleasures: Couples, Conversation, and Comedy*. Columbus: Ohio State University Press.

## Filmography

- Altman, Robert. *The Long Goodbye*. 1973. USA.
- . *California Split*. 1974. USA.
- . *Nashville*. 1975. USA.
- Fleming, Victor. 1939. *The Wizard of Oz*. USA.
- Hitchcock, Alfred. 1960. *Psycho*. USA.
- Lang, Fritz. 1933. *The Testament of Dr. Mabuse (Das Testament des Dr. Mabuse)*. Germany.
- Van Sant, Gus. 1998. *Psycho*. USA.