General Introduction: Cinema as Event

Rick Altman

A dozen years ago, I edited an issue of *Yale French Studies* entitled *Cinema/Sound*. Featuring essays on the role of the sound track in film theory, history, and analysis, *Cinema/Sound* served as a catalyst for further work on film sound. Though *Cinema/Sound* is now out of print, many of the individual articles have gone on to influence scholars in the United States and around the world. The introduction and almost half of the articles have been reprinted, with several translated into a number of languages, including Russian.

As influential as *Cinema/Sound* may have been, a decade's distance reveals the limitations of the articles that it contains. With few exceptions, these articles treat cinema as a series of self-contained texts, divorced from material existence and the three-dimensional world. Heavily marked by the project of semiotics, most of the articles aim at describing the properties of sound, the relationship between image and sound, or the functioning of sound in a particular textual situation. Treatment of the audience is limited to the experience of film-viewing; contemporary culture is alluded to only when it constitutes a film's specific subject matter; sound technology is treated as if it were used only for films. Published in 1980, *Cinema/Sound* clearly bears the stamp of its text-oriented era.

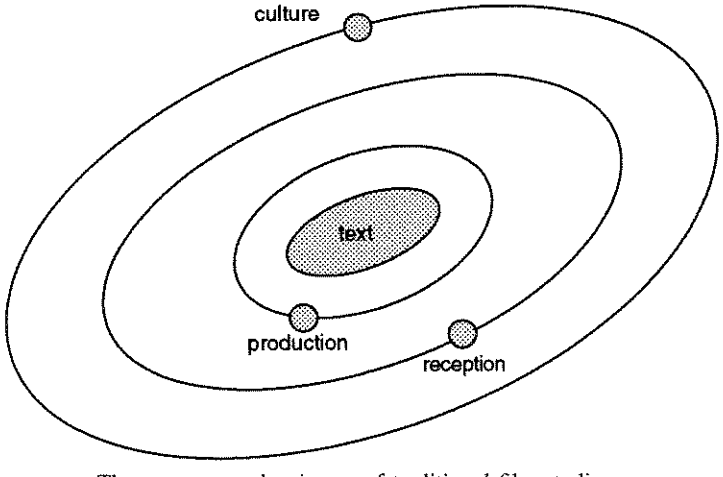
In retrospect, the cost of *Cinema/Sound*'s text-based strategies becomes clear. Though the volume was conceived as a rehabilitation of the sound track, in all its diversity, *Cinema/Sound* actually stresses only a very narrow range of sound-oriented concerns. The sounds of silent films are hardly mentioned; sound technology is almost entirely neglected; no attention is paid to non-narrative, non-feature, or non-western films. More important still, sound itself is most often treated as if it were an ideal conveyor of linguistic or musical information, received by an ahistorical

audience in a generic viewing situation, with no particular moviegoing purpose. While it sensitized film scholars to the importance of the sound track, *Cinema/Sound* masked certain very real problems of the period's film scholarship.

For the present volume, I propose a different model, a new way of thinking about cinema in general and the sound track in particular. Building on recent theoretical developments, this new approach radically extends the range of critical discourse appropriate to film studies, while offering a new coherence among the various types of scholarship currently devoted to cinema.

For decades, film has been regularly defined as a text, an autonomous aesthetic entity most closely related to other autonomous aesthetic entities. During this period, film theory stressed relationships internal to individual films or characteristic of cinema as a whole. Film history typically sorted films according to textual similarity and assessed the evolution of the resultant generic or thematic categories. Film analysis was built on the tacit assumption that differing audiences nevertheless shared the same basic film-viewing experience, regardless of differences in gender, class, or viewing situation. In recent years, this text-oriented model has begun to waver in the face of discursive approaches, feminist theory, cultural studies, and other critical methods sensitive to a broader notion of what film is and how it affects human activities. Considered as a text, each film appears as a self-contained, centered structure, with all related concerns revolving around the text like so many planets.

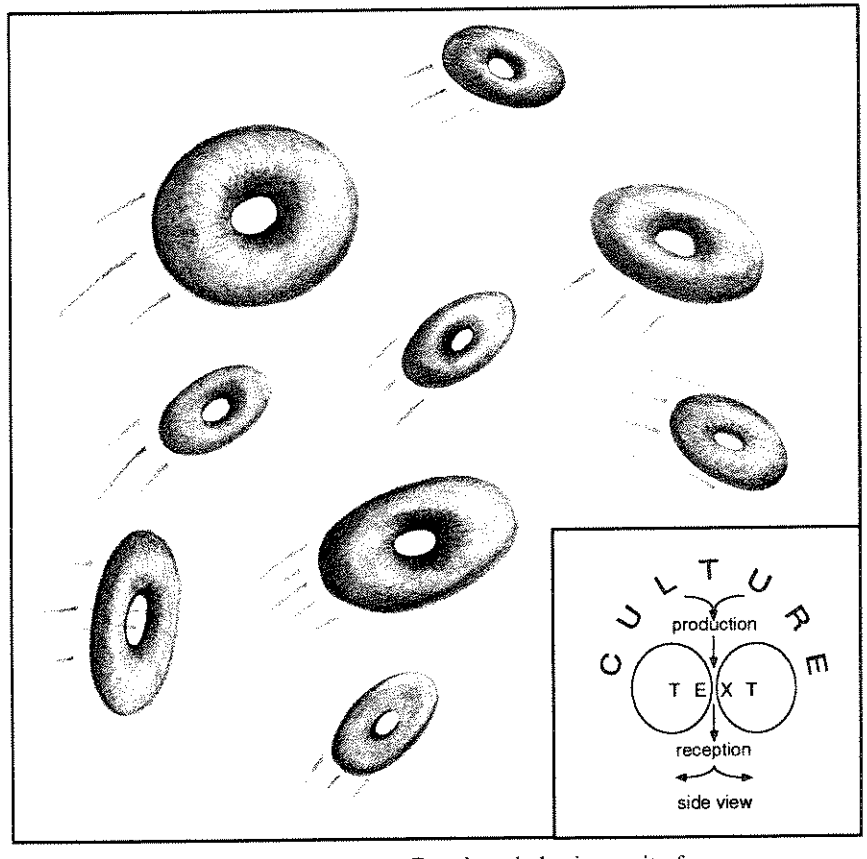
In opposition to the notion of film as text, I have found it helpful to conceive of *cinema as event*. Viewed as a macro-event, cinema is still



The text-centered universe of traditional film studies.

seen as centered on the individual film, but according to a new type of geometry. Floating in a gravity-free world like doughnut-shaped spaceships, cinema events offer no clean-cut or stable separation between inside and outside or top and bottom.

In this three-dimensional Moebius strip world, the textual center is no longer the focal point of a series of concentric rings. Instead, like the pinhole at the center of an hourglass, it serves as a point of interchange between two "V" shapes, one representing the work of production, the other figuring the process of reception. Beginning as a subset of culture at large, one "V" progressively narrows as the work of film production runs its course, first broadly, with diverse ideas and scripts, sets and rushes, technicians and rewrites, until eventually the work of production has been resolved into a single narrow product: the text. The process of reception then broadens out again, eventually reaching the point where it



The geometry of cinema events: Doughnut holes in gravity-free space.

is indistinguishable from the culture in general. In a gravity-free world, however, this hourglass system is entirely reversible. Just as production flows through the text toward reception, so reception regularly influences production.

Each "V" opens out onto an infinite cultural space, containing other cinema events, that eventually bends back around toward the opposite "V". In other words, this Moebius vessel fails to distinguish between inside and outside, though it does have two distinct domains of interchange: the narrow textual isthmus connecting the two "V"s, and the indeterminate peripheral culture which offers any number of avenues of interchange between the open ends of the two "V"s.

Because this new type of geometry does not allow for clear distinctions between inside and outside, between top and bottom, the event that is cinema cannot be identified as privileging one particular aspect of the system. Instead, the cinema event is constituted by a continuing interchange, neither beginning nor ending at any specific point. No fixed trajectory characterizes this interchange, nor is it possible to predict which aspect of the system will influence which other aspect.

Seen as a macro-event, cinema is conveniently characterized by an even dozen attributes: multiplicity, three-dimensionality, materiality, heterogeneity, intersection, performance, multi-discursivity, instability, mediation, choice, diffusion, and interchange. In the paragraphs that follow, while treating in greater detail these various aspects of cinema as event, I will show how the study of film sound in particular is affected by this new approach to the phenomenon that we call cinema.

Multiplicity

By concentrating on cinema's major product, the individual film, traditional approaches to cinema have sought to bring a semblance of unity to a complex phenomenon. Ignoring differences among release prints (both planned and accidental), critics have fastened on the film itself as cinema's common factor. 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.

By stressing the single moment of apparent unity between two periods of multiplicity, critics have effectively neutralized much of cinema's complexity. In doing so, they have systematically concentrated on the uniformity of the image (itself compromised by the difference between film and television formats), thus neglecting such essential variations in

the sound track as 1) three decades of live, unstandardized accompaniment of "silent" films, 2) simultaneous release of silent and sound versions during the late twenties and early thirties, and 3) parallel distribution of magnetic and optical track versions during the fifties and sixties, as well as mono, stereo, and surround versions in the seventies and eighties.

Three-Dimensionality

Symbolic of the attempt to base the definition of cinema on the coherence of the individual film is the care with which film theaters have concentrated attention on a two-dimensional screen. As in Plato's cave, movie theaters hold our bodies in a fixed position in relation to the screen; complemented by carefully arranged lighting, this reduced mobility serves to convince us that film-viewing is limited to the experience of the two-dimensional rectangle before us. Even before 1910, newly built theaters were often engineered to include a ground-floor projection room, thus avoiding a keystoned image and the resultant recognition that the image is the product of a three-dimensional projection system (rather than an automatically produced replica of the seemingly two-dimensional celluloid original).

While the prestige of the image may be enhanced by this emphasis on two-dimensionality, sound is poorly served indeed, for sound cannot exist in a two-dimensional context. Though conventional speaker placement attempts to identify sound sources with the two-dimensional area of the screen, sound occurs only in the three-dimensional volume of the theater at large. Because sound is always recorded in a particular three-dimensional space, and played back in another, we are able to sense the spatial cues that give film sound its personalized spatial signature.

Materiality

Rather than conceiving of cinema as a unified chain of film images, we might instead stress cinema's material existence. Though the material history of painting and sculpture provides ample models for this type of approach (due to the material nature of their medium), note that the publishing history of literature and music offers little encouragement in this direction (since their medium has usually been judged, wrongly, not to be material). Lyric poetry may well have been written for oral delivery to a specific group on a set occasion, but critics have for centuries treated poems as texts made up of words alone. However dependent the novel may have been on developments in the printing industry, critical discourse has systematically abstracted the novel's aesthetic existence from its materiality, overtly privileging the former over the latter while actively neglect-

ing reception conditions. When critics read the filmic text ideally, as music scholars typically read the musical score or literary scholars regularly interpret the literary text, they break all ties with the text's material conditions of existence.

Conceived as a series of events, cinema reveals rather than dissimulates its material existence. From the complexity of its financing and production to the diversity of its exhibition, cinema must be considered in terms of the material resources that it engages. From the standpoint of sound, this shift is of capital importance, for it removes cinema from the customary, purely visual definition. As a material product, cinema quickly reveals the location and nature of its sound track(s), the technology used to produce them, the apparatus necessary for reproduction, and the physical relationship between loudspeakers, spectators, and their physical surroundings. Such an approach encourages us to move past the imaginary space of the screen to the spaces and sounds with which cinema must compete—the kids in the front rows, the air conditioner hum, the lobby cash register, the competing sound track in the adjacent multiplex theater, passing traffic, and a hundred other sounds that are not part of the text as such, but constitute an important component of cinema's social materiality.

Heterogeneity

For many years the film image/text has been the source of a fundamental paralysis in cinema studies. Ostensibly analyzing *the* film, cinema critics have been at pains both to homogenize the lived experience of film-viewing and to avoid undermining that homogeneity. Rather than recognize the legitimate existence of multiple versions of a film, based on diverse social and industrial needs (censorship, standardized length, colorization, foreign-language dubbing, etc.), critics have regularly made a fetish of locating the "original" version. Instead of attending to the variety of exhibition spaces where a given film is projected, or to the diversity of audiences present, or to the various social contexts in which the film is seen, critics typically mold a film's reception to fit a single mode, apparently "neutral" in nature, but typically covertly reflecting each critic's particular reception stance. Text-based criticism often finds a place for decisions and personnel deemed to have contributed to producing the image, but it has rarely known what to do with the non-filmic components of film exhibition: live acts sharing a film program, commercial tie-ins, ticketing policies, seating practices, theater acoustics, intermission activities, popcorn availability, sales of residual products (sheet music, records, videotapes, T-shirts, etc.), and many more.

Recognition of the heterogeneous nature of the cinema experience not

only opens the field to consideration of a broad spectrum of objects, processes, and activities, but has an especially direct impact on the study of sound. As soon as we move away from film as a single, homogeneous phenomenon, we become aware of the heterogeneous chain of objects and spaces which serve as a vehicle for sound. What kind of sound head is installed? On what sort of projector? What type of amplifier is used? With what speaker system? Where are the speakers located? How are they aimed? What are the characteristics of the house acoustics? Do they change according to audience size? Can they be modified? Throughout the history of cinema, exhibitors have paid close attention to these problems, but their concerns have rarely been shared by the scholarly community. Without attention to these matters it is not possible to explain why theater seats evolved from hardback uprights to plush armchairs, why theater architecture and outfitting changed radically shortly after the conversion to sound, or why Tomlinson Holman and Lucasfilm felt the need to develop the THX theater sound system.

Intersection

Just as Foucault proposes replacing history by an archaeology in which the individual strands making up any single event are teased out and separately followed up, we can usefully conceive of cinema events as the intersection of many separate lines of endeavor, throughout the production, reception, and cultural spheres. In the past, little attention has been paid to a film's technical credits; the laborers have systematically been passed over in favor of the architect and contractor. Conception has been preferred to execution, to the point where critics have apparently convinced themselves that film technicians do no more conceiving or take no more initiative than ditchdiggers. Belied at every point by the intellectual quality of cinema's technical journals, the assumption that technicians execute rather than conceive has led critics to neglect the collaborative nature of film production.

Yet the cinema experience results from the intersection of more than just production personnel. For any given film exhibition to take place, the activity of many groups and individuals is required, from cinema architects and screen manufacturers to film distributors and projectionists. Conceived as an intersection, each viewing event involves *lines* of activity, intersecting during the event but beginning beforehand and continuing afterwards. The theater is not there by chance; it was conceived by a firm experienced in designing concert halls or multipurpose recreation centers (the difference between two such firms going a long way to explain the disparity in the acoustical dimension of two cinema events). The musical accompaniment does not happen in an instant separated from time; on the

contrary, it is the result (in sound as well as silent film) of an entire industry with evolving ties to other music industries and cultural precedents. To understand the event is to understand the complexity of this contribution.

Even the spectator must not be conceived as present by chance. While recent theories of subject formation typically posit a single, specific explanation of spectator presence or desire (often psychoanalytic in nature), they rarely attempt to take into account the multiplicity of motives that bring diverse audience members to the theater. While it would be absurd to treat all viewers as separate individuals, sharing no common interests or cultural positions, the notion of the cinema event as intersection has the definite virtue of emphasizing the trajectory bringing each spectator to the theater. And not just to the theater, but away from the theater as well. For an understanding of cinema depends just as much on a knowledge of the activities that cinema engenders or promotes as it does on the desires that bring the audience in. The vogue of the theme song, for example, especially strong in the late twenties and early fifties, cannot be explained by reference to textual evidence alone (although theme songs certainly do have textual ramifications); to understand this phenomenon we need to investigate at least three intersecting lines: the purchase of music companies by Hollywood studios, the growth of a hit-parade approach to radio programming, and the tendency of spectators to perpetuate their experience of a particular film by purchasing sheet music in the twenties or a record in the fifties (the difference between the two revealing the disparity between the active nature of singing around the piano and the passivity of listening to a record).

Performance

Textual approaches to cinema are based on the notion of an unvarying text, thus negating the scandal of variety, neutralizing cinema's need for a spatio-temporally specific projection, and effacing cinema's heritage as a performing art. While standardized presentation has long been the dream of cinema producers, it has never been fully realized in practice. One of the reasons for this is that exhibitors have every reason to employ presentational differences as a prime method of product differentiation. This strategy was especially salient throughout the silent period, when differences in accompaniment (piano, organ, orchestra, lecturer, voices behind the screen, effects, etc.) served as an important method of individualization, along with the other films, acts, and music on the program, not to mention the ushers' costumes, the theatrical décor, giveaway programs, and what have you. However much producers would like to think of film as an ideal image, automatically conveyed to the ultimate ideal consumer, we all know that the film must pass through the hands of a projectionist,

whose performance is constantly open to criticism. While current theaters tend toward what we might call the zero degree of performance (standardized spaces, automatic projection, a program limited to the feature film), even the drab multiplex theater usually constructs the marquee or the lobby as a performance space, featuring still photos, cut-out stand-up characters, and other presentational devices.

Instead of considering the elaborate showmanship and diverse accompaniments of the silent era as an anomaly, or the presentation acts of the thirties as an outmoded practice, we need to recognize that film is always the product of performance (more or less self-conscious, more or less complex, more or less commodified). Basing their conception of cinema on a consistent attempt to dissimulate cinema's performance orientation, critics have regularly neglected important aspects of earlier film exhibition. Silent films are interpreted without reference to their musical accompaniment; feature films are treated independently of the full program of music, shorts, and newsreels that originally accompanied them; the conversion-to-sound practice of switching the sound from one speaker to another (music to the pit speaker, dialogue to the screen speaker) is consistently forgotten; the difference between projection practices in Latin America and New York is entirely bypassed; the problematic location of surround speakers receives no commentary; the reverberation differences between television sets with speakers on the front or on the side are elided. Cinema will recover some of its richness when we learn to remember that for most of its history it was a performance-oriented medium—less spectacularly so than vaudeville, perhaps, but performance-oriented nevertheless.

Multi-Discursivity

After the story-oriented sixties and early seventies, the "discovery" of cinema's discursive nature was one of the most important theoretical processes of the late seventies and eighties. Following Metz and others, recent critics have been quick to recognize the discursive investment of cinematic texts. Typically conceived as addressed by a cinema industry to an undifferentiated audience, films are considered as employing discursivity to construct subjectivity, to propagate ideology, or to create a situation of hegemony. As compared to the cinema-as-text approach, this recognition of discursivity clearly constitutes an improvement. However, the current notion of discursivity, typically collapsed into the singular, fails to capture the complexity of cinema's existence as event. Who addresses the audience in cinema? Is it the culture? the industry? the writers? the director? the actors? the exhibitor? Certainly no single answer is appropriate here, nor will the same answer be appropriate for different

films, or even for the "same" film exhibited in divergent ways in differing places.

On the contrary, the complexity of the cinema experience derives from cinema's extraordinary ability to serve as the intersection of a variety of discourses, framed by diverse groups and addressed to populations varying from single individuals to the entire culture. Sharing the same space and time, these discourses commonly hide one another, with a given film-viewing thus successively revealing little bits of each individual discourse. A film does not carry a single message, unified, unilinear, and univocal. Instead, it is more like a *scarred palimpsest*, at various points revealing diverse discursive layers, each one recorded at a different point in time.

This recognition of the text's layered, potentially contradictory nature offers a new opportunity for attention to sound's discursive contributions. Much ink has flowed recently over the development of narrative images during the nineteen hundreds and the early teens; perhaps it is time now to recognize that the contemporary rejection of short, coherent musical forms (primarily the lapidary popular song) contributes heavily to the development of a unified narrative editing style. In many filmmaking styles, differing types of sound make contradictory discursive appeals. Throughout the thirties, in Europe as well as the United States, films merged rough regional or lower-class speech with the newly popularized, mellifluous radio accents. Hollywood regularly contrasts hyper-intelligible dialogue lacking any spatial markers with point-of-audition sound carrying appropriate volume and reverberation shifts. Television sound often transfers into its narrative programs the volume differentials that typically exist between commercials and public service announcements (or other strictly informative messages), thus splitting the program itself between informer and advertiser of its own information. In a world where sound is commonly taken as an unproblematic extension of the image, within a comfortably unified text, the concept of multi-discursivity is bound to enfranchise sound, concentrating attention on its ability to carry its own independent discourses.

Instability

In spite of cinema's historical connections to theater and the performing arts in general, critics have preferred to emphasize cinema's debt to the novel. Treating a film as the heir apparent to novelistic prose has made it easy for critics to construct cinema as a minimally material object that easily maintains its identity from decade to decade. But film is not made of language, nor is it printed with movable type. Cinema will never have its Gutenberg, because its very existence depends on its multi-discursive performance orientation. Standardization of print made the novel appear to

escape from its material grounding; eventually, standardization of theaters, projection, seating, advertisement, and so forth could possibly reduce cinema's debt to the material differences of performance. For the time being, however, differing discursive investments maintain cinema's material dimension and its performance base. Because of this aspect of the cinema event, what we refer to as "the film" is fundamentally unstable in nature.

It's not just that we almost never see and hear a film as it was originally seen and heard; in fact, we would be hard put to identify what the phrase "originally seen and heard" actually means, since there never was a single original. For silent film music, do we mean Manhattan's Rialto, with its standing orchestra and staff of arrangers, or Thomas Brown's Iowa City Nickeldom, with the first Wurlitzer unit orchestra west of the Mississippi? Is it the proper Eastern style, eschewing rag and comic effects, or is it the broader Western approach, with its syncopated rhythms and aural jokes? Is it the downtown theater with a four-man orchestra, or the rural weekends-only theater with a young girl practicing her recital pieces at the piano? However similar the image in all these cases, the cinema events involved are anything but stable. Surprisingly, the conversion to sound changed little of this. Just because the sound track happens to be inscribed down the side of the film, there is no guarantee of standardized performance. Add to this equation the radical changes in dynamic and frequency response between a first-run theater and a portable television, and the instability of the cinema event (and thus of the cinema text) becomes all too clear.

Mediation

As long as cinema scholars were laboring to establish cinema as an autonomous art and cinema studies as an independent intellectual domain, there was good cause, rhetorically speaking, to play down cinema's debt to other media. Consequently, there has been a regrettable undertheorizing of the relationship between cinema and the extraordinary variety of media to which it is related. Because the cinema event includes the spectator and, by extension, the spectator's experience of other media, we must conclude that one measure of a film's success derives from spectator evaluations based on a set of preestablished notions about what constitutes reality, acceptable ending points, moral behavior, entertainment, and so forth. In other words, the values and standards associated with cinema cannot be described independently of the models through which they are mediated.

The mediation factor is especially important in the case of sound, because sound technology has changed so often over the past century.

Whereas the film image has undergone little more than successive tinkerings (along with the more important developments in color and image shape), film sound has been revolutionized many times, each time in connection with contemporary developments affecting other entertainment and communication industries besides cinema. Once we recognize the mediated nature of the cinema event, film appears caught up in a complex web of potential models. Besides vaudeville and melodrama, the late nineteenth-century concert hall provided a model for silent film's handling of music and effects. In the early and mid-twenties, radio served as a regular model for cinema sound, while later on in the decade the phonographic record provided an ineluctable model for a film sound technology which is based, after all, on phonographic records. (For an overview of mediation in the U.S. film industry during this period, see the introduction to Section Two, "Historical Speculations.") If musicals in the fifties were constantly criticized for slavishly imitating Broadway plays, it is partly because Hollywood was actually emulating the original cast albums of the resurgent long-playing record industry. For years, critics have discussed cinema's tendency to imitate its own previous successes, yet the intertextual motif within film is far more prevalent on the image side; sound, on the other hand, regularly finds its models outside the film medium, whence the necessity to expand the definitional limits of the cinema event.

Choice

Just as the mediation factor expands our notion of the cinema event, so does the phenomenon of choice, whether operative on the production side or the reception side. When a financier chooses to back a show, where is the money *not* going? When a sound man selects a microphone, what types is he implicitly rejecting? When an exhibitor purchases a sound-filtering screen, what were his choices? When a spectator chooses to spend an hour's wage on a film, what were the alternatives? The road not taken is just as much a part of the cinema event as cinema itself. Would the exact "same" movie really be the same in 1915, 1940, 1965, and 1990? No, because its rivals would respectively be vaudeville, radio, television, and video, or perhaps barbershop singing, Bing Crosby ten-inch 78-rpm records, Top 40 seven-inch 45s, or music videos (among others), with each bringing out a different aspect of the film.

Now, while this example is simplified in the extreme, it does serve to highlight the importance of cinema as part of a differential system, in the strong sense of Saussure's semiotics: there are no positive terms, only differences. The same logic applies, then, to the spectator's choices about whether to spend money on the cinema or on baseball, fast food, a new tie, or a Saturday night special. To be sure, we can hardly analyze all

spectator choice patterns, even for a single showing of a single film; we must, however, find some place for these choices in our theories about how cinema works and what it means.

Diffusion

What does cinema facilitate? What are its residual effects? What kind of afterlife does the cinema event have? Sheet music and sing-alongs or Smurf glasses and trips to Disneyworld? For too long, assuming that cinema constitutes a world of its own, we have turned a blind eye to cinema's impact on the urban landscape or living room design, and a deaf ear to its influence on music preferences and dialogue delivery styles. In the concluding chapter of my book on *The American Film Musical* (Altman 1987), I labeled this aspect of cinema its "operational" component. Especially strong in the musical, which achieved diffusion throughout the culture in the form of various operational strategies, cinema's operational aspect leads in a number of largely unexpected directions. As we can see from any Hollywood pressbook of the twenties through the fifties, diffusion of the cinema event hardly took place by accident.

Indeed, some of Hollywood's most successful strategies involve attempts to capitalize on cinema's ability to be diffused throughout the culture. When Erno Rapée wrote his first two hit theme songs in the late twenties ("Diane" and Charmaine"), he was simply carrying out his role as music director and intermittent composer. Soon, however, every studio was looking for the shot in the arm that could be provided by a hit song. When the major studios snapped up all available music publishing houses, the circle was closed; now the publishers could provide publicity for the films, and vice versa. It is through strategies such as this that the culture is marked by the diffusion of cinema's residual effects.

Interchange

It is tempting to assume that all cinema events take place in a predictable downward direction through the center of the cinema event hourglass: production distills multiple inputs into a single text, which is in turn received by an expanding set of spectators. Indeed, this is the way that cinema has traditionally been studied. However, this approach neglects the cinema event's gravity-free Moebius strip nature. The production-text-reception continuum appears to be the "inside" of cinema, with everything else on the outside, yet the unexpected construction of the cinema event suggests that "outside" and "inside" are so continuous as to be indistinguishable. We are accustomed to analyzing the interchanges that take place through the intermediary of the text; we must now become more

attuned to the interchanges between the production-text-reception system and the culture(s) at large.

While this interchange is too large a topic to cover in any detail here, we easily note the important role that sound is destined to play in this area. Sound's ability to diffuse the cinema event throughout the culture is matched by sound's equal capacity to infuse cinema with elements of the culture's soundscape. Through the mediation of the culture's other sound technologies—live and recorded music, radio, television, and many others—film sound is in a constant state of interchange with the culture at large. Standards of intelligibility developed for the telephone invaded Hollywood by the early thirties; Hollywood frequency response and dynamic range set expectations for the radio and record industries; film music now fills our living rooms and shopping malls. Radio-enforced standardization of speech around a middle-class, non-regional model has had an enormous impact on the social ramifications of speech patterns, while cinema has given increased meaning to the smallest sound events of everyday life. Today, political writers learn their trade from cinema scriptwriters; the politicians try to deliver their one-liners with the panache of movie comedians; and now television and cinema have begun to edit dialogue in imitation of political sound bites. Everywhere we turn, we find sound providing a perpetual and highly charged interchange between cinema and its culture(s).

Cinema as event, replacing cinema as text: this will be the watchword of the nineties, as we shall see in many of the essays that make up this collection.

1

The Material Heterogeneity of Recorded Sound

Rick Altman

When we understood cinema as a text, we borrowed our terminology and our methodology from previously established textual domains. An understanding of cinema as event requires new terms and models for a new type of multi-dimensional analysis. It seems especially appropriate to begin this retooling process with the development of a new vocabulary for sound analysis, for sound itself is particularly event-oriented itself. Whereas image analysis has given us many terms and techniques fully consonant with a textual approach to cinema, sound's heterogeneity has much to offer to an event-oriented aesthetic. In order to reap sound's harvest, however, we must take a new approach to sound, replacing the idealist models offered by musical analysis with increased sensitivity to sound's three-dimensional materiality.

Current approaches to film sound systematically borrow a musical model. The most influential introductory film textbook of the last decade defines the acoustic properties of sound as *loudness, pitch,* and *timbre.*[1] This definition is based on the apparent assumption that all film sounds have the nature of musical notes, that is, they are single phenomena, produced instantaneously, emitted from a point source, and perceived in an immediate and direct fashion. With a definition like this one, we can explain many aspects of film sound, such as contrasts or confluences in volume, frequency, and tone.

In fact, since the terminology is borrowed from the realm of music, we find that with these terms we can handle almost any of the types of analysis typically practiced on a musical score. We note Hitchcock's suspenseful diminuendo from a loud slam to muffled scratching, the harmony of Orson Welles' bass and Joseph Cotten's tenor, the melodic gifts of Cary Grant and Katharine Hepburn, the awkward timbre of Zasu Pitts and Jerry Lewis,

or the varied instrumentation of the "Symphony of Sounds" with which Rouben Mamoulian opens *Love Me Tonight*. It we could notate all film sounds according to the musical criteria of loudness, pitch, and timbre, then these three criteria would suffice for the analysis of film sound.

Yet this is precisely what we cannot do. While all film sounds have loudness, pitch, and timbre, not a single sound in cinema can be adequately described with this musical terminology. In fact, not even musical sounds can be fully described with musical terminology. More appropriate for describing musical scores than individual performances, musical terminology pays little attention to the details of any particular performance, concentrating instead on the common factors joining all performances of the same score. If I attend three concerts of Mozart's "Little Night Music," one in a well-upholstered salon, another in a large concert hall, and a third in a city park, I am in one sense hearing the "same" music three times, that is, music that is represented by a single, identical score. Yet how different are the sounds that reach my ears during the three concerts!

Musical notation assumes that each sound is single, discrete, uniform, and unidimensional. Stressing the formal concerns of music's internal, self-referential aspect, musical notation diverts attention from sound's discursive dimensions, concealing the fact that sound is in reality multiple, complex, heterogeneous, and three-dimensional. As a concept, middle C exists independently of space and time, in the abstract notion of a sound of approximately 262 cycles per second. As a reality, however, no two versions of middle C are identical, because of the different temporal and spatial circumstances in which they originate and are heard. The middle C located on the first line below the G clef may be only a concept, but the sound that we hear with our ears—whether on the street or in a movie theater—is a heterogeneous event that carries its own temporal and spatial dimensions and constitutes a full-fledged narrative. When we listen to recorded sound we are therefore always listening to a particular account of a specific event.

In order to respect the discursive complexity that is characteristic of all sound events, we can no longer continue to depend on a fundamentally conceptual terminology that remains insensitive to sound's phenomenality. Instead we must have a terminology capable both of respecting sound's heterogeneous nature and of figuring the narrative component built into the very process of recording and reproducing sound. This article proposes such a terminology, based on a schematic but systematic review of the physical phenomenon that we call sound.

Sound Events: The Production of Sound

What is sound? What happens when a sound is made? While this is hardly a technical treatise, it will nevertheless be useful to recall the

manner in which sounds are produced. Three elements are required for the production of any sound. First, there must be *vibration*, such as that of the vocal cords or a violin string. Second, the vibration must take place in a *medium* whose molecules can be set in motion, such as air, water, or a railroad rail (sound cannot be transmitted through a vacuum). Third, the transmitting medium must absorb and transmit the original vibrations in the form of *changes in pressure*. When a violin string is plucked or bowed, the molecules of the surrounding medium are compressed, with the pressure passed on from one molecule to the next. When the string reaches the end of its travel, maximum *compression* is achieved. As the string starts back, the molecules rush back to fill up the void left by the departing string. When the string reaches the end of its travel in the opposite direction, maximum *rarefaction* occurs. In order to create a specific, recognizable note this process must be repeated in rapid succession hundreds or even thousands of times a second. For example, the G string on a violin causes the surrounding air to go through 196 *compression/rarefaction cycles* per second, commonly expressed as a frequency of 196 cps or 196 Hz. In other words, what we call the musical note G below middle C is in fact a series of rapid changes in pressure.

Even taking the three-dimensional nature of sound events into account, however, this description vastly oversimplifies the situation. Whereas an electronic tone generator is capable of producing pure tones, all musical instruments produce notes that combine a *fundamental* frequency (such as the violin's 196 Hz G string) with a series of partials: *harmonics* (tones whose frequency is a whole number multiple of the fundamental) and *overtones* (tones whose frequency is related to the fundamental according to a more complex formula). Depending on the instrument and the way it is played, the combination of harmonics and overtones can vary tremendously. When played in such a way as to emphasize the upper harmonics, for example, the violin sounds harsh and strident, while a mellow tone results from stressing instead the lower harmonics. If the oboe, trumpet, flute, and cello sound so recognizably different, it is primarily because they produce radically different combinations of partials.

While few people are trained to hear harmonics and overtones, most listeners will rapidly recognize their absence, as when music is played through the telephone or over an old record player with limited frequency response. While the loss of these partials reduces our pleasure in listening to music, it may have an even more radical effect on other sounds. Spoken language becomes far more difficult to understand, voices and familiar sound effects may become harder to differentiate, even our ability to judge the distance and direction of a sound source may be impaired. In other words, the composite nature of sounds is hardly limited to music.

In fact, most of the events that we think of as a single sound are not

singular at all. The musical model of tone generators and violin strings is extremely misleading. If a violin note could be produced by a violin string alone, then Stradivarius would never have become a household name. Every violin note is a complex event combining the vibrations of a string, a wooden case, and the air trapped inside that case. Each of these three contributes to the overall tone of the note played.

For what we call a sound is typically made up not only of multiple frequencies, but actually has multiple different fundamentals produced over a period of time. Think of the following familiar sounds: a refrigerator, snoring, a lawnmower, the wind, a squeaky door. We think of each as a single sound, but none is actually single in the way that an A-440 produced by a tuning fork is unitary. Each of these sounds constitutes an event taking place in time, involving multiple separate sounds organized in a familiar, recognizable fashion. Given the importance of rhythmic and melodic elements for our recognition of each of these sounds, it would be more appropriate to compare them to musical phrases than to individual notes.

Yet even individual notes have a temporal dimension. Returning for a moment to our violin string, consider the difference between plucking and bowing the string. In one case the sound starts suddenly, reaching its full volume extremely rapidly; in the other case the violinist seems to be sneaking up on the note, teasing the molecules into moving rather than suddenly shoving them. Whether violent or peaceful, this initiation of the sound event is termed the *attack*. It is followed by the *sustain*. How long is the note hold? How long does it stay at full volume? Finally, the sound fades away. This stage is called the *decay*, implying not only a temporal measure but also a qualitative one. Compare, for example, the decay of a plucked string that is simply allowed to spend its own energy and the decay of a plucked string instantaneously dampened by a finger.

As parts of the *sound envelope*, the stages of attack, sustain, and decay apply equally to any sound event. Contrast, for example, the smooth attack of Orson Welles' opening voice-over in *The Magnificent Ambersons* to the sharp attack of Georgie Minafer's dialogue. How essential to the *soundscape* of *The Wizard of Oz* is the gulf separating Margaret Hamilton's staccato attack and nearly instantaneous decay from Judy Garland's ability to ease in and out of speech! Anyone who has ever tried to edit dialogue will understand just how important the elements of the sound envelope are for the establishment of auditory realism. Even when the initial or final words of a sentence are perfectly comprehensible, they create an uneasy feeling whenever part of their attack or decay has been cut off in the editing process.

The production of sound is thus a material event, taking place in space and time, and involving the disruption of surrounding matter. This doesn't

mean that we have to be molecular physicists or sound engineers to understand sound, but it does suggest a very precise basis for our description of sound events. It is no longer sufficient to analyze a musical score or a written text to understand the effects of a particular performance event. Recognizing the extent to which sound sets matter in motion—albeit invisibly—we readily see the importance of developing a vocabulary and a methodology appropriate to the complex materiality of sound. Instead of describing just a sound's loudness, pitch, and timbre, we stress the extent to which every sound event includes multiple sounds, each with its particular fundamental and array of partials, each with its characteristic sound envelope, each possessing its own rhythm within the sound event's overall temporal range.

The Sound Narrative: The Story of a Sound Event

In order to understand sound as it is produced, we need to recognize the material heterogeneity of sound events. Sound production is only part of the story, however, for sound, like the proverbial tree falling in the forest, must be heard in order to take on its narrative and social significance. By offering itself up to be heard, every sound event loses its autonomy, surrendering the power and meaning of its own structure to the various contexts in which it might be heard, to the varying narratives that it might construct. Beginning as the vibration that induces molecular movement, sound is not actualized until it reaches the ear of the hearer, which translates molecular movement into the sensation of sound. Just as the sound event necessarily introduces a temporal dimension into the production of every sound, so the process of perception always guarantees sound's spatial nature.

When we speak of language, we implicitly agree to disregard certain aspects of linguistic discourse as somehow sub-linguistic. Fred Astaire and Ginger Rogers may make something of the difference between ee-ther and eye-ther, but no normal user of the English language shows such a concern. Regional accents and personal idiosyncracies produce recognizable differences, but these are not taken to be differences in language. Whether it's ee-ther or eye-ther, it's still the English word "either." Our understanding of sound works in a very similar fashion. We know that our neighbor's lawn mower sounds very different when it's mowing on the near or the far side of the house (and *vive la différence!*), yet that difference does not change our nomenclature. Whether the sound comes from the near side or the far side of the neighbor's house, it is still the sound of the neighbor's lawn mower. The sounds are different, but the *name* of the sounds is not (Metz). Systematically, the name of a sound

refers to the production of sound and not to its consumption, to the object making the sound rather than the person perceiving it.

Yet the hearing process necessarily involves important variables that often outweigh the sound itself in importance. It doesn't take children long to learn that the word "Boo!" does not by itself produce surprise. When a child jumps out from behind a chair at the other end of the room and shouts "Boo!" the reaction is likely to be mild indeed. When my ten-year-old suddenly emerges from beneath my desk, on the other hand, she can be assured of a good return on her "Boo!", however quietly it may be spoken. To be sure, the sharp attack of the letter "b" contributes to the effect. (If perchance you are not convinced of this, try to scare someone with the pastoral attack of a "Moo!") Still more important, though, is the proximity effect obtained by a good surprise. The effect is dissipated if the booer holds her hand in front of her mouth or looks away from the booee. The reason for this is very simple: the surprise is created largely by the sudden arrival of a zone of sound pressure on the ear. Anything that diminishes the sharpness of this experience (standing too far away, whispering instead of shouting, facing away from the booee, or uttering the "Boo!" before emerging from the hiding place) spoils the effect. Having learned to distinguish between various versions of the "same" sound, our ears tell us how to react not on the basis of the sound event alone, but also according to our perceived relation to that sound event.

How does a sound event contribute to hearing? And what are we actually perceiving when we hear? In the previous section, I explained the molecular basis for sound's characteristic compression/rarefaction cycle. Vibration creates pressure, which is communicated through a medium. At the other end of sound's path, the human ear collects that pressure and transforms its mechanical energy into electrical impulses that the brain understands as sound. Sensitive to frequency (pitch), amplitude (loudness), and many other factors, the human ear is a marvelously sensitive organ capable of very minute distinctions. The ear hears not only a sound's fundamental frequencies, but its harmonics and overtones as well, thus facilitating the distinction between male and female voices or French horns and saxophones. Through the ear's ability to sense not only pressure but the rate of changes in pressure as well, we are able to measure even minute differences in the sound envelope, and thus to distinguish between individual voice patterns.

The ear must do far more than this, however, for until now I have assumed that sound arrives directly to the ear, in a single pencil of pressure. This is precisely not the case. Imagine an actress standing in the center of a stage in a large auditorium, 150 feet wide and 200 feet deep. Since sound travels at about 1130 feet per second at 70 degrees Fahrenheit, the actress's voice takes approximately one-eighth of a second to reach a member of the audience sitting three-quarters of the way back in the orchestra.

But what happens to the sound that goes straight to the spectator on the edge of the auditorium? Certainly it doesn't die there; it must eventually reach the ears of other audience members as well. Radiating out like a cone from the actress's mouth, the sound pressure soon films up the entire auditorium, bouncing off the walls, the floor, and the ceiling, and bending around audience members, chairs, and posts until it is finally completely absorbed.[2] The notion that sound travels in a straight line from sound event to hearing ear is thus radically incomplete. In addition to *direct sound*, there is also a great deal of *reflected sound* or *reverberation*, produced by the sound that reaches the hearing ear only after bouncing off one or more surfaces. In a large room, the delay between the arrival of direct sound and the arrival of the last reflections can be quite long. When the full effect of three-dimensional reverberation is considered, delays of multiple seconds may easily be encountered.

Contrary to popular assumptions, even apparently instantaneous sounds thus have a considerable temporal dimension. Our notation systems for sound reinforce a received notion that separately produced sounds are also perceived separately. As they are printed, Hamlet's words "To be or not to be" provide a blueprint for sounds that are clearly separate and sequential. As they are perceived, however, the direct sound of one word is often heard before the reflected sound of the previous word ceases. Musical notation systematically distinguishes between *melody* (sequential sounds) and *harmony* (simultaneous sounds). Yet the sounds notated as sequential are heard as overlapping, thus confounding the received distinction. The reflected sounds of the first beat of the measure continue to be heard as the direct sound of later beats reaches our ears.

Such a distinction might easily appear purely academic and theoretical. Our ears know, however, that this is not just a question of splitting hairs. Who has not been in a large auditorium, cafeteria, or gymnasium and had trouble making out the speaker's words. The master of ceremonies may be saying "The winning numbers are seventeen, forty-three, fourteen, and seventy-two," but what we actually hear is more like "The win—— num—— seven—— four—— four—— seven——." Because we know what to expect in this context, we easily complete the opening words, but strain as we might, there is no understanding the all-important numbers, for the reflected sound of the first part of each word is bouncing all around the cafeteria—off chairs, tables, floor, walls, and ceiling—long enough to obscure the direct sound of the second half of each number. We are all aware of the difficulty of understanding a telephone message with the competition of a nearby conversation; our ears know that speakers in large halls often provide their own competition.

The subsistence of reflected sound does more than block understanding, however. Our ears are marvelously tuned instruments, extraordinarily capable of making fine distinctions of which we remain largely unaware.

Imagine that there are two actors on our stage, one facing the audience, the other facing backstage. The lights are low; the audience cannot always be sure of seeing which actor's lips are moving. Yet we never have any doubt whatsoever about who is speaking. Our ears tell us.

The first actor, facing the audience, sends a strong ray of sound directly to each spectator (as well as an infinite number of rays that reach individual spectators as reflected rather than direct sound). The words pronounced by actor two, away from the audience, are prevented by the actor's own head from reaching spectators directly. In order to be heard at all, these sounds must rebound off the set or backdrop, thus taking up to three times as long to reach spectators. Fortunately, these spectators are also auditors. Their ears rapidly process this data and easily distinguish between the words that are being spoken directly to them and the words that have to bounce around the theater before arriving. This ability to measure the *ratio of direct to reflected sound* provides one of our most important capacities: the ability to distinguish between sounds that are being spoken to us and those that are meant for others. Imagine John Wayne walking down a line of new recruits standing at attention. The script might read "Johnson, straighten up. Jackson, button that top button. Jones, get that chin down. Altman, where'd you learn to tie a tie?" and so forth. The name at the beginning of each sentence, apparently spoken to gain each recruit's attention in turn, is actually quite redundant, for as Wayne moves down the line there is a change in the ratio of direct to reflected sound heard by each recruit. I need no course in acoustics to know when it's my turn, when I am the one being addressed.

In other words, the fact that a "single" sound reaches our ears over a period of time permits us to reconstitute certain facts about the circumstances surrounding the production of that sound. What our ears are doing is a form of narrative analysis. They are analyzing the narrative produced by sound pressure, in all of its complexity, in order to ascertain how, by whom, and under what conditions that sound pressure was produced. To be sure, some people have ears that are better trained in this process of narrative analysis than others, but we have *all* developed over the years a great deal of expertise in this area. We use the delay between visual information and the first arrival of direct sound to determine the distance of the sound source. The difference in the characteristics of sound arriving at our two ears permits us to locate the sound source laterally. The ratio of reflected to direct sound helps us to decide whether the speaker is facing us or not. Combined with other information, this ratio also helps us recognize the size of the room in which the words are spoken. By noting how long the reflected sound lasts, we refine our conclusions about the originating space.

In fact, we regularly draw still other conclusions from the other aspects

of sound's itinerary. During the course of its picaresque journey from production to perception, sound not only takes many specific courses, it sets in motion a particular medium and is reflected off particular surfaces. Imagine the difference that would be made by staging the preceding example of two onstage actors in two different theaters, one a plush Broadway theater with a velvet backdrop and the other a high school gymnasium with a concrete wall at the back of the stage. Just like a tennis ball thrown toward the back of the stage, the actor's words will in one case be muffled by the backdrop while in the other case they will shoot off of it at nearly their original velocity and volume. While few people are aware of the theory underlying such differences, our ears are surprisingly attentive to them. They seem to know that certain surfaces reflect different frequencies better than others, that some surfaces absorb more sound and dampen specific frequencies more than others, that some environments will continue to reflect sound almost indefinitely, while others will restrict reflected sound to a minimum. In this way, even before we look out the window, our ears tell us that it has snowed during the night. They help us distinguish between the recording of a junior high concert made in the gymnasium and the next day's recording of the "same" concert made in an upholstered, acoustically treated auditorium. They help us to get the "feel" of every room we enter, without ever touching any of the room's surfaces.

The fact that we come equipped with two functioning ears each makes still more information available to us. Because all sounds that are not exactly equidistant from both ears arrive at our ears one after the other, and under slightly different conditions, our ears are able to localize sound laterally as well as in terms of distance. Especially when aided by a radar-like rotation of the head, our own personal sonar gives us varied information about our soundscape.

Our ears are so good at decoding sound that it would be a shame to deprive our terminology of our ears' expertise. Without entering the specialized worlds of acoustics, audio engineering, and otology, we must nevertheless find ways of respecting not only sound's material heterogeneity, but also the cleverness of our ears in analyzing the auditory narratives that it constitutes. Constantly delayed, dampened, reinforced, overlapped and recombined, sound provides us with much of the information we need to understand its origins and its itineraries—but the existing terminology clearly does not.

The Sound Record: Recording the Story of a Sound Event

Every sound initiates an event. Every hearing concretizes the story of that event. Or rather, it concretizes a particular story among the many that could be told about that event. When the baseball broke the window, I

was outside, more than a little worried; I heard the sound of the break directly, with little reflected sound, since there are no walls and ceiling outdoors to keep the reverberation going. My father was sitting in his favorite chair, right next to the broken window; he was subjected not only to the direct sound of the impact but also to a roomful of reflected sound. My mother was ironing in the back room; she thought something had broken, but the muffled reflected sound that reached her didn't specify whether it was a window, a vase, a car headlight, or something else still. Doing her homework in the second floor back bedroom, my sister hardly knew anything had happened. At least until she heard my father bellow.

All four of us heard the "same" sound, yet all four of us heard a "different" sound. Or, to put it in a more useful fashion, each of us heard a different narrative of the same event. Sound's existence as both event and narrative immensely complicates—and enriches—our understanding. Usually discussed as the most transparent of classical narratives, sound is in fact a *Rashomon* phenomenon, existing only in the separate stories of various perceivers of the original event. Potentially important apropos of any sound and its perception, this fact takes on special significance in all media that make use of recorded sound. For what the record contains is not the sound event as such but a record of a particular hearing, a specific version of the story of the sound event. Every recording is thus signed, as it were, with the mark of the particular circumstances in which it was heard. A recording of the shattering window made next to my father's easy chair will be signed in a different way from a recording of the "same" event made next to my sister's desk. Every recording carries the elements of this *spatial signature*, carried in the audible signs of each hearing's particularities. Even when those signs are contradictory or have been tampered with, even when they seem not to match the visual data provided with the sound record, they still carry information that is narrative and spatial in nature.

The situation is immensely complicated by the fact that sound records never convey exactly the same information that a given auditor would experience. Far from arresting and innocently capturing a particular narrative, the recording process simply extends and complicates that narrative. Just as the upholstery of a particular soundscape has an impact on the sound narrative, so the way in which sound is collected and entered into memory becomes part and parcel of the overall sound phenomenon.

Even in the simplest of sound collection systems, decisions regarding the location of the microphone carry enormous importance, especially when the sound is to accompany a related image. Should sound collection take place in the same room as the sound to be recorded? At what distance? Under what acoustic conditions? Or should sound collection be in a remote location, thus reducing volume, dampening certain frequencies, and increasing the ratio of reflected to direct sound? This approach will certainly convince auditors that they are not located in the same sound space as the speaker. In fact, if the reverb level is high enough and the image slightly out of focus, the sound may even appear to have been collected in a time frame different from its production.

The process of editing further complicates the question of microphone location. Should the microphone location be changed every time the camera is moved and the shot changes? Or should sound logic remain entirely independent of image logic? To what extent is consistency of sound collection needed? Must sound collection decisions be subordinated to narrative concerns? Under what conditions may the volume and spatial characteristics of synchronized sound be modified during the editing process? Are there special volume and reverberation requirements for sound effects recorded separately from dialogue? All these and many other questions are implied by the simple necessity of choosing a microphone location.

Since the very beginnings of sound cinema, filmmakers have been convinced that intelligibility is one of the most important requisites of recording speech. Indeed, nowhere else are the stakes of microphone location so clear. Imagine that we are recording a sentence spoken by a woman to the man she is facing. While she is speaking a child walks silently past, catching the woman's attention and causing her to turn away from her interlocutor. Now, in order to maximize the intelligibility of the woman's words we might legitimately decide to "pan" the microphone with her, so that she is always talking directly into the mike, maximizing direct sound and thus intelligibility. Note, however, that this decision robs the sound track of its spatial characteristics. Instead of telling us that the woman turned away from her initial position, the sound track implies that she continued to face in the same direction.

If, instead, we choose to retain the initial microphone position throughout, the sound track will exhibit a faithful spatial signature, but it will almost certainly reduce our ability to understand the final parts of the woman's sentence. We will realize that the woman has turned her head while talking, but, like the man to whom she speaks, we may miss some of her words. Recording choices, as we easily see from this example, govern our perception of particular sound events. Far from simply recording a specific story of a specific sound event, the sound engineer actually has the power to create, deform, or reformulate that event. In the example just illustrated, the sound engineer must choose to allow either deformation of the dialogue or mistaken perception on the part of the auditor.

Nor is microphone location the only variable available to the sound engineer. The microphone itself makes many choices regarding the type,

amount, and source of sound that will be collected. It is perhaps useful, in an image-oriented world, to think of the microphone as a "sound-camera," a collection device for sound that shares many of the characteristics of familiar image-collection devices. Just as cameras may have wide-angle or telephoto lenses, changing the angle of image collection and thus the apparent distance of the object filmed, so microphones vary from omnidirectional to narrowly focused, thus changing both the angle of sound collection and the apparent distance of the sound source. In addition, the change in the ratio of direct to reflected sound that accompanies a change in microphone may also affect perception of room size and other characteristics.

Microphones also vary in their sensitivity to specific sound frequencies. The familiar carbon microphone in our telephones has an extremely limited frequency response. Sound heard over the telephone thus always sounds dull and lifeless. Close-miking with a telephone mike (or stripping the sound of appropriate frequencies in postproduction) thus gives the impression that all sounds presented are being heard through a telephone. Since no microphone is equally sensitive to all frequencies, the choice of a microphone fairly assures that some sounds will be boosted, while others will be dampened.

Many other microphone characteristics may come into play as well. It is often assumed that every microphone produces a faithful sound record. Actually, *no* microphone produces an entirely faithful sound record. Not only does every microphone have its own particular *directional characteristics* (omnidirectional, bidirectional, cardioid, shotgun and so on), but every microphone also has its own particular frequency response, sound configuration, and power requirements. In addition, many microphones produce unwanted sounds of various types (hum, pop, hiss, buzz, crackle and so on) in a wide variety of situations (loud sound signal, wind pressure, close sound source, vibration and so on).

Recorded sound thus always carries some record of the recording process, superimposed on the sound event itself. Added to the story of sound production we always find the traces of sound recording as well, including information on the location, type, orientation, and movement of the sound collection devices, not to mention the many variables intervening between collection and recording of sound (amplification, filtering, equalization, noise reduction, and so forth). Indeed, the recording system itself provides one of the most important determinants of sound characteristics; as such it not only provides a record of sound, it also participates in the overall sound narrative. Think for example of the differing frequency responses of 78 rpm records and digital compact disks. It is so difficult to compare musical performances recorded on these two radically different technologies that the masterworks of Toscanini and Furtwangler seem diminished without the wonders of digital remastering (which is none other than an attempt to restore the frequencies to which pre-war disk recording was not sensitive).

To record is thus to recall to mind, as the dictionary would have it, but like most mnemonic devices, sound recordings must heighten some aspects of the original phenomenon at the expense of others. So-called recordings are thus always representations, interpretations, partial narratives that must nevertheless serve as our only access to the sounds of the past.

Sound Reproduction: Playing the Record of the Story of a Sound Event

But how can we gain access to those sounds? A recording, as we all know, is not a sound. Without some sort of playback device, a recording can only sit silently on the shelf. And as long as it sits on the shelf, it has only one space: the space of the recording of the original sound event. My record of Oistrakh and Rostropovich playing the Brahms Double Concerto with George Szell and the Cleveland Philharmonic Orchestra was recorded in Severance Hall. Once I put the record on my stereo and set the needle down, however, the Concerto becomes very Double indeed. Not only do I hear the fabulous acoustics of the Cleveland Orchestra's home concert hall, but at the same time I have to put up with the less than ideal acoustics of my own living room. Every sound I hear is thus double, marked both by the specific circumstances of recording *and* by the particularities of the reproduction situation.

The late twenties were a time of particularly intense reflection on this problem. Throughout the twenties, movie theaters had grown increasingly large and ornate. While the desire to accommodate growing numbers of higher class patrons was an important factor in the rise of the picture palace, theater acoustics played a part as well. Silent films depended heavily on music chosen from familiar nineteenth-century sources. Now, the nineteenth century had little use for chamber music or small baroque organs, preferring instead large orchestras and enormous choirs, along with the long-lasting reverberations and high indirect-to-direct sound ratios characteristic of spacious concert halls or churches. The music that silent cinema inherited from late romantic composers was thus expected to sound as if it were being played in large, enclosed spaces. Ornate picture palaces, with their multiple levels, private boxes, rows of fluted columns, and endless plaster moldings, were thus a perfect environment for the sounds of silent cinema.

When synchronized dialogue came to cinema, however, a new set of requirements was rapidly imposed. The words had to be comprehensible; to that end the amount and duration of reflected sound had to be kept

extremely low. Studios thus rushed to create acoustically treated sets, open-ended studios, and other devices designed to limit reverberation and maximize intelligibility. The biggest stumbling block of all turned out, however, to have nothing to do with film production. It was film exhibition that caused all the problems. Built to maximize reverberation, those drafty barns called picture palaces made it nearly impossible to understand the spoken word. When played over the loudspeakers of a huge, hard-surfaced Roxy with three second decay time, even intimate scenes recorded to give the impression of small, private spaces sounded as if they were set in cavernous public halls. Carefully recorded speech turned into the same auditory mush that had become the trademark of romantic church organs and mighty theater Wurlitzers alike. Only careful redesign and costly acoustic treatment were able to solve this problem.

Even with the practical problem solved, however, the theoretical difficulty remains. Which acoustics am I listening to? The Hollywood sound stage or the Rialto? Severance Hall or my living room? For that matter, which sound am I listening to? The original sound event or its loudspeaker reproduction? In order to understand sound—cinema sound in particular—we must recognize both the narrative and the represented nature of sound as it reaches our ears in the movie theater. The sound system plays the record of the story of an event. At every point in that chain, new variables enter, new elements of uncertainty. Sound heads, amplifiers, leads, loudspeakers, and theater acoustics all force new auditory data on the audience, just as the recording process itself had earlier introduced an implicit viewpoint.

Hearing Events: Hearing the Record of the Story of a Sound Event

Just as sound events remain only hypothetical sound sources until they are actualized by a hearer, so the playing of a sound record takes on meaning only in the presence of an audience. Yet the process of hearing a recording differs significantly from listening to a live sound event. This should come as no surprise to anyone who has contemplated the difference between a photograph and the scene that it represents. When we look directly at a scene we gain a sense of depth from our binocular vision, by rotating our head, or by moving to the left or right. If we want to know what's underneath a chair we have but to lean down. In order to get a clearer view of a specific object, we need only adjust the focus of our eyes. Yet all of this is to no avail when we view a photograph. No amount of rotating, moving, leaning, or adjusting will deliver information that the photograph lacks. We may have two eyes, but we might as well all be Cyclops when it comes to sensing distance in a photograph, for here the concept of distance is encoded through size, masking, and detail rather

than sensed by the parallax implicit in binocular vision. Without requiring any special education, we have all learned to use our eyes radically differently when we view three-dimensional space and when we view a two-dimensional representation of that space.

A similar situation obtains with the sound representations that we call recordings. When attempting to locate a crying child we normally call heavily on our binaural hearing system to provide cues regarding lateral location. When we listen to a recording of a crying child, no such localization is possible. However much we might rotate our heads or change positions, we remain unable to make use of the directional information that was present when the sound was produced, but which is no longer available in the recording (unless it is in stereo, and even then the location of microphones and speakers plays just as important a role as the location of the original sound source). For listening to the sound pouring out of a loudspeaker is like hearing a lawn mower through an open window: wherever the lawn mower may actually be, it always appears to be located on the side of the house where the open window is.

When we listen for a crying child, we are marvelously effective at cutting out extraneous sounds and concentrating on the cries that we recognize as those of our own child. Dubbed the *cocktail party effect* by Colin Cherry, the process of selective auditory attention is far more difficult when we are listening to recorded material. Whereas live sound provides an extraordinary number of variables, each permitting and promoting selective attention, recorded sound folds most of those variables into a single, undifferentiated source. In a live situation, we easily differentiate among the various sound sources surrounding us, but with recorded sound no such clear distinctions are possible.

Live sound situations reveal the actual relationship between the sound producer and perceiver, while recordings suggest only an apparent relationship. If I sit in an auditorium and listen with my eyes closed to a series of speeches, I remain constantly aware of the speakers' location. I know what direction they are facing, how loud they are speaking, and what tones of voice they are using. When I listen to a recording of the same meeting, I can no longer locate the speakers. Nor can I be sure of their original body positions, volume, or tones. Depending on the type, location, and movement of the microphone(s) used in the recording process, the 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.

What happens in the course of a hearing event is thus not the expected detective activity wherein the hearer searches the recorded sound track for clues permitting reconstitution of the original sound event. Instead, we

follow the trail that has been laid for us all the way to an apparent sound event having all the aural guarantees of reality but only partial correspondence to the original sound event. Indeed, it is the partial nature of the relationship that makes hearing events so fascinating. If there were no connection between the apparent sound event and the original sound source, recorded sound would not have its extraordinary capacity for ideological impact. It is precisely because recorded sound seems to reproduce an original phenomenon that recordings attract and hold audiences so readily. Between the illusion of reproduction and the reality of representation lies the discursive power of recorded sound.

We hear recordings with the same ears we use for live sound. We reach conclusions about the evidence provided by recordings in the same way that we interrogate and evaluate live sound. We constitute apparent sound events just as we directly perceive live sound events. Yet recordings systematically fail to justify our confidence in them. Most listeners have learned to concentrate on the aspects of sound events that are most faithfully rendered by recordings and to pay little attention to the aspects introduced or transformed by the recording process. A proper theory of sound will accept no such selective deafness. It will pay special attention to those very points where confusion is possible, recognizing in such moments of imprecision, indecision, or incoherence the very place where sound seizes the opportunity to take an active role in the definition and exploitation of culture.

Sound Terminology: Talking about Hearing the Record of the Story of a Sound Event

Often called "distortions," on the theory that sound recording is a science of reproduction rather than an art of representation, the variables introduced by sound's material heterogeneity, along with the system constituted to record (that is, represent) it, lie at the very heart of film sound. Though they may constitute distortions for the sound engineer, the marks of the sound narrative and the recording process that appear as part of any sound record constitute the very text of the sound analyst, the fundamental signs of the sound semiotician, the basic facts of the sound historian.

Central to the interpretation of film sound is the fact that multiple moments and operations must be carried simultaneously by the same final sound track. The characteristics of sound production, sound recording, sound reproduction, and audience perception are all superimposed in a single experience. When we hear any particular film sound, how do we know to whom to attribute it? Which part of the sound chain has produced, selected, highlighted, or masked it? Does a decrease in the ratio of direct to indirect sound mean that the character has turned away, an obstacle has been introduced, the microphone has been moved, the sound engineer has fiddled with the dials, or the spectator has shifted her position?

To study film sound is to take seriously the multiplicity of possible determinants of any given audience perception. As a complex representation of a complex sound event, cinema sound offers sound designers infinite possibilities for creation and obfuscation. As such, it also offers theoreticians and critics of cinema sound fascinating opportunities to recognize and analyze the techniques, conventions, codes, and ideological investments of the sound chain. This work is only beginning. It will move more quickly if we adopt a vocabulary that reflects the material, heterogeneous nature of sound presented here.