

## Introduction: Four and a Half Film Fallacies

*Rick Altman*

Logically, every theory of cinema should address the problem of film sound. Practically speaking, such has hardly been the case. On the contrary, a surprising number of theoreticians blithely draw conclusions about the nature of cinema simply by extrapolating from the apparent properties of the moving image. If this were just a question of oversight, the problem would be rapidly corrected. In fact, the theoreticians who overlook sound usually do so quite self-consciously, proposing what they consider strong arguments in favor of an image-based notion of cinema. Indeed, some of these arguments have reached the level of truisms, uninterrogated assumptions on which the entire field is based. In the pages that follow, I propose to reopen the cases of these arguments, cross-examining the very assumptions that have guided cinema theory over the years.

### **The Historical Fallacy**

The late twenties' worldwide conversion to synchronized sound was received by many film-makers as an affront (Clair, Eisenstein, and Pudovkin, among others). Intent on exacting satisfaction, they found a clever method of disenfranchising the offending sound track. Cinema was cinema before the sound track was added, they said, so sound cannot be a fundamental component of the cinematic experience. Historically, sound is an add-on, an afterthought, and thus of secondary importance.

Ironically, it is precisely because of insufficient historical knowledge and reflection that these avengers err. As a purely historical argument, the notion of sound-as-afterthought cannot stand careful scrutiny. Apparently convinced that "silent" film had always conformed to the mid-twenties model of standardized organ or orchestral accompaniment, sound's critics

set up an all-or-nothing opposition that has been perpetuated by generations of critics. On one side an ethereal cinema of silence, punctuated only by carefully chosen music; on the other side, the talkies, with their incessant, anti-poetic dialogue. Too heavily dependent on the practice of the twenties, this is an unacceptable assessment of the first thirty years of cinema history.

Here, for example, is a pop quiz that is not likely to be passed by sound's detractors. In what year did the following editorial appear? "In our opinion the singing and talking moving picture is bound sooner or later to become a permanent feature of the moving picture theater." 1926? 1927? 1928? Wrong by a wide margin. This 1910 *Moving Picture World* editorial came at the height of sound film's expansion. Cameraphone, Chronophone, Cinephone and dozens of other competing systems were not only invented in this period; during the end of the century's first decade they were installed in hundreds of theaters across Europe and from coast to coast in the United States. Their competition came, by the way, not from silent films, with or without musical accompaniment, but from road shows with extremely sophisticated and carefully synchronized effects (a technique originated by Lyman Howe), and from the many "wheels" (vaudevillelike circuits) of human-voice-behind-the-screen companies, with colorful names like Humanovo, Actologue, Humanophone, Humanoscope, Natural Voice Talking Pictures, Ta-Mo-Pic, and Dram-o-tone. In short, the world did not wait until *Don Juan* and *The Jazz Singer* to discover the entertainment (and financial) value of synchronized sound. From a purely historical point of view, the notion that sound is a Johnny-come-lately add-on to a thirty-year-old silent medium simply will not stand.

Even if the historical information had been correct, however, the claims of sound's early critics would still have been fallacious. Though their appeal is apparently to history, these unconditional lovers of "silent" cinema actually close themselves off from history, refusing to recognize that the identity and form of the media are in no sense fixed. Why do we identify the human appendix as vestigial? Because we recognize that it is possible for evolution to redefine the structure and even the nature of the human body. How can we tell when one system has given way to another? This we can do only by analyzing the functioning of the system. The fact that one element appeared before or after another carries no weight in this evaluation. At stake here is the very ability to take into account historical change in theoretical arguments. It is regularly assumed that a single term (like cinema) covers a single object. If our theories are to become sufficiently sensitive to historical concerns, we must abandon that assumption, recognizing instead that historical development regularly occurs within an apparently single object, thus often hiding under a single name

two or more historically distinct objects. In other words, even if silent film were the object that sound's detractors claim, the sound-as-afterthought argument would still not hold up. Cinema changes, and the action of sound is one of the prime reasons for that change.

### The Ontological Fallacy

Though they continue to influence cinema studies, historical fallacy arguments were especially popular among the filmmakers of the late twenties and early thirties. Later in the thirties, a new argument appeared in the writings of such influential critics as Rudolf Arnheim and Bela Balazs. Eschewing historical arguments, they make the formal case that the image without sound still constitutes cinema, while sound without an image is no longer cinema. Clearly assuming that cinema is a firm, unchanging category, immune to history, these critics present their arguments as logical and permanent. Indeed, so strong is the apparent appeal of this ontological claim that it regularly reappears in the writings of current theoreticians.

Two primary considerations undermine the ontological argument. The first is a practical concern relating to the way in which ontological critics use their claims. Like the historical case, the ontological argument seeks to disenfranchise sound, to prove that sound has (or should have) little effect on overall film structure. Even if we were to accept the notion that film is a fundamentally image-oriented medium, this conclusion begs the larger question of the relationship between ontology and structure. May we affirm with confidence that an object's structure can be predicted from its nature? The answer to this question depends on the way in which we construe the term "nature." If nature is defined through structure, that is, if all claims about the nature of a class of objects are derived from analysis of the structures characteristic of the objects, then we can treat nature as predictive of structure. But this is precisely what sound's ontological critics do not do. On the contrary, they base their claims about cinema on a single surface aspect rather than on a careful inspection of the structure of actual films and the system that produces them. Indeed, the acerbic vocabulary and prescriptive exhortations of these critics suggests that they are more interested in influencing the structure of future films than they are in analyzing the structure of existing ones.

What about the truth value of ontological critics' claims? The problem with these apparently rock solid claims, I would suggest, is that they are actually built on sand. Presented as absolute and unchanging, appeals to the nature of cinema appear to be independent of history. In spite of appearances, however, the evidence actually offered is all historically specific. To say that a particular configuration would not be recognized

as cinema, while another would, is to affirm that *in the present conditions* these conclusions would be reached. Present conditions, though, have to do with the way in which a given configuration has been used, and not with some transhistoric category. The ontological argument, it turns out, is only falsely ontological.

Even in the absence of a properly ontological argument, however, the historical claim would remain: cinema has indeed been exploited as a visual medium, to the point where audiences identify the medium with the image rather than the sound. To the extent that it represents a carefully documented historical argument, this position has a certain amount of merit (however unontological it may be). Even when the ontological argument is reduced to its historical evidence, however, hesitations must still remain. For the historical circumstances assumed by ontological critics have not always obtained. During the many periods when cinema was heavily marked by its relation to the music industry, for example, music accompanied by a blank screen has regularly been recognized as cinema: the long overtures to the early Vitaphone sound-on-disk features, the introduction of a film's theme song before the images or its continuation after the post-credits (as in *Nashville*), and the use of a totally black screen in recent music videos. These examples hardly prove that cinema is regularly taken as a sound-based medium, but they do suggest the *historical* possibilities of cinema as an audio-visual medium, in which sound-oriented proclivities regularly confront image-based tendencies, thus producing a varied history belying claims of a solely image-oriented ontology.

Since first pointing out the ontological fallacy in *Cinema/Sound*, I have become aware of an even more problematic appeal to ontology in the study of sound. Surprisingly, this dependency on ontological arguments comes not from the enemies of sound, but from its greatest defenders. In their celebrated book on *Composing for the Films*, Theodor Adorno and Hanns Eisler attribute to hearing a privileged relation to pre-individualistic collective times; music thus has a pre-capitalistic nature, being more direct and more closely connected to the unconscious. While their other arguments are by and large well attuned to historical differences, this approach to hearing and sound edges dangerously close to an ontological claim, apparently capable of predicting sound's nature in any given situation, but actually able only to locate sound's action in certain past situations.

A similar danger lurks in the work of Mary Ann Doane, Kaja Silverman, Michel Chion, Claudia Gorbman, and other critics who have leaned heavily on the psychoanalytic theories of Guy Rosolato and Didier Anzieu, who characterize the voice as archaic, based on the notion that we hear the soothing voice of the mother from the womb long before we are able to see. It is not surprising that such a transhistoric proposal, apparently

predictive of sound's role in any situation whatsoever, should lead to such conclusions as Doane's claim that "the aural illusion of position constructed by the very approximation of sound perspective and by techniques which spatialize the voice and endow it with 'presence' *guarantees* the singularity and stability of a point of audition, *thus* holding at bay the potential trauma of dispersal, dismemberment, difference" (Doane 1980b, 171; my emphasis). Yet we know from actual listening that very few films construct an approximation of sound perspective. Can it then be said that "the subordination of the voice to the screen as the site of the spectacle's unfolding makes vision and hearing work together in manufacturing the 'hallucination' of a fully sensory world" (Doane 1980b, 171) The problem here is that an apparently ontological claim about the role of sound has been allowed to take precedence over actual analysis of sound's functioning. (In my article "Sound Space," later in this volume, I suggest a different approach to the same question, based not on an assumption of unity and concordance, but on a perceived conflict between sound scale and image scale.) While it would be unreasonable to cut short speculation on the sources of sound's attraction, it is essential that such speculation not be taken as a prescription, as a binding assumption about the way sound must work in all cases. If we are fully to restore a sense of sound's role in creating our sense of the body, we must depend on historically grounded claims and on close analyses of particular films rather than on ontological speculations that presume to cover all possible practices.

### The Reproductive Fallacy

In spite of the fact that, as a storage medium, sound recording lags behind the image by tens of thousands of years, recorded sound has from its very beginnings held a great fascination for critics. Whereas the image, however carefully rendered, clearly reduces a three-dimensional original to two dimensions, sound appears to reproduce the original faithfully, in its full three-dimensionality. By no means limited to early admirers of the newfangled technology, this position was until recently held by a majority of sound critics (many of whose pronouncements are quoted at the beginning of Jim Lastra's article, below). By and large, critics remain convinced that sound is literally *reproduced* by a high quality recording and playback system, in spite of Alan Williams' demonstration of the contrary in his *Cinema/Sound* article ("Is Sound Recording Like a Language?").

Sound, it is worth recalling, cannot be construed independently of the volume of air (or other medium) in which it is heard. Typically, we notate sound (through writing or musical notation) as if sounds were ideal entities. But volume, frequency, and timbre cannot exist independently of several material factors which preclude reproduction as such. To be sure, in some

sense a G# is a G#, whether it is played at home or on stage, but that does not make the two sounds identical. By restricting our description of sounds to familiar musical terminology, we have bamboozled ourselves into believing that sound itself is restricted to those characteristics. Does the G# have a slow attack? a long decay? an echo? reverberation? Does it bounce around like a superball in a hollow cavity? Or does it rapidly lose its force, like a beanbag hitting a pillow? If all we want to know about a sound is that it was a G#, then all G#s are the same, but if we care about the material differences between two sounds, and the spatial configurations that cause them, then we must recognize that no recording can possibly reproduce an original sound.

Recordings do not reproduce sound, they represent sound. According to the choice of recording location, microphone type, recording system, postproduction manipulation, storage medium, playback arrangement, and playback locations, each recording proposes an interpretation of the original sound. To be sure, one of the common strategies involved in this process is an attempt to convince the audience that they are listening not to a representation but to a reproduction. We must not, however, be taken in by advertisements for "high fidelity" sound. The notion of "fidelity" is not a measure of success in reproduction, but a way of assessing a recording's adherence to a set of evolving conventions, like the parallel standards established for such culturally important qualities as "realism," "morality," or "beauty." The concept of fidelity is thus a strange hybrid of engineers' aspirations and ideology, serving to mask recording's representational nature.

Considered as a reproduction, recording seems to fall under the aegis of technology and engineering. Construed as a representation, however, sound inherits the double mantle of art. Simultaneously capable of misrepresentation and of artistically using all the possibilities of representation, sound thus recovers some of the fascination lost to its reputation as handmaiden of the image. Indeed, it is recording's very ability to manipulate sound that makes it so amply worthy of our interest.

### **The Nominalist Fallacy**

In order to show that recording cannot possibly reproduce the original sound, critics (Williams, Levin, Altman) have regularly made the following points: (1) sound exists as pressure within a volume; (2) it is impossible to collect all the sound of a particular performance, since it disperses differently into the various parts of the theater or other surrounding space; (3) even at a live performance, different spectators hear different sounds, depending on where they are seated and which way they and the performers are turned; (4) sound systems always enforce a particular set of values in

selecting microphone type and location, frequency response, volume levels, and many other recording and playback characteristics; (5) playback involves the same set of differences and choices involved in recording.

Within this apparently coherent argument lurks a potential danger. Stressing the material nature of sound in order to counter fundamentally idealist assumptions, this approach fragments sound to the point where the emission of a single sound apparently gives rise to the perception of a multiplicity of different sounds. By concentrating on the differences between the sound as heard in the orchestra and the "same" sound as heard from the balcony, this argument has rendered the important service of sensitizing critics to the materiality, complexity, and context-based nature of sound. At the same time, however, these defenses against the reproductive fallacy have failed to address the problem of the communicative language used by auditors having heard the "same" sound to overcome the fact that they actually perceived physically different sounds.

This is an old problem, closely identified with the weaning of philosophy from theology in the latter part of the Middle Ages. When I pick up two different rocks and call them "rocks," what is the status of the name that I attribute to them? Is the name itself real? Or is the name just a convenient label? To put it another way, is the shared category to be understood as actually existing, or are the objects themselves the only things that exist? Is there such a thing as the category "rock," or are there only objects, on which for the sake of convenience we confer names (such as "rock") which have no existence independent of the objects they represent? The traditional position, usually identified with Plato and Augustine, is termed "realism," because it takes the general category as real; the radical position, championed by William of Ockham and generally thought to have been instrumental in paving the way for Renaissance individualism, is known as "nominalism," because it considers that the general category is just a convenient name. Especially concerned to recognize individual difference (and thus the value of the created world), the nominalists accused the realists of subordinating the entirety of creation to a set of preexisting universals.

This is precisely where we stand today with regard to sound. As Jim Lastra demonstrates so well in his article, below, the critics of the reproductive fallacy have edged dangerously close to an ultra-nominalism in which differing auditor perceptions make a single original sound appear like so many different rocks with no common identity save their common name. The very names used to identify sounds are suspect, disrespectful of sound's material heterogeneity. Yet we do discuss the film as we file out of the theater. In spite of the fact that we have literally, really heard different sounds, we still manage to find a common ground on which to base our conversation.

At this point in time, the study of sound shares the position of reception studies. I once witnessed an interchange that says a great deal about the project of reception studies. After demonstrating that neither the author nor the text can possibly determine readings, that each reader may read the text in differing ways, Tony Bennett opened the floor to questions. Said Paul Hernadi: "This is all well and good, but if what you say is true, how did I understand what you just said?" Taking an ultra-nominalist stance (which, by the way, he has toned down since), Bennett laid such heavy emphasis on our freedom of interpretation from textual constraints that he jeopardized the very notion of understanding. Even today, reception studies need to concentrate more fully on the bridges, the terms, the categories, the reading formations that permit a Paul Hernadi to understand a Tony Bennett.

A similar situation holds in sound studies. While not abandoning for a moment the notion that every auditor of the "same" performance actually hears different sounds, we need actively to interrogate the cultural phenomena that permit us to compose sentences, frame ideas, and ultimately communicate about the sounds which are heard. A decade or two ago, it would no doubt have been politically essential to defend at all costs the free play of the signifier; today it seems far more important to remember with Saussure that signification can occur only through the repression of the signifier and to call for increased sensitivity to the many strategies adopted by various cultures to assure the repression of sound's differences in favor of language's communicative value.

#### **Indexicality: Half a Fallacy Working on the Other Half**

Inherited from photography, one of the most deeply ingrained notions about cinema is that it depends primarily on recording. Unlike painting or writing, it is commonly supposed, cinema uses motion picture photography and sound recording to fix and retain in memory a physical image of the pro-filmic scene. Whereas representational painting is based largely on iconic resemblances, and writing is built around symbolic relationships (according to the terminology of Charles Sanders Peirce), cinema is thought to depend especially strongly on indexical connections, that is, those revealing a particularly close existential relationship between the represented item and its representation (such as that which exists when light rays bouncing off an object expose motion picture film, or when sound waves either drive the stylus of a disk recorder or, once transformed into light, expose the sound track portion of the film). This close connection of course creates an iconic relationship between the pro-filmic object or sound and its filmic representation (that is, the object and its representa-

tion have the same shape), but critics and theoreticians have consistently stressed the indexical ties over the iconic resemblance. In particular, André Bazin's realist criticism has been especially influential in popularizing an indexical approach to cinema. For Bazin, cinema is like a mold that takes each scene's impression; cinema thus works like a death mask or the Shroud of Turin, recording as it were by sacred contact.

During the early history of photography, photographs could be produced only through indexical relationships. As photography matured, however, photographers discovered various methods of "correcting" nature, typically adding painted-in iconic details to the photograph's indexical base. From the very start, however, such techniques were roundly condemned. Offering ontological arguments about photography's nature, critics insisted that certain types of retouching constituted a highly undesirable form of "cheating." The moralistic tone of this argument carried over intact to motion picture. In spite of early cinema's non-indexical display of color, "silent" cinema's iconic and symbolic approach to sound, or the late twenties' creation of the falsely indexical playback system, critics and theorists continued to stress the ability of the motion picture camera, like that of its still picture cousin, to take snapshots of reality. As cinema developed, to be sure, the scenes recorded by the camera began to depend on an increasing amount of manipulation (through set design, costuming, makeup, and so forth), and the final film image was increasingly constructed from a combination of separate images (through mattes, background projections, and other processes). For all this prestidigitation, however, the basic assumption was never jeopardized: cinema is primarily an indexical medium, directly dependent on the photographic recording of each pro-filmic scene (or scene fragment).

Throughout the history of cinema, the image has by and large corresponded to this indexical, recording-oriented definition (with the exception of special cases like scratch-animation, computer graphics, or electronically generated images). Cinema sound, on the other hand, has rarely been the result of straightforward indexical recording. Long before the cinema industry converted to sound, the market for telephonic communication, phonograph records, public address systems, and radio entertainment had led engineers to investigate the possibilities of "enhancing" sound in order to achieve greater volume, presence, or intelligibility, while reducing unwanted characteristics. At first, engineers concentrated on the recording process itself, laboring to increase the indexical fidelity of the sound recording apparatus. Soon, however, sound's capacity for post-recording transformation became apparent. Why *record* the reverberation associated with sounds produced in a large hard-walled room if you can simply process the reverb in later? Throughout the thirties, parallel

developments in electronics and film sound led to the creation of myriad devices designed to produce final-release sound differing radically from the sound originally recorded on the set.

Little by little, the indexical nature of film sound became compromised by the ability of acoustic networks and electronic circuits to alter or simulate sound. Once the sole province of high-end sound production facilities (such as those found in New York, Hollywood, and a few other important centers around the world), the electronic revolution has now made it possible to produce all the music and effects for a film sound track without recording a single cricket or musical instrument. For a decade, film sound has been heavily influenced by digital systems like MIDI and the Synclavier. Even the most inexpensive films feature sound tracks that are no longer primarily recorded. In cinema, television, and disk production, sound has definitely surpassed the era of indexicality.

Today, the customary electronic manipulation and construction of sound has begun to serve as a model for the image. Though the film image currently depends primarily on a chemical (and thus indexical) technology, the electronic nature of the television image provides a different model, whose influence is increasingly felt in the cinema world. In order to create Mickey Mouse, Walt Disney had to do more than just make thousands of drawings; he had to record them with the same (indexical) motion picture process used for live actors. Today's animators work in an entirely different fashion. While most still depend in part on drawing, they make heavy use of electronics and its ability to produce iconic relationships without depending on the indexicality of recording. A similar shift has taken place in color technology. Whether two- or three-strip, whether additive or subtractive, traditional color processes systematically depended on indexical relationships: the color was fixed by the existential contact of the object with the film. Today, films are colorized by an electronic process that owes nothing to recording. It is only a matter of time before the plunging prices of all electronic processes turn colorization from a postproduction technique into a production device.

In short, the recording medium that cinema once was has now been massively transformed and risks ultimate obsolescence. However accurate it may once have been to understand cinema as recording its object by sacred contact, we must recognize that three-quarters of a century of electronics has radically desacralized cinema, substituting circuitry for direct contact, constructed iconicity for recorded indexicality, and the infinite imagined possibilities of the keyboard for the restricted immediacy of recording reality. Not so very long ago, treating cinema as *écriture* was a radical move; the technology itself is now turning this metaphor into a reality. Once, cinema was recorded with a camera; now, it is increasingly written with a keyboard.

So far, it is only half a fallacy to treat cinema as a recording medium. By the end of the century, however, cinema will be well on its way toward full digitalization. The end of the indexical era looms large. Perhaps it is time to revise our theories and our vocabulary to take this transformation into account. More experienced in this domain than the image, sound must lead the way.

## 2 Sound Space

*Rick Altman*

“The real can never be represented; representation alone can be represented. For in order to be represented, the real must be known, and knowledge is always already a form of representation.” From this claim, which I made in an earlier article on the role of technology in the history of representation (1984), we can deduce several essential principles for the writing of cinema history. In particular, we readily conclude that the “reality” which each new technology sets out to represent is in large part defined by preexistent representational systems. In order for its new mode of representing to achieve acceptance, photography had to conform not to reality as such, but to the visual version of this reality imposed by a certain style of painting and engraving. In the same manner, the early years of sound cinema were marked by a heavy debt to contemporary arbiters of sound representation: radio, theater, phonography, and public address.

Such a theory is hardly devoid of problems, however. While it helps us to understand how a nascent technology leans on preexisting forms, it remains all too static, offering little insight into the processes whereby a new form of representation is liberated from its models, eventually offering to subsequent technologies its own representational norm. A constant source of debate during the Hollywood thirties, the problem of sound space provides a particularly clear test case, a unique occasion where a change in representational norms is carefully discussed, documented, and even quantified by contemporary technicians.

The single most important question occupying Hollywood sound technicians during the late twenties and early thirties was this: what relationship should obtain between image scale and sound scale? Disarmingly simple, this question in fact implies a complex series of related problems. What type of microphone should be used? Where should it be placed? May it

be moved during a take? Is it appropriate to make multiple image and/or sound takes simultaneously? What sound take should be paired with what image? What volume level should be used? Is it appropriate to mix multiple takes? Under what circumstances must reverberation be added? And many others. Indeed, as the most perceptive of sound technicians recognized from the start, the question of sound scale foregrounds to an unexpected extent problems of audience identification, of spectator pleasure, and of subject placement.

Concentrating on the broad range of problems implied by the question of sound scale, this essay will be divided into three parts. First, I will trace through its various stages the industry’s desire for a match between image scale and sound scale. Second, I will attempt to reach some general conclusions about the standard sound practices which evolved in Hollywood during the thirties. Third, I will consider the ramifications of the representational system thus constituted for the placement of the hearing and seeing subject.

### **The Dream: Correlating Sound and Image**

The question of sound scale, as contemplated by Hollywood technicians of the early sound years, must be seen as part of a longer chain of attempts to assure sound localization. Roughly stated, the three main approaches involve:

1. manipulation at the place of exhibition, largely through speaker placement and switching mechanisms (1927–31);
2. manipulation during production, especially of microphone choice and placement, along with control of sound levels during editing (1929–present);
3. development of multi-channel technology, eventually including stereophonic localization capability (1930–present).

While the latter approach falls largely outside the scope of this study, some information on the first stage will provide appropriate background for the second stage, on which the remainder of this article will concentrate.

That early attempts to localize cinema sound should have concentrated on the movie house itself is hardly surprising. After all, for a quarter of a century movie theater owners had been designing the sound accompaniment to silent pictures. Even before *The Jazz Singer*, Lee DeForest was insisting that loudspeakers be located in the orchestra pit “to simulate a fifty-piece orchestra” (DeForest, 72). As soon as the spoken word became a staple of sound cinema, a second horn was added to the standard orchestra pit speaker, this one above the screen and facing out, whereas



the orchestra horn typically aimed straight up (Rainey, Wilcox, Peck, Hopkins 1930a and 1930b). The presence of two speakers, each destined to reproduce a different type of sound, of course required a switching mechanism. The projectionist, who needed to know the sound track like an orchestra conductor, would simply switch the sound output to the appropriate speaker each time there was a change from music to dialogue or back. This of course assumes sound tracks where music and dialogue are not mixed, a requirement assured by the difficulty of mixing separate tracks before late 1930, by which time the separate speaker arrangement had largely disappeared.

But if sound could be localized either on screen or in the orchestra pit, then it could also, at least theoretically, be reproduced at different spots on the screen, depending on where the person speaking is portrayed. Never to my knowledge realized, this project, reported by J. C. Kroesen in July, 1928, demonstrates the extent to which early technicians assumed the necessity of tying sound to the image. "The screen," said Kroesen, "should be divided and so arranged that sound will be reproduced only at or as near the point of action as possible" (Kroesen, 8). Pity the poor projectionist trying desperately to operate the switching mechanism!

Nor was lateral placement the only possibility afforded by careful speaker location. In an October, 1930, discussion, the Society of Motion Picture Engineers reviewed the possibilities afforded by the recently introduced multiple sound track technology. Realistic offstage effects can be produced, a Mr. Ross pointed out, "by employing separate sound film having a plurality of sound tracks, each related to a group of loudspeakers located at points from which the sounds are to be produced. . . ." In this manner, Ross suggested, depth localization as well as lateral placement may be assured. "Any sound," he said, "that one might wish to produce from points other than the immediate foreground depicted on the screen may be handled in this manner. Loudspeakers may be placed at remote portions of the stage or auditorium. There are decided advantages for this arrangement which are quite evident to anyone who has tried it."<sup>1</sup> Ross's protestations notwithstanding, Hollywood was not yet ready for the multiple-channel solution.

Exhibitors were quite ready, however, to dispense with the need for manual switching. Neither better instructions nor the short-lived introduction of automatic switching through control tracks could stem the tide. By 1931, the dream of sound localization through speaker placement was a dead letter. In his description of the Western Electric Reproducing System, Bell engineer S. K. Wolf clearly opposed the old-fashioned system ("when it was desired that reproduced music should simulate accompaniment by a theatre orchestra") to what he called "modern practice," which recognized the fact that "in most instances the sound is desired to come from



positions on important business," while the practice of mixing multiple mikes made the sound "run throughout as though heard from the indefinite position described above" (325). Cass concluded: "Since it is customary among humans to attempt to maintain constant the distance between the eye and the ear, these organs should move together from one point to another in order to maintain our much mentioned illusion [of reality]" (325). Demonstrating nothing short of missionary zeal in his attempts to save the spectator from monstrosity, Cass failed to consider the possibility, to which we shall return later, that spectators do not remain from age to age the same, that even the body and its functions are culturally determined, and that spectators who live long enough with monstrosity learn to consider it not only beautiful, but even, eventually, normal.

Before the normality of a many-eared spectator could be contemplated, however, numerous other attempts at codifying methods for assuring an appropriate image/sound match would appear. The key figure here was again Maxfield, the most powerful voice within the Bell/Western Electric/ERPI complex that dominated the Hollywood sound scene until the early thirties. In a series of articles which reiterated numerous times the same points, Maxfield hammered home the need to limit most takes to a single microphone (Maxfield 1929, 1930a, 1930b, 1931). Thus solving the problem posed by Cass, Maxfield showed how the use of a single mike, placed near the camera's line of sight, automatically coordinates sight and sound by providing a sound record of the characters' movements toward and away from the camera. Characters approaching the camera automatically approach the microphone as well, thus matching closer image scale to closer sound scale; conversely, the character who speaks from the background demonstrates distant characteristics in sound and image alike.

While the use of a single, stable microphone assures the matching of sound scale with image scale within a shot, a different set of guidelines was necessary in order to assure proper matching of succeeding shots. These guidelines, first proposed by Maxfield in 1931 (in the graphic, empirical style characteristic of Bell's scientific pretensions in the early thirties), and reiterated in 1938, indicate proper microphone placement in varying image situations (Maxfield 1931, Maxfield 1938). Throughout the decade, Maxfield had been a pioneer in isolating the aspects of sound which determine the spectator's perception and evaluation of specific sound phenomena. Demonstrating that volume alone is an insufficient marker of distance, Maxfield had earlier revealed the importance of reverberation (or more accurately, the ratio of reflected sound to direct sound) for determining perception of sound scale. In addition, Maxfield showed that the focusing capabilities of the listening binaural human subject, permitting humans selectively to cut out a certain percentage of reflected sound, have no parallel in the monaural sound collection system of cinema.

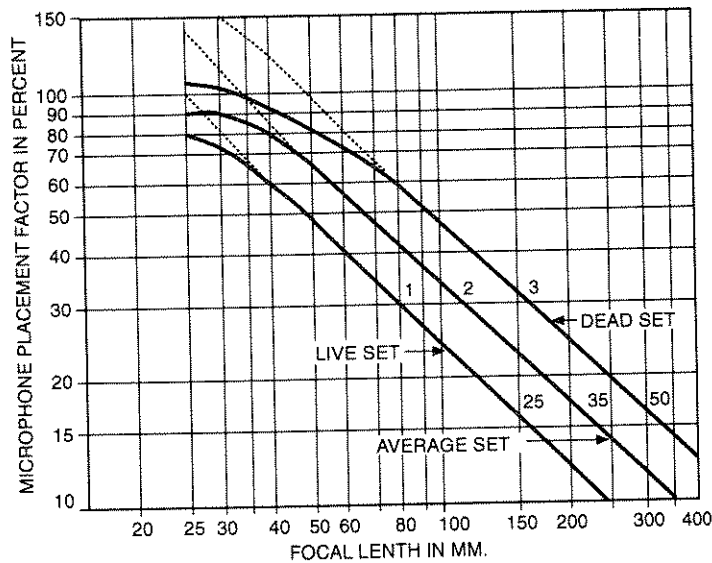
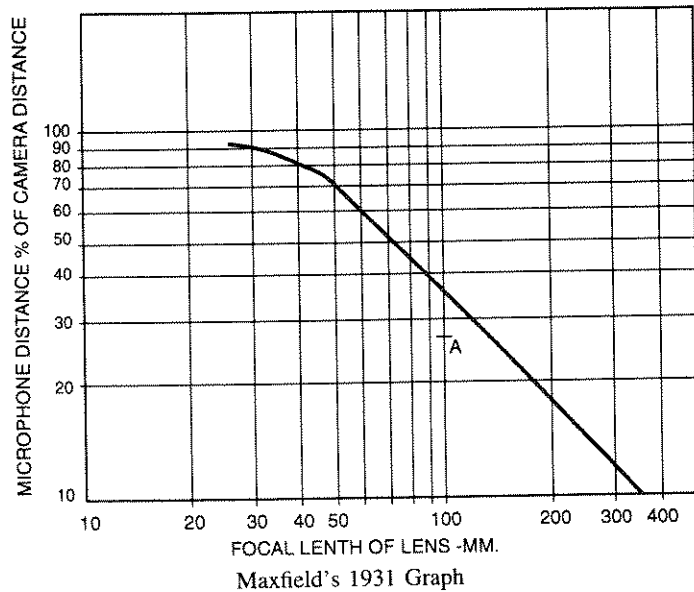
A monaural correction factor is thus built into all Maxfield's careful determinations.<sup>3</sup> Taken as a whole, Maxfield's enormously influential writings provided a complete program for matching sound scale to image scale, within and between shots.

### The Reality: Mismatching Image Scale and Sound Scale

With Maxfield's influential articles reprinted in journals of all sorts throughout the thirties, referred to by one Bell author after another, and imitated throughout the industry, one might well assume that Hollywood practice followed Maxfield's strictures to the letter. A careful survey of the period's sound practices is far from bearing out such an assumption. Unfortunately, space does not permit full treatment of this important question. A few examples, along with a rough sketch of the period's general penchants, must suffice. As in every period, examples abound of atypical practices, but here I will stress instead what I take to be the accepted norm from the late twenties through the Second World War. An appropriate starting point might be Rouben Mamoulian's *Applause*, one of the few films of 1929 to be universally praised for its revolutionary approach to the sound track. My purpose here, however, is not to dwell on Mamoulian's many innovations, like the subjective use of sound levels in the opening parade scene, but to show that even *Applause* neglects the careful matching of sound scale to image scale.

The first stage scene in *Applause*, for example, exhibits a sound track of uniform volume and reverberation characteristics. The sound track's uniformity is hardly matched by the image track, however, which reveals a heavily edited theater scene, combining shots of varying scales and angles. The "Doctor in the house" routine which calls the scene to a close, for example, is clearly shot with a single microphone, while two cameras are churning out images of different scales. Once edited together, the two simultaneous camera takes produce a scene typical of the period. Perhaps it is fitting to remark here that the term editing, entirely appropriate for the images, is less so for the sound, since the sound take used is apparently continuous and uncut. In fact, it would be perfectly correct to say that the contemporary practice of using a single microphone system synchronized to two or three cameras fairly begged early editors to use a continuous sound track as the bench mark to which they edited the various images.<sup>4</sup>

For obvious economic and technical reasons, multiple-camera shooting remained the rule throughout the early sound period. Soon, however, an important change took place in the type of sound record associated with the multiple-camera arrangement. The condenser mikes used in the late twenties required extremely close placement in order to provide a distinct dialogue record (Hunt, 482). Often, the mike had to be so close to the



Relation between focal length of camera and microphone position, for achieving acoustic perspective.

Maxfield's 1938 Graph

speaker that it could not be kept out of the field in a medium shot, thus resulting in the common practice of handling action in medium shot, while flashing into close-up for the sound record (thereby avoiding revelation of the microphone, which would be visible in the medium shot of the same scene). During this period, where a series of dialogue locations were built into a single shot, preference was often given to a multiple-microphone setup, with a mixer choosing the clearest sound record. More intimate scenes easily accommodated the increasingly widespread choice of single-miking.

In the early thirties, however, new microphones became available; lighter, more compact, and requiring no amplification stage near the mike, these new units made the microphone boom far more practical (Altman 1985b, Altman 1986a). Whereas the twenties often used what Dreher dubbed "prop pick-ups" (Dreher 1929b), microphones which had to be hidden in a prop in order to get close enough to the speaker to achieve acceptable sound quality, the thirties adopted the mobile mike, suspended from a boom which could be moved silently about, always pausing at the appropriate point to capture a perfect rendition of lines which otherwise might have turned out garbled or fuzzy. Furthermore, the mobile boom made it relatively easy to stay out of the camera field while remaining at proper distance for sound recording.<sup>5</sup> In short, the combination of multiple-camera shooting and single-miking with a mobile boom made for an ideal combination, for two related reasons. First, the boom simplified the sound problems inherent in the multiple-camera arrangement, thus preserving an important economy factor (Hunt, 481-82). Second, the boom changed radically the character of the sound "in the can." With a single immobile mike, such as that championed at the turn of the decade by Maxfield, the spatial characteristics of the pro-filmic scene were already inscribed on the sound track. A character receding or turning away from the mike was recorded with a higher ratio of reflected to direct sound; similarly, the size of the room had its effect on volume, reverberation, and frequency characteristics. With the new system, however, the microphone is perpetually kept within approximately the same distance of the speaker, thus canceling out nearly all the factors which the earlier system retained.

Coupled with devices for adding reverberation, voice equalization, effort equalization, and so forth, in which the mid-thirties abound, this new approach assured Hollywood both the economic benefits and the requisite control associated with a system permitting the *construction* of a sound track rather than the direct *recording* of already constructed sounds. Parallel to the many image-treatment processes which permitted the Hollywood of the thirties to exercise control over the image while reducing the cost of its production, sound construction processes serve to enhance the ability of the boomed mike to provide a clean, clear, continu-

ous sound record, oblivious to image scale but attuned to dialogue intelligibility, story continuity, and freedom of action.<sup>6</sup>

Perhaps most telling of all is the 1938 article in which Maxfield reiterated his strictures regarding microphone placement. Still insisting on a careful matching of image and sound scale, Maxfield again explained his chart providing proper microphone placement, but this time his instructions were interspersed with remarks reflecting years of experience watching technicians use his chart. These remarks reveal a fascinating tendency:

It has been the author's experience, and that of some of the microphone men with whom they have discussed the problem, that unless some such guide is used there is a tendency to set the close-up takes correctly and to make the microphone positions for the long-shot and semi-long-shot takes decidedly too close. The use of the curve, of course, helps to keep the judgment of the operator calibrated. (Maxfield 1938, 672)

Now, Maxfield's original microphone distance chart was based on so-called empirical data, generated by the experience of the first three years of sound film. In 1931, Maxfield explained that his data came from his own records of "several pictures with which the writer was associated" (Maxfield 1931, 74). In the seven or eight years since Maxfield's data were first collected, however, something had evidently changed. Whereas the 1931 chart was derived from actual experience, producing the straight-line function presented in both charts, the 1938 article clearly admitted that the chart must be used to *control* and *rectify* the intuitions of technicians, who without such a guide would always tend to set microphones at a distance producing close-up sound quality.

In other words, the "gut reaction" of sound technicians has changed over the course of the decade. Their intuition in 1938 clearly reflected the changing practice of the thirties; having internalized a new standard, the technicians no longer sought to match sound scale to image scale through "correct" microphone placement, but instead sought to produce a continuous sound track of nearly level volume and unbroken close-up characteristics. Throughout the thirties it was for the clarity of their sound tracks that sound technicians had been praised and rewarded, rather than for their spatial realism.

What once appeared as monstrosity had now become the norm. In fact, Maxfield himself recognized the extent to which careful scale-matching had disappeared. "There are occasions," he admitted, "when it is necessary to use several cameras on the same scene simultaneously. Where acoustic perspective has no dramatic importance, a single close-up track can be used for all the picture takes, the sound being dubbed at slightly lower level for the long-shot scenes" (672). This capitulation by the champion

of scale-matching coincides exactly with numerous contemporary remarks by top sound men. To quote but one, stereo pioneer W. H. Offenhauser stated categorically that "it is our practice . . . to record all our sound with the microphone placed close to the sound source" (Offenhauser, 146).

Why this striking change? Sound was not yet in its teens and already sound technicians had reversed their position about sound space, not only in theory but also in practice. It is no exaggeration to claim that this reversal represented a fundamental turnabout in human perception. We often give lip service to the notion that cinema teaches us to see and to hear, that the media determine our very notion of reality. Yet we are rarely privileged to isolate the moment when and the process whereby our perception changes.

During the early years of sound cinema, theoreticians regularly insisted that sound be treated according to the model provided by the human body. While these appeals to nature carry strong rhetorical value, they frequently disguise other, more important models. According to the theory elaborated by Maxfield and followed by many early sound men, the sound track must carry, independently from the image, all the information necessary to reconstruct the "real" space of the scene (that is, the one represented by the image). In this approach we easily recognize the technique of a representational system with a decade's experience in creating sound space. Maxfield and his colleagues may have stressed nature and the body, but their method owed more to radio technique than to the fixed distance between the eyes and the ears. For where had Hollywood found its sound technicians? By far the majority, like Carl Dreher, had come from the radio studios. The early years of sound cinema were thus heavily marked by the version of reality offered by other modes of representation—first silent cinema, then radio.

"The real can never be represented; representation alone can be represented." Up to now, this theory would appear confirmed: the audio technique of early sound cinema refers to other systems of representation. Not only silent cinema for the location of loud speakers and radio for sound perspective, but other models as well. Due to space limitations I will outline only one of these here: the acoustics of large public spaces (church, palace, concert hall), with their continuous reverberation and long decay time. In spite of the importance of limiting reverberation in order to assure dialogue intelligibility, Hollywood sound technicians regularly insisted on reproducing music with a high degree of reverb, corresponding to the large reverberant halls in which we are accustomed to hearing nineteenth-century "serious" music played.

Silent cinema, radio, the concert hall: can history be built on the simple notion that each new representational system derives its initial task from

a previous system? Certainly not, for while such a theory helps us to understand the early logic of a new representational technology, it fails to explain eventual modifications in the representational use of that technology. In order to understand the alterations in Maxfield's claims from 1931 to 1938, we need to be able to explain how the construction of sound tracks changed over that period.

A few examples should help us to understand these changes. The opening scene from Paramount's 1939 *Union Pacific* provides a particularly representative case, directed as it is by one of the period's most conservative and exemplary directors, Cecil B. DeMille. This debate on the floor of the U.S. Senate exhibits a sound track of uniform volume and reverberation characteristics, with differences in sound level attributable to the senators' rhetoric or intensity rather than to any technical considerations. Yet this uniform sound track is matched to a heavily edited image track, revealing shots of radically differing scales. In particular, one camera movement stands out for its clear demonstration of the lack of concern for sound/image scale-matching: during one speech, the camera tracks constantly back, reducing the scale of the actor, while the senator's speech level remains unchanged, with neither volume nor reverb varying in the slightest. We note, therefore, that the choice of image to accompany any particular sound is in no case dependent on the spatial characteristics of the sound. Instead, the choice depends entirely on the *narrative* characteristics of the sound. The sound track remains uniform throughout, displaying medium close-up characteristics from beginning to end. The image changes scale repeatedly, however, matching the dramatic effect of the words uttered. The constant-level sound track thus serves to anchor a pasted-up, discontinuous image sequence which remains obedient to narrative concerns.

A second example comes from the same year's *Only Angels Have Wings*, made for Columbia by Howard Hawks, another filmmaker known for thematic rather than technical innovations, and who might thus be expected to reflect the industry's standard pre-war practice.<sup>7</sup> At the end of the opening sequence, Jean Arthur is invited by Noah Beery and a fellow pilot to have a drink. Their conversation is interrupted by the arrival of the restaurant owner. During the ensuing scene there is a cut-in to a medium close-up of the owner, without any accompanying change in sound level. In the exterior conversation that follows, we witness another cut which eloquently testifies to the period's presuppositions about the sound/image match. In the *Applause* and *Union Pacific* scenes, image cuts always occur between speeches or during pauses, like punctuation in a paragraph. While no change in sound level or characteristics is noticeable, the practice clearly sets up a rhythmic correspondence between sound and

image, one which might just as well have been used to establish a scale match between sound and image.

In the exterior conversation from *Only Angels Have Wings*, however, the cut is made right smack in the middle of a phrase. Something different is going on here. Far from matching sound scale to image scale—the dream of technicians and theoreticians alike in the early thirties—Hawks uses the uniformity and continuity of the medium-close-up sound track to cover over a cut. Now, this technique obviously assumes a system in which no match between sound scale and image is sought. Whereas *Union Pacific*'s practice of making image cuts in the silences between phrases could have attenuated the effect of a change in sound level, the cut during a speech in *Only Angels Have Wings* would create a naked juxtaposition of the two levels if there were to be a match in scale, thus revealing the processes of image editing and sound mixing alike, thereby foregrounding an apparatus which Hollywood would rather hide. That cutting during dialogue has become routine by the late thirties reveals the extent to which the uniform sound track has become the rule, unmatched to and independent of the image.<sup>8</sup>

A second example from the same film further illustrates this fact. When Noah Beery flies off to his death in the following scene, numerous shots of the plane accompany a homogeneous, uniform-level sound track of the plane's engine noise. With one exception, the sound level remains the same, whether we see the plane in long shot or just the pilot in medium shot. As the camera closes in on the plane, no change in sound ties the sound track to the image scale. Only when an internal auditor is implied does the sound scale match the image, a situation which occurs when Cary Grant and Jean Arthur listen to the plane disappear down the far end of the runway, the fading of the motor sound replicating its growing distance from the listeners. (More on this special situation in my final section.)

That the practices illustrated by *Union Pacific* and *Only Angels Have Wings* represent the industry standard and continue through the forties is clearly demonstrated by the authoritative comments of one of Hollywood's most knowledgeable and influential sound men, John G. Frayne:

To insure high intelligibility in a sound-stage pickup it is customary practice to place the microphone as close to the actor as possible, the distance usually being limited only by the camera angle of the scene. . . . no attempt is ordinarily made in practice to try to obtain the same acoustic as visual perspective of the scene. . . . "Panning" of the microphone by the "boom" man is an accepted technique in production recording and is omitted only if the physical location of the actors makes it impossible for the boom man to keep up with the action.

Panning from one actor to another or following the movements of an actor tends to keep a constant relationship between microphone and speaker with respect both to distance and to orientation of the sound pressure axis. Thus a constant-frequency characteristic is preserved, and a change in tonal inflection of the actor as he moves around in the sound field is avoided. . . . (Frayne, 52–53)

No longer is there any question of matching sound scale to image scale, unless it is to show how any indication of sound scale can be avoided. We have moved a long way from the repeated demands for scale-matching during sound's early years.

### Cinema's Bifurcated Subject: Seeing/Hearing

Why did early technicians' calls for scale-matching fall, as it were, on deaf ears? Or, rather, why is it that early thirties proponents of auditory perspective had by the end of the decade abandoned their dedication to the creation of sound/image proportionality? What factors determined the sound level practices which dominate Hollywood's studio years? Two related considerations come immediately to mind. A third is perhaps less obvious.

Whether expressed in terms of "story continuity" (Mueller), the "business of the play" (Dreher) or the creation of a persuasive illusion (DeForest, Cass, Maxfield, Dreher), the criterion of intelligibility of dialogue retained its primary importance throughout the period under study (Altman 1985b, Altman 1986a). For Harold B. Franklin, speaking in 1930 as head of Fox West Coast Theaters for exhibitors everywhere, sound cinema's greatest advantage was the "ability to present every word so clearly and distinctly that no one need strain to hear what is being said, at least when recording and reproducing is properly conducted. A whisper is clearly audible from the front row in the orchestra to the last row in the balcony" (Franklin, 302). Franklin thus echoed one of the creators of the medium, Lee DeForest, for whom "one of the great advantages of Phonofilm is that, in common with the 'Public Address' system, the voice of the screen image is far more distinct and clear in the far reaches of the house and gallery than would be the normal human voice of a speaker on the stage."<sup>9</sup> For DeForest, cinema was thus an improved megaphone, a mechanical aide for the hoarse actor and the carnival barker.

That the ideal of intelligibility might contradict the desire for a faithful matching of sound scale to image scale occurred to many early theoreticians of sound editing. While Bell's Maxfield tried to dodge the problem, however, by blithely asserting that proper scale-matching produces intelligibility (Maxfield 1931, 71), RCA's Dreher openly faced the problem,

recognizing that there are two potentially contradictory requirements of good recording: "(1) intelligibility of dialog" and "(2) naturalness, or acoustic fidelity to the original rendition," within which category he included the need to retain the spatial characteristics of the original, and thus a sound/image match.<sup>10</sup> Influencing sound technology throughout the thirties (especially the development of the microphone boom, sound collector, and directional mikes), the ideal of intelligibility remained a central factor throughout the period. Indeed, as I have written elsewhere, this insistence on intelligibility at the expense of fidelity to the pro-filmic situation suggests that the referent of Hollywood sound is not the pro-filmic scene at all, but a narrative constructed as it were "behind" that scene, a narrative that authorizes and engenders the scene, and of which the scene itself is only one more signifier (Altman 1985b, Altman 1986a).

A second, related, consideration regards changing standards of reality during the early years of sound cinema. In the late twenties and early thirties, as we have seen, the reality standard constantly held up to the cinema sound track was daily life in the real world. For inventors and engineers like DeForest, Miller, and Dreher, sound cinema would succeed only if it was "natural;" other theoretically inclined technicians, like Physioc, Maxfield, and Cass, stressed instead the natural coordination of the eyes and ears within the overall system of the human body. In calling for a careful matching of sound scale to image scale, early theoreticians clearly assumed that sound cinema needed to match a reality code derived from daily life, where small-scale people—distant individuals—have small-scale voices, and close-up people have close-up voices. Competing with this daily life model, expressed in terms of scale-matching, there ran throughout the decade another model, this one expressed in terms of intelligibility.

We find here once again the familiar opposition of intelligibility to naturalness (or acoustic fidelity), but it was rewritten in terms of differing codes of reality. On the one side, daily life, on the other the medium that taught the audiences of the twenties and thirties to expect visual narrative to provide intelligible dialogue. I speak of course of the theater, that old enemy of "pure cinema," back to haunt the faithful once again. For if sound cinema continued to practice intelligibility in spite of repeated appeals for acoustic fidelity, it was because cinema continued to find in the theater a long-consecrated code of reality applicable to audiovisual narratives. Not even the naturalist theories of an Antoine could radically alter the theater's commitment to understandable dialogue, achieved through such devices as the stage whisper, playing toward the audience, and the declamatory style (which would be replaced only when sound cinema's superior ability to assure intelligibility led theater to borrow the cinema's technological means for amplifying dialogue<sup>11</sup>). To call for



intelligibility in the language of the thirties' cinema technicians is thus to call for adherence to the theater as code of reality. With the theater of the period itself stressing textual comprehension more than ever,<sup>12</sup> it is hardly surprising that Hollywood felt the need to follow suit, abandoning the image/sound match in favor of intelligibility, the everyday life model in favor of a code of reality provided by the theater.

A third consideration involves nothing less than the subject placement implied by the dominant sound model adopted by Hollywood during the thirties. In order to elucidate this process, I must at this point expand on the model described earlier. At one point during the scenes I have described, an impression of auditory perspective is created by a change in volume and reverberation levels. In *Only Angels Have Wings*, when Noah Beery's plane disappears into the night at the end of the runway, the next shot, a two-shot of Cary Grant and Jean Arthur, is matched to the dwindling sound of an airplane motor in the distance. Situations like these, which have mistakenly caused some critics to see auditory perspective as a common aspect of thirties' sound tracks, are indeed common throughout the Hollywood studio years, but they must not be confused with the scale-matching discussed earlier.<sup>13</sup> In both these cases we are dealing with what might be called "point-of-audition" sound, a clumsy term whose only merit is to recall unflinching the "point-of-view" shot.

Frequently used to establish spatial relationships among neighboring spaces which cannot be presented visually in a single master shot, point-of-audition sound is identified by its volume, reverb level, and other characteristics as representing sound as it would be heard from a point within the diegesis, normally by a specific character or characters. In other words, point-of-audition sound always carries signs of its own fictional audition. As such, point-of-audition sound always has the effect of luring the listener into the diegesis not at the point of enunciation of the sound, but at the point of its audition. Point-of-audition sound thus relates us to the narrative not as external auditors, identified with the camera and its position (such as would have been the case with Maxfield's acoustic perspective), nor as participant in the dialogue (the standard situation of the "intelligible" approach), but as internal auditor. When, in 1938, Maxfield alluded to the potential "dramatic importance" of auditory perspective, he was referring to those situations where the auditory perspective involved is that of a character. Whereas in 1931 he was wholly concerned with the perspective of the external auditor, by 1938 Maxfield—and Hollywood as a whole—showed increased interest in the internal auditor.<sup>14</sup>

We are asked not to hear, but to identify with someone who will hear for us. Instead of giving us the freedom to move about the film's space at will, this technique locates us in a very specific place—the body of the character who hears for us. Point-of-audition sound thus constitutes the

perfect interpellation, for it inserts us into the narrative at the very intersection of two spaces which the image alone is incapable of linking, thus giving us the sensation of controlling the relationship between those spaces.

What is it then that is happening during those numerous moments, exemplified by the long legislative scene from *Union Pacific*, where no such identification is called for, where we find a sound track of uniform level with no spatial characteristics? First, something important is clearly *not* happening here: the auditor is at no point made aware of the sound track as sound track by the radical changes in volume which would have to accompany a careful sound/image match. This initial negative consideration has numerous ramifications, not the least of which is the dissimulation of the sound apparatus itself. The construction of a uniform-level sound track, eschewing any attempt at matching sound scale to image scale, thus takes its place alongside the thirties' numerous invisible image-editing devices within the overall strategy of hiding the apparatus itself, thus separating the spectator from the reality of the representational situation, thereby making that spectator more available for reaction to the subject-placement cues provided by the fiction and its vehicle.

Just as the lack of sudden changes in sound level duplicates the self-effacing effect of contemporary image-editing, so the reverberation characteristics of standard practice sound place the auditor in a manner quite similar to familiar spectator-placement methods. According to the familiar subject-placement arguments advanced by Pleynet, Baudry, and Comolli apropos of the perspective image, we spectators are built into the picture as source and consumer. Perspective images are always made for us; they present a sumptuous banquet with an empty chair awaiting the honored spectator-guest. Now, in order to achieve the continuous close-up sound quality characteristic of Hollywood's standard practice, the microphone must be brought quite close to the speaker, cutting out unwanted set noises while—and this is the important concern for the present argument—also radically reducing the level of reverberation.

But what is sound without reverberation? On the one hand, to be sure, it is close-up sound, sound spoken by someone close to me, but it is also sound spoken *toward* me rather than away from me. Sound with low reverb is sound that I am meant to hear, sound that is pronounced *for me*. Like the perspective image, therefore, the continuous-level, low-reverb sound track comforts the audience with the notion that the banquet is indeed meant for them. The choice of reverberless sound thus appears to justify an otherwise suspect urge toward eavesdropping, for it identifies the sound we want to hear as sound that is made for us. While the image is carefully avoiding signs of discursivity in order better to disguise Hollywood's underlying discourse, the sound track overtly adopts the

discursive approach of low-reverb sound in order better to draw us into a fabricated narrative.

Hollywood cinema thus established, in the course of the thirties, a careful balance between a “forbidden” image, which we watch as voyeurs, and “sanctioned” dialogue, which appears to be addressed directly to the audience. In terms of movement, a similar complementarity was achieved. The image displaces us incessantly, offering us diverse angles on objects located at radically different distances. Our voyeurism consists precisely in this mobility. Yet we flit about at our own peril, constantly risking dizziness. Just as we are about to lose our balance, however, the sound track holds out its hand, offering continuity of scale as an effective stabilizer. Indeed, if we take the risk of flying about at all, it is certainly in large part because we know that our bodies are anchored by sound, and by the single, continuous experience that it offers. It is thus the sound track that provides a base for visual identification, that authorizes vision and makes it possible. The identity of Hollywood spectators begins with their ability to be auditors.

While cinema’s perspective image carries a built-in spectator spot, an interpellative position ready-made for the theatrical spectator, the varied-scale editing practices developed during the silent period move the spectator around at a dizzying pace. Far from inheriting a single place, the spectator must fight to integrate the multiple positions allotted by the film into a single unified home. While this wanderlust is partially cured by a learned, and thus historically grounded, ability to insert shots of various scales into a coherent *Gestalt* of filmic space, it is only with the aid of a continuous-level sound track that the spectator finds a comfortable home. By holding the auditor at a fixed and thus stable distance from all sound sources (except those treated, for previously discussed reasons, through a point-of-audition approach), Hollywood uses the sound track to anchor the body to a single continuous experience. Along with the narrow dynamic range allowed for background music, this process serves to constitute more completely the spectator’s unconscious self-identify as *auditor*, thus providing a satisfying and comfortable base from which the eyes can go flitting about, voyeuristically, satisfying our visual desires without compromising our unity and fixity.

#### A New Model of Technological History

To summarize, we may say that the development of a stable Hollywood audio/visual representational system is best understood according to a tripartite historical model:

1. *Multiple identities derived from pre-existing reality codes.* In its early years, sound cinema endeavored to be all things to all people. It

offered the amplification of public address systems, the performers of vaudeville, the diversity of radio programs, the music of “silent” cinema, the acoustics of the concert hall, and the dialogue consciousness of theater. Never imitating all of these reality codes at once, sound cinema nevertheless gained much of its identity from a clear ability to serve purposes and offer experiences defined by pre-existing representational systems. This period witnessed varied attempts, typically modeled on pre-existing representational systems, to represent space through sound.

2. *Jurisdictional struggle.* For a number of years after Hollywood’s conversion to sound, cinema remained the site of an ongoing jurisdictional struggle. To which union would sound projectionists and engineers belong? Where would sound technicians fit in the studio structure? By reference to which model would differences of approach be adjudicated? While many solutions were reached through industry-wide compromise (such as the adoption of the intelligibility-oriented theatrical reality code outlined above), others required open warfare among Hollywood personnel (such as the late twenties/early thirties battle between sound and image technicians<sup>15</sup>) or the deployment of new apparatus (such as the directional microphones developed by RCA in the early thirties<sup>16</sup>). In some arenas, this jurisdictional struggle was over by the end of the twenties (for example, the decision to base music volume on “objective” stable radio or silent film orchestra levels rather than on the volatile and “subjective” levels of parade music or marching bands), while in others it dragged out well into the thirties (for example, the differentiation between the concert hall as the acoustic model for high music reverb levels and the drawing room as the basis for reduced dialogue reverb—even when both types of sound are represented by the image as produced in the same space). During this period, dialogue volume was commonly modeled on theatrical intelligibility, with primary attention devoted not to space but to speech and to the narrative content of which speech is the primary vehicle.
3. *Development of new reality codes based on technological specificity.* Laboring hard to emulate diverse already-existing representational systems, Hollywood directors and sound technicians discovered only very slowly the special capabilities of cinema sound technology. While point-of-audition sound appeared very early in the history of sound cinema, it did not become integrated into a general system of representation until most of Hollywood’s jurisdictional problems had been solved. Indeed, curiously, it never received the kind of careful and extended theoretical discussion devoted during the late twenties and early thirties



to the realism versus intelligibility debate. Nevertheless, the mid-thirties were an important crucible for the development of a new audiovisual identity for cinema subjects, dependent both on homogeneous dialogue levels (authorizing, as we have seen, the classic symbiosis of forbidden image and sanctioned sound, as well as the comfortable fit of varying image scale and stable sound scale), and on point-of-audition sound (enforcing sound-based identification with specific characters). Together, the intelligibility system (borrowed early on from theatrical and telephonic precedents) and the point-of-audition system (elaborated during the thirties out of cinema-specific possibilities) constituted a new mode of cinematic unity and a new subject position for the Hollywood audience.

In order to understand Hollywood's conversion to sound we must grasp the many ways in which Hollywood attempted to model cinema sound on other existing uses of sound. If we want to understand Hollywood's standard representational system, however, we must do more than this. We must reckon, as I have tried to do in this essay, with Hollywood's jurisdictional struggles and with the new sound structures developed by Hollywood during the course of the thirties.

Theorists and historians have always concentrated heavily on image space in their attempts to define Hollywood classical narrative. Perhaps the arguments presented in this article will make it impossible in the future to discuss Hollywood's standard mode of representation without appropriate consideration of sound space.

## 3

## Reading, Writing, and Representing Sound

*James Lastra*

What we hear from the screen is not an image of the sound, but the sound itself which the sound camera has recorded and reproduced again . . . there is no difference in dimension and reality between the original sound and the recorded and reproduced sound.

Béla Balázs  
*Theory of the Film*

. . . in a photograph, the original is as present as it ever was. Sound can be perfectly copied . . . the record reproduces the sound.

Stanley Cavell  
*The World Viewed*

And it is true that in the cinema—as in all talking machines—one does not hear an image of the sounds, but the sounds themselves. Even if the procedure for recording the sounds and playing them back deforms them, they are reproduced and not copied.

J. -L. Baudry  
"The Apparatus"

There is no ontological difference between hearing a violin in a concert hall and hearing it on a sound track in a movie theater.

Gerald Mast  
*Film/Cinema/Movie*

Auditory aspects, provided 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 gunshot heard in a film from a gunshot heard in the street.

Christian Metz  
"Aural Objects"