A Statement (1928)  
S. M. Eisenstein, V. I. Pudomn, and G. V. Alexandrov

The dream of a sound film has come true. With the invention of a practical sound film, the Americans have placed it on the first step of substantial and rapid realization. Germany is working intensively in the same direction. The whole world is talking about the silent thing that has learned to talk.

We who work in the U.S.S.R. are aware that with our technical potential we shall not move ahead to a practical realization of the sound film in the near future. At the same time we consider it opportune to state a number of principal premises of a theoretical nature, for in the accounts of the invention it appears that this advance in films is being employed in an incorrect direction. Meanwhile, a misconception of the potentialities within this new technical discovery may not only hinder the development and perfection of the cinema as an art but also threaten to destroy all its present formal achievements.

At present, the film, working with visual images, has a powerful effect on a person and has rightfully taken one of the first places among the arts.

It is known that the basic (and only) means that has brought the cinema to such a powerfully effective strength is **MONTAGE.** The affirmation of montage, as the chief means of effect, has become the indisputable axiom on which the worldwide culture of the cinema has been built.

The success of Soviet films on the world's screens is due, to a significant degree, to those methods of montage which they first revealed and consolidated.

Therefore, for the further development of the cinema, the important moments will be only those that strengthen and broaden the montage methods of affecting the spectator. Examining each new discovery from this viewpoint, it is easy to show the insignificance of the color and the stereoscopic film in comparison with the vast significance of **SOUND.**

Sound recording is a two-edged invention, and it is most probable that its use will proceed along the line of least resistance, i.e., along the line of satisfying simple curiosity.

In the first place there will be commercial exploitation of the most salable merchandise, **talking films.** Those in which sound recording will proceed on a naturalistic level, exactly corresponding with the movement on the screen, and providing a certain "illusion" of talking people, of audible objects, etc.

A first period of sensations does not injure the development of a new art, but it is the second period that is fearful in this case, a second period that will take the place of the fading virginity and purity of this first perception of new technical possibilities, and will assert an epoch of its automatic utilization for "highly cultured dramas" and other photographed performances of a theatrical sort.

To use sound in this way will destroy the culture of montage, for every **adhesion** of sound to a visual montage piece increases its inertia as a montage piece, and increases the independence of its meaning-and this will undoubtedly be to the detriment of montage, operating in the first place not on the montage pieces but on their **JUXTAPOSITION.**

Only a **CONTRAPUNTAL USE** of sound in relation to the visual montage piece will afford a new potentiality of montage development and perfection.

The first experimental work with sound must be directed along the line of its distinct **NON-SYNCHRONIZATION WITH THE VISUAL IMAGES.** And only such an attack Will give the necessary palpability which will later lead to the creation- of an **ORCHESTRAL COUNTERPOINT** of visual and aural images.

This new technical discovery is not an accidental moment in film history but an organic way out of a whole series of impasses that have seemed hopeless to the cultured cinematic avant-garde.
The first impasse is the subtitle and all the unavailing attempts to tie it into the montage composition, as a montage piece (such as breaking it up into phrases and even words, increasing and decreasing the size of type used, employing camera movement, animation, and so on). The second impasse is the explanatory pieces (for example, certain inserted close-ups) that burden the montage composition and retard the tempo.

The tasks of theme and story grow more complicated every day; attempts to solve these by methods of "visual" montage alone either lead to unsolved problems or force the director to resort to fanciful montage structures, arousing the fearsome eventuality of meaninglessness and reactionary decadence.

Sound, treated as a new montage element (as a factor divorced from the visual image), will inevitably introduce new means of enormous power to the expression and solution of the most complicated tasks that now oppress us with the impossibility of overcoming them by means of an imperfect film method, working only with visual images.

The CONTRAPUNTAL METHOD of constructing the sound film will not only not weaken the international cinema but will bring its significance to unprecedented power and cultural height. Such a method for constructing the sound film will not confine it to a national market, as must happen with the photographing of plays, but will give a greater possibility than ever before for the circulation throughout the world of a filmically expressed idea.
Asynchronism as a Principle of Sound Film (1929, excerpt)

V. I. Pudovkin

The technical invention of sound has long been accomplished, and brilliant experiments have been made in the field of recording. This technical side of sound filmmaking may be regarded as already relatively perfected, at least in America. But there is a great difference between the technical development of sound and its development as a means of expression. The expressive achievements of sound still lie far behind its technical possibilities. I assert that many theoretical questions whose answers are clear to us are still provided in practice only with the most primitive solutions. Theoretically, we in the Soviet Union are in advance of Western Europe and [the] U.S.A.

Our first question is: What new content can be brought into the cinema by the use of sound? It would be entirely false to consider sound merely as a mechanical device enabling us to enhance the naturalness of the image. Examples of such most primitive sound effects: in the silent cinema we were able to show a car, now in sound film we can add to its image a record of its natural sound; or again, in silent film a speaking man was associated with a title, now we hear his voice. The role which sound is to play in film is much more significant than a slavish imitation of naturalism on these lines; the first function of sound is to augment the potential expressiveness of the film's content.

If we compare the sound to the silent film, we find that it is possible to explain the content more deeply to the spectator with relatively the same expenditure of time. It is clear that this deeper insight into the content of the film cannot be given to the spectator simply by adding an accompaniment of naturalistic sound; we must do something more. This something more is the development of the image and the sound strip each along a separate rhythmic course. They must not be tied to one another by naturalistic imitation but connected as the result of the interplay of action. Only by this method can we find a new and richer form than that available in the silent film. Unity of sound and image is realized by an interplay of meanings which results, as we shall presently show, in a more exact rendering of nature than its superficial copying. In silent film, by our editing of a variety of images, we began to attain the unity and freedom that is realized in nature only in its abstraction by the human mind. Now in sound film we can, within the same strip of celluloid, not only edit different points in space, but can cut into association with the image selected sounds that reveal and heighten the character of each wherever in silent film we had a conflict of but two opposing elements, now we can have four.

A primitive example of the use of sound to reveal an inner content can be cited in the expression of the stranding of a town-bred man in the midst of the desert. In silent film we should have had to cut in a shot of the town; now in sound film we can carry town-associated sounds into the desert and edit them there in place of the natural desert sounds. Uses of this kind are already familiar to film directors in Western Europe, but it is not generally recognized that the principal elements in sound film are the asynchronous and not the synchronous; moreover, that the synchronous use is, in actual fact, only exceptionally correspondent to natural perception. This is not, as may first appear, a theoretical figment, but a conclusion from observation.

For example, in actual life you, the reader, may suddenly hear a cry for help; you see only the window; you then look out and at first see nothing but the moving traffic. But you do not hear the sound natural to these cars and buses; instead you hear still only the cry that first startled you. At last you find with your eyes the point from which the sound came; there is a crowd, and someone is lifting the injured man, who is now quiet. But, now watching the man, you become aware of the din of traffic passing, and in the midst of its noise there gradually grows the piercing signal of the ambulance. At this your attention is caught by the clothes of the injured man: his suit is like that
of your brother, who, you now recall, was due to visit you at two o'clock. In the tremendous
tension that follows, the anxiety and uncertainty whether this possibly dying man may not indeed
be your brother himself, all sound ceases and there exists for your perceptions total silence. Can it
be two o'clock? You look at the clock and at the same time you hear its ticking. This is the first
synchronized moment of an image and its caused sound since first you heard the cry.
Always there exist two rhythms, the rhythmic course of the objective world and the tempo and
rhythm with which man observes this world. The world is a whole rhythm, while man receives
only partial impressions of this world through his eyes and ears and to a lesser extent through his
very skin. The tempo of his impressions varies with the rousing and calming of his emotions,
while the rhythm of the objective world he perceives continues in unchanged tempo.
The course of man's perceptions is like editing, the arrangement of which can make

...
London, May 1929. Today there is no individual, no company, no financial coalition capable of stopping the triumphant march of the talking film. The industrialists of the American cinema maintain that the public has clearly manifested its liking for talkies, and that they have done no more than meet the public's wishes. But if the public suddenly got tired of its new toy, the same docile industrialists would certainly refuse to pander further to its whims. For meanwhile the talkies have become one of the biggest business undertakings of our age, to which, banks and public utility companies with interests on an imperial scale have linked their fate. So many thousand million dollars have been invested in this enterprise that from now on any and every means will be used to ensure its success. The talking film exists, and those skeptics who prophesy a short reign for it will die themselves long before it's over. It is too late for those who love the art of moving pictures to deplore the effects of this barbaric invasion. All they can do is try to cut their losses.

The talking film is not everything. There is also the sound film-on which the last hopes of the advocates of the silent film are pinned. They count on the sound film to ward off the danger represented by the advent of talkies, in an effort to convince themselves that the sounds and noises accompanying the moving picture may prove sufficiently entertaining for the audience to prevent it from demanding dialogue, and may create an illusion of "reality" less harmful for the art than the talking film.

However, we have grounds to fear that this solution will only half-satisfy the public. If there is almost universal agreement about the advantages of a mechanical musical accompaniment over the improvisations of a cinema orchestra, opinions vary as far as noises accompanying the action are concerned. The usefulness of such noises is often questionable. If at first hearing they are surprising and amusing, very soon they become tiresome. After we have heard a certain number of sound films, and the first element of surprise has worn off, we are led to the unexpected discovery that the world of noises seems far more limited than we had thought.

Although the talkies are still in their first, experimental stage, they have already, surprisingly enough, produced stereotyped patterns. We have barely "heard" about two dozen of these films, and yet we already feel that the sound effects are hackneyed and that it is high time to find new ones. Jazz, stirring songs, the ticking of a clock, a cuckoo singing the hours, dance-hall applause, a motorcar engine, or breaking crockery—all these are no doubt very nice, but become somewhat tiresome after we have heard them a dozen times in a dozen different films.

We must draw a distinction here between those sound effects which are amusing only by virtue of their novelty (which soon wears off), and those that help one to understand the action, and which excite emotions which could not have been roused by the sight of the pictures alone. The visual world at the birth of the cinema seemed to hold immeasurably richer promise. . . . However, if imitation of real noises seems limited and disappointing, it is possible that an interpretation of noises may have more of a future in it. Sound cartoons, using "real" noises, seem to point to interesting possibilities.

Unless new sound effects are soon discovered and judiciously employed, it is to be feared that the champions of the sound film may be heading for a disappointment. We shall find ourselves left with the "hundred per cent talkie," as they say here, and that is not a very exhilarating prospect.

Of all the films now showing in London, Broadway Melody is having the greatest success. This new American film represents the sum total of all the progress achieved in sound films since the appearance of The Jazz Singer two years ago. For anyone who has some knowledge of the complicated technique of sound recording, this film is a marvel. Harry Beaumont, the director,
and his collaborators (of whom there are about fifteen, mentioned by name in the credit titles, quite apart from the actors) seem to delight in playing with all the difficulties of visual and sound recording. The actors move, walk, run, talk, shout, and whisper, and their movements and voices are reproduced with a flexibility which would seem miraculous if we did not know that science and meticulous organization have many other miracles in store for us. In this film, nothing is left to chance. Its makers have worked with the precision of engineers, and their achievement is a lesson to those who still imagine that the creation of a film can take place under conditions of chaos known as inspiration.

In Broadway Melody, the talking film has for the first time found an appropriate form: it is neither theater nor cinema, but something altogether new. The immobility of planes, that curse of talking films, has gone. The camera is as mobile, the angles are as varied as in a good silent film. The acting is first-rate, and Bessie Love talking manages to surpass the silent Bessie Love whom we so loved in the past. The sound effects are used with great intelligence, and if some of them still seem superfluous, others deserve to be cited as examples.

For instance, we hear the noise of a door being slammed and a car driving off while we are shown Bessie Love's anguished face watching from a window the departure which we do not see. This short scene in which the whole effect is concentrated on the actress's face, and which the silent cinema would have had to break up in several visual fragments, owes its excellence to the "unity of place" achieved through sound. In another scene we see Bessie Love long thoughtful and sad; we feel that she is on the verge of tears; but her face disappears in the shadow of a fade-out, and from the screen, now black, emerges a single sob.

In these two instances the sound, at an opportune moment, has replaced the shot. It is by this economy of means that the sound film will most probably secure original effects. We do not need to hear the sound of clapping if we can see the clapping hands. When the time of these obvious and unnecessary effects will have passed, the more gifted filmmakers will probably apply to sound films the lesson Chaplin taught in the silent films, when, for example, he suggested the arrival of a train by the shadows of carriages passing across a face. (But will the public, and, above all, the filmmakers, be satisfied with such a discreet use of sound? Will they not prefer an imitation of all the noises to an intelligent selection of a few useful ones?)

Already in the films we are shown at present, we often feel that in a conversation it is more interesting to watch the listener's rather than the speaker's face. In all likelihood American directors are aware of this, for many of them have used the device quite often and not unskillfully. This is important, for it shows that the sound film has outgrown its first stage, during which directors were intent on demonstrating, with childish persistence, that the actor's lips opened at exactly the same moment as the sound was heard—short, that their mechanical toy worked beautifully. It is the alternate, not the simultaneous, use of the visual subject and of the sound produced by it that creates the best effects. It may well be that this first lesson taught us by the birth pangs of a new technique will tomorrow become this same technique's law.

Whenever the most faithful devotees of the silent cinema undertake an impartial study of talking films, they inevitably lose some of their assurance right at the start, for, at its best, the talkie is no longer photographed theater. It is itself. Indeed, by its variety of sounds, its orchestra of human voices, it does give an impression of greater richness than the silent cinema. But are such riches not in fact quite ruinous to it? Through such "progressive" means the screen has lost more than it has gained. It has conquered the world of voices, but it has lost the world of dreams. I have observed people leaving the cinema after seeing a talking film. They might have been leaving a music hall, for they showed no sign of the delightful numbness which used to overcome us after a passage through the silent land of pure images. They talked and laughed, and hummed the tunes they had just heard. They had not lost their sense of reality.
In Audio-Vision, the French composer-filmmaker-critic Michel Chion presents a reassessment of the audiovisual media since sound’s revolutionary debut in 1927 and sheds light on the mutual influences of sound and image in audiovisual perception.

Chion expands on the arguments from his influential trilogy on sound in cinema—La Voix au cinema, Le Son au cinema, and La Toile trouée—while providing an overview of the functions and aesthetics of sound in film and television. He considers the effects of evolving audiovisual technologies such as widescreen, multi-track sound, and Dolby stereo on audio-vision, influences of sound on the perception of space and time, and contemporary forms of audio-vision embodied in music videos, video art, and commercial television. His final chapter presents a model for audiovisual analysis of film.

Walter Murch, who contributes the foreword, has been honored by both the British and American Motion Picture Academies for his sound design and picture editing. He is especially well-known for his work on The Godfather, The Conversation, and Apocalypse Now.

"Michel Chion is the leading French cinema scholar to study the sound track. . . . I know of no writer in any language to have published as much in this area, and of such uniformly high quality, a he."

ALAN WILLIAMS
RUTGERS UNIVERSITY

MICHEL CHION is an experimental composer, a director of short films, and a critic for Cahiers du cinema. He has published books on screenwriting, Jacques Tati, David Lynch, and Charlie Chaplin, in addition to his four books on film sound.

CLAUDIA GORBMAN is a Professor in the Liberal Studies Program at the University of Washington, Tacoma.

Jacket design: John Costa
Printed in U.S.A.
FOUR
THE AUDIOVISUAL SCENE

IS THERE AN AUDITORY SCENE?

"The Image" = The Frame

Why in the cinema do we speak of "the image" in the singular when a film has thousands of them (only several hundred if it's shots we're counting, but these too are ceaselessly changing)? The reason is that even if there were millions, there would still be only one container for them, the frame. What "the image" designates in the cinema is not content but container: the frame.

The frame can start out black and empty for a few seconds (Ophuls's Le Plaisir, Preminger's Laura) or even for several minutes (Duras's L'Homme Atlantique). But it nevertheless remains perceivable and present for the spectator as the visible, rectangular, delimited place of the projection. The frame thus affirms itself as a preexisting container, which was there before the images came on and which can remain after the images disappear (end credits reaffirm this role in a certain way). 1

What is specific to film is that it has just one place for images—as opposed to video installations, slide shows, Sound and Light shows, and other multimedia genres, which can have several. This fact, and no other, accounts for why we speak of the image in the singular.

Let us recall that in the first years of the cinematograph people sought to soften the hard borders of the frame, through iris, masking, or haloing, similar to such effects in photography. But these techniques were abandoned little by little, and, aside from the rare experiment with changing frame dimensions within a single film (Max Ophuls in Lola Montès), the principle of the full-frame image came to dominate in 99 percent of movies. Similarly, the occasional experiment with multiscreen cinema—Abel Gance's Napoleon, Michael Wadleigh's Woodstock, or even Paul Morrissey's Forty Deuce—have not spawned many descendants, and as exceptions they prove the rule of the classical frame.

There Is No Auditory Container for Sounds

What is the corresponding case for sound? The exact opposite. For sound there is neither frame nor preexisting container. We can pile up as many sounds on the soundtrack as we wish without reaching a limit. Further, these sounds can be situated at different narrative levels, such as conventional background music (nondiegetic) and synch dialogue (diegetic)—while visual elements can hardly ever be located at more than one of these levels.
At once. So there is no auditory container for film sounds, nothing analogous to this visual container of the images that is the frame.

What do sounds do when put together with a film image? They dispose themselves in relation to the frame and its content. Some are embraced as synchronous and onscreen, others wander at the surface and on the edges as offscreen. And still others position themselves clearly outside the diegesis, in an imaginary orchestra pit (nondiegetic music), or on a sort of balcony, the place of voiceovers. In short, we classify sounds in relation to what we see in the image, and this classification is constantly subject to revision, depending on changes in what we see. Thus we can define most cinema as "a place of image's, plus sounds." with sound being "that which seeks its place." This relation differs from that of television, as we will see later on.

If we can speak of an audiovisual scene, it is because the scenic space has boundaries, it is structured by the edges of the visual frame. Film sound is that which is contained or not contained in an image; there is no place of the sounds, no auditory scene already preexisting in the soundtrack—and therefore, properly speaking, there is no soundtrack.

But Jean-Marie Straub’s and Danièle Huillet’s highly idiosyncratic 1969 film Othon (which acts out a Roman tragedy by Corneille on modern-day Roman locations) demonstrates what a sound scene or an auditory container-of-sounds might be in a monaural film. We’d have to agree that the sounds are the actors’ voices declaiming their lines, and that the container would be the urban hum of distant traffic in which the voices and lines are heard. Actors in Othon often give long monologues offscreen, and yet such voices are not perceived as the traditional offscreen voice entirely determined by the image. Their voices seem to be "in the same place" as voices of actors we do see, a space defined by the background noise. A related effect can be felt in another film of the same year, Jacques Rivette’s La Religieuse. Here, the reverb around voices, which results from direct sound (as with Straub and Huillet), has a similar role of enveloping and homogenizing the voices, inscribing them in a space like the medium of city traffic noise does in Othon. The price each film pays is a relative loss of intelligibility. Generally speaking, certain effects of the "spatial signature," as Rick Altman calls it, can provide the framework for an auditory scene.

At least all this holds true until the arrival of Dolby, which now creates a space with fluid borders, a sort of superscreen enveloping the screen—the superfield, which I expand upon in a later chapter. But the superfield does not altogether upset the structure we have described, even if it has set it trembling on its base.

How THE IMAGE "MAGNETIZES" SOUND IN SPACE

What does a sound typically lead us to ask about space? Not "Where is it?"—for the sound "is" in the air we breathe or, if you will, as a perception it’s in our head—but rather, "Where does it come from?" The problem of localizing a sound therefore most often translates as the problem of locating its source.

Traditional monaural film presents a strange sensory experience in this regard. The point from which sounds physically issue is often not the same as the point on the screen where these sounds are supposed to be coming from, but the spectator nevertheless does perceive the sounds as coming from these "sources" on the screen. In the case of footsteps, for example, if the character is walking across the screen, the sound of the footsteps seems to follow his image, even though in the real space of the movie theater, they continue to issue from the same stationary loudspeaker. If the character is offscreen, we perceive the footsteps as if they are outside the field of vision—an "outside" that’s more mental than physical.
Moreover, if under particular screening conditions the loudspeaker is not located behind the screen, but placed somewhere else in the auditorium or in an outdoor setting (e.g., at the drive-in), or if the soundtrack resonates in our head by means of earphones (watching a movie on an airplane), these sounds will be perceived no less as coming from the screen, in spite of the evidence of our own senses.

This means that in the cinema there is spatial magnetization of sound by image. When we perceive a sound as being offscreen or located at screen right this is a psychological phenomenon, at least if a monaural projection is involved.

During the first years of multitrack sound, attempts at real spatialization were made—that is, really locating the sound on the left side of the screen if its source was shown there. The problem with these efforts is precisely that they ran into this psychological phenomenon of spatialization. Mental spatialization had been a blessing for the sound film, since it allowed movies to function for well over forty years without problems. We only need imagine the mess if sounds had to issue from the points where their sources on the screen were shown: one would have to install veritable beehives of speakers behind and around the screen. Not to mention, of course, the headaches of sound matching that would have resulted.

In using Dolby today filmmakers have learned the lesson from these first efforts in realistic spatialization and their "in-the-wing effects" (see p. 83). Today's multitrack mixes very often strike a compromise between psychological localization and real localization.

Note that sound coming from another point than the screen is "magnetizable" only if the sound itself maintains a basic spatial stability. If it constantly moves back and forth among loudspeakers, the image will have a harder time absorbing it, and the sound takes on a centrifugal force of its own that resists visual "attraction."

Even in the classic case of a single loudspeaker, there is one real sonic dimension that the sound cinema capitalized on in its infancy, and neglected later: depth, the sensation of distance from the source. The ear detects depth from such indices as a reduced harmonic spectrum, softened attacks and transitions, a different blend of direct sound and reflected sound, and the presence of reverberation. The factor of depth has figured importantly in experiments with sound perspective in some films.6 Let us note, however, that sound perspective was not so much a true depth, necessarily situating the sound source to the rear of the spatial plane of the screen, as a distance interpreted by the spectator in various different directions, depending on what she or he saw on the screen and could infer about the place of the source. In other words, a distant sound can be interpreted as being distantly to the left, far to the right, far behind the spectator, far to the rear of the screen; in other words, always localized in space depending on mental factors.

Thus to mental localization, determined more by what we see than by what we hear (or rather by the relationship between the two), we may oppose the absolute spatialization made possible by multitrack film sound.

**THE ACOUSMATIC**

Acousmatic, a word of Greek origin discovered by Jerome Peignot and theorized by Pierre Schaeffer, describes "sounds one hears without seeing their originating cause."7 Radio, phonograph, and telephone, all which transmit sounds without showing their emitter, are acousmatic media by definition. The term acousmatic music has also been coined; composer Francis Bayle, for example, uses
it to designate concert music that is made for a recorded medium, intentionally eliminating the possibility of seeing the sounds' initial causes.

What can we call the opposite of acousmatic sound? Schaeffer proposed "direct," but since this word lends itself to so much ambiguity, we shall coin the term visualized sound—i.e., accompanied by the sight of its source or cause.

In a film an acousmatic situation can develop along two different scenarios: either a sound is visualized first, and subsequently acousmatized, or it is acousmatic to start with, and is visualized only afterward. The first case associates a sound with a precise image from the outset. This image can then reappear with greater or lesser distinctness in the spectator's mind each time the sound is heard acousmatically. It will be an "embodied" sound, identified with an image, demythologized, classified.⁸

The second case, common to moody mystery films, keeps the sound's cause a secret, before revealing all. The acousmatic sound maintains suspense, constituting a dramatic technique in itself. A theatrical analogy to this treatment of sound might be to announce and then to delay a stage entrance; think of Tartuffe, who finally enters during the third act of Molière's play. The cinema gives us the famous example of M; for as long as possible the film conceals the physical appearance of the child-murderer, even though we hear his voice and his maniacal whistling from the very beginning. Lang preserves the mystery of the character as long as he can, before "de-acousmatizing" him.⁹

A sound or voice that remains acousmatic creates a mystery of the nature of its source, its properties and its powers, given that causal listening cannot supply complete information about the sound's nature and the events taking place.

If it's fairly common in films to see evil, awe-inspiring, or otherwise powerful characters introduced through sound before they are subsequently thrown out to the pasture of visibility, de-acousmatic-

tized. Odile Larere has discussed the example of Visconti's Conversation Piece, where the intruders who disturb the lovely universe of the hero, the old professor played by Burt Lancaster, systematically make their entrance on the soundtrack before being visible.⁸

The opposition between visualized and acousmatic provides a basis for the fundamental audiovisual notion of offscreen space.

THE QUESTION OF OFFSCREEN SPACE
Onscreen, Offscreen, Nondiegetic

The question of offscreen sound has long dominated an entire field of thinking and theorizing about film sound, and it occupies a central place in my first two books on sound as well. Although we can see now that it seems to have been privileged at the expense of other avenues of investigation, it has yet to lose its importance as a central problem—even if the recent evolution of film sound, involving mainly multitrack sound and the "superfield" it establishes, has modified some of its basic traits.

In the narrow sense offscreen sound in film is sound that is acousmatic, relative to what is shown in the shot: sound whose source is invisible, whether temporally or not. We call onscreen sound that whose source appears in the image, and belongs to the reality represented therein.

Third, to designate sound whose supposed source is not only absent from the image but is also external to the story world, I shall use the term nondiegetic. This is the widespread case of voiceover commentary and narration and, of course, musical underscoring.

Do Exceptions Disprove the Rule?
In Le Son au cinéma I presented onscreen, offscreen, and nondiegetic as three zones of a circle, wherein each communicates with the other two:
But in recent years, the distinction onscreen-offscreen-nondeictic, which arises from very basic considerations, has often been denounced as obsolete and reductive. Critics have problematized it with increasing fervor, because of the exceptions and special cases it doesn't seem to account for. For example, where should we situate sounds (usually voices) that come from electrical devices located in the action and that the image suggests or directly shows: telephone receivers, radios, public-address speakers? And what to do with a character who speaks with her back to us, so we don't actually see her speak? Is her voice acousmatic (offscreen)? And what can we say about the so-called internal voice of a character who can be seen in the image—the voices of his conscience, of his memory, of his imaginings and fantasies?

What about Amy Heckerling's *Look Who's Talking*, where an adult voice accompanies the facial expressions of a baby, and articulates the baby's thoughts and feelings when the baby obviously doesn't have the physical and intellectual ability to do so? The voice is definitely connected to the present of the action, but it is not visualizable; so it seems unconcerned with these distinc-
tions, being tied to the image via the loosest of synchronization. And finally how should we classify general background sounds such as birdsongs and wind, heard with natural exteriors? It seems rather ridiculous to characterize them as offscreen, on the basis that we don't "see" the little birds chirping or the wind blowing.

These exceptions, though distressing, do not by any means cancel out the validity or interest of a basic distinction between onscreen, offscreen, and nondiegetic sound, or of the basic division between acousmatic and visualized.

A Topological and Spatial Perspective

Anyone who brings up such exceptions in order to claim the categories useless or trivial is throwing out the baby with the bathwater. Why reject a valuable distinction simply because it isn't absolute? It is a mistake to see things in a binary, all-or-nothing logic. These distinctions only have meaning from a geographical, topological, and spatial perspective, analogous to zones among which one finds many shadings, degrees, and ambiguities. Of course we must continue to refine and fill in our typology of film sound. We must add new categories—not claiming thereby to exhaust all possibilities, but at least to enlarge the scope, to recognize, define, and develop new areas.

Ambient Sound (Territory-Sound)

Let us call ambient sound sound that envelops a scene and inhabits its space, without raising the question of the identification or visual embodiment of its source: birds singing, churchbells ringing. We might also call them territory sounds, because they serve to identify a particular locale through their pervasive and continuous presence.
Figure 2-51. PRIMARY COLOR THEORY. For psychological reasons all colors of the visible spectrum can be reproduced by the combination of three so-called primary colors. The "additive" primaries are red, blue, and green. The "subtractive" primaries are magenta (red-blue), cyan (blue-green), and yellow. All the colors of the spectrum together produce white light, just as complete absence of color yields black. If both magenta and yellow are subtracted from a beam of white light, the result is red. If both red and green beams are added together, the result is yellow, and so on. There is no mathematical reason for this to be; it's a psychological conundrum.

imbalances. Moreover, digital color correction now goes a long way to freeing the cinematographer from these limitations.

The Soundtrack

The recording of sound is roughly parallel to the recording of images: the microphone is, in effect, a lens through which sound is filtered; the recorder compares roughly with the camera; both sound and picture are recorded linearly and can be edited later. But there is one significant difference: because of the contrasting manners in which we perceive them, sound must be recorded continuously, while pictures are recorded discretely. The concept of "persistence of vision" does not have an aural equivalent, which is one reason why we don't have "still sounds" to compare with still pictures. Sound must exist in time.

A corollary of this is that we cannot apply sound recording devices to aural information in the same way we can apply cinematography to visual information. Film can stretch or compress time, which is useful scientifically, but sound must exist in time, and it is pointless to compress or
stretch it. Sound was digitized as early as the 1970s. Digital recordings can be played back at a faster or slower rate, but when we change the speed of an analog recording we change the quality of the sound as well.

The union of sound and image, the original dream of the inventors of cinema, was delayed for technological and economic reasons until the late 1920s. So long as image was recorded in a linear, discontinuous mode and sound was recorded in a circular, continuous mode, the problem of synchronization of sound and image was insurmountable. Lee De Forest's audion tube, invented in 1906, made it possible for the first time to translate sound signals into electrical signals. The electrical signals could then be translated into light signals that could be imprinted on film. Then the two prints—sound and image—being parallel, could easily be "married" together so that they were always and forever synchronous, even if the film broke and had to be spliced. This was essentially the German Tri-Ergon system that was patented as early as 1919. This optical sound system has existed more or less unchanged to the present.

For twenty years after the sound film was born in 1926, filmmakers were hampered by the bulky and noisy electromechanical equipment necessary to record sound on the set. Even though portable optical recorders were soon available, recording on location was discouraged. In the late forties, however, the technology of film took another quantum leap with the development of magnetic recording. Tape is easier to work with than film, more compact, and, thanks to transistors, the recording devices themselves are now small and lightweight. Magnetic tape, in general, also produces a much better quality signal than an optical soundtrack does. Today, magnetic recording has entirely replaced optical recording on the set, although the optical soundtrack is still more common than the magnetic soundtrack in theaters. (See Figure 2-70.) There is good reason for this: optical soundtracks can be printed quickly and easily along with image tracks, while magnetic soundtracks must be recorded separately. Developments in optical soundtrack technology, moreover, suggest that some of the advantages that magnetic recording now enjoys over optical recording might be matched: variable-density and variable-hue optical soundtracks could eliminate the effects of rough handling, providing a higher fidelity, and could also be adapted to stereophonic and multiphonic systems. Because of its advantages in handling, editing, and mixing, however, magnetic tape—and now digital disk storage—remain the mediums of choice on the set and in the laboratory.
Figure 2-52. THE DOLBY EFFECT. A certain amount of basic surface noise is inherent in any recording medium (A). It presents no problem when the level of the recorded signal is high enough, but it masks out the weaker parts of the signal. The Dolby system boosts the weaker signal during recording (B), then reduces it to its proper level during playback, thus reducing the recorded surface noise along with it (C).

You may think that digital distribution media like DVD have obviated these problems. Far from it. MPEG-2 and MPEG-4, the standard encoding systems for DVD and HDTV, keep the sound signal separate from the video signal resulting often in quite noticeable synchronization errors.

The microphone, the lens of the sound system, acts as the first gate through which the signal passes. Unlike the optical lens, however, it also translates the signal into electrical energy, which can then be recorded magnetically on tape or disk. (Playback systems work exactly the reverse: magnetic potential energy is translated into electrical energy which is further translated into physical sound by the loudspeaker.) Since sound is being recorded on a tape physically separate from the filmed image, there must be some method of synchronizing the two. This is accomplished either by a direct mechanical linkage, or by electrical cable connections that carry a timed impulse, or by a crystal sync generator, which produces a precisely timed pulse by using crystal clocks. This pulse regulates the speeds of the two separate motors, keeping them precisely in sync. The sound record is then transferred to magnetically coated film, where the sprocket holes provide the precise control over timing that is necessary in the editing process. Finally, the print of the film that is projected carries the signal, usually in the optical mode, but sometimes magnetically. Stereophonic and “quintaphonic” 5.1 sound systems common to 70 mm systems almost always use magnetic tracks.

The variables that contribute to the clear and accurate reproduction of sound are roughly comparable to the variables of filmstock. The factor of amplitude can be compared to the exposure latitude of filmstock: the amplitude is the measure of the strength of a signal. Tape, recorder, and microphone working in concert should be able to reproduce a wide range of amplitudes, from very soft to very loud.
Next in importance is the frequency range, directly comparable to the scale of hues reproducible in color film. The normal range of frequencies to which the ear responds is 20 to 20,000 Hertz (cycles per second). Good high-fidelity recording equipment can reproduce this range adequately, but optical playback systems have a much more limited range of frequency response (100 to 7,000 Hertz, on the average).

The recording medium and the equipment should also be able to reproduce a wide range of harmonics, those subtones that give body and life to music and voices. The harmonics of sound can be compared to the tonal qualities of an image. The signal should be free from wow, flutter, and other kinds of mechanical distortion, and the equipment should have a satisfactory response time: that is, the ability to reproduce sounds of short duration without mushiness. This is the "resolution" of the sound signal.

While stereoscopic images are subject to special psychological and physical problems that significantly reduce their value, stereophonic sound is relatively free of these problems and therefore highly desirable. We are used to hearing sound from every direction. Although we engage in selective attention to sounds, we don’t focus directly on a sound the way we focus on an image. Film sound should have the ability to reproduce the total aural environment, the "soundstage."

In the 1970s, the assimilation of multitrack recording techniques developed in the music industry expanded the horizons of the art of film sound—for example, such sophisticated soundtracks as Coppola’s *The Conversation* (1974) and Altman’s *Nashville* (1975), produced on an eight-track system. The application of Dolby techniques of noise reduction and signal enhancement in the mid-seventies greatly increased the potential fidelity of film sound as well. Roughly comparable to the flashing of filmstock, the Dolby electronic circuitry reduces the background noise inherent in even the best tape stock, thereby significantly enhancing the latitude. It does this by selecting out the area of the sound spectrum in which the noise occurs and boosting the signal level in that area during recording. When the signal is reduced to normal levels during playback, the noise is reduced along with the audio signal.

For years film sound technology had lagged behind home audio technology. Many films were released with monophonic tracks long after stereo became the standard for records. Beginning in the 1980s, however, theater owners began to pay more attention to quality sound reproduction. Dolby and Sony led the way with ingenious schemes for recording multitrack sound on basic filmstock. George Lucas’s THX program set the stan-
standard for advanced sound systems. By the mid-1990s sophisticated sound reproduction had become a major marketing advantage for theater chains.

Post-Production

Film professionals divide the process of their art into three phases: pre-production, shooting, and post-production. The first phase is preparatory—the script is written, actors and technicians hired, shooting schedules and budgets planned. In another art, this period of preparation would be relatively uncreative. But Alfred Hitchcock, for one, regarded this period of the film process as paramount: once he had designed the film, he used to say, its execution was comparatively boring. Moreover, in this most expensive of arts, intelligent and accurate planning often spells the difference between success and failure. It must be clear by now that making films is a complicated business—so much so that modern systems design has had a measurably positive effect on the process. The elaborate, carefully organized systems Stanley Kubrick created for his film projects, for example, were one of the more intriguing aspects of his work.

Nearly all the discussion in this chapter on film technology has so far centered on the second phase of film production: shooting. Yet there is a sense in which this area of the process can be seen as preparatory, too. Shooting produces the raw materials that are fashioned into finished products only in the third stage of the process. Editing is often regarded as the fulcrum of film art, since it is in this process that film most clearly separates itself from competing arts. The theory of film editing will be discussed in Chapters 3 and 5; here we will outline the practice and describe the equipment involved. Three jobs generally proceed more or less concurrently during post-production: editing; sound mixing, augmentation, and looping (or ADR and Foley work); and laboratory work, opticals, and special effects. A film could conceivably be edited, mixed, and printed within a few hours; assuming both the soundtrack and the picture in their raw state were satisfactory, the editing would be simply a matter of splicing a few takes end to end. But very few films are this simple, and post-production work often takes longer than the actual shooting of the film. Although it is called “post-production,” the work often begins during the shoot and runs concurrently.
Actions/Interactions:
The Source Beyond the Source

One of the constant points of examination in the domain of film music is the distinction between diegetic, or "source," music and nondiegetic music, the latter generally taking the form of a score composed especially for the film in question. Diegetic music theoretically comes from a source within the diegesis—a radio, a phonograph, a person singing, an orchestra playing—and the characters within the film can theoretically hear that music. Nondiegetic music theoretically exists for the audience alone and is not supposed to enter in any way into the universe of the filmic narrative and its characters. (This distinction, of course, has little meaning in the film musical, which is not under consideration in this book.) The very existence of the nondiegetic score is of course an anomaly. Some directors, in fact, have balked against its use. John Ford, for instance, never liked the idea of having one of his characters lost in the middle of the desert while somehow a fifty-piece orchestra can be heard. Other directors have made fun of the nondiegetic score and its ambiguities. At one point in Bananas, director/star Woody Allen has been told that he has been invited to have dinner with the president of the Latin American dictatorship he finds himself in. Dreamily, he lies back on the bed in his hotel room and exclaims, "Dinner with the president!" several times. A harp flour- ish, apparently from the nondiegetic music track, surrounds Allen with an aura that reinforces, for the audience, the dreamy nature of his musings. But Allen has apparently heard that same music: looking confused, he gets up, opens a closet door, and discovers a harp and a local resident, who explains that he needed a place to practice. Later on, during that dinner with the president, Allen further makes fun of the diegetic/nondiegetic distinction by hav- ing a group of musicians performing a string quartet on the balcony of the
dining room. But the musicians have no instruments and instead mime the
gestures of playing "invisible" music on their invisible instruments, which
does not stop the Woody Allen character from growling at the musicians, as
he leaves the dining room, "Could you keep it down? I have a headache." At
the beginning of Mel Brooks's Hitchcock spoof, High Anxiety, the director/
star is being taken in a limo to the quasi-Spellbound institution that he is sup-
posed to head. As costar Ron Carey, explaining the strange goings-on in the
institution, mentions the words "foul play," a stereotypically violent musical
outburst is heard, apparently, as in the Woody Allen film, on the nondiegetic
music track. But, as in Bananas, the characters react to this seemingly invis-
ible music, only to immediately see its "source" pass by in the form of a bus
filled with members of the "Los Angeles Symphony." Later in the film,
Brooks likewise makes fun of certain invisible visual devices by having the
camera break the glass in a French door as it performs a gorgeous, Hitchcock-
ian tracking shot from outdoors into a dinner scene.

Both Bananas and High Anxiety comedically remind audiences of the cin-
ematic givens they willingly accept in the area of music: a) that there will al-
most always be an invisible nondiegetic score to boost the affective impact at
key moments, and b) that source music, which is almost always laid in on the
soundtrack after the film has been shot, is being or can be heard by the char-
acters in the narrative, whereas in fact the actors and actresses on the set gen-
erally hear no music at all. In an interesting reversal, director Ken Russell,
during the making of The Music Lovers, decided that his stars, Richard Cham-
berlain and Glenda Jackson, needed music to get their affective juices flowing
during the shooting of the scene on the train, during which the homosexual
Tchaikovsky discovers that he cannot consummate the marriage to his nym-
phomanic wife. Although the nondiegetic music heard in the final cut of
The Music Lovers comes from Tchaikovsky's Sixth Symphony, Russell had
his actors listen to the even more violent The Execution of Stepan Razin by
Shostakovich. In the days of the accompanied silent film, source music, gen-
erally on a solo instrument, could be presented only as a near or not-so-near
illusion. And when it was attempted, it was undermined first by the difficulty
of split-second coordination between the music and the action and second by
the fact that it could not stand apart particularly well from the nonstop non-
diegetic score. In the early days of the talkies, as we have seen, the lack of
post-recording techniques created almost the opposite situation: nonmusical
films had little music, and what there was usually came from a source. This
led to an aesthetics of realism in which Hollywood obviously did not really
have its heart, as was borne out by the gusto with which Hollywood plunged
into nondiegetic music once the score could be added to the film after the
shooting. Certain directors have revived this aesthetics of realism from time to
time, including Orson Welles in the 1958 Touch of Evil and Peter Bogdan-
ovich in his 1971 The Last Picture Show. These directors fortified the seamy,
steamy atmosphere of their films by using mostly diegetic music throughout. On the other hand, the universe of Hitchcock's *Psycho* is so closed off that the audience hears not a single note of diegetic music. But these practices, of course, reveal a diegetic = real/nondiegetic = unreal prejudice that ignores some of the deeper implications of the musical image. Perhaps the ultimate statement in this area was made by Luis Buñuel in his scathing *Le Journal d'une femme de chambre* (Diary of a Chambermaid), which uses not one note of music, diegetic or nondiegetic, throughout the entire film, thereby depriv ing the audience of any chance to indulge in real/unreal distinctions. One might also note the mild experiments to be found in films such as Richard Fleischer's 1952 *The Narrow Margin*, in which strategically placed steam-engine whistles provide that film's only "music."

Although the film can create certain illusions that fortify audience acceptance of a given piece of music as diegetic—the sound quality can be diminished and rendered hissy and scratchy for a scene in which a 78 r.p.m. recording is played, for instance—the diegetic tipoff generally comes from an object such as a radio or a phonograph in the visual diegesis. The ear, then, is generally not afforded the luxury of being able to make a major distinction between diegetic and nondiegetic musical cues: to the ear, both are aural/musical images. And, in fact, numerous movies have sequences in which the music shifts diegetic–nondiegetic alliances with no apparent justification, practical, artistic, philosophical, or otherwise. In Hitchcock's *North by Northwest*, for instance, a kind of bland music accompanies the opening dialogue between Cary Grant and Eva Marie Saint in the dining car. Although no source is seen, the audience presumes that this is a typical example of the use of Muzak (the music, in fact, was composed by André Previn and used in an earlier MGM film, *Designing Woman*). As the conversation amorously heats up, however, the "Love Theme" composed by Bernard Herrmann for the nondiegetic music track takes over without missing a beat.