

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 min-

utes (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-ing, 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 Montes*), 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."² 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 Daniele 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.³

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.⁴ 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."⁵ 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.⁶

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 Moliere'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.⁷

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.

It is 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-acousma-

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

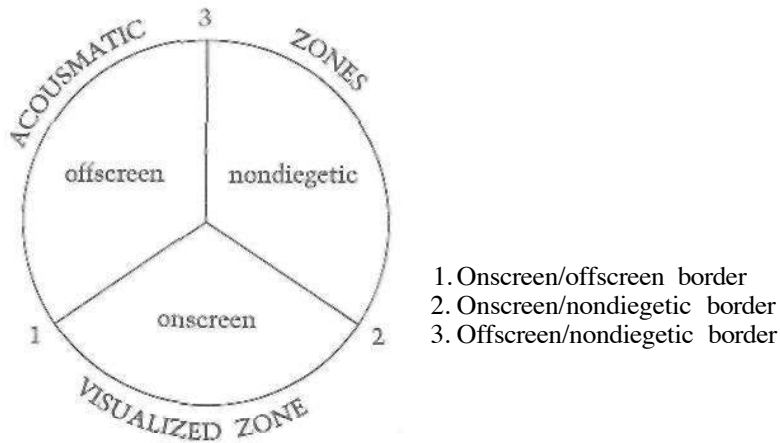
The question of offscreen sound has long dominated an entire field of thinking and theorizing about film sound, and it occupies a central place in my first two books on sound as well. Although we can see now that it seems to have been privileged at the expense of other avenues of investigation, it has 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 temporarily 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-nondiegetic, 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?¹⁰

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.

Internal Sound

Internal sound is sound which, although situated in the present action, corresponds to the physical and mental interior of a character. These include physiological sounds of breathing, moans or heartbeats, all of which could be named *objective-internal* sounds. Also in this category of internal sounds are mental voices, memories, and so on, which I call *subjective-internal* sounds.

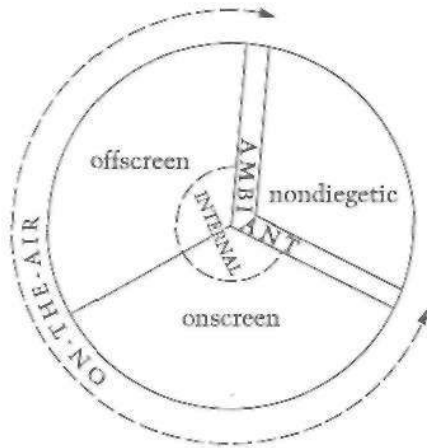
Bruce Willis's voice in *Look Who's Talking* gives us an interesting case of an internal voice, partly externalized through gesture. The film establishes it as not being heard by the other characters. In the voice of the adult that the baby will become, it tells us what the baby might be thinking, even as this voice is associated with the gestures in a way that is faithful to codes of realism regarding the baby's physical abilities.

"On-the-Air" Sound

I shall refer to sounds in a scene that are supposedly transmitted electronically as *on-the-air*—transmitted by radio, telephone, amplification, and so on—sounds that consequently are not subject to "natural" mechanical laws of sound propagation. In fact, to an ever greater degree, these sounds from television sets, clock radios, and intercoms are taking on a unique status in the films they appear in. Sometimes we hear them in sound closeup—clear and sharp, as if the film's loudspeaker were directly plugged into the radio, telephone, or phonograph depicted on the screen. At the other extreme they can be identified in the setting by acoustical traits to produce an effect of distancing, reverb, and the particular tone color of the speakers or whatever their onscreen source is. Between these two cases lie infinite degrees of variation. On-the-air sounds, usually situated in the scene's real time, enjoy the freedom of crossing boundaries of cinematic space.

A particular case of on-the-air sound is that of recorded or broadcast music. Depending on the particular weight given by such factors as mixing, levels, use of filters, and conditions of music recording—i.e., whether the emphasis is on the sound's *initial source* (the real instruments that play, the voice that sings) or on the *terminal source* (the speaker present in the narrative whose material presence is felt through use of filters, static, and reverb), the sound of on-the-air music can transcend or blur the zones of onscreen, offscreen, and nondiegetic. It can also be read, to greater or lesser degrees, as screen music or pit music. Road movies such as Barry Levinson's *Rain Man* constantly play with this oscillation. As early as 1975 George Lucas's *American Graffiti*—with the help of its sound designer Walter Murch—explored the entire gamut of possibilities between these two poles. The film was based on the simple setup of placing its characters in their cars for much of the action, all listening to a single rock-and-roll station.

The same problem exists for *dialogue* presented in the diegesis as recorded: does it refer to the time of its production or to the time at which we are hearing it? Imagine a scene in a film where a man is listening to a taped interview. If the sound being listened to has technical qualities of directness and presence, it refers back to the circumstances of its original state. If it has aural qualities that highlight its "recordedness," and if there is emphasis on the acoustic properties of the place where it is being listened to in the diegesis, we tend to focus on the moment where the recording is being heard. In *The Passenger* there is a sequence where Jack Nicholson listens to the recording of a conversation he had with a man he met by chance. Antonioni shuttles from one position to the other and in this way leads into a flashback. The interview Nicholson is listening to becomes real, becomes the scene of the interview itself.



So our tripartite circle becomes more complicated, but also richer. Through the very exceptions we introduce, it continues to illustrate the dimensions and oppositions involved:

- the opposition between acousmatic and visualized,
- the opposition between objective and subjective or real and imagined,
- the differences between past, present, and future.

It is important to think of the circle as consisting of interlocking sectors. In fact, this would probably be expressed much better by a topological model in three dimensions. We also return to the question of the *source*, which conditions such distinctions. First, the idea of sound source must be relativized and "unpacked," since "the" sound source is usually a multifaceted phenomenon. And second, the making and even the conception of a film and its screenplay are likely to emphasize one of these facets to a greater or lesser degree.

Place of the Sound, Place of the Source

Spatially speaking, a sound and its source are two different entities. In a film the emphasis may fall on one or the other, and the onscreen-offscreen question will pose itself differently, according to which thing—the sound or its cause—the spectator reads as being "in" the image or "outside" it. For sound and cause, though quite distinct, are almost always confused. But surely this confusion is inscribed also at the very heart of our experience itself, like an unsettling knot of problems.

For example, the sound of a shoe's heel striking the floor of a reverberant room has a very particular source. But as sound, as an agglomerate of many reflections on different surfaces, it can fill a big a volume as the room in which it resonates. In fact, no matter how precisely a sound's source can be identified, the sound in itself is by definition a phenomenon that tends to spread out, like a gas, into whatever available space there is.

In the case of ambient sounds, which are often the product of multiple specific and local sources (a brook, bird songs), what is important is the space inhabited and defined by the sound, more than its multisource origin. The same goes for films of musical performances. Depending on choices in the editing and technical directing of sound and image, the emphasis can fall either on the specific material source of the sound (the instrument, the singer) or on the sound as it fills the auditory space, considered independently from the source.

The more reverberant the sound, the more it tends to express the space that contains it. The deader it is, the more it tends to refer to its material source. The voice represents a special case. In a film, when the voice is heard in sound closeup without reverb, it is likely to be at once the voice the spectator internalizes as his or her own and the voice that takes total possession of the diegetic

space. It is both completely internal and invading the entire universe. This is what I have called the *I-Voice*.¹¹ Of course the voice owes this special status to the fact that it is the original, definitive sound that both fills us and comes from us.

In the play of onscreen and offscreen space, background music also stands out as a type of exception that proves the overall rule.

THE EXCEPTION OF MUSIC

I have given the name *pit music* to music that accompanies the image from a nondiegetic position, outside the space and time of the action. The term refers to the 'classical opera's orchestra pit. I shall refer as *screen music*, on the other hand, to music arising from a source located directly or indirectly in the space and time of the action, even if this source is a radio or an offscreen musician.

These ideas were developed in chapters on music in *Le Son au cinéma*. They correspond to a distinction that has long been noted, with a variety of names. Some say nondiegetic for the first and diegetic for the second, or commentative and actual, or objective and subjective. For music I prefer to rely on terms that simply designate the place where each (supposedly) comes from. A music cue inscribed in the action can of course be just as "commentative" as a nondiegetic music cue, as in Siodmak's *Abschied*, where the protagonists' neighbor is a pianist whose music accompanies and punctuates their emotional states. *Rear Window* conclusively demonstrates this as well.

Once this distinction is established it is relatively simple to describe ambiguous or mixed cases. Consider the case of screen music framed by a pit music cue with ampler orchestration: someone plays a piano in the action, to the accompaniment of the pit orchestra. This occurs in many musicals; an example that comes to mind is in Raoul Walsh's *The King and Four Queens*. In

another kind of case music begins as screen music and continues as pit music by separating from the action. Or, inversely, a grand pit music cue can narrow into screen music being played by an instrument onscreen, for example in older movies, when opening credit music segues into the start of the action.

Not to mention the numerous cases in current films where music established as on-the-air freely circulates between the two levels. In *Taxi Driver* Bernard Herrmann's main theme, heard as pit music throughout much of the film, crops up as the music on a phonograph to which the pimp (Harvey Keitel) and his young hooker (Jodie Foster) dance.

Music as Spatiotemporal Turntable

All music in a film, especially pit music, can function like the spatiotemporal equivalent of a railroad switch. This is to say that music enjoys the status of being a little freer of barriers of time and space than the other sound and visual elements. The latter are obliged to remain clearly defined in their relation to the diegetic space and to a linear and chronological notion of time.

Another way to put it is that music is cinema's *passe-muraille*, capable of instantly communicating with the other elements of the action.¹² For example, it can accompany from the nondiegetic realm a character who is onscreen. Music can swing over from pit to screen at a moment's notice, without in the least throwing into question the integrity of the diegesis, as a voiceover intervening in the action would. No other auditory element can claim this privilege. Out of time and out of space, music communicates with all times and all spaces of a film, even as it leaves them to their separate and distinct existences.

Music can aid characters in crossing great distances and long stretches of time almost instantaneously. This use of music is fair-

ly frequent, ever since the beginning of sound. In King Vidor's *Hallelujah* protagonist Zeke moves through several locales during the singing of one spiritual, "Going Home": a boat on the Mississippi, the roof of a train, a prairie. We can recognize here the embryonic structure of the music video, which, governed by musical form (its only constraint being to include points of synchronization here and there to solder the music and image together) allows the image to wander at will through time and space. In the music video there really no longer exists an audiovisual scene anchored in coherent time and space.

In Vidor's film music gives the characters winged feet; it functions to contract both space and time. In general, however, we can say that music makes space and time pliable, subject to contraction *or* distention. In suspense scenes, it is music that makes us accept the convention of a frozen moment, eternalized by editing.

And in the long confrontations in Sergio Leone's films, where characters do little but pose like statues staring at each other, Ennio Morricone's music is crucial in creating the sense of temporal immobilization. True, Leone also tried to stretch time without the help of music. Notably, at the opening of *Once Upon a Time in the West*, he made do with the occasional creaking of a weather vane or a noria. But there, the plot situation—a long period of waiting and inaction—was chosen to justify the immobility of the characters. At any rate, Leone developed this sort of epic immobility with reference to opera and by generally using music overtly on the soundtrack.

RELATIVE OFFSCREEN SPACE AND ABSOLUTE OFFSCREEN SPACE

The term *offscreen sound* is deceptive; it might lead us to think that the sound itself has some intrinsic quality. We only have to close

our eyes at a film or look away from the screen to register the obvious: without vision, offscreen sounds are just as present—at least as well-defined acoustically speaking—as onscreen sounds. Nothing allows us to tell the two apart. Acousmatized and reduced to an ensemble of sounds that certainly constitute a soundtrack worthy of the name, the film completely changes. I already cited the example of certain scenes of *Mr Hulot's Holiday*: listen to the sounds without the image and they reveal a different character.

Thus sound's "offscreenness," in monaural cinema, is entirely a product of the combination of the visual and aural. It is really a *relation* of»what one hears to what one sees, and exists only in this relation; consequently it requires the simultaneous presence of both elements.

Without the image, the sound of numerous great films of the past is meaningless. In particular, the magical voices that fascinated us would atrophy or become prosaic. The voices of Norman's mother in *Psycho*, Dr. Mabuse in *The Testament of Dr. Mabuse*, or Marguerite Duras in *L'Homme Atlantique* would no longer be extraordinary if they ceased to interact with a screen where they encountered the void of their presence.

Multitrack Cinema's "In-the-Wings Effect" and "Offscreen Trash"

Characteristic of real spatialization and of early multitrack sound film experiments, and generally avoided since then, the "in-the-wings effect" is produced whenever a sound linked to a cause likely to appear onscreen, or which has just exited, lingers in one of the offscreen loudspeakers to one side. Examples are the footsteps of a character approaching or leaving, the engine of a car that has just gone offscreen or that is about to appear, or the voice of a protagonist just out of view.

At these times we have the feeling, which is disconcerting to our normal sense of spectatorship, that we're being encouraged to believe that the audiovisual space is literally being extended into the theater beyond the borders of the screen, and that, over the exit sign or above the door to the restrooms, the characters or cars are *there*, preparing their entrance or completing their exit.

Sometimes this in-the-wings effect cannot be attributed to the direction and mixing of the film, but is simply created by an aberrant placement of speakers in the theater. Sometimes it is indeed due to an attempt by the sound engineers or the director to exploit the effect of *absolute offscreen space*, an effect made possible by multitrack.

Slowly, this practice has been dropped. Sounds of entrances and exits are now rendered with greater discretion and subtlety, or they are opportunely drowned in the sound mix (numerous ambient sounds, music) so as to avoid the sense of the nearby off-stage wings.

Certainly, the in-the-wings effect created a nagging problem by violating the conventions of continuity editing and making sound matching problematic. But maybe it could have gained more permanent admittance into film practice had it been systematized along with some partial adjustments in editing conventions—just as the superfield of the multitrack cinema was able to strike a compromise with traditional editing. So perhaps it was a mistake to have given it up so quickly.

The *offscreen trash* is a particular case of passive offscreen space (see below) that results from multitrack sound. It is created when the loudspeakers outside the visual field "collect" noises—whistles, thuds, explosions, crashes—which are the product of a catastrophe or a fall at the center of the image. Action and stunt movies often draw on this effect. Sometimes poetic, sometimes intentionally comic, the "offscreen trash" momentarily gives an

almost physical existence to objects at the very moment they are dying. A modern action movie like John McTiernan's *Die Hard*, a veritable feast of glass-breaking and deflagration taking place in a tower where a man fights terrorists, is filled with such effects.

Active and Passive Offscreen Sound

I shall give the name *active offscreen sound* to acousmatic sound that raises questions—What is this? What is happening?—whose answer lies offscreen and which incite the look to go there and find out. Such sound creates a curiosity that propels the film forward, and it engages the spectator's anticipation: "I'd like to see his face when the other character says that to him." The sounds in active offscreen space necessarily issue from objects that could be identified by sight. Active offscreen sound is used frequently in traditional sound-image editing, bringing objects and characters into a scene by means of sound, then showing them. Films like *Psycho* are based entirely on the curiosity aroused by active offscreen sound: this mother we keep hearing, what does she look like?

Passive offscreen sound, on the other hand, is sound which creates an atmosphere that envelops and stabilizes the image, without in any way inspiring us to look elsewhere or to anticipate seeing its source. Passive offscreen space does not contribute to the dynamics of editing and scene construction—rather the opposite, since it provides the ear a *stable place* (the general mix of a city's sounds), which permits the editing to move around even more freely in space, to include more close shots, and so on, without disorienting the spectator in space. The principal sounds in passive offscreen space are *territory sounds* and *elements of auditory setting*.

Dolby multitrack has naturally favored the development of passive offscreen space over active. Why? The answer may be in the fact that active offscreen space mobilizes identifiable, single

sources—a human body, an object—and in multitrack sound, real (no longer mental) localization of offscreen sound poses the problem of the too realistic "in-the-wings" effect I have already mentioned. If you want to avoid this effect, it is hardly advisable to employ insistent offscreen sounds that pose enigmas and demand to be de-acousmatized, for, logically speaking, this sound should be situated outside the field of the screen. The entrance of Roy Batty, the antagonist in *Blade Runner*, would have been done by the sound of his voice or his footsteps if the film had been recorded in mono. In the actual film this character is almost always present in the image at the same time as his voice. It is as if we were in a perpetual present. In the traditional monaural cinema, on the other hand, offscreen sound demands its resolution from the center of the image, from the very heart of the image, and thus can be called active.

But as early as 1954 *Rear Window* included much passive offscreen sound: city noise, apartment courtyard sounds, and radio, which, full of reverb, cued the ear into the contextual setting of the scene without raising questions or calling for the visualization of their sources.¹³

EXTENSION

Recall the fixed images, like photographs, in Bergman's *Persona*: shots of the park, a hospital wall, and a pile of dirty snow. Over these shots we heard churchbells and no human sound; these created the impression of a small slumbering village.

Let us take away Bergman's sounds and replace them with, say, the sound of the ocean. We see the same pile of snow, the same grillwork, but offscreen space takes on a salty sea smell. If we now remove the ocean sound and instead dub in a crowd of voices and footsteps, the offscreen space becomes a busy street. Nothing pre-

vents us from taking these same images and beginning with a nearby sound (e.g., footsteps in the snow), then bringing in other sounds that suggest a larger space—car sirens—and so on: someone walks by, the siren passes and fades into the distance, faraway churchbells begin to ring. On one static long take we can thus infinitely dilate the offscreen space imagined and evoked by the soundtrack. And we can shrink it just as easily, in which case we will retain the memory of the vast space evoked at the beginning.

Extension of the the sound environment is our designation for the degree of openness and breadth of the concrete space suggested by sounds, beyond the borders of the visual field, and also within the visual field around the characters.

We can speak of *null extension* when the sonic universe has shrunk to the sounds heard by one single character, possibly including any inner voices he or she hears. At the other end of the spectrum we might call *vast extension* the arrangement wherein, for example, for a scene taking place in a room, we not only hear the sounds in the room (including those offscreen) but also sounds out in the hallway, traffic in the street nearby, a siren farther away, and so on. Ambient extension has no absolute limit except those of the universe—if, of course, sounds could ever be found that were capable of maximally dilating the perception of space surrounding the action.

Obviously what is interesting in the cinema is not only the extension that remains the same throughout a scene and even throughout a film but also contrasts and variations in extension from one scene to another, or even within one and the same scene. Sound designer Walter Murch alludes to variation in extension (without using this term) in describing his work on Coppola's *The Conversation* and *Apocalypse Now*.¹⁴

Dolby stereo, having dramatically increased the possibilities of layering sounds and deploying them in wide concentric spaces,

encourages experimentation with extension. Almost forty years ago *Rear Window*—where everything is seen from a flat in a Greenwich Village courtyard apartment house—made magnificent use of variations in extension. Sometimes it lets us hear the big city thrumming outside this courtyard that the film never leaves. At other times the soundtrack eliminates the larger cityscape entirely, so as to reconcentrate the spectator on the apartment itself, which then becomes for our couple, Grace Kelly and James Stewart, a theater stage cut off from its surroundings. At the very end of the film, the extension becomes extremely narrow, focussing on a single point, like a lone spotlight pursuing a character on a stage—the footsteps' of the killer in the stairway, which Stewart can hear approaching ...

The final scene of *Children of a Lesser God* achieves a similar tightening of spatial extension. As the two estranged lovers reunite in the cool night air, we perceive more and more faintly the noise of a disco dance going on nearby; then it fades down entirely.

Although variations in extension can also consist of sudden contrasts between one scene and the next, they are generally executed in such a way as not to be noticed as a technical manipulation. The occasions on which they *are* made obvious usually contribute toward some emotional effect. This is not like reframings, for example, which are tolerated as technical and coded.

Some films adopt a single fixed strategy for spatial extension and maintain it throughout. In Lang's *M* extension is generally quite limited. All we hear during a conversation scene is what the characters onscreen are saying; almost never do we hear ambient sounds outside the frame. On the other hand, certain modern films adopt a consistently vast extension: think for example of *Blade Runner*, where rumblings of the city behind characters in the frame constantly remind the viewer of the presence of a huge spatial context.

In fact one of the trickiest things in Dolby stereo films is narrowing extension down to one sound or one point in space, since this necessitates silencing several loudspeakers. The final effect of *Rear Window*, for example, would be very difficult in multitrack sound.

Varying extension to the point of absolute silence is of course used for achieving effects of subjective sound. The suppression of ambient sounds can create the sense that we are entering into the mind of a character absorbed by her or his personal story. A good example occurs in the scene in Bob Fosse's *All That Jazz* where the protagonist has a heart attack.

POINT OF AUDITION

Spatial and Subjective Point of Audition

The notion of a point of audition is a particularly tricky and ambiguous one. Several scholars (Francois Jost in particular) have approached this subject, and I myself devoted a chapter to it in *Le Son au cinéma*—where, to tell the truth, I raised more questions than I provided answers. It might be useful to return to it here with greater precision.

Let us first note that critics have come up with the concept of a point of audition based on the model of point of view. Here begins the problem, since cinematic point of view can refer to two different things, not always related:

1. The place from which I the spectator see; from what spatial location the scene is presented—from above, from below, from the ceiling, from inside a refrigerator. This is the strictly *spatial* designation of the term.
2. Which character in the story is (apparently) seeing what I see. This is the *subjective* designation.

In most shots of a modern-day film the camera's point of view is not that of a specific character. Which does not mean that it is necessarily arbitrary: it tends to obey certain laws and constraints. For example, the camera will rarely be located where the eye of a normal human character couldn't be (on the ceiling, in a closet, etc.). Or it only shoots along certain privileged axes, excluding others (e.g., Bergman's *After the Rehearsal*, which takes place on a theater stage, excludes the fourth side, which is the auditorium, the theater seats).

The notion of point of view in this first spatial sense rests on the possibility of inferring fairly precisely the position of an "eye" based on the image's composition and perspective.

Let us recall too that point of view in the subjective sense may be a pure effect of editing. If I cut from a shot of a character looking out the window to a shot of an exterior scene, it is highly likely that the second shot will be perceived as the character's point of view, as long as the information in shot B doesn't contradict anything in shot A.

Now, by comparison, let us examine the notion of a point of audition. This too can have two meanings, not necessarily related:

1. A spatial sense: from where do I hear, from what point in the space represented on the screen or on the soundtrack?
2. A subjective sense: which character, at a given moment of the story, is (apparently) hearing what I hear?

In the first definition, we should start by noting that the specific nature of aural perception prevents us, in most cases, from inferring a point of audition in space based on one or more sounds. This is because of the omnidirectional nature of sound (which, unlike light, travels in many directions) and also of listening (which picks up sounds in the round), as well as of phenomena involving sound reflection.

Consider a violinist playing in the center of a large round room, her audience grouped in various places against the wall. Most of the listeners, even those standing at diametrically opposite points of the room, will hear roughly the same sound, with slight differences in reverberation. These differences, related to the acoustics of the space, are not sufficient to locate specific points of audition. Every *view* of the violinist, on the other hand, can immediately situate the point from which she is being looked at.

So it is not often possible to speak of a point of audition in the sense of a precise position in space, but rather of a place of audition, or even a zone of audition.

In the second, subjective sense of point of audition, we find the same phenomenon as that which operates for vision. It is the *visual* representation of a character in closeup that, in simultaneous association with the hearing of sound, identifies this sound as being heard by the character shown.¹⁵

The classic example of audiovisual counterpoint cited in Eisenstein's manifesto—the image of a man on lookout duty, and the creaking of a character's boots offscreen—is of the type that is commonplace today. The question is not what characteristics of distance, color, and reverberation *in terms of sound* allow us to infer that the sound is heard by character X. For it is the image that always creates the point of audition, which in this case is worthy of the term *point*.

A special case of point of audition is one defined by sounds that "don't carry," supposedly of such a nature that one must be right up close in order to hear them. Upon hearing these sounds or indices of proximity (e.g., breathing in a voice), the spectator can locate the point of audition as that of a character in the scene—provided of course that the image, the editing and the acting all confirm the spectator's hunch. Phone conversations are the most common example. When the spectator hears the voice of the

unseen person speaking clearly in sound closeup, with its characteristic filtering, she or he can identify the point of audition as being that of the character seen receiving the call. Unless, of course, we are in a situation of on-the-air, which unhooks the sound from its point of departure or arrival and accordingly renders the notion of point of audition no longer pertinent.

Frontal Voice, Back Voice

In some special cases it is nevertheless possible to attribute a direction to what is heard. A sound's high frequencies actually travel in a more directional manner than the low; and when someone speaks to us with his back turned we perceive fewer of the voice's high harmonics and find the voice less present. We can therefore speak of an audible difference between the *frontal voice* and the *back voice*.

In certain films shot in direct sound we can hear variations in a voice's color, due to the fact that a character turns away now and then from the microphone, which is generally above his head. These fluctuations in tone color help to give a particular kind of life to direct sound, and they also function as "materializing indices" (see chapter 5).

Note, however, that, first, there is no law against simulating or reconstituting such variations during postsynchronization, by moving the actor or the mike. (See for example the postsynch of *L'Homme blesse* by Patrice Chereau.) Second, conversely, the mike during shooting can be arranged so as to follow the actor constantly "in front," particularly when the actor wears a lavalier mike.

If the cinema usually employs the frontal voice, with the most treble allowed by the equipment, it is for a reason: high frequencies are crucial for intelligibility.

However, when the spectator hears a back voice, he or she cannot automatically infer the shot's point of audition from this: for one thing because in most cases this is a momentary effect, not stable and pronounced enough. For another, the spectator does not associate the point of audition with any mental representation of a microphone.

Blind Spots in Theorizing the Mike/Ear

This important question of the "scotomization" of the role of the microphone applies not only to the voice but also more generally to all sounds in a film.¹⁶ And not only to the cinema but equally to most radiophonic, musical, and audiovisual creations that rely on sound recording. The camera, though excluded from the visual field, is nonetheless an active character in films, a character the spectator is aware of; but the mike must remain excluded not only from the visual and auditory field (microphone noises, etc.) but also from the spectator's very *mental representation*. It remains excluded, of course, because everything in movies, including films shot in direct sound, has been designed to this end. This naturalist perspective remains attached to sound, but it is a perspective from which the image—60s and 70s theories on the "transparency" of mise-en-scene notwithstanding—has long been liberated. The naturalist conception of sound continues to infuse real experience and critical discourse so completely that it has remained unnoticed by those who have referred to it and critiqued this same transparency on the level of the image.

We might locate the reasons for such a difference of status between image and sound, in different technical, aesthetic, physiological, and ideological problems, by asking which ones serve as alibis or coverups for which others. We must, for example, explore the implications of the fact that the ears are not disposed

directionally like the eyes. Or the technical possibility, unexplored in the image but utilized in soundtracks since the coming of sound, of *mixing* sounds recorded simultaneously by several mikes placed at different locations: what becomes of the mike/ear then?

But perhaps this is beside the point. After all, the camera has little to do with our eyes (to begin with, it is monocular), which hasn't prevented it from becoming the agent of the look. So the problem lies rather in ways of thinking. To disengage sound thinking and its technical and aesthetic applications from its naturalist rut might well take many years. A concern which lies at the heart of our project.

FIVE

THE REAL AND THE RENDERED

THE ILLUSION OF UNITY

A common perspective to which we made reference in the preceding chapter, which might be called naturalist, postulates that sounds and images start out in "natural harmony." Proponents of this approach seem surprised not to find it working in the cinema; they attribute the lack of this natural audiovisual harmony to technical falsifications in the filmmaking process. If people would only use the sounds recorded during shooting, without trying to improve on them, the argument goes, this unity could be found.

Such is of course rarely the case in reality. Even with so-called

direct sound, sounds recorded during filming have always been enriched by later addition of sound effects, room tone, and other sounds. Sounds are also eliminated during the very shooting process by virtue of placement and directionality of microphones, soundproofing, and so on. In other words, the processed food of location sound is most often skimmed of certain substances and enriched with others. Can we hear a great ecological cry—"give us organic sound without additives"?

Occasionally filmmakers have tried this, like Straub in *Trop tot trop tar A}*. The result is totally strange. Is this because the spectator isn't accustomed to it? Surely. But also because reality is one thing, and its transposition into audiovisual two-dimensionality (a flat image and usually a monaural soundtrack), which involves radical sensory reduction, is another. What's amazing is that it works at all in this form. Indeed, we tend to forget that the audiovisual tableau of reality the cinema furnishes us, however refined it may seem, remains strictly (on the level of reproduction) that which a sketched representation of a human, with a circle for the head and sticks for the arms and legs, is to an anatomical drawing by Albrecht Dürer. There is really no reason for audiovisual relationships thus transposed to appear the same to us as they are in reality, and especially for the original sound to ring true.

We might go far as to say that all the conventions of rendering, sound effects, and so forth, which we shall examine further on, consist of accommodations and adjustments, taking into account the audiovisual transposition in order to try to conserve a certain sense of realism and truth in their new representational context.

This does not mean that it is wrong to aspire to a better simulacrum. Quite to the contrary, an experiment like Douglas Trumbull's *Showscan*, which not only uses high-definition 70-millimeter film but also considerably improves on frames per second—sixty rather than the usual flickering of twenty-four—such

experiments should be taken seriously. Trompe-l'oeil, as well as "trompe-roreille," is a worthy art (even though the notion of trompe-l'oeil does not really involve the exact reproducing of an impression in the same way).

But when we say in disappointment that "the sound and image don't go together well," we should not blame it exclusively on the inferior quality of the reproduction of reality. For this situation merely echoes a phenomenon we are generally blind to. In concrete experience itself, independent from cinema, they sometimes don't go together either.

The most familiar example is the "mismatch" of an individual's voice and face when we have had the experience of getting to know one of them well before discovering the other. We never fail to be surprised, even shocked, when we complete the picture. Consider also the children's books that teach the noises animals make: as if there were the slightest connection, aside from the connection created by purely Pavlovian training, between the sound a duck makes and what it looks like, or the onomatopoeic words for the duck's call in different languages.

Basically, this question of the unity of sound and image would have no importance if it didn't turn out, through numerous films and numerous theories, to be the very signifier of the question of human unity, cinematic unity, unity itself. The proof is dualistic films based on a carefully planned de-acousmatization . . . which is often eluded at the last minute.² It is not I but the cinema that, via films like *Psycho* and *India Song*, tells us the impossible and desired meeting of sound and image can be an important thing.

Strangely, the disjunctive and autonomist impulse that predominates in intellectual discourse on the question ("wouldn't it be better if sound and image were independent?") arises entirely from the unitary illusion we have described: the false unity this

thinking denounces in the current cinema implicitly suggests a true unity existing elsewhere.

Also, disjunctive ideology usually takes as given the technical fidelity of the recording. Fidelity, of course, is a problematic notion from the start.

QUESTIONS OF SOUND REPRODUCTION

A sound recording's *definition*, in technical terms, is its acuity and precision in rendering of detail. Definition is a function of the width of the frequency band (which allows us to hear frequencies all the way from extreme low to extreme high) as well as its dynamic range (amplitude of contrasts, from the weakest levels to the strongest). It is particularly through gains in high frequencies that sound has progressed in definition; high frequencies reveal a new multitude of details and information, contributing to an effect of greater presence and realism.

I am speaking of *definition* (a precise and quantifiable technical property, just like definition or sharpness in a photographic or video image) and not of *fidelity*. The latter is a tricky term; strictly speaking it would require making a continuous close comparison between the original and its reproduction, which normally would be quite difficult to physically arrange. Someone who listens to an orchestra on a sound system in his living room is not likely to be able to compare it with some orchestra playing at his doorstep. It should be known, in fact, that the notion of high fidelity is a purely commercial one, and corresponds to nothing precise or verifiable.

However, it happens that today *definition is* (mistakenly) *taken as proof of fidelity*, when it's not being confused with fidelity itself.

In the "natural" world sounds have many high frequencies that so-called hi-fi recordings do capture and reproduce better than

they used to. On the other hand, current practice dictates that a sound recording should have more treble than would be heard in the real situation (for example when it's the voice of a person at some distance with back turned). No one complains of nonfidelity from too much definition! This proves that it's definition that counts for sound, and its hyperreal effect, which has little to do with the experience of direct audition. For the sake of rigor, therefore, we must speak of high definition and not high fidelity.

In the cinema sound definition is an important means of expression with multiple consequences. First, a more defined sound, containing more information, is able to provide more *materializing indices*. And second, it lends itself to a more lively, spasmodic, rapid, alert mode of listening, particularly to agile phenomena that occur in the higher frequencies (e.g., a feeling of temporal acceleration, very distinct in recent films).

Isolation and Disconnection of Sound Properties: The Example of THX

Everyone over the age of thirty today has intense memories of certain film viewing experiences; for me it was our weekly screenings at boarding school in the small town of Creil. Very eclectic programming—from American and Russian war films to low-budget westerns, with Italian neorealist films thrown in between—characterized these 16-millimeter screenings, which took place in the high school's assembly hall. Two strong aural memories stay with me. One is the wavering sound—especially noticeable during musical passages—caused by the projector's uneven speed. The other is a cavernous resonance, owing to the poor quality of sound reproduction, but also because of the reverberation that the room's acoustics gave to the actors' voices. These conditions, which we might call caricatural, were hardly more

than a slight exaggeration of the conditions under which most movies were seen in those days.

Whoever goes to a modern theater graced with the label THX (created by George Lucas) will find the exact opposite: stable sound, extremely well defined in high frequencies, powerful in volume, with superb dynamic contrasts, and also, despite its strength and the probably large theater space, a sound that does not seem very reverberant at all. One finds in THX theaters the realization of the modern ideal of a great "dry" strength.

New movie houses whose acoustics are conceived or overhauled with luxury sound projection in mind have indeed mercilessly vanquished reverb through the choice of building materials and architectural planning. The result is that the sound feels very present and very neutral, but suddenly one no longer has the feeling of the real dimensions of the room, no matter how big it is. So what results is the enlargement, without any modification in tone, of a good home stereo sound.

At the start of each show some of these movie theaters play a short that consists of a title reminding you that "you are in a THX theater/" during which you hear an electronic sound effect for about thirty seconds: a bunch of glissandi falling toward the low bass register, spiraling spatially around the room from speaker to speaker, ending triumphantly on an enormous chord. And it's all at an overwhelming volume that leads the audience instinctively to react by applauding in a sort of physical release.

We may note two characteristics of this sound demo that typify current taste. First, the bass sound that the glissando ends on is clean of all distortion and secondary vibrations, even though very low sounds in the real world have the necessary consequence of causing small objects to vibrate—for example, a passing semi truck sets the furniture or the dishes to shaking. What the demo short is doing to stir the audience's admiration, far from any idea of fidelity, is showing off the technical capacity to isolate and

purify the sound ingredients. Second, one finds no trace in the demo of the reverberation that normally accompanies and muddles loud sounds in an enclosed space.

In "real life" audio characteristics always vary in association with each other: if the volume of a sound event increases, the sound changes nature, color, resonance. In the sonic world before electronic amplification the presence of reverb prolonging the sound marked a change in spatial properties, just as the presence of secondary vibrations from the principal sound signaled a change to greater intensity. Here, on the contrary, volume aside (volume **too* has become a sound property as isolated or independent as others), the sound event remains as clear and distinct as if we heard it on the small speaker of a compact home stereo. So that in the type of sound projection preferred today in movie houses, where the real size of the auditorium is immaterial, amplification no longer has a true scale of reference. Amplified sound remains the same at all volume levels, without any traits that might mark the crossing of a threshold. The big THX theaters no longer give us collective sound in the old style; it's inflated personal stereo sound.

Phonogeny and Technical Mediation

The sound media (recording, talking pictures, radio) of the twenties through forties subscribed to a certain notion we have virtually forgotten today: phonogeny. Phonogeny refers to the rather mysterious propensity of certain voices to sound good when recorded and played over loudspeakers, to inscribe themselves in the record grooves better than other voices, in short to make up for the absence of the sound's real source by means of another kind of presence specific to the medium.

This notion rose to great popularity when sound engineers of the first talkies, men who came from the recording and radio

102 • • • The Audiovisual Contract

industries (from where else could they have come?), attempted to valorize phonogenic criteria, decreeing that actor X had a terrific voice while actor Y was deplorably unphonogenic. Pagnol recounts one of the best-known examples of these dangerous verdicts: during the shooting of *Marius* a Western Electric engineer declared Raimu's voice impossible to record! Not that sound engineers were always mistaken in this regard; many actors of the era were afraid they might not have the accursed phonogeny it took to survive in the age of the microphone.

Obviously this notion arose from the technical conditions that prevailed at the time: the equipment was less sensitive and precise than that of today. So certain voices did have timbres favored by the technology. They articulated themselves clearly through the microphone's filter; in a word, they "meshed" well with the sensitive part of the system.

In retrospect we might say that the voices of a Gerard Depardieu or Catherine Deneuve wouldn't have been judged phonogenic enough according to the criteria of those days: not clearly pronounced enough or sonorous enough.³ (The verdict on Raimu, someone who surely didn't lack resonance, doubtless came from the fact that he had *too much*.) In this sense, the idea of phonogeny was nothing other than the idea of the adaptability of a certain type of voice and diction to the technical conditions of recording and reproduction. Which was by no means absurd. But even then the term also carried a strongly irrational charge. One said of a voice that it was phonogenic the way we say someone is cool or sexy, referring to some ineffable impact on people in terms of communication or seductiveness.

The criterion of phonogeny emerged, of course, as an analogy of photogeny, which was strongly operative during the era of movie stars. Unlike the former, the latter has survived to our day, and it is not uncommon to see filmmakers declare of a woman

that she is not beautiful but incredibly photogenic. But the equivalent notion for sounds and the voice in particular has completely died out.

So it all seems as if we're implicitly convinced that the means of sound collection and reproduction has become transparent, rendering unnecessary any advance screening for "recordability" of a sound from the acoustic event to playback. In this, of course, we are mistaken. The most highly perfected digital recordings are certainly quantitatively richer in detail than those of yesteryear, but they are no less colored by the technical process—perhaps even more so. We simply need ten or twenty years to elapse before we can perceive this.

People who once talked about phonogeny (even if the aim of their discourse was to blindly apply the notion and issue verdicts that history would later prove wrong) were therefore actually more conscious than we are: they understood that the sound heard at the end of the process is the product of a preexisting reality plus conditions of reproduction. This end product is a specific reality: neither the neutral transmission of a sound event, nor an entire fabrication by technical means.

This leads us to wonder what the disappearance of the notion of phonogeny is the symptom of. Perhaps it signals an important mutation, to our total everyday immersion in *mediated acoustical reality* (sound is relayed by amplifiers and loudspeakers). The new sound reality has no difficulty supplanting unmediated acoustical reality in strength, presence, and impact, and bit by bit it is becoming the standard form of listening. It's a form of listening that is no longer perceived as a reproduction, as an image (with all this usually implies in terms of loss and distortion of reality), but as a more direct and immediate contact with the event. When an image has more presence than reality it tends to substitute for it, even as it denies its status of image.

Let us imagine that in our everyday visual experience certain things were to appear in windows and on screens with incomparably more vivid color, brightness, and definition than what we perceive directly. Then reality seen with the "naked eye"—with no technical mediation—would become lifeless, ill-defined, distant. This is exactly the impression felt by people at a harpsichord concert who know this instrument only from recordings or radio: "You can't hear a thing!"

Curiously, the more unmediated acoustical reality loses its value as real experience and the less it is the lived standard to which we compare what we hear, the more it becomes the abstract reference we call on conceptually—for example, regarding the notion of acoustic fidelity that the cinema demands. The more we use recorded and/or transmitted sound, the more we mythify its contrary: a natural acoustical experience that we actually have less and less frequently.

The fading of the idea of phonogeny has another origin as well. Phonogeny still had meaning in an era during which people had learned to speak and to project their voices in a predominantly mechanical environment. Today, the manner in which people speak is just as strongly influenced by voices heard on television, radio, and films as by voices heard "naturally"; in this environment it is difficult to compare a natural voice to a mediated one. For natural voices are not only produced but also heard by unconscious comparison with mediated ones, which tend to stand out more. It is also in light of this new fact of our experience that the notion of phonogeny needs to be rethought—but first, rediscovered.

Silences of Direct Sound

Many people consider *location sound* not only the sole morally acceptable solution in filmmaking but also the one that simplifies

everything, since it eliminates the problem of having to make choices. Rohmer's films are mentioned most often when critics wish to vaunt the virtues of location sound and to present it as a simple, obvious, rigorous, and irresistible option. However, Rohmer's choices for direct sound consciously involve sacrifices and difficulties. The location sound is subjugated to the artist's purposes. In the acoustic climate of *The Aviator's Wife*, for example, shot in Paris on location, we notice a great neutrality, and an almost complete silence of the setting. There is nothing to disturb our concentration on the characters and their lines. All events that usually intrude on sound recording in a city or in the country have been completely eliminated, but the life that can come with them has also been eliminated in the process.

When Philippe Marlaud and then Mathieu Carriere go up the stairs leading to Marie Riviere's little gabled apartment at the beginning of the film, we hear no sound but their footsteps. When Marie Riviere opens her windows we hear an anonymous hum of the city, very general and diffused, through the windows____And when characters speak in her bedroom nothing gets in the way of their lines, even though Lord knows how many intrusions occur in a real apartment in the city.

It seems obvious that Rohmer in his soundtracks sniffs out and chases away any anecdotal sound—whether it's the kind that's recorded accidentally by the location mike, such as beeps and honks, sirens, shouts, or other noises, or whether it's the kind of anecdotal sound usually added intentionally later in the mix, in accordance with certain codes, habits, and cliches. For example, for going up the stairs, Rohmer could have put stairwell noises and radio sounds to suggest the time of day, 7:00 A.M. When Marie Riviere opens her windows, many films set in Paris normally mix in the cooing of pigeons to complete the picture, to flesh out the city and its rooftops, and to enliven the soundtrack

with a conventional decorative touch. Likewise, in Rohmer's film, when Philippe Marlaud walks down the street his footsteps are heard in an indistinct hum of traffic; no specific noise, no burst of conversation comes to break the atmosphere. The risk of a spontaneous intrusion of sound when you film with location sound is that it might give a sudden specific and undesired meaning to a word of dialogue or to an actor's gesture. A motorscooter accelerating offscreen, a radio or TV whose sound comes through the windows might not just drown out the actor's line but inflect it, make it mean something else.

And this is precisely the risk that Straub and Huillet accepted the run of in *Othon*, when a backfiring motorcycle almost grotesquely accentuated a line of Corneille. In general, however, the filmmakers chose shooting locations that were sufficiently removed from traffic, so the traffic for the most part consisted of a continuous generalized hum.

In *The Aviator's Wife*, in order to obtain this well-blended sonic environment with nothing standing out, this environmental silence enveloping the characters (also very noticeable in the long sequence shot in the park at Les Buttes-Chaumont), Rohmer had to shoot at precise hours in the day (which by no means always matched the supposed time of the action), and to reject any takes marred by interruptions. In a word, it was only through a process of choice and elimination that Rohmer could construct the sonic milieu he envisioned. Thus, the notion of direct sound involves no less of a reconstruction (even if by simple subtraction) than the notion of postsynchronized sound.

The most noticeable and anecdotal sound in this film is Marie's sink vibrating when she turns on the faucet. Even this noise was not introduced or kept on the soundtrack for any purely comical or picturesque effect in itself. It has a precise function in the screenplay: a suitor uses the necessary repairs in her apartment as

a pretext to plague this young woman (who will continue to reject him). Besides, the noise does not go unnoticed; Mathieu Carriere refers to it directly. It is thus integrated into the dialogue, digested by the screenplay.

The risk of direct sound and location shooting is that unplanned visuals and sounds that are not integrated in the fabric of the screenplay can intrude, take on an autonomous existence. Some filmmakers love and even provoke such intrusions; others go so far as to simulate them. Still others avoid them: even Godard's work allows no anecdotal intrusion of sound to distract from his aims. Thus, direct sound is not so "open," not so synonymous with "simple and obvious" solutions as it may seem.

Sound Truth and Sound Verisimilitude

A fourth question regarding sound reality, the question of verisimilitude, is a terribly ambiguous and complicated one. Let us merely examine several aspects of it.

First of all, sound that rings true for the spectator and sound that *is* true are two very different things. In order to assess the truth of a sound, we refer much more to codes established by cinema itself, by television, and narrative-representational arts in general, than to our hypothetical lived experience. Besides, quite often we have no personal memory we might refer to regarding a scene we see. If we are watching a war film or a storm at sea, what idea did most of us actually have of sounds of war or the high seas before hearing the sounds in the films?

And in the case of scenes we might experience in everyday life (with Rohmer, for example), hardly ever have we paid distinct and focused attention to these sounds either. We only retain impressions of such sounds if they carry material and emotional significance; those that don't interest or surprise us get eliminat-

ed from memory. Everyday reality hardly puts us in a position to listen to its sounds for themselves and to focus on their intrinsic acoustical qualities, since context so strongly influences our perception. In order to really listen to sounds one would have to recreate the auditory analogy of the visual *camera obscura* that allowed painters to observe nature and to apprehend its true values of light and color.

The codes of theater, television, and cinema have created very strong conventions, determined by a concern for the *rendering* more than for literal truth. We are all thoroughly familiar with these conventions, and they easily override our own experience and substitute for it, becoming our reference for reality itself.

For one thing, film as a recording art has developed specific codes of realism that are related to its own technical nature. Of two war reports that come back from a very real war, the one in which the image is shaky and rough, with uneven focus and other "mistakes," will seem more true than the one with impeccable framing, perfect visibility, and imperceptible grain. In much the same way for sound, the impression of realism is often tied to a feeling of discomfort, of an uneven signal, of interference and microphone noise, etc. These effects can of course be simulated in the studio during postproduction and orchestrated: *Alien*, for example, uses effects of acoustic discomfort to heighten its sense of realism.

For another thing, when the spectator hears a so-called realistic sound, he is not in a position to compare it with the real sound he might hear if he were standing in that actual place. Rather, in order to judge its "truth," the spectator refers to his memory of this type of sound, a memory resynthesized from data that are not solely acoustical, and that is itself influenced by films.

Naturally criteria for auditory verisimilitude differ according to the specific competences and experience of the individual. A

nature lover wrote a letter to the editor of *Telerama* (France's *TV Guide*), having seen Bertrand Blier's *Trop belle pour toi*. He expressed shock in hearing birds that in his experience cannot sing either in the season in which the story takes place (winter) nor in the location shown (the area of Beziers). This evidence exposed the fabricated nature of the soundtrack and its trumped-up sound effects, and prevented him from "believing" the scene.

Let us note that this demand for realism has its blind spots and limitations. For one thing, it is altogether conceivable that through a fluke of nature the birds in question really were singing during the shooting. For another, the same spectator who's finicky about sounds might be indifferent to aberrant light (incoherent lighting setups given the light sources depicted) that might disturb the photography specialist. In other words, every film is premised on the acceptance of rules of the game—not the least of which is agreeing to see flat images in depth!

RENDERING AND REPRODUCTION

What Is a Rendering?

In considering the realist and narrative function of diegetic sounds (voices, music, noise), we must distinguish between the notions of *rendering* and *reproduction*. The film spectator recognizes sounds to be truthful, effective, and "fitting not so much if they reproduce what would be heard in the same situation in reality, but if they render (convey, express) the feelings associated with the situation. This occurs at a barely conscious level, for filmviewers (in which we must include most critics and theoreticians) have little more than a fairly crude and immediate understanding of the cinema's figurative nature.

Leonardo da Vinci made a remark in his *Notebooks* that synthesizes the problem quite well: "If a man jumps on the points of his

feet, his weight does not make any sound."⁴ Here we have the creator of the Mona Lisa stating with wonderment that sound does not render a person's weight, as if it had a special vocation to do so. In other words; the assumption is that a sound should constitute a microcosm of the whole event, with the same characteristics of speed, matter, and expression. Everyone continues to have this same expectation of sound, centuries after Leonardo, despite the recordings of acoustical reality that can be made today, which should disabuse us of our misconceptions.

In truth the question is a complex one, even at the very level of language. Consider a scene in Truffaut's *The Bride Wore Black*. Claude Rich plays a recording for his friend Jean-Claude Brialy in which we can hear a subtle sound of some kind of friction, unidentifiable and periodic. It leaves Brialy perplexed—he can't tell what it is. Rich then identifies the sound as that of a woman's stockings as she crosses her legs. He specifies that it was recorded without the knowledge of the woman in question. He adds that the lady was wearing nylons: "I tried it with silk stockings, but that didn't give a good rendering at all." What does he mean by "rendering," this character presented to us as a ladies' man?

If I understand correctly, in playing his recording the character was not so interested in getting his friend to identify the real source. If he were, he could have said, "But you couldn't tell that it was the sound of stockings." Rather, he wished to convey an effect or feeling associated with the sound source—an effect of sensuality, eroticism, intimacy, contact. This is why the nylon stockings, even though a more common material, proved more to his taste than silk stockings in the rendering of the recording.

The pleasure-seeker played by Claude Rich conducted an experiment demonstrating that a sound doesn't necessarily "render" its source: silk does not make a noise that self-evidently relates the sensuality, the luxury, and the tactile pleasure of silk.

But he also demonstrated (without articulating the consequences) that the noise of nylon stockings itself needs to be accompanied by a verbal explanation in order to become evocative, i.e., to give a "rendering." Truffaut himself must have reached this conclusion when he faced the problem of producing the sound we hear in the film (which is no doubt the work of a Foley artist).

Thus we can say that in this filmic example one of the two lessons of the experiment—that the nylons give a better rendering than the silk stockings—overshadows the other—that in order to refer to their source, both sounds need to be identified verbally. One blots out the other, rather than the two mutually indicating each other or dispelling the common illusion out of which both arise—the illusion of a natural narrativity of sounds.

So common belief lends a double property to sound: not only do we believe that sound can "objectively" and single-handedly indicate its source but also that it evokes impressions linked to this source. For example, of the sound of a caress we normally say that skin is rubbing on skin, and also we say it with sensuality, and not clinically. This is really magical thinking, as when it is believed that making an image of a person takes away his or her soul.

In a scene in *Children of a Lesser God* William Hurt questions his deaf-mute lover, who is reading his lips and gestures. Consumed with curiosity about what she must be feeling, he asks her what a sound she can't hear is like—for example, what does a wave sound like? Her answer is to run her open hands along her own body, miming a caress. He, in love (but also obsequious toward the deaf), says, enraptured, that yes, that's what waves sound like. Even though what Marlee Matlin was miming had nothing at all to do with the sound of waves, but rather described the wave in general, or rather the wave-and-my-body.

What Is Rendered Is a Clump of Sensations

Why is this so, and why should sounds "render" their sources all by themselves—a belief that sound-effects people are obviously completely disabused of? No doubt because sounds are neither experienced objectively nor named, and through a magnetism related to all the vagueness and uncertainty surrounding them, sounds "attract" affects for which they are not especially responsible.⁵

It might be believed that the question of rendering boils down to that of translating one order of sensation into another. For example, in Truffaut's sequence, rendering would involve "transliterating" tactile sensations into auditory sensations: the rustling of nylon stockings would have to render the touch of legs sheathed in silk.

But in reality rendering involves perceptions that belong to no sensory channel in particular. When Leonardo da Vinci marveled that sound does not render the fall of a human body, he was thinking not only about the body's weight but also its mass as well as the sensation of falling, the jolt it causes to the person falling, and so forth. In other words, he was thinking about something that cannot be reduced to one simple sensory message. This is surely why, in most films that show falling, we are given to hear (in contradiction to real-life experience) great crashes whose volume has the duty of "rendering" weight, violence, and pain.

In fact, most of our sensory experiences consist of these clumps of agglomerated sensations.

It is morning; I open the shutters of my bedroom window. All at once I am hit with images that stun me, a violent sensation of light on my corneas, the heat of the sun if it's a nice day out, and outdoor noises that get louder as the shutters open. All this comes upon me as a whole, not dissociated into separate elements.

I have already cited in *La Toile trouee* the example of the car that zooms by while you stand on the curb: your sudden impression is composed of the sound that comes from a distance and takes a certain amount of time to disappear once the car has passed, the perception of a vibration of the ground, the vehicle's path across your field of vision, sensations of air movement, changes in temperature, and so on.

On screen, the audiovisual channel has to do all the work of transmitting these two scenes: the filmmaker must "render" them by the sole means of image and sound. Sound especially will be called upon to render the situation's violence and suddenness. In the *lived experience* of these two sample scenes, the changes in volume when we open the shutters or when the car rushes by are progressive and relative, even modest. In any case they're not surprising: before opening the window or seeing the car, we already hear their sounds. But the cinema systematically exaggerates the contrast of intensity. This device of exaggerating contrast is a kind of white lie committed even in films that use direct sound. Sometimes a sound will be made to arise suddenly out of complete silence, at the exact moment of the window-opening or the car's passing. The point is that the sound here must tell the story of a whole a rush of composite sensations and not just the auditory reality of the event.

This issue of rendering illustrates certain problems in representation—pictorial problems in the classical sense—largely ignored by film analysis, which has taken the cinema's figurative dimension as something settled and agreed upon. In sidestepping the difficulties such analysis of course has a much easier task; it can move directly on to the *terra firma* of narratological problems, territory already well scouted by literary studies.

If for our part we consider these questions to be important—not without shocking some of our students—it is because we

believe that in addressing them the cinema, reproblematising itself as a simulacrum, can find new vitality.

Materializing Sound Indices (M.S.I.)

A sound of voices, noise, or music has a particular number of *materializing sound indices*, from zero to infinity, whose relative abundance or scarcity always influences the perception of the scene and its meaning. Materializing indices can pull the scene toward the material and concrete, or their sparsity can lead to a perception of the characters and story as ethereal, abstract and fluid.

The materializing indices are the sound's details that cause us to "feel" the material conditions of the sound source, and refer to the concrete process of the sound's production. They can give us information about the substance causing the sound—wood, metal, paper, cloth—as well as the way the sound is produced—by friction, impact, uneven oscillations, periodic movement back and forth, and so on. Among the most common noises surrounding us there are some that are poor in materializing indices, which, when heard apart from their source (acousmatized), become enigmas: a motor noise or creaking can acquire an abstract quality, deprived of referentiality.

In many musical traditions perfection is defined by an absence of m.s.i.s. The musician's or singer's goal is to purify the voice or instrument sound of all noises of breathing, scratching, or any other adventitious friction or vibration linked to producing the musical tone. Even if she takes care to conserve at least an exquisite hint of materiality and noise in the release of the sound, the musician's effort lies in detaching the latter from its causality. Other musical cultures—some African traditions, for example—strive for the opposite: the "perfect" instrumental or vocal performance enriches the sound with supplementary noises, which bring

out rather than dissimulate the material origin of the sound. From this contrast we see that the composite and culture-bound notion of *noise* is closely related to the question of materializing indices.

The apportioning of materializing sound indices is controlled either at the source, by the ways in which noises are produced during filming as well as how they are recorded, or during mixing and in postsynchronization. In the audiovisual contract the way m.s.i.s are apportioned plays a preeminent part in mise-en-scene, dramatization, and the film's structuration.

A sound of a footstep, for example, can contain a minimum of materializing sound indices (abstract footstep sounds like the unobtrusive clicking in serial TV dramas) or, on the contrary, many details of texture, giving the impression of leather and cloth, and cues about the composition of what is being walked on—gravel crunching, a squeaking wood floor. Either option may be chosen in connection with any image, and synchresis predisposes the spectator to hear either one and accept the sounds he hears. At the beginning of *Mon Oncle* when the Arpel family gets up in the morning, the little boy's footsteps on the cement in the yard make a pleasant and concrete rustling, while those of his father, a large, uptight, and unhappy man, only produce a thin, unrealistic "ding."

Materializing sound indices frequently consist of *unevennesses* in the course of a sound that denote a resistance, breach, or hitch in the movement or the mechanical process producing the sound. An m.s.i. in a voice might also consist of the presence of breathing noise, mouth and throat sounds, but also any changes in timbre (if the voice breaks, goes off-key, is scratchy). For the sound of a musical instrument, m.s.i.s would include the attack of a note, unevennesses, friction, breaths, and fingernails on piano keys. An out of tune chord in a piano piece or uneven voicing in a choral piece have a materializing effect on the sound heard. They return

the sound to the sender, so to speak, in accentuating the work of the sound's emitter and its faults instead of allowing us to forget the emitter in favor of the sound or the note itself.

In the communion mass in *he Plaisir* Ophuls sets up a contrast between the very materialized vocalizing of the priests (very dense, throaty, off-key voices) and the faultless and pure voices of the little communicants. We don't even see the children singing, unlike the officiants from whose crude perspiring faces we are not spared.

Take one image and compare the effect of a music cue played on a well-tuned piano with the effect of a cue played on a slightly out of tune piano with a few bad keys. We tend to read the first cue more readily as "pit music," while with the second, even if the instrument isn't identified or shown in the image, we will sense its concrete presence in the setting.

The effects of spatial acoustics (the sensation of distance between sound source and microphone and the presence of a characteristic reverberation that exposes the sound as produced in a concrete space) can also contribute toward materializing sound. But not systematically: for a certain type of unrealistic reverberation, not commensurate with the place shown in the image, can also be coded as dematerializing and symbolizing.

Reinforcement with materializing indices (or, on the other hand, erasing them) contributes toward the creation of a universe, and can take on metaphysical meaning. Bresson and Tarkovsky have a predilection for materializing indices that immerse us in the here-and-now (dragging footsteps with clogs or old shoes in Bresson's films, agonized coughing and painful breathing in Tarkovsky's). Tati, by suppressing m.s.i.s, subtly gives us an ethereal perception of the world: think of the abstract, dematerialized klunk of the dining room's swinging door in *Mr. Hulot's Holiday*.

In films shot with direct sound the changes in voice color resulting from sound recording conditions (e.g., microphone placement at a person's face or back) also count as materializing indices, since they localize the voice in question in a concrete space and they anchor the sound in a more tangible quadrant of reality.

Examples of Rendering: *The Bear* and *Who Framed Roger Rabbit*

We might recall that in anticipation of the Christmas holidays the autumn 1988 movie season saw animal species temporarily steal the show from the usual human stars. On our left, there was a cartoon rabbit who chatted with real characters and became ensconced in carnal human space, projecting shadows, smashing against real walls, manipulating solid objects. This was of course Robert Zemeckis's *Who Framed Roger Rabbit*. To our right, several hundred pounds of decidedly real and endangered wildlife were framed as artfully as John Wayne or Gerard Depardieu might be, in *The Bear* by Jean-Jacques Annaud. As is customary, the publicity campaigns surrounding these two releases revealed several "secrets" about the production. Rarely, however, was there any mention of the problems that might have been involved in creating the films' soundtracks.

How, for example, did *Roger Rabbit's* soundmen conceive of the sounds their rabbit-hero would make? Apparently—but this is only a hypothesis, based on my audio-viewing of the film—they started out with what is paradoxical about the cartoon form itself. It is graphic, ostentatiously *drawn*, but at the same time it's modeled in three dimensions, through a play of shadows and volumes added onto a basically flat nature. And they must have wondered what sounds could convincingly make this creature seem to

maneuver in a concrete universe when it walked, slid, or banged into something.

In the noisy world of the traditional animated cartoon no one ever had to ask these questions. Filmmakers used stylized synch sound effects analogous to those in the circus, presenting sounds that followed the action as sonic symbols for impacts and movements without specifying what substances the moving beings were made out of. We should note that certain comedy directors like Jacques Tati and Blake Edwards have enjoyed treating humans in a similar way.

But we find the contrary in *Roger Rabbit*, whose soundtrack subtly attempts to give material solidity to a graphic being. The noises of the cartoon characters' bodies remain light, and the single moment in the film involving a sound effect designed to be consciously noticed is where the pulpy Jessica rubs up against the human detective played by Bob Hoskins. When the latter's very concrete skull bumps into the voluminous cartoon breasts, we hear a hollow clang, which never fails to get a laugh. But there are many other moments when the cartoon characters, especially Roger Rabbit, subtly make noises of friction and contact, conveying the impression of a thin hollow elastic material, like inflated plastic.

So through sound, the effects experts of *Roger Rabbit* let us know that the "toons" are hollow, lightweight beings. If the spectator pays little attention to these sounds, that doesn't mean by any stretch of the imagination that she doesn't hear them, or isn't influenced by them in her perception of the images. Watching the screen, she believes she simply sees what in fact she hears-sees, owing to the phenomenon I have described as added value.

It is not certain that the various creators and technicians who worked with Jean-Jacques Annaud on *The Bear* had such a very different plan from the creators of the American movie. Of course they had a different point of departure—real-life (trained) bears

being filmed as human actors would be. Only the crew of *The Bear* knew that you can't just film shots of a bear and thereby automatically convey the bear's strength, its odor, weight, and *animality*: and they knew to draw on sound to aid in rendering all these qualities.

As we know, Annaud chose not to keep the direct sound he obtained during filming. There are obvious material reasons for such a choice, not the least of which is that the beasts could only be directed by means of profuse injunctions and vociferations from their trainers offscreen. In France, where just about everyone toots his or her horn for direct sound, most filmmakers still go ahead and partially or wholly redo sound after shooting anyway; ordinarily, though, this fact is hidden like a shameful thing. So let us give credit to Annaud for his honesty, which helped him adopt a widespread practice with few qualms. Thus we learned that the animal cries of his film had been redone in a zoo, and were to various degrees edited and sometimes even dubbed by humans—particularly the noises that help express in the baby bear's throat an entire range of anthropocentric emotions: all this under the supervision of a sound designer named Laurent Quaglio.

Another behind-the-scenes artisan who seems to have played a key part in the "rendering" of the bear was the eminent Foley artist Jean-Pierre Lelong, who recreated the animal's footsteps in the studio. He is probably responsible for the undeniable success of the first appearance of Bart, the big bear. The impression of a crushing mass results largely from the cavernous sounds we hear in synchrony with the monster's stride.

But at other moments (for example, when Bart meets the hunter), the very realism of the cinematography—i.e., true closeups rather than shots taken from a distance by telephoto lenses—makes us realize how the ferocious roaring was added on to the image afterwards so as to force the meaning. There is

also a lack of naturalness in the sound mix as a whole, a muddy consistency that colors Philippe Sarde's orchestral score as well. It is hard to tell at what stage of the mixing or editing this problem was introduced. Most probably, it's one of those problems of technical coordination from which the French cinema tends to suffer.

Sound in Animation: Putting Sound to Movements

In a book on the development of musicality in children, two scholars, Francois Delalande and Bernadette Celeste, investigated a common yet misunderstood phenomenon. They studied the vocalizing with which children at play punctuate their movements of objects, dolls, toy cars, and so on.⁶ I am not speaking of dialogues children have with their toy friends, but sound effects they produce orally to accompany these activities. Delalande and Celeste's observations on this subject are highly interesting for our own inquiry. For their research also applies to the questions involved in putting sounds into films, in particular the matter of how to do sound for cartoons and other kinds of animated film.

Delalande and Celeste determined that sometimes these vocal productions "partake in a code of expression of feelings" (the descending "oooh" from a little girl when the toy character she's playing with gets a tear in her dress), and sometimes, particularly in boys' play, noises such as "brrrrr," "vrrrr," "bzzhhhh," and other labial and pharyngeal vocalizations function as sound effects and punctuations to accompany movements of vehicles, and actions of their robots and their machines. Where can they possibly come from, these codes governing such spontaneous sound productions whose job is to bring immobile objects to life?

Delalande and Celeste attempted to map out the functions of

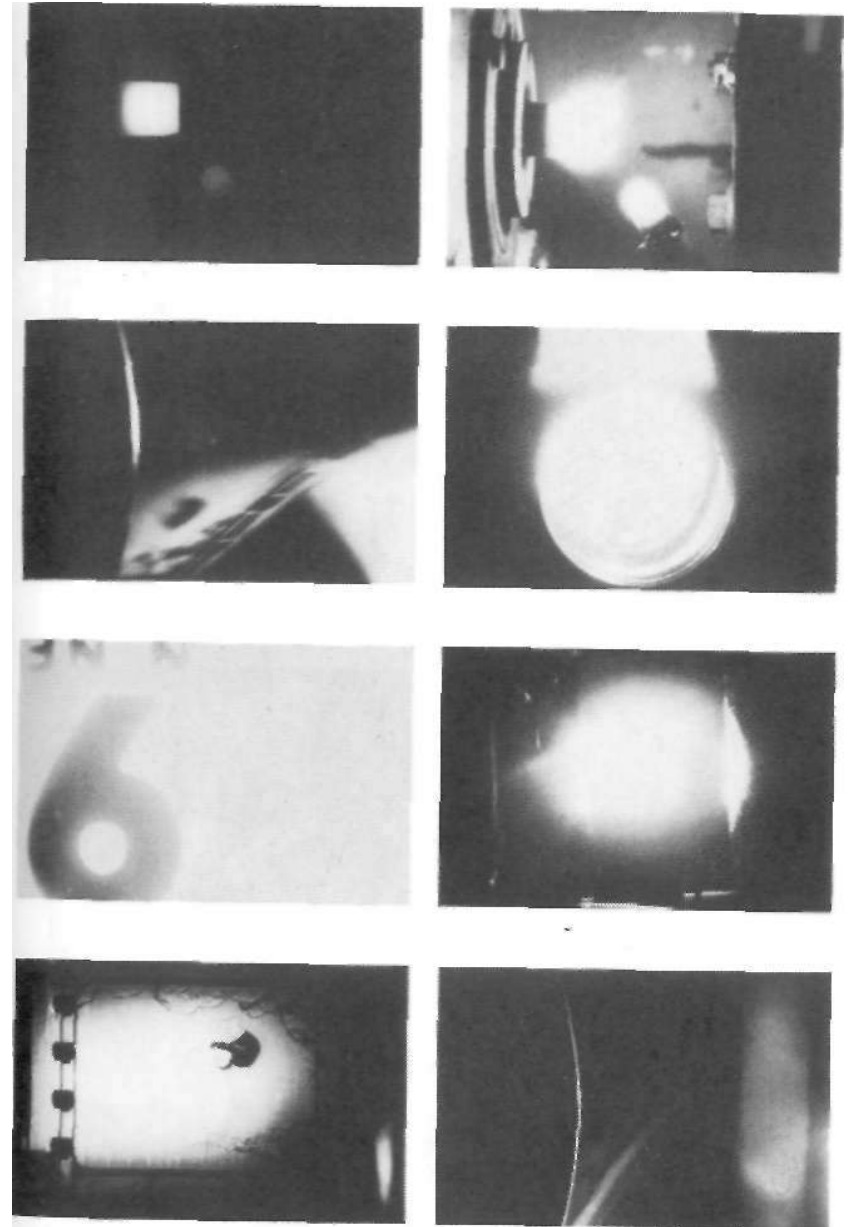
such vocal expressions in children. These include the "representation" of movements and dynamics of characters and machines involved in playing. Such representation is not so much in line with a strategy of literal reproduction, as in terms of a "mechanical and even mainly kinematic (movement-oriented) symbolism." The point is not to imitate the noise produced by the thing, but to evoke the thing's movement by means of isomorphism, that is, by "a similarity of movement between the sound and the movement it represents." When one of the boys they observed at play stops rolling his little car around, he makes a sliding sound with his mouth that recalls an airplane diving. "The descending part of the sound probably represents the slowing down of the vehicle." The sound here conveys movement and its trajectory rather than the timbre of the noise that supposedly issues from a car. "The substance of the sound has nothing to do with resemblance, it's the sound's trajectory that does." It is not difficult to see that this kind of relation between sound and movement is the very same relation used in the animated film, especially the cartoon.

Let us return to the famous and common procedure of using an ascending musical figure to accompany the climbing of a hill or a flight of stairs . . . even though the sounds of the character's footsteps do not themselves go up any scale of pitches. What is being imitated here is the trajectory and not the sound of the trajectory, drawing on a universal spatial symbolism of musical pitches. Sound is applied to most visual movements in this manner, and the animated film is the privileged province of this sound-image relation.

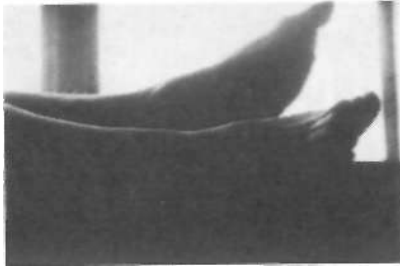
The animated film also provided the reference for *mickeymousing*, the name for a process of music-image pairing that's employed in the nonanimated cinema as well. Mickeymousing consists in following the visual action in synchrony with musical trajectories (rising, falling, zigzagging) and instrumental punctu-

ations of action (blows, falls, doors closing). This device, which I have already mentioned in connection with *The Informer*, has been criticized for being redundant, but it has an obvious function nonetheless. Try watching a Tex Avery cartoon without the sound, especially without the musical part. Silent, the visual figures tend to telescope, they do not impress themselves well in the mind, they go by too fast. Owing to the eye's relative inertia and laziness compared to the ear's agility in identifying moving figures, sound helps to imprint rapid visual sensations into memory. Indeed, it plays a more important role in this capacity of aiding the apprehension of visual movements than in focusing on its own substance and aural density.'

Many variations on this theme are possible. Take Tex Avery's *What Price Freedom*, the story of a tender idyll between a male flea and the wandering dog who offers him shelter in his fur, until the day when a female flea, on another dog... you can guess the rest. The various gags and actions of the film are accompanied and punctuated by the musical figures you would expect; when the flea jumps, a music cue jumps with it, as in the circus. But sometimes, by means of diabolical touches here and there, the real and the corporeal reemerge in the soundtrack. When a big city dog crushes the flea under his heel, we hear a slight but realistic crushing sound, like the one I mentioned in *The Skin*. Disturbing. Or when the wandering dog is happy to find Homer the flea, who returns with a big family; the cartoon animal pants with pleasure at the prospect of lodging everyone, and the panting is concrete, realistic, canine. Animality in Avery's cartoons is never very far away. And sound—ineffable and elusive sound—so clear and precise in our perception of it, and at the same time so open-ended in all it can relate—infiltrates the reassuring, closed and inconsequential universe of the cartoon like a drop of reality, a tiny, anxiety-producing drop of reality.



Persona (Ingmar Bergman, 1967). Stills from the prologue analyzed in chapter 10.



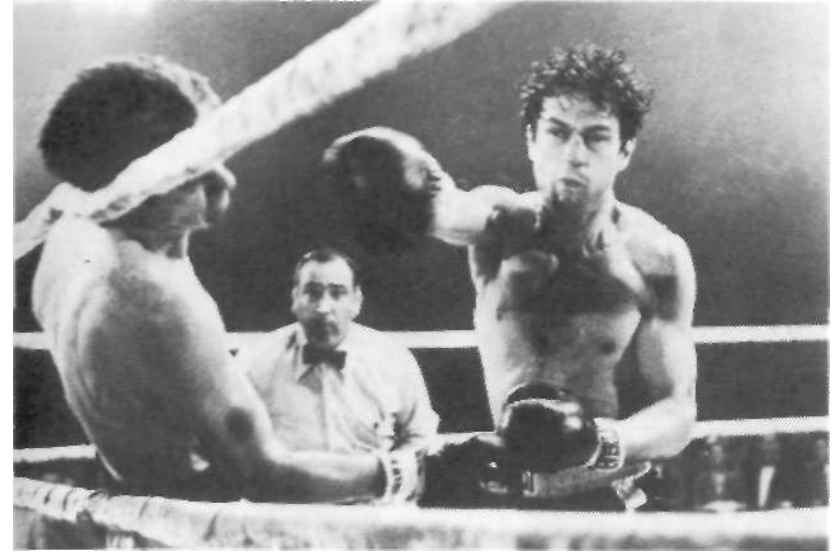
A production still from *The Invisible Man* (James Whale, 1933).



Sven Wollter, Erland Josephson, Filippa Franzen, and Susan Fleetwood in *The Sacrifice* (Andrei Tarkovsky, 1987).



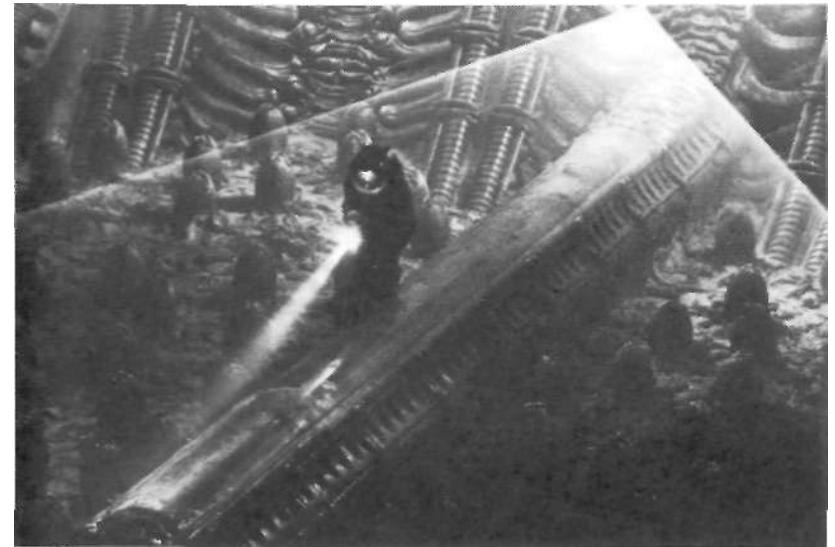
Anne Brumangne, Adriano Apra, and Anthony Pensabene in *Othon* (Jean Marie Straub and Daniele Huillet, 1969).



Robert DeNiro in *Raging Bull* (Martin Scorsese, 1980).



Marlee Matlin and William Hurt in *Children of a Lesser God* (Randa Haines, 1987).



John Hurt in *Alien* (Ridley Scott, 1979).