AUDIO-VISION

• • •

SOUND ON SCREEN

Michel Chion

• • •

edited and

translated by

Claudia Gorbman

with a foreword by

Walter Murch

COLUMBIA UNIVERSITY PRESS • NEW YORK

1994

FOREWORD

WALTER MURCH

• • •

We gestate in Sound, and are born into Sight
Cinema gestated in Sight, and was born into Sound.

We begin to hear before we are born, four and a half months after conception. From then on, we develop in a continuous and luxurious bath of sounds: the song of our mother's voice, the swash of her breathing, the trumpeting of her intestines, the timpani of her heart. Throughout the second four-and-a-half months, Sound rules as solitary Queen of our senses: the close and liquid world of uterine darkness makes Sight and Smell impossible, Taste monochromatic, and Touch a dim and generalized hint of what is to come.

Birth brings with it the sudden and simultaneous ignition of the other four senses, and an intense competition for the throne that Sound had claimed as hers. The most notable pretender is the darting and insistent Sight, who dubs himself King as if the throne had been standing vacant, waiting for him.

Ever discreet, Sound pulls a veil of oblivion across her reign and withdraws into the shadows, keeping a watchful eye on the braggart Sight. If she gives up her throne, it is doubtful that she gives up her crown.

In a mechanistic reversal of this biological sequence, Cinema spent its youth (1892—1927) wandering in a mirrored hall of voiceless images, a thirty-five year bachelorhood over which Sight ruled as self-satisfied, solipsistic King—never suspecting that destiny was preparing an arranged marriage with the Queen he thought he had deposed at birth.

This cinematic inversion of the natural order may be one of the reasons that the analysis of sound in films has always been peculiarly elusive and problematical, if it was attempted at all. In fact, despite her dramatic entrance in 1927, Queen Sound has glided around the hall mostly ignored even as she has served us up her delights, while we continue to applaud King sight on his throne. If we do notice her consciously, it is often only because of some problem or defect.

Such self-effacement seems at first paradoxical, given the power of sound and the undeniable technical progress it has made in the last sixty-five years. A further examination of the source of this power, however, reveals it to come in large part from the very handmaidenly quality of self-effacement itself: by means of some mysterious perceptual alchemy, whatever virtues sound brings to the film are largely perceived and appreciated by the audience in *visual* terms—the better the sound, the better the image. The

French composer, filmmaker, and theoretician Michel Chion has dedicated a large part of *Audio-Vision* to drawing out the various aspects of this phenomenon—which he terms *added value*—and this alchemy also lies at the heart of his three earlier, as-yet-untranslated works on film sound: *Le Son au cinéma, La Voix au cinéma,* and *La Toile trouée.* It gives me great pleasure to be able to introduce this author to the American public, and I hope it will not be long before his other works are also translated and published.

It is symptomatic of the elusive and shadowy nature of film sound that Chion's four books stand relatively alone in the landscape of film criticism, representing as they do a significant portion of everything that has ever been published about film sound from a theoretical point of view. For it is also part of Sound's effacement that she respectfully declines to be interviewed, and previous writers on film have with uncharacteristic circumspection largely respected her wishes.

It is also characteristic that this silence has been broken by a European rather than an American—even though sound for films was an American invention, and nearly all of the subsequent developments (including the most recent Dolby SR-D digital soundtrack) have been American or Anglo-American. As fish are the last to become aware of the water in which they swim, Americans take their sound for granted. But such was—and is—not the case in Europe, where the invasion of sound from across the Atlantic in 1927 was decidedly a mixed blessing and something of a curse: not without reason is chapter 7 of *Audio-Vision* (on the arrival of sound) ironically subheaded "Sixty Years of Regrets."

There are several reasons for Europe's ambivalent reaction to film sound, but the heart of the problem was foreshadowed by Faust in 1832, when Goethe had him proclaim:

> It is written that in the Beginning was the Word!
> Hmm . . . *already* I am having problems.

The early sound films were preeminently *talking* films, and the Word—with all of the power that language has to divide nation from nation as well as conquer individual hearts—has long been both the Achilles' heel of Europe as well as its crowning glory. In 1927 there were over twenty different languages spoken in Europe by two hundred million people in twenty-five different, highly developed countries. Not to mention different dialects and accents within each language and a number of countries such as Switzerland and Belgium that are multilingual.

Silent films, however, which blossomed during and after the First World War, were Edenically oblivious of the divisive powers of the Word, and were thus able—when they so desired—to speak to Europe as a whole. It is true that most of these films had intertitle cards, but these were easily and routinely switched according to the language of the country in which the film was being shown.

Even so, title cards were generally discounted as a necessary evil and there were some films, like those of writer Carl Mayer (*The Last Laugh*), that managed to tell their story without any cards at all and were highly esteemed for this ability, which was seen as the wave of the future.

It is also worth recalling that at that time the largest studio in Europe was Nordisk Films in Denmark, a country whose population of two million souls spoke a language understood nowhere else. And Asta Nielsen, the Danish star who made many films for Ufa Studios in Germany, was beloved equally by French and German soldiers during the 1914–18 war—her picture decorated the trenches on both sides. It is doubtful that the French poet Apollinaire, if he had heard her speaking in German, would have written his ode to her—

She is all!
She is the vision of the drinker and the dream of the lonely man!

—but since she hovered in shimmering and enigmatic silence, the dreaming soldiers could imagine her speaking any language they wished and make of her their sister or their lover according to their needs.

So the hopeful spirit of the League of Nations, which flourished for a while after the War That Was Supposed to End All Wars, seemed to be especially served by many of the films of the period, which—in their creative struggle to overcome the disability of silence—rose above the particular and spoke to those aspects of the human condition that know no national boundaries: Chaplin was adopted as a native son by each of the countries in which his films were shown. Some optimists even dared to think of film as a providential tool delivered in the nick of time to help unite humanity in peace: a new, less material tower erected by a modern Babel. The main studios of Ufa in Germany were in fact located in a suburb of Berlin named Neubabelsberg (new Babel city).

Thus it was with a sense of queasy forboding that many film lovers in Europe heard the approaching drumbeat of Sound. Chaplin held out, resisting a full soundtrack for his films until—significantly—*The Great Dictator* (1938). As it acquired a voice, the Tool for Peace began more to resemble the Gravedigger's Spade that had helped to dig the trenches of nationalist strife.

There were of course many more significant reasons for the rise of the Great Dictators in the twenties and thirties, and it is true that the silent film had sometimes been used to rally people around the flag, but it is nonetheless chilling to recall that Hitler's ascension to power marched in lockstep with the successful development of the talking film. And, of course, precisely because it did emphasize language, the sound film dovetailed with the divisive nationalist agendas of Hitler, Stalin, Mussolini, Franco, and others. Hitler's first public act after his victory in 1933

was to attend a screening of *Dawn*, a sound film about the German side of the 1914—18 conflict, in which one of the soldiers says, "Perhaps we Germans do not know how to live; but to die, that we know how to do incredibly well."

Alongside these political implications, the coming of sound allowed the American studios to increase their economic presence in Europe and accelerated the flight of the most talented and promising continental filmmakers (Lubitsch, Lang, Freund, Wilder, Zinnemann, etc.) to distant Hollywood. Neubabelsberg suffered the same fate as its Biblical namesake. To further sour the marriage, the first efforts at sound itself were technically poor, unimaginative, and expensive—the result of American patents that had to be purchased. Early sound recording apparatus also straitjacketed the camera and consequently impoverished the visual richness and fluidity that had been attained in the mature films of the silent era. Nordisk Films collapsed. The studios that were left standing, facing rising production costs and no longer able to count on a market outside the borders of their own country, had to accept some form of government assistance to survive, with all that such assistance implies. Studios in the United States, on the other hand, were insulated by an eager domestic audience three times the size of the largest single European market, all conveniently speaking the same language. As the United States was spared the bloodshed on its soil in both world wars, it was spared the conflict of the sound wars and, in fact, managed to profit by them.

Sixty-five years later, the reverberations of this political, cultural, and economic trauma still echo throughout Europe in an unsettled critical attitude toward film sound—and a multitude of aesthetic approaches—that have no equivalent in the United States: compare Chion's description of the French passion for "location" sound at all costs (Eric Rohmer) with the Italian reluctance to use it under any circumstances (Fellini). This is not to say

that Chion, as a European, shares the previously mentioned regrets—just the opposite: he is an ardent admirer and proponent of soundtracks from both sides of the Atlantic—but as a European he is naturally more sensitive to the economic, cultural, political, and aesthetic ramifications of the marriage of Sight and Sound. And since the initial audience for his books and articles has also—until now—been European, part of his task has been to convince his wary continental readers of the artistic merits of film sound (the French word for sound effect, for instance, is *bruit*—which translates as "noise," with all of the same pejorative overtones that the word has in English) and to persuade them to forgive Sound the guilt by association of having been present at the bursting of the silent film's illusory bubble of peace. American readers of this book should therefore be aware that they are—in part—eavesdropping on the latest stage of a family discussion that has been simmering in Europe, with various degrees of acrimony, since the marriage of Sight and Sound was consummated in 1927.

Yet a European perspective does not, by itself, yield a book like *Audio-Vision*: Chion's efforts to explore and synthesize a comprehensive theory of film sound—rather than polemicize it—are largely unprecedented even in Europe. There is another aspect to all this, which the following story might illuminate.

In the early 1950s, when I was around ten years old, and inexpensive magnetic tape recorders were first becoming available, I heard a rumor that the father of a neighborhood friend had actually acquired one. Over the next few months, I made a pest of myself at that household, showing up with a variety of excuses just to be allowed to play with that miraculous machine: hanging the microphone out the window and capturing the back-alley reverberations of Manhattan, Scotchtaping it to the shaft of a swing-arm lamp and rapping the bell-shaped shade with pencils,

inserting it into one end of a vacuum cleaner tube and shouting into the other, and so forth.

Later on, I managed to convince my parents of all the money our family would save on records if we bought our own tape recorder and used it to "pirate" music off the radio. I now doubt that they believed this made any economic sense, but they could hear the passion in my voice, and a Revere recorder became that year's family Christmas present.

I swiftly appropriated the machine into my room and started banging on lamps again and resplicing my recordings in different, more exotic combinations. I was in heaven, but since no one else I knew shared this vision of paradise, a secret doubt about myself began to worm its way into my preadolescent thoughts.

One evening, though, I returned home from school, turned on the radio in the middle of a program, and couldn't believe my ears: sounds were being broadcast the likes of which I had only heard in the secrecy of my own little laboratory. As quickly as possible, I connected the recorder to the radio and sat there listening, rapt, as the reels turned and the sounds became increasingly strange and wonderful.

It turned out to be the *Premier Panorama de Musique Concrète*, a record by the French composers Pierre Schaeffer and Pierre Henry, and the incomplete tape of it became a sort of Bible of Sound for me. Or rather a Rosetta stone, because the vibrations chiseled into its iron oxide were the mysteriously significant and powerful hieroglyphs of a language that I did not yet understand but whose voice nonetheless spoke to me compellingly. And above all told me that I was not alone in my endeavors.

Those preadolescent years that I spent pickling myself in my jar of sound, listening and recording and splicing without reference to any image, allowed me—when I eventually came to

film—to see through Sound's handmaidenly self-effacement and catch more than a glimpse of her crown.

I mention this fragment of autobiography because apparently Michel Chion came to his interest in film sound through a similar sequence of events. Such a "biological" approach—sound first, image later—stands in contrast not only to the way most people approach film—image first, sound later—but, as we have seen, to the history of cinema itself. As it turns out, Chion is a brother not only in this but also in having Schaeffer and Henry as mentors (although he has the privilege, which I lack, of a long-standing personal contact with those composers), and I was happy to see Schaeffer's name and some of his theories woven into the fabric of *Audio-Vision*. At any rate, I suspect that a primary emphasis on sound for its own sake—combined in Chion's case with a European perspective—must have provided the right mixture of elements to inspire him to knock on reclusive Sound's door, and to see his suitor's determination rewarded with armfuls of intimate details.

What had conquered me in 1953, what had conquered Schaeffer and Henry some years earlier, and what was to conquer Chion in turn was not just the considerable power of magnetic tape to capture ordinary sounds and reorganize them—optical film and discs had already had something of this ability for decades—but the fact that the tape recorder combined these qualities with full audio fidelity, low surface noise, unrivaled accessibility, and operational simplicity. The earlier forms of sound recording had been expensive, available to only a few people outside the laboratory or studio situations, noisy and deficient in their frequency range, and cumbersome and awkward to operate. The tape recorder, on the other hand, encouraged play and experimentation, and that was—and remains—its preeminent virtue.

For as far back in human history as you would care to go, sounds had seemed to be the inevitable and "accidental" (and therefore mostly ignored) accompaniment of the visual—stuck like a shadow to the object that caused them. And, like a shadow, they appeared to be completely explained by reference to the objects that gave them birth: a metallic clang was always "cast" by the hammer, just as the smell of baking always came from a loaf of fresh bread.

Recording magically lifted the shadow away from the object and stood it on its own, giving it a miraculous and sometimes frightening substantiality. King Ndombe of the Congo consented to have his voice recorded in 1904, but immediately regretted it when the cylinder was played back and the "shadow" danced, and he heard his people cry in dismay, "The King sits still, his lips are sealed, while the white man forces his soul to sing!"

The tape recorder extended this magic by an order of magnitude, and made it supremely democratic in the bargain, such that a ten-year-old boy like myself could think of it as a wonderful toy. Furthermore, it was now not only possible but easy to change the original sequence of the recorded sounds, speed them up, slow them down, play them backward. Once the shadow of sound had learned to dance, we found ourselves able to not only listen to the sounds themselves, liberated from their original causal connection, and to layer them in new, formerly impossible recombinations (*Musique Concrète*) but also—in cinema—to reassociate those sounds with images of objects or situations that were different, sometimes astonishingly different, than the objects or situations that gave birth to the sounds in the first place.

And here is the problem: the shadow that had heretofore either been ignored or consigned to follow along submissively behind the image was suddenly running free, or attaching itself mischievously to the unlikeliest things. And our culture, which is not an

"auditive" one, had never developed the concepts or language to adequately describe or cope with such an unlikely challenge from such a mercurial force—as Chion points out: "There is always something about sound that bypasses and surprises us, no matter what we do." In retrospect, it is no wonder that few have dared to confront the dancing shadow and the singing soul: it is this deficiency that Michel Chion's *Audio-Vision* bravely sets out to rectify.

The essential first step that Chion takes is to assume that there is no "natural and preexisting harmony between image and sound"—that the shadow is in fact dancing free. In his usual succinct manner, Robert Bresson captured the same idea: "Images and sounds, like strangers who make acquaintance on a journey and afterwards cannot separate."

The challenge that an idea like this presents to the filmmaker is how to create the right situations and make the right choices so that bonds of seeming inevitability are forged between the film's images and sounds, while admitting that there was nothing inevitable about them to begin with. The "journey" is the film, and the particular "acquaintance" lasts within the context of that film: it did not preexist and is perfectly free to be reformed differently on subsequent trips.

The challenge to a theoretician like Chion, on the other hand, is how to define—as broadly but as precisely as possible—the circumstances under which the "acquaintance" can be made, has been made in the past, and might best be made in the future. This challenge Chion takes up in the first six chapters of *Audio-Vision* in the form of an "Audiovisual Contract"—a synthesis and further extension of the theories developed over the last ten years in his previous three books. I should mention that as a result this section has a structural and conceptual density that may require closer attention than the second part (chapters 7–10: "Beyond Sounds and Images"), which is more freely discursive.

In the course of drawing up his contract, Chion quickly runs into the limits of ordinary language (English as well as French) to describe certain aspects of sound. This is to be expected, given the fact that we are trying to trap a shadow behind the bars of a contract, but in the process Chion forges a number of original words that give him at least a fighting chance: *synchresis, spatial magnetization, acousmatic sound, reduced listening, rendered sound, sound "en creux," the phantom of the Acousmêtre,* and so on—even *audio-vision* itself, which acquires a new meaning beyond the obvious.

Some of these terms represent concepts that will be familiar to those of us who work in film sound, but which we have either never had to articulate or for which we have developed our own individual shorthand—or for which we resort to grunts and gestures. It was a pleasure to see these old friends dressed up in new clothes, so to speak, and to have the opportunity to reevaluate them free of old or unstated assumptions. By the same token, other of Chion's ideas are, for me, completely new and original ways of thinking about the subject—in that regard I was particularly impressed by the concept of the "Acousmêtre." But the real achievement of *Audio-Vision* is—beyond simply naming and describing these isolated ideas and concepts—that it manages to synthesize them into a coherent whole whose overall pattern makes it accessible to interested nonprofessionals as well as those who have experience in the craft.

We take it for granted that this dancing shadow of sound, once free of the object that created it, can then reattach itself to a wide range of other objects and images. The sound of an ax chopping wood, for instance, played exactly in sync with a bat hitting a baseball, will "read" as a particularly forceful hit rather than a mistake by the filmmakers. Chion's term for this phenomenon is *synchresis*, an acronym formed by the telescoping together of the two words *synchronism* and *synthesis*: "The spontaneous and irresistible mental fusion, completely free of any logic, that happens between a sound and a visual when these occur at exactly the same time."

It might have been otherwise—the human mind could have demanded absolute obedience to "the truth"—but for a range of practical and aesthetic reasons we are lucky that it didn't: the possibility of reassociation of image and sound is the fundamental stone upon which the rest of the edifice of film sound is built, and without which it would collapse.

This reassociation is done for many reasons: sometimes in the interests of making a sound appear more "real" than reality (what Chion calls *rendered sound)*—walking on cornstarch, for instance, records as a better footstep in snow than snow itself; sometimes it is done simply for convenience (cornstarch, again) or necessity—the window that Gary Cooper broke in *High Noon* was not made of real glass, the boulder that chased Indiana Jones was not made of real stone, or morality—the sound of a watermelon being crushed instead of a human head. In each case, our species' multimillion-year habit of thinking of sound as a submissive shadow now works in a filmmaker's favor, and the audience is disposed to accept, within certain limits, these new juxtapositions as the truth.

But beyond all practical considerations, this reassociation is done—should be done, I believe—to *stretch* the relationship of sound to image wherever possible: to create a purposeful and fruitful tension between what is on the screen and what is kindled in the mind of the audience—what Chion calls sound *en creux* (sound "in the gap"). The danger of present-day cinema is that it can crush its subjects by its very ability to represent them; it doesn't possess the built-in escape valves of ambiguity that painting, music, literature, radio drama, and black-and-white silent film automatically have simply by virtue of their sensory incom-

pleteness—an incompleteness that engages the imagination of the viewer as compensation for what is only evoked by the artist. By comparison, film seems to be "all there" (it isn't, but it seems to be), and thus the responsibility of filmmakers is to find ways within that completeness to refrain from achieving it. To that end, the metaphoric use of sound is one of the most fruitful, flexible, and inexpensive means: by choosing carefully what to eliminate, and then reassociating different sounds that seem at first hearing to be somewhat at odds with the accompanying image, the filmmaker can open up a perceptual vacuum into which the mind of the audience must inevitably rush.

It is this movement "into the vacuum" (or "into the gap," to use Chion's phrase) that is in all probability the source of the added value mentioned earlier. Every successful metaphor—what Aristotle called "naming a thing with that which is not its name"—is seen initially and briefly as a mistake, but then suddenly as a deeper truth about the thing named and our relationship to it. And the greater the metaphoric distance, or gap, between image and accompanying sound, the greater the value added—within certain limits. The slippery thing in all this is that there seems to be a peculiar "stealthy" quality to this added value: it chooses not to acknowledge its origins in the mind.

The tension produced by the metaphoric distance between sound and image serves somewhat the same purpose, creatively, as the perceptual tension produced by the physical distance between our two eyes—a three-inch gap that yields two similar but *slightly different* images: one produced by the left eye and the other by the right. The brain is not content with this close duality and searches for something that would resolve and unify those differences. And it finds it in the concept of depth. By adding its own purely mental version of three-dimensionality to the two flat images, the brain causes them to click together into one image

with depth added. In other words, the brain resolves the differences between the two images by imagining a dimensionality that is not actually present in either image but added as the result of a mind trying to resolve the differences between them. As before, the greater the differences, the greater the depth. (Again, within certain limits: cross your eyes—exaggerating the differences—and you will deliver images to the brain that are beyond its power to resolve, and so it passes on to you, by default, a confusing double image. Close one eye—eliminate the differences—and the brain will give you a flat image with no confusion, but also with no value added.)

There really is of course some kind of depth out there in the world: the dimensionality we perceive is not a hallucination. But the *way* we perceive it—its particular flavor—is uniquely our own, unique not only to us as a species but to each of us individually. And in that sense it *is* a kind of hallucination, because the brain does not alert us to the process: it does not announce, "And now I am going to add a helpful dimensionality to synthesize these two flat images. Don't be alarmed." Instead, the dimensionality is fused into the image and made to seem as if it is coming from out there rather than "in here."

In much the same way, the mental effort of fusing image and sound in a film produces a "dimensionality" that the mind projects back onto the image as if it had come from the image in the first place. The result is that we see something on the *screen* that exists only in our minds, and is in its finer details unique to each member of the audience. It reminds me of John Huston's observation that "the real projectors are the eyes and ears of the audience." Despite all appearances, we do not *see* and *hear* a film, we *hear/see* it—hence the title of Chion's book: *Audio-Vision*. The difference is the time it takes: the fusion of left and right eye into three dimensions takes place instantly because the distance

between our eyes does not change. On the other hand the metaphoric distance between the images of a film and the accompanying sounds is—and should be—continuously changing and flexible, and it takes a good number of milliseconds (or sometimes even seconds) for the brain to make the right connections. The image of a door closing accompanied simply by the sound of a door closing is fused almost instantly and produces a relatively flat "audio-vision"; the image of a half-naked man alone in a Saigon hotel room accompanied by the sound of jungle birds (to use an example from *Apocalypse Now*) takes longer to fuse but is a more "dimensional" audio-vision when it succeeds.

I might add that, in my own experience, the most successful sounds seem not only to alter what the audience sees but to go further and trigger a kind of *conceptual resonance* between image and sound: the sound makes us see the image differently, and then this new image makes us hear the sound differently, which in turn makes us see something else in the image, which makes us hear different things in the sound, and so on. This happens rarely enough (I am thinking of certain electronic sounds at the beginning of *The Conversation*) to be specially prized when it does occur—often by lucky accident, dependent as it is on choosing exactly the right sound at exactly the right metaphoric distance from the image. It has something to do with the time it takes for the audience to "get" the metaphors: not instantaneously, but not much delayed either—like a good joke.

The question remains, in all of this, why we generally perceive the product of the fusion of image and sound—the audio-vision—in terms of the image. In other words, why does King Sight still sit on his throne?

One of Chion's most original observations—the phantom Acousmêtre—depends for its effect on delaying the fusion of sound and image to the extreme, by supplying only the sound—

almost always a voice—and withholding the image of the sound's true source until nearly the very end of the film. Only then, when the audience has used its imagination to the fullest, as in a radio play, is the real identity of the source revealed, almost always with an accompanying loss of imagined power: the wizard in *The Wizard of Oz* is one of a number of examples cited, along with Hal in *2001* and the mother in *Psycho*. The Acousmêtre is, for various reasons having to do with our perceptions (the disembodied voice seems to come from everywhere and therefore to have no clearly defined limits to its power), a uniquely cinematic device. And yet . . .

And yet there is an echo here of our earliest experience of the world: the revelation at birth (or soon after) that the song that sang to us from the very dawn of our consciousness in the womb—a song that seemed to come from everywhere and to be part of us before we had any conception of what "us" meant—that this song is the voice of another and that she is now separate from us and we from her. We regret the loss of former unity—some say that our lives are a ceaseless quest to retrieve it—and yet we delight in seeing the face of our mother: the one is the price to be paid for the other.

This earliest, most powerful fusion of sound and image sets the tone for all that are to come. One of the dominant themes of my experience with sound, ever since that first encounter at age ten, has been continual discovery—the exhilaration forty years later of coming upon new features of a landscape that has still not been entirely mapped out. Chion's contributions here and in his previous books combine a serious attempt to discover the true coordinates and features of this continent of sound with the excitement of those early explorers who have forged their own path through the forests and return with tales of wonderful things seen for the

first time. For all that Chion pursues the goal of a coherent theory, though, perhaps his theory's greatest attribute is its recognition that within that coherence there is no place for completeness—that there will always be something about sound that "bypasses and surprises us," and that we must never entirely succeed in taming the dancing shadow and the singing soul.

ONE

PROJECTIONS OF SOUND ON IMAGE

. . .

The house lights go down and the movie begins. Brutal and enigmatic images appear on the screen: a film projector running, a closeup of the film going through it, terrifying glimpses of animal sacrifices, a nail being driven through a hand. Then, in more "normal" time, a mortuary. Here we see a young boy we take at first to be a corpse like the others, but who turns out to be alive—he moves, he reads a book, he reaches toward the screen surface, and under his hand there seems to form the face of a beautiful woman.

What we have seen so far is the prologue sequence of Bergman's *Persona*, a film that has been analyzed in books and

university courses by the likes of Raymond Bellour, David Bordwell, Marilyn Johns Blackwell. And the film might go on this way.

Stop! Let us rewind Bergman's film to the beginning and simply *cut out the sound*, try to forget what we've seen before, and watch the film afresh. Now we see something quite different.

First, the shot of the nail impaling the hand: played silent, it turns out to have consisted of three separate shots where we had seen one, because they had been linked by sound. What's more, the nailed hand in silence is abstract, whereas with sound, it is terrifying, real. As for the shots in the mortuary, without the sound of dripping water that connected them together we discover in them a series of stills, parts of isolated human bodies, out of space and time. And the boy's right hand, without the vibrating tone that accompanies and structures its exploring gestures, no longer "forms" the face, but just wanders aimlessly. The entire sequence has lost its rhythm and unity. Could Bergman be an overrated director? Did the sound merely conceal the images' emptiness?

Next let us consider a well-known sequence in Tati's *Monsieur Hulot's Holiday*, where subtle gags on a small bathing beach make us laugh. The vacationers are so amusing in their uptightness, their lack of fun, their anxiety! This time, let's cut out the visuals. Surprise: like the flipside of the image, another film appears that we now "see" with only our ears; there are shouts of children having fun, voices that resonate in an outdoor space, a whole world of play and vitality. It was all there in the sound, and at the same time it wasn't.

Now if we give Bergman back his sounds and Tati his images, everything returns to normal. The nailed hand makes you sick to look at, the boy shapes his faces, the summer vacationers seem quaint and droll, and sounds we didn't especially hear when there was only sound emerge from the image like dialogue balloons in comics.

Only now we have read and heard in a different way.

Is the notion of cinema as the art of the image just an illusion? Of course: how, ultimately, can it be anything else? This book is about precisely this phenomenon of *audiovisual illusion*, an illusion located first and foremost in the heart of the most important of relations between sound and image, as illustrated above with Bergman: what we shall call *added value*.

By *added value* I mean the expressive and informative value with which a sound enriches a given image so as to create the definite impression, in the immediate or remembered experience one has of it, that this information or expression "naturally" comes from what is seen, and is already contained in the image itself. Added value is what gives the (eminently incorrect) impression that sound is unnecessary, that sound merely duplicates a meaning which in reality it brings about, either all on its own or by discrepancies between it and the image.

The phenomenon of added value is especially at work in the case of sound/image synchronism, via the principle of *synchresis* (see chapter 3), the forging of an immediate and necessary relationship between something one sees and something one hears. Most falls, blows, and explosions on the screen, simulated to some extent or created from the impact of nonresistant materials, only take on consistency and materiality through sound. But first, at the most basic level, added value is that of text, or language, on image.

Why speak of language so early on? Because the cinema is a vococentric or, more precisely, a verbocentric phenomenon.

Value Added by Text[1]

In stating that sound in the cinema is primarily vococentric, I mean that it almost always privileges the voice, highlighting and setting the latter off from other sounds. During filming it is the voice that is collected in sound recording—which therefore is

almost always voice recording—and it is the voice that is isolated in the sound mix like a solo instrument—for which the other sounds (music and noise) are merely the accompaniment. By the same token, the historical development of synch sound recording technology, for example, the invention of new kinds of microphones and sound systems, has concentrated essentially on speech since of course we are not talking about the voice of shouts and moans, but the voice as medium of verbal expression. And in voice recording what is sought is not so much acoustical fidelity to original timbre, as the guarantee of effortless intelligibility of the words spoken. Thus what we mean by vococentrism is almost always verbocentrism.

Sound in film is voco- and verbocentric, above all, because human beings in their habitual behavior are as well. When in any given sound environment you hear voices, those voices capture and focus your attention before any other sound (wind blowing, music, traffic). Only afterward, if you know very well who is speaking and what they're talking about, might you turn your attention from the voices to the rest of the sounds you hear. So if these voices speak in an accessible language, you will first seek the meaning of the words, moving on to interpret the other sounds only when your interest in meaning has been satisfied.

Text Structures Vision

An eloquent example that I often draw on in my classes to demonstrate value added by text is a TV broadcast from 1984, a transmission of an air show in England, anchored from a French studio for French audiences by our own Léon Zitrone[2]. Visibly thrown by these images coming to him on the wire with no explanation and in no special order, the valiant anchor nevertheless does his job as well as he can. At a certain point, he affirms, "Here

are three small airplanes," as we see an image with, yes, three little airplanes against a blue sky, and the outrageous redundancy never fails to provoke laughter.

Zitrone could just as well have said, "The weather is magnificent today," and that's what we would have seen in the image, where there are in fact no clouds. Or: "The first two planes are ahead of the third," and then everyone would have seen *that*. Or else: "Where did the fourth plane go?"—and the fourth airplane's absence, this plane hopping out of Zitrone's hat by the sheer power of the Word, would have jumped to our eyes. In short, the anchor could have made fifty other "redundant" comments; but their redundancy is illusory, since in each case these statements would have guided and structured our vision so that we would have seen them "naturally" in the image.

The weakness of Chris Marker's famous demonstration in his documentary *Letter from Siberia*—already critiqued by Pascal Bonitzer in another context[3]—where Marker dubs voiceovers of different political persuasions (Stalinist, anti-Stalinist, etc.) over the same sequence of innocuous images, is that through his exaggerated examples he leads us to believe that the issue is solely one of political ideology, and that otherwise there exists some neutral way of speaking. The added value that words bring to the image goes far beyond the simple situation of a political opinion slapped onto images; added value engages the very structuring of vision—by rigorously framing it. In any case, the evanescent film image does not give us much time to look, unlike a painting on a wall or a photograph in a book that we can explore at our own pace and more easily detach from their captions or their commentary.

Thus if the film or TV image seems to "speak" for itself, it is actually a ventriloquist's speech. When the shot of the three small airplanes in a blue sky declares "three small airplanes," it is a puppet animated by the anchorman's voice.

Value Added by Music

Empathetic and Anempathetic Effects

In my book *Le Son au cinéma* I developed the idea that there are two ways for music in film to create a specific emotion in relation to the situation depicted on the screen.[4] On one hand, music can directly express its participation in the feeling of the scene, by taking on the scene's rhythm, tone, and phrasing; obviously such music participates in cultural codes for things like sadness, happiness, and movement. In this case we can speak of *empathetic music*, from the word empathy, the ability to feel the feelings of others.

On the other hand, music can also exhibit conspicuous indifference to the situation, by progressing in a steady, undaunted, and ineluctable manner: the scene takes place against this very backdrop of "indifference." This juxtaposition of scene with indifferent music has the effect not of freezing emotion but rather of intensifying it, by inscribing it on a cosmic background. I call this second kind of music *anempathetic* (with the privative *a-*). The anempathetic impulse in the cinema produces those countless musical bits from player pianos, celestas, music boxes, and dance bands, whose studied frivolity and naiveté reinforce the individual emotion of the character and of the spectator, even as the music pretends not to notice them.

To be sure, this effect of cosmic indifference was already present in many operas, when emotional pitch was so high that it froze characters into inaction, provoking a sort of psychotic regression. Hence the famous operatic convention of madness, with the dumb little music that a character repeats while rocking back and forth. . . . But on the screen the anempathetic effect has taken on such prominence that we have reason to consider it to be intimately related to cinema's essence—its mechanical nature.

For, indeed, all films proceed in the form of an indifferent and automatic unwinding, that of the projection, which on the screen and through the loudspeakers produces simulacra of movement and life—and this unwinding must hide itself and be forgotten. What does anempathetic music do, if not to unveil this reality of cinema, its robotic face? Anempathetic music conjures up the mechanical texture of this tapestry of the emotions and senses.

Finally, there also exist cases of music that is neither empathetic nor anempathetic, which has either an abstract meaning, or a simple function of presence, a value as a signpost: at any rate, no precise emotional resonance.

The anempathetic effect is most often produced by music, but it can also occur with noise—when, for example, in a very violent scene after the death of a character some sonic process continues, like the noise of a machine, the hum of a fan, a shower running, as if nothing had happened. Examples of these can be found in Hitchcock's *Psycho* (the shower) and Antonioni's *The Passenger* (an electric fan).

Influences of Sound on the Perception of Movement and Perception of Speed

Visual and auditory perception are of much more disparate natures than one might think. The reason we are only dimly aware of this is that these two perceptions mutually influence each other in the audiovisual contract, lending each other their respective properties by contamination and projection.[5]

For one thing, each kind of perception bears a fundamentally different relationship to motion and stasis, since sound, contrary to sight, presupposes movement from the outset. In a film image that contains movement many other things in the frame may remain fixed. But sound by its very nature necessarily implies a

displacement or agitation, however minimal. Sound does have means to suggest stasis, but only in limited cases. One could say that "fixed sound" is that which entails no variations whatever as it is heard. This characteristic is only found in certain sounds of artificial origin: a telephone dial tone, or the hum of a speaker. Torrents and waterfalls can produce a rumbling close to white noise too, but it is rare not to hear at least some trace of irregularity and motion. The effect of a fixed sound can also be created by taking a variation or evolution and infinitely repeating it in a loop. As the trace of a movement or a trajectory, sound thus has its own temporal dynamic.

Difference in Speed of Perception

Sound perception and visual perception have their own average pace by their very nature; basically, the ear analyzes, processes, and synthesizes faster than the eye. Take a rapid visual movement—a hand gesture—and compare it to an abrupt sound trajectory of the same duration. The fast visual movement will not form a distinct figure, its trajectory will not enter the memory in a precise picture. In the same length of time the sound trajectory will succeed in outlining a clear and definite form, individuated, recognizable, distinguishable from others.

This is not a matter of attention. We might watch the shot of visual movement ten times attentively (say, a character making a complicated arm gesture), and still not be able to discern its line clearly. Listen ten times to the rapid sound sequence, and your perception of it will be confirmed with more and more precision.

There are several reasons for this. First, for hearing individuals, sound is the vehicle of language, and a spoken sentence makes the ear work very quickly; by comparison, reading with the eyes is notably slower, except in specific cases of special training, as for deaf people. The eye perceives more slowly because it has more to do all at once; it must explore in space as well as follow along in time. The ear isolates a detail of its auditory field and it follows this point or line in time. (If the sound at hand is a familiar piece of music, however, the listener's auditory attention strays more easily from the temporal thread to explore spatially.) So, overall, in a first contact with an audiovisual message, the eye is more spatially adept, and the ear more temporally adept.

Sound for "Spotting" Visual Movements and for Sleight-of-Hand

In the course of audio-viewing a sound film, the spectator does not note these different speeds of cognition as such, because added value intervenes. Why, for example, don't the myriad rapid visual movements in kung fu or special effects movies create a confusing impression? The answer is that they are "spotted" by rapid auditory punctuation, in the form of whistles, shouts, bangs, and tinkling that mark certain moments and leave a strong audiovisual memory.

Silent films already had a certain predilection for rapid montages of events. But in its montage sequences the silent cinema was careful to simplify the image to the maximum; that is, it limited exploratory perception in space so as to facilitate perception in time. This meant a highly stylized visual mode analogous to rough sketches. Eisenstein's *The General Line* provides an excellent example with its closeups in the cream separator sequence.

If the sound cinema often has complex and fleeting movements issuing from the heart of a frame teeming with characters and other visual details, this is because the sound superimposed onto the image is capable of directing our attention to a particular visual trajectory. Sound even raises the possibility of sleight-of-

hand effects: sometimes it succeeds in making us see in the image a rapid movement that isn't even there.

We find an eloquent example in the work of sound designer Ben Burtt on the *Star Wars* saga. Burtt had devised, as a sound effect for an automatic door opening (think of the hexagonal or diamond-shaped automatic doors of sci-fi films), a dynamic and convincing pneumatic "shhh" sound. So convincing, in fact, that, in making *The Empire Strikes Back*, when director Irving Kershner needed a door-closing effect he sometimes simply took a static shot of the closed door and followed it with a shot of the door open. As a result of sound editing, with Ben Burtt's "psssht," spectators who have nothing before their eyes besides a straight cut nevertheless think they see the door slide open. Added value is working full steam here, in accordance with a phenomenon specific to sound film that we might call faster-than-the-eye.

Deaf people raised on sign language apparently develop a special ability to read and structure rapid visual phenomena. This raises the question whether the deaf mobilize the same regions at the center of the brain as hearing people do for sound—one of the many phenomena that lead us to question received wisdom about distinctions between the categories of sound and image.

The Ear's Temporal Threshold

Further, we need to correct the formulation that hearing occurs in continuity. The ear in fact listens in brief slices, and what it perceives and remembers *already* consists in short syntheses of two or three seconds of the sound as it evolves. However, within these two or three seconds, which are perceived as a gestalt, the ear, or rather the ear-brain system, has minutely and seriously done its

investigation such that its overall report of the event, delivered periodically, is crammed with the precise and specific data that have been gathered.

This results in a paradox: we don't hear sounds, in the sense of recognizing them, until shortly after we have perceived them. Clap your hands sharply and listen to the resulting sound. Hearing—namely the synthesized apprehension of a small fragment of the auditory event, consigned to memory—will *follow* the event very closely, it will not be totally simultaneous with it.

Influence of Sound on the Perception of Time in the Image

Three Aspects of Temporalization

One of the most important effects of added value relates to the *perception of time in the image*, upon which sound can exert considerable influence. An extreme example, as we have seen, is found in the prologue sequence of *Persona*, where atemporal static shots are inscribed into a time continuum via the sounds of dripping water and footsteps. Sound temporalizes images in three ways.

The first is temporal animation of the image. To varying degrees, sound renders the perception of time in the image as exact, detailed, immediate, concrete—or vague, fluctuating, broad.

Second, sound endows shots with temporal linearization. In the silent cinema, shots do not always indicate temporal succession, wherein what happens in shot B would necessarily follow what is shown in shot A. But synchronous sound does impose a sense of succession.

Third, sound *vectorizes* or dramatizes shots, orienting them toward a future, a goal, and creation of a feeling of imminence

and expectation. The shot is going somewhere and it is oriented in time. We can see this effect at work clearly in the prologue of *Persona*—in its first shot, for example.

Conditions Necessary for Sound to Temporalize Images

In order to function, these three effects depend on the nature of the sounds and images being put together.

First case: *the image has no temporal animation or vectorization in itself.* This is the case for a static shot, or one whose movement consists only of a general fluctuating, with no indication of possible resolution—for example, rippling water. In this instance, sound can bring the image into a temporality that it introduces entirely on its own.

Second case: *the image itself has temporal animation* (movement of characters or objects, movement of smoke or light, mobile framing). Here, sound's temporality *combines* with the temporality already present in the image. The two may move in concert or slightly at odds with each other, in the same manner as two instruments playing simultaneously.

Temporalization also depends on the type of sounds present. Depending on density, internal texture, tone quality, and progression, a sound can temporally animate an image to a greater or lesser degree, and with a more or less driving or restrained rhythm.[6] Different factors come into play here:

1. *How sound is sustained.* A smooth and continuous sound is less "animating" than an uneven or fluttering one. Try accompanying an image first with a prolonged steady note on the violin, and then with the same note played with a tremolo made by rapidly moving the bow. The second sound will cause a more tense and immediate focusing of attention on the image.

2. *How predictable the sound is as it progresses.* A sound with a regular pulse (such as a basso continuo in music or a mechanical ticking) is more predictable and tends to create less temporal animation than a sound that is irregular and thus unpredictable; the latter puts the ear and the attention on constant alert. The dripping of water in *Persona* as well as in Tarkovsky's films provide good examples: each unsettles our attention through its unequal rhythm.

However, a rhythm that is too regularly cyclical can also create an effect of tension, because the listener lies in wait for the possibility of a fluctuation in such mechanical regularity.

3. *Tempo.* How the soundtrack temporally animates the image is not simply a mechanical question of tempo. A rapid piece of music will not necessarily accelerate the perception of the image. Temporalization actually depends more on the regularity or irregularity of the aural flow than on tempo in the musical sense of the word. For example, if the flow of musical notes is unstable but moderate in speed, the temporal animation will be greater than if the speed is rapid but regular.

4. *Sound definition.* A sound rich in high frequencies will command perception more acutely; this explains why the spectator is on the alert in many recent films.

Temporalization also depends on the *model of sound-image linkage* and on the *distribution of synch points* (see below). Here, also, the extent to which sound activates an image depends on how it introduces points of synchronization—predictably or not, variously or monotonously. Control over expectations tends to play a powerful part in temporalization.

In summary, for sound to influence the image's temporality, a minimum number of conditions are necessary. First, the image must lend itself to it, either by being static and passively receptive

(cf. the static shots of *Persona*) or by having a particular movement of its own (microrhythms "temporalizable" by sound). In the second case, the image should contain a minimum of structural elements—either elements of agreement, engagement, and sympathy (as we say of vibrations), or of active antipathy—with the flow of sound.

By visual *microrhythms* I mean rapid movements on the image's surface caused by things such as curls of smoke, rain, snowflakes, undulations of the rippled surface of a lake, dunes, and so forth—even the swarming movement of photographic grain itself, when visible. These phenomena create rapid and fluid rhythmic values, instilling a vibrating, trembling temporality in the image itself. Kurosawa utilizes them systematically in his film *Dreams* (petals raining down from flowering trees, fog, snowflakes in a blizzard). Hans-Jürgen Syberberg, in his static and posed long takes, also loves to inject visual microrhythms (smoke machines in *Hitler*, the flickering candle during Edith Clever's reading of Molly Bloom's monologue, etc.), as does Manoel de Oliveira (*Le Soulier de satin*). It is as if this technique affirms a kind of time proper to sound cinema as a recording of the microstructure of the present.

Sound Cinema is Chronography

One important historical point has tended to remain hidden: we are indebted to synchronous sound for having made cinema an art of time. The stabilization of projection speed, made necessary by the coming of sound, did have consequences that far surpassed what anyone could have foreseen. Filmic time was no longer a flexible value, more or less transposable depending on the rhythm of projection. Time henceforth had a fixed value; sound cinema guaranteed that whatever lasted x seconds in the

editing would still have this same exact duration in the screening. In the silent cinema a shot had no exact internal duration; leaves quivering in the wind and ripples on the surface of the water had no absolute or fixed temporality. Each exhibitor had a certain margin of freedom in setting the rhythm of projection speed. Nor is it any accident that the motorized editing table, with its standardized film speed, did not appear until the sound era.

Note that I am speaking here of the rhythm of the finished film. Within a film there certainly may be material shot at nonstandard speeds—accelerated or slow-motion—as seen in works of Michael Powell, Scorsese, Peckinpah, or Fellini at different points in sound film history. But if the speed of these shots does not necessarily reproduce the real speed at which the actors moved during filming, it *is* fixed in any case at a precisely determined and controlled rate.

So sound temporalized the image: not only by the effect of added value but also quite simply by normalizing and stabilizing film projection speed. A silent film by Tarkovsky, who called cinema "the art of sculpting in time," would not be conceivable. His long takes are animated with rhythmic quiverings, convulsions, and fleeting apparitions that, in combination with vast controlled visual rhythms and movements, form a kind of hypersensitive temporal structure. The sound cinema can therefore be called "chronographic": written in time as well as in movement.

Temporal Linearization

When a sequence of images does not necessarily show temporal succession in the actions it depicts—that is, when we can read them equally as simultaneous or successive—the addition of realistic, diegetic sound imposes on the sequence a sense of real time,

like normal everyday experience, and above all, a sense of time that is linear and sequential.

Let us take a scene that occurs frequently enough in silent film: a crowd reacting, constructed as a montage of closeups of scowling or grinning faces. Without sound the shots that follow one another on the screen need not designate actions that are temporally related. One can quite easily understand the reactions as being simultaneous, existing in a time analogous to the perfect tense in grammar. But if we dub onto these images the sounds of collective booing or laughter, they seem magically to fall into a linear time continuum. Shot B shows someone who laughs or jeers *after* the character in shot A.

The awkwardness of some crowd scenes in the very earliest talkies derives from this. For example, in the opening company dinner of Renoir's *La Chienne*, the sound (laughter, various verbal exchanges among the partygoers) seems to be stuck onto images that are conceived as inscribed in a kind of time that was not yet linear.

The sound of the spoken voice, at least when it is diegetic and synched with the image, has the power to inscribe the image in a real and linearized time that no longer has elasticity. This factor explains the dismay of many silent filmmakers upon experiencing the effect of "everyday time" at the coming of sound.

Synchresis, which we shall discuss at greater length in chapter 3, is a powerful factor in linearizing and inscribing images into real time.

Vectorization of Real Time

Imagine a peaceful shot in a film set in the tropics, where a woman is ensconced in a rocking chair on a veranda, dozing, her chest rising and falling regularly. The breeze stirs the curtains and the bamboo windchimes that hang by the doorway. The leaves of the banana trees flutter in the wind. We could take this poetic shot and easily project it from the last frame to the first, and this would change essentially nothing, it would all look just as natural. We can say that the time this shot depicts is real, since it is full of microevents that reconstitute the texture of the present, but that it is not vectorized. Between the sense of moving from past to future and future to past we cannot confirm a single noticeable difference.

Now let us take some sounds to go with the shot—direct sound recorded during filming, or a soundtrack mixed after the fact: the woman's breathing, the wind, the chinking of the bamboo chimes. If we now play the film in reverse, it no longer works at all, especially the windchimes. Why? Because each one of these clinking sounds, consisting of an attack and then a slight fading resonance, is a finite story, oriented in time in a precise and irreversible manner. Played in reverse, it can immediately be recognized as "backwards." Sounds are vectorized.

The same is true for the dripping water in the prologue of *Persona*. The sound of the smallest droplet imposes a real and irreversible time on what we see, in that it presents a trajectory in time (small impact, then delicate resonance) in accordance with logics of gravity and return to inertia.

This is the difference, in the cinema, between the orders of sound and image: given a comparable time scale (say two to three seconds), aural phenomena are much more characteristically vectorized in time, with an irreversible beginning, middle, and end, than are visual phenomena.

If this fact normally eludes us, it is because the cinema has derived amusement from exceptions and paradoxes by playing on what's visually irreversible: a broken object whose parts all fly back together, a demolished wall that reconstructs, or the inevitable gag of the swimmer coming out of the pool feet first

and settling upon the diving board. Of course, images showing actions that result from nonreversible forces (gravity causes an object to fall, an explosion disperses fragments), is clearly vectorized. But much more frequently in movies, images of a character who speaks, smiles, plays the piano, or whatever are reversible; they are not marked with a sense of past and future. Sound, on the other hand, quite often consists of a marking off of small phenomena oriented in time. Isn't piano music, for example, composed of thousands of little indices of vectorized real time, since each note begins to die as soon as it is born?

Stridulation and Tremolo: Naturally or Culturally Based Influence

The temporal animation of the image by sound is not a purely physical and mechanical phenomenon: cinematic and cultural codes also play a part in it. A music cue or a voiceover that is culturally perceived as not "in" the setting will not set the image to vibrating. And yet, the phenomenon still has a noncultural basis.

Take the example of the string tremolo, a device traditionally employed in opera and symphonic music to create a feeling of dramatic tension, suspense, or alarm. In film we can get virtually the same result with sound effects: for example, the stridulation of nocturnal insects in the final scene of Randa Haines's *Children of a Lesser God*. This ambient sound, however, is not explicitly coded as a "tremolo"; it is not in the official repertoire of standard devices of filmic writing. Nevertheless it can have on the dramatic perception of time exactly the same effect of concentrating attention and making us sensitive to the smallest quivering on the screen, as does the tremolo in the orchestra. Sound editors and mixers frequently do utilize such nocturnal ambient sounds, and

parcel out the effect like orchestra conductors, by their choices of certain sound-effects recordings and the ways they blend these to create an overall sound. Obviously the effect will vary according to the density of the stridulation, its regular or fluctuating quality, and its duration—just as for an orchestral effect.

But what exactly is there in common, for a film spectator, between a string tremolo in a pit orchestra, which the viewer identifies as a cultural musical procedure, and the rustling of an animal, which the viewer perceives as a natural emanation from the setting (without dreaming, of course, that the latter could have been recorded separately from the image and expertly recomposed)? Only an acoustic identity: that of a sharp, high, slightly uneven vibrating that both alarms and fascinates. It appears, then, that we have a universal and spontaneous effect operating here. It is also, however, a very fragile effect, which the slightest thing—bad sound balance, a spectator's loss of confidence in the audiovisual contract due to a fault in production—suffices to compromise.

This also holds true for all effects of added value that have nothing of the mechanical: founded on a psychophysiological basis, they operate only under certain cultural, aesthetic, and emotional conditions by means of a general interaction of all elements.

Reciprocity of Added Value: The Example of Sounds of Horror

Added value works reciprocally. Sound shows us the image differently than what the image shows alone, and the image likewise makes us hear sound differently than if the sound were ringing out in the dark. However for all this reciprocity the screen remains the principal support of filmic perception. Transformed

by the image it influences, sound ultimately reprojects onto the image the product of their mutual influences. We find eloquent testimony to this reciprocity in the case of horrible or upsetting sounds. The image projects onto them a meaning they do not have at all by themselves.

Everyone knows that the classical sound film, which avoided showing certain things, called on sound to come to the rescue. Sound *suggested* the forbidden sight in a much more frightening way than if viewers were to see the spectacle with their own eyes. An archetypal example is found at the beginning of Aldrich's masterpiece, *Kiss Me Deadly*, when the runaway hitchhiker whom Ralph Meeker picked up has been recaptured by her pursuers and is being tortured. We see nothing of this torture but two bare legs kicking and struggling, while we hear the unfortunate woman's screams. There's a typical use of sound, we might say. Of course—as long as it's clear that what makes the screams so terrifying is not their own acoustic properties but what the narrated situation, and what we're allowed to see, project onto them.

Another traumatic aural effect occurs in a scene in *The Skin*, by Liliana Cavani (1981, based on Malaparte's novel). An American tank accidentally runs over a little Italian boy, with—if memory does not fail me—a ghastly noise that sounds like a watermelon being crushed. Although spectators are not likely to have heard the real sound of a human body in this circumstance, they may imagine that it has some of this humid, viscous quality. The sound here has obviously been Foleyed in, perhaps precisely by crushing a melon.

As we shall see, the figurative value of a sound in itself is usually quite nonspecific. Depending on the dramatic and visual context, a single sound can convey very diverse things. For the spectator, it is not acoustical realism so much as synchrony

above all, and secondarily the factor of verisimilitude (verisimilitude arising not from truth but from convention), that will lead him or her to connect a sound with an event or detail. The same sound can convincingly serve as the sound effect for a crushed watermelon in a comedy or for a head blown to smithereens in a war film. The same noise will be joyful in one context, intolerable in another.

In Franju's *Eyes Without a Face* we find one of the rare disturbing sounds that the public and critics have actually remarked upon after viewing: the noise made by the body of a young woman—the hideous remains of an aborted skin-transplant experiment—when surgeon Pierre Brasseur and his accomplice Alida Valli drop it into a family vault. What this flat thud (which never fails to send a shudder through the theater) has in common with the noise in Cavani's film is that it transforms the human being into a thing, into vile, inert, disposable matter, with its entrails and osseous cavities.

But it is an upsetting noise also in that within the film's rhythm it constitutes an *interruption of speech*, a moment where the two perpetrators' speech is absent. At the cinema or in real life certain sounds have this resonance because they occur at a certain place: in a flow of language, where they make a hole. A ghastly example of this idea can be seen in Tarkovsky's *Andrei Rublov*. A Russian prince emerges from being tortured by the Tatars; he is covered with bandages, which hide his mutilated body and leave nothing visible but his lips. Abandoned on a bed, he curses his torturers; but just after, the torturer's hand brings a ladle full of boiling oil which is poured down his throat. This action is masked from view by the back of the torturer, who has mercifully (or rather cleverly) interposed himself at that moment between the spectator and the victim's head. What we hear is the atrocious sound of gargling, which makes the skin crawl. All the same, as with the crushing

sound mentioned above, this could be the same sound Peter Sellers might make as he gargles in a Blake Edwards comedy.

Here, the effect of the sound is so strong because it represents human speech felled at its physical core: what has been destroyed are a larynx and a tongue, which have just spoken.

NOTES

1. Projections of Sound on Image

1. Chion's terminology, referring as it does to the register of political economy, is based on a pun: *value added by text* plays on the *value added tax* imposed on purchases of goods and services in France and the rest of the EC.—Trans.

2. Léon Zitrone: French TV anchor, a household word in France since the early years of television. Zitrone did commentary for horse races, figure skating, official ceremonies such as royal weddings, and air shows.

3. Pascal Bonitzer, *Le Regard et la voix*, pp. 37–40. See also the opening of chapter 8, on film sound, in David Bordwell and Kristin Thompson, *Film Art: An Introduction*.

4. Michel Chion, *Le Son au cinéma*, chapter 7, "La Belle Indifférente," pp. 119–42, especially pp. 122–26.

5. Throughout this book I use the phrase *audiovisual contract* as a reminder that the audiovisual relationship is not natural, but a kind of symbolic contract that the audio-viewer enters into, agreeing to think of sound and image as forming a single entity.

6. Here by *density* I mean the density of sound events. A sound with marked and rapid modifications in a given time will temporally animate the image in a different way than a sound that varies less in the same time.