

AUDIO-VISION

• • •

SOUND ON SCREEN

Michel Chion

• • •

edited and

translated by

Claudia Gorbman

with a foreword by

Walter Murch

COLUMBIA UNIVERSITY PRESS • NEW YORK

Columbia University Press wishes to express its appreciation of assistance given by
the government of France through Le Ministère de la Culture in the preparation of
the translation.

. . .



Columbia University Press
New York Chichester, West Sussex
L'Audio-Vision © 1990 Editions Nathan, Paris

Copyright © 1994 Columbia University Press
All rights reserved

Library of Congress Cataloging-in-Publication Data
Chion, Michel

[Audio-vision. French]

Audio-vision: sound on screen/Michel Chion; edited and translated by
Claudia Gorbman; with a foreword by
Walter Murch.

p. cm

Includes bibliographical references and index.

ISBN 0-231-07898-6

ISBN 0-231-07899-4 (pbk.)

1. Sound motion pictures. 2. Motion pictures—Sound effects. 3.
Motion pictures—Aesthetics. I. Gorbman, Claudia. II. Murch,

Walter, 1943— . III. Title.

PN1995.7.C4714 1994

791.43'024—dc20 93-23982

CIP

Casebound editions of Columbia University Press books are printed on
permanent and durable acid-free paper

c 10 9 8 7 6 5 4 3 2 1

p 10 9 8 7 6 5 4 3 2 1

CONTENTS

Foreword by Walter Murch vii • Preface xxv

PART ONE • THE AUDIOVISUAL CONTRACT	1
1 PROJECTIONS OF SOUND ON IMAGE	3
2 THE THREE LISTENING MODES	25
3 LINES AND POINTS: HORIZONTAL AND VERTICAL PERSPECTIVES ON AUDIOVISUAL RELATIONS	35
4 THE AUDIOVISUAL SCENE	66
5 THE REAL AND THE RENDERED	95
6 PHANTOM AUDIO-VISION	123
 PART TWO • BEYOND SOUNDS AND IMAGES	 139
7 SOUND FILM—WORTHY OF THE NAME	141
8 TELEVISION, VIDEO ART, MUSIC VIDEO	157
9 TOWARD AN AUDIOLOGOVISUAL POETICS	169
10 INTRODUCTION TO AUDIOVISUAL ANALYSIS	185

Notes 215 • Glossary 221 • Bibliography 225 • Index 229

8-9
45
18-72

PART 1

. . .

THE
AUDIOVISUAL
CONTRACT

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

well, 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.

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 TEXT¹

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 Zitrone². 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 context³—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 train-

ing, 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 "pssst," 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 bam-

boo 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.

TWO

THE THREE

LISTENING MODES

. . .

CAUSAL LISTENING

When we ask someone to speak about what they have heard, their answers are striking for the heterogeneity of levels of hearing to which they refer. This is because there are at least three modes of listening, each of which addresses different objects.¹ We shall call them *causal listening*, *semantic listening*, and *reduced listening*.

Causal listening, the most common, consists of listening to a sound in order to gather information about its cause (or source). When the cause is visible, sound can provide supplementary information about it; for example, the sound produced by an