

In *Audio-Vision*, the French composer-filmmaker-critic Michel Chion presents a reassessment of the audiovisual media since sound's revolutionary debut in 1927 and sheds light on the mutual influences of sound and image in audiovisual perception.

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

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

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

ALAN WILUAMS

RUTGERS UNIVERSITY

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

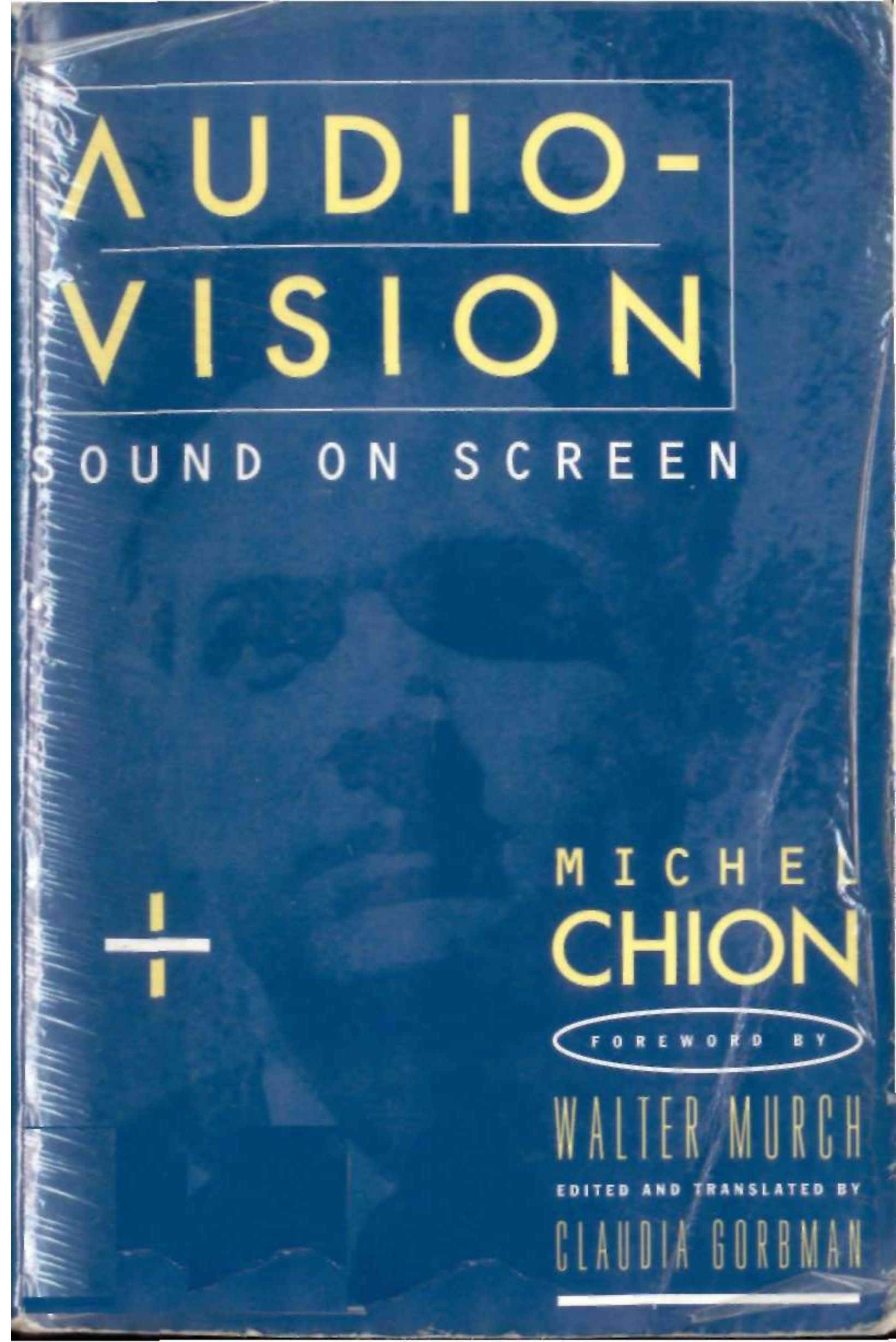
CLAUDIA GORBMAN is a Professor in the Liberal Studies Program at the University of Washington, Tacoma.

Jacket illustration: *Eratormad* by David Lynch, 1976.

Jacket design: John Costa

Printed in U.S.A.

COLUMBIA UNIVERSITY PRESS
NEW YORK



(...)

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-Jiirgen 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.

THREE
LINES AND POINTS:
HORIZONTAL AND
VERTICAL
PERSPECTIVES ON
AUDIOVISUAL
RELATIONS

• • •

Harmony or Counterpoint?

The arrival of sound in the late twenties coincided with an extraordinary surge of aestheticism in silent film, and people took passionate interest in comparing cinema with music. This is why they came up with the term *counterpoint* to designate their notion of the sound film's ideal state as a cinema free of

redundancy where sound and image would constitute two parallel and loosely connected tracks, neither dependent on the other.

Remember that in the language of Western classical music counterpoint refers to the mode of composition that conceives of each of several concurrent musical voices as individuated and coherent in its horizontal dimension. *Harmony* concerns the vertical dimension, and involves the relations of each note to the other notes heard at the same moment, together forming chords; harmony governs the conduct of the voices in the way these vertical chords are obtained. Training in classical composition involves learning both disciplines; and most musical works in the Western classical tradition combine these two dimensions, which are closely associated, to varying degrees.

If there exists something one can call audiovisual counterpoint, it occurs under conditions quite different from musical counterpoint. The latter exclusively uses notes—all the same raw material—while sound and image fall into different sensory categories. If there's any sense at all to the analogy, audiovisual counterpoint implies an "auditory voice" perceived horizontally in tandem with the visual track, a voice that possesses its own formal individuality.

What I wish to show is that films tend to exclude the possibility of such horizontal-contrapuntal dynamics. Quite to the contrary: in the cinema, harmonic and vertical relations (whether they be consonant, dissonant, or neither, à la Debussy) are generally more salient—i.e., the relations between a given sound and what is happening at that moment in the image. So to speak about counterpoint in the cinema is therefore to borrow a notion somewhat wrongheadedly, applying an intellectual speculation rather than a workable concept.

As proof we might note that historically, film studies quickly became muddled by this analogy, often to the point of using it entirely the wrong way. Many cases being offered up as models of

counterpoint were actually splendid examples of *dissonant harmony*, since they point to a momentary discord between the image's and sound's figural natures. If we, too, sometimes make use of the musical analogy, we need to be careful: the term harmony doesn't take into account the specificity of audiovisual phenomena either.¹

Our investigation of the horizontal and vertical aspects of the audiovisual sequence, to which this chapter is devoted, underscores their interdependence and their dialectical relationship. For instance, films characterized by a sort of horizontal freedom—the typical example is the music video, whose parallel image and sound tracks often have no precise relation—also exhibit a vigorous perceptual solidarity, marked by points of synchronization that occur throughout. These synch points—to return to the musical analogy—provide the harmonic framework of the audiovisual system.

* Audiovisual Dissonance

Audiovisual counterpoint, which film aestheticians seem perennially to advocate, plead for, and insist upon, occurs on television every day, even though no one seems to notice. You can find it especially in replays of sports events, when the image goes its own way and the commentary goes another. I like to use the example of the coverage of a certain bicycle race in Barcelona. The image shows the racers from a helicopter. The soundtrack consists of a dialogue between the TV reporters and some cyclists not participating in the race. It is obvious that those speaking are not watching the images, nor are they saying anything remotely about them. Image and sound follow two totally different tracks for two minutes; the only thing giving any sense to the cohabitation of these two universes is the topic of cycling. And yet no one who views this clip notices its obvious counterpoint.

Why not? It is not enough if the sound and image differ in nature (the content of each, their spatial characteristics, etc.). Audiovisual counterpoint will be noticed only if it sets up an opposition between sound and image on a precise point of meaning. This kind of counterpoint influences our reading, in postulating a certain linear interpretation of the meaning of the sounds. Take for example the moment in Godard's *First Name Carmen* when we see the Paris metro and hear the cries of seagulls. Critics identified this as counterpoint, because the seagulls were considered as signifiers of "seashore setting" and the metro image as a signifier of "urban setting." This is what I mean by a linear interpretation: it reduces the audio and visual elements to abstractions at the expense of their multiple concrete particularities, which are much richer and full of ambiguity. Thus this counterpoint reduces our reading to a stereotyped meaning of the sounds, drawing on their *codedness* (seagulls = seashore) rather than their own sonic substance, their specific characteristics in the passage in question.

So the problem of counterpoint-as-contradiction, or rather of audiovisual dissonance, as it has been used and touted in films like Robbe-Grillet's *L'Homme qui merit*, is that counterpoint or dissonance implies a prereading of the relation between sound and image. By this I mean that it forces us to attribute simple, one-way meanings, since it is based on an opposition of a rhetorical nature ("I should hear X, but I hear Y"). In effect it imposes the model of language and its abstract categories, handled in yes-no, redundant-contradictory oppositions.

There exist hundreds of possible ways to add sound to any given image. Of this vast array of choices some are wholly conventional. Others, without formally contradicting or "negating" the image, carry the perception of the image to another level. And audiovisual dissonance is merely the inverse of convention, and thus pays homage to it, imprisoning us in a binary logic that has only remotely to do with how cinema works.

For an example of true *free counterpoint* consider the amazing resurrection scene in Tarkovsky's film *Solaris*. The hero's former wife, who committed suicide, comes back to him in flesh and blood on a space station, thanks to mysterious forces summoned forth by a brain-planet. Driven to despair on realizing that she is a nonhuman artifact, she kills herself yet again by swallowing liquid oxygen. The hero embraces her frozen body. But pitilessly the ocean-brain resuscitates her, and we see her body shaken with convulsions that are no longer those of agony or pleasure but of returning to "life." Over these images Tarkovsky had the imagination to dub sounds of breaking glass, which yield a phenomenal effect. We do not hear them as "wrong" or inappropriate sounds. Instead, they suggest that she is constituted of shards of ice; in a troubling, even terrifying way, they render both the creature's fragility and artificiality, and a sense of the precariousness of bodies.

The Predominance of the Vertical (There Is No Soundtrack)

In *La Voix au cinema* I declared my position on the question of sound in the cinema, stating what ought to be obvious—that there is no soundtrack.²

Of course, no one would deny that in the purely technical sense of the word there does exist a sound channel that runs the length of the film. But this does not necessarily mean that the sounds of the film constitute a coherent entity.

If I do occasionally use the term *soundtrack* it is in a technical way, to designate empirically the simple end-to-end aggregation of all sounds in a film—inert and with no active autonomous meaning. In current parlance the idea of the soundtrack derives from a purely mechanical analogy with the image track; the latter is indeed a valid concept. The image track owes its being and its unity to the presence of a frame, a space of the images in which the spectator is invested.

By stating that *there is no soundtrack* I mean first of all that the sounds of a film, taken separately from the image, do not form an internally coherent entity on equal footing with the image track. Second, I mean that each audio element enters into *simultaneous vertical relationship* with narrative elements contained in the image (characters, actions) and visual elements of texture and setting. These relationships are much more direct and salient than any relations the audio element could have with other sounds. It's like a recipe: even if you mix the audio ingredients separately before pouring them into the image, a chemical reaction will occur to separate out the sounds and make each react on its own with the field of vision.

In the simplest and strongest relation, that of *offscreen sound*, the confrontation of sound with image establishes the sound as being offscreen, even as this sound is heard coming from the surface of the screen. Take away the image, and the offscreen sounds that were perceived apart from other sounds, purely by virtue of the visual exclusion of their source, become just like the others. The audiovisual structure collapses, and the sounds make a completely new one together. A film deprived of its image and transformed into an audio track proves altogether strange—provided you *listen* and refrain from imposing the images from your memory onto the sounds you hear. Only at this point can we talk about a soundtrack.

Therefore, there is no image track and no soundtrack in the cinema, but a *place* of images, plus sounds.

SOUND AND IMAGE IN RELATION TO EDITING

Sound Editing Has Not Created a Specific Unit

Sounds, like film images, are editable. By this I mean that they are recorded on strips of tape or film that can be cut, assembled, and moved around at will.

For the image it is this very fact of editing-construction that created the specific unit of cinema, the shot. The shot is a unit of greater or lesser pertinence for film analysis (depending on who has made the film and how), but is nevertheless quite convenient for doing breakdowns of films. Even if we do not consider shot 67 to be a structural narrative unit in itself but "only" a shot, i.e., the length of film between two splices, it is a great help to be able to say that the interesting, pertinent, significant element we are discussing can be found between the middle of shot 66 and the end of shot 68. The shot has the enormous advantage of being a neutral unit, objectively defined, that everyone who has made the film as well as those who watch it can agree on.

We can instantly see that no such condition obtains for sound: the editing of film sounds has created no specific sound unit. Unlike visual cuts, sound splices neither jump to our ears nor permit us to demarcate identifiable units of sound montage.

The cinema isn't the only place this occurs. Sounds have been edited since it became technically possible in radio (about sixty years ago) and in phonograph and tape recording. In none of these instances, regardless of whether images are involved, has the notion of an "auditory shot" or unit of sound montage emerged as a neutral, universally recognizable unit.

Possibility of Inaudible Sound Editing "

On one hand, as we know, a film's "soundtrack" often consists of several layers independently recorded and mixed, which then overlap one another. Imagine a film resulting from mixing three layers of images in superimposition: only with great difficulty could one locate cuts. (This is the case in certain moments of Gance's *Napoleon* and Vertov's *Man with a Movie Camera*.)

On the other hand, the very nature of recorded sound events allows us to join one recorded sound with another in editing

without the join being noticed. A film dialogue can be crawling with inaudible splices, impossible for the listener to detect. While, as we know, it is very difficult to invisibly join two shots filmed at different times—the cut jumps to our eyes. (Hitchcock made *Rope* look like a "one-shot film" by a simple trick, shooting the back of a character or a dark piece of furniture, so darkness fills the frame at the end and beginning of each reel.) Also, of course, auditory cuts can be quite distinctly heard. So both audible and inaudible editing are possible with sound.

In current practice the mixing of soundtracks consists essentially in the art of smoothing rough edges by degrees of intensity. This fact in itself already makes it impossible to adopt any unit of sound editing as a unit of perception or as a unit of film language.

However, some people view current practices not as "natural" but as the embodiment of a particular ideological and aesthetic position characteristic of the dominant cinema, conforming to the desire to bury the traces of work in order to give the film an appearance of continuity and transparency. Many analyses of this sort appeared in the sixties and seventies, invariably concluding with the call for a cinema of demystification based on discontinuity. Very few directors actually answered the call except Godard. Godard was one of the rare filmmakers to cut sounds as well as images, thereby accentuating jumps and discontinuities, in greatly restricting inaudible editing with its gradations of intensity and all the fades, dissolves, and other transitions always employed in editing sound in film.

Does an Audible Slice of Sound Make a "Sound Shot"?

Godard unmasks conventional sound editing all the more in the way he avoids the usual practice of mixing many tracks at once

such that our attention is not grabbed by breaks and cuts in the sonic flow—in some of his films he limits the tracks to two. The result is that our attention can follow the thread of the sonic discourse, and it can hear unadorned all the ruptures, since the latter are made audible. Godard's films set up the most frank and radical conditions to apprehend what could be called a sound shot.

For example, in the beginning of *Hail Mary* (*Je vous salue Marie*) we can plainly hear the cuts that demarcate the slices of sound: a fragment of a Bach prelude played on a piano, shouts of a women's basketball team playing in a gym, offscreen voices, and so forth. For the listener, however, these perfectly demarcated sound slices do not add up to create a sense of units. Sound perception, which always occurs in time, merely jumps across the obstacle of the cut and then moves on to something else, forgetting the form of what it heard just before. The sound segment, especially if it lasts any time at all, does not synthesize into any particular bloc or totality in our perception.

Note that the same holds true for visual shots when they involve constantly mobile framing. Vision under these conditions occurs more along the flow of time, since it has no stable spatial referent. In the case of a sequence composed of static (or less constantly moving) shots, we can identify each shot by a certain composition, mise-en-scene, and perspective, and so we find it easy to represent this spatial arrangement in our memory.

On the other hand—even in the case of a stable sonic background cut up into little fragments as by Godard—it is inevitably sequential, temporal perception that still dominates for sound, at least for sounds of some duration.

And, above all, you cannot create an abstract and structural relationship between two successive sound segments (e.g., a fragment of bird calls or of music) the way you can between shots (a

character looks offscreen—cut to what he is looking at; or an establishing shot—cut to a detail in the scene). If you try something like this with the soundtrack, the abstract relation you wish to establish gets drowned in the temporal flow. What strikes the listener instead is the dynamics of the break itself between the two fragments. The explanation of this mystery is that when we talk about a shot we are lumping together the shot's space and its duration, its spatial surface and its temporal dimension. While for sound pieces the temporal dimension seems to predominate, and the spatial dimension not to exist at all.

So that when the audiovisual contract is in force, governing the copresence of visual and auditory channels, visual cuts continue to provide the reference point for perception. If Godard's sound cuts "fracture" the shot's continuity, as some scholars poetically put it, they're hardly creating more than a hairline crack in a glass pane that remains essentially intact.

Necessary Conditions for a Place of Sounds

The example of *Hail Mary* is interesting for additional reasons. Godard imposed the rule not to use more than two audio tracks at any given time, as a personal constraint on himself; but the spectator is not thereby automatically aware of the two separate tracks. In fact, the only way really to notice these two tracks would be to assign each one to a different spatial source in the theater. With each attached to its own loudspeaker one would get the feeling of a real *place of the sound*, of a sonic container of sounds. Not only would the sounds have to issue from a source clearly distinct from the auditory space of the screen but in addition they would have to avoid any synchronizing with the visuals in order not to fall prey to the effect of spatial magnetization by the image (see chapter 4), which is generally the stronger!

Units, But Not Specific Ones

Does this mean that a film's soundtrack constitutes a continuous flow without breaks for the listener? Not at all, for we can still discern units. But such units—sentences, noises, musical themes, "cells" of sound—are exactly of the same type as in everyday experience, and we identify them according to criteria specific to the different types of sounds heard. If the scene has *dialogue*, our hearing analyzes the vocal flow into sentences, words—hence, linguistic units. Our perceptual breakdown of noises will proceed by distinguishing sound events, the more easily if there are isolated sounds. For a piece of *music* we identify the melodies, themes, and units of rhythmic patterns, to the extent that our musical training permits. In other words, we hear as usual, in units not specific to cinema that depend entirely on the type of sound and the chosen level of listening (semantic, causal, reduced).

The same thing obtains if we are obliged to separate out sounds in their *superimposition* and not in their succession. In order to do so we draw on a multitude of indices and levels of listening: differentiating masses and acoustic qualities, doing causal listening, and so on.

This explains why the specifically cinematic visual unit of the shot remains by far the most salient, and why the composition of the soundtrack is subordinate to the shot.

Sonic Flow: Internal Logic and External Logic

The flow of a film's sound is characterized by how well related, how fluidly or imperceptibly connected the various sound elements are, and whether they are successive and superimposed—or on the contrary, whether they have degrees of discontinuity, and are punctuated by breaks that interrupt one sound suddenly with another.

The spectator's general impression of sonic flow will result not from characteristics of editing and mixing conceived separately but from all the elements combined. Jacques Tati, for example, uses extremely pointed sound effects, recorded separately and inserted into the soundtrack's continuum at specific places. Heard in succession, they would make for a halting and fragmented soundtrack if not for his use of continuous background sounds to tie the whole thing together. Think for example of his "phantom" background voices (playing on the beach in *Mr. Hulot's Holiday*, hawking goods in the market in *Mon Oncle*), which act as connective tissue, nicely concealing the breaks that inevitably occur when one constructs an extremely fragmentary and discontinuous soundtrack.

I shall call *internal logic* of the audiovisual flow a mode of connecting images and sounds that appears to follow a flexible, organic process of development, variation, and growth, born out of the narrative situation itself and the feelings it inspires. Internal logic tends toward continuous and progressive modifications in the sonic flow, and makes use of sudden breaks only when the narrative so requires. I shall call *external logic* that which brings out effects of discontinuity and rupture as interventions external to the represented content: editing that disrupts the continuity of an image or a sound, breaks, interruptions, sudden changes of tempo, and so on.

Films like Ophüls's *Earrings of Madame de*, Fellini's *La Dolce Vita*, or Randa Haines's *Children of a Lesser God* adopt an internal logic. The sound swells, dies, reappears, diminishes, or grows as if cued by the characters' feelings, perceptions, or behaviors. Films such as Scott's *Alien*, Lang's *M*, or Godard's *Nouvelle Vague* obey an external logic, with marked effects of transitions and breaks.

Using external logic does not necessarily mean achieving critical distanciation—as critics often propose with regard to

Godard. In *Alien*, for example, the frequent jolts to sound continuity and the jerky progression of the visual and sound tracks—characteristic of external logic—serve to reinforce the tension of the action. It is true that in this case we have a science fiction film where audio transmission by radio and phone, with their unpredictable fading in and out, is itself present as a concrete element in the screenplay, and directly motivates many of these effects. (For example, we see characters throw switches, turn on video screens and work at control panels, thereby acting as manipulators of sounds and images themselves.) Generally speaking, the modern action-adventure film engages external logic quite often.

But in a contemplative film like *The Goalie's Anxiety at the Penalty Kick* by Wenders, sound as well as image use external logic in response to a wholly different "literary" impulse, involving an existential fragmentation into "impressions," little sensory haikus.

SOUND IN THE AUDIOVISUAL CHAIN

Unification

The most widespread function of film sound consists of unifying or binding the flow of images. First, in temporal terms, it unifies by bridging the visual breaks through sound overlaps. Second, it brings unity by establishing atmosphere (e.g., birdsongs or traffic sounds) as a framework that seems to contain the image, a "heard space" in which the "seen" bathes. And third, sound can provide unity through nondiegetic music: because this music is independent of the notion of real time and space it can cast the images into a homogenizing bath or current.

This function of the unifying sound bath, where sound temporally and spatially overflows the limits of shots on the screen,

has come under attack by what one might call the differentialist school of criticism, which favors the idea of sound and image working in separate zones. Strangely, this approach neglects to criticize the same impulse toward unity when it is applied to the *image*. I am referring to the pursuit of visual continuity that prevails for cinematography in almost all films, whether silent or sound (including the films of Godard, Duras, and Syberberg), and that takes great pains with matching and balance of light and of color to make a well coordinated whole. Upon seeing a film consisting of four hundred to five hundred shots, would we be ready to perceive it as a succession of five hundred perfectly distinct units, as some experimental filmmakers have attempted?

Punctuation

The function of punctuation in its widest grammatical sense (placement of commas, semicolons, periods, exclamation points, question marks, and ellipses, which can not only modulate the meaning and rhythm of a text but actually determine it as well), has long been a central concern of theater directing. The play's text is approached as a sort of continuum to be punctuated with bits of stage business already indicated to some degree by the stage directions but also worked out during rehearsal: pauses, intonation, breathing, gestures, and the like.

The silent cinema adopted the traditional methods of punctuating scenes and dialogues (for the cinema did, after all, have dialogues). And it also naturally borrowed narrative techniques from opera, which used a great many punctuative musical effects by drawing on all the resources of the orchestra.

The silent cinema had multiple modes of punctuation: gestural, visual, and rhythmical. Intertitles functioned as a new and specific

kind of punctuation as well. Beyond the printed text, the graphics of intertitles, the possibility of repeating them, and their interaction with the shots constituted so many means of inflecting the film.

So synchronous sound brought to the cinema not the *principle* of punctuation but increasingly subtle means of punctuating scenes without putting a strain on the acting or the editing. The barking of a dog offscreen, a grandfather clock ringing on the set, or a nearby piano are unobtrusive ways to emphasize a word, scan a dialogue, close a scene.

Punctuative use of sound depends on the initiative of the editor or the sound editor. They make decisions on the placement of sound punctuation based on the shot's rhythm, the acting, and the general feel of the scene, working with the sounds imposed on them or chosen by them. (In rare cases, the director makes such decisions him or herself, and some sound punctuation is already determined at the screenwriting stage.)

Naturally, music can play a major punctuative role. It certainly did in the silent era, but in a less precise, more approximate way, owing to the much looser methods of synchronizing music with image. And so it is not very surprising if certain early sound films dared to employ music in an unabashedly punctuative manner. John Ford's *The Informer*, with its score by Max Steiner, provides a good example.

Music as Symbolic Punctuation: *The Informer*

Hailed as an event in the film world on its release in 1935, John Ford's *Informer* appeared for many years on numerous Best Ten lists of all-time greatest films. Today, Fordians no longer tend to pay much attention to it at all. Its expressionism jars with the serenity and loftiness of vision they normally associate with their favorite director. However, despite the way it has aged (its prin-

cial weakness resides in the writing and performances of the female parts), it remains a very strong work.

Curiously enough, *The Informer* does not hold its place in film history so much for its own merits as for the legendary story of some gulps of beer.

Not only in the West—to paraphrase a line from *The Man Who Shot Liberty Valance*—do they print the legend when it is more popular than reality; it's the same everywhere. How are legends born? Out of a lack, a void that needs filling. Legends get going only when room has been made for them.

The same goes for the anecdote, often cited as the height of ridiculousness in film music aesthetics, about a drinker's swallowing that *The Informer's* composer, in a frenzy of mickeymousing, went as far as to imitate musically. As early as 1937 we find the composer Maurice Jaubert claiming in a frequently quoted article: "In *The Informer*, the technique [of synchronism] is carried to its highest point of perfection, in the cue which is supposed to imitate the sound of coins falling to the ground, and even—with a suggestive little arpeggio—the trickling of a glass of beer down the gullet of a drinker."³

We can trace the lineage of this story through books on film music—including my own *Le Son au cinéma*, where, not having yet seen the film, I followed tradition and passed it on (which one should never do).⁴ But when Ford's biographer Tag Gallagher tells the same anecdote and uses almost the same words ("beer gurgling in a man's throat"), I doubt whether he too is indebted to Jaubert. We are witness here to the spontaneous formation of a legend.

The reality that can be verified by watching a tape of the movie is that there is indeed a moment during which the protagonist is drinking, accompanied by a musical theme. But, first, it's a glass of whiskey and, second, the music at that moment couldn't be less

imitative. Far from being a so-called suggestive little descending arpeggio, it is actually a resolute melody played on a solo French horn, ending with an upward jump of a diminished fifth, conveying something at once heroic and interrogative. It just can't be taken for gurgling. In fact the viewer recognizes it as one of the score's main themes—it is the first five notes of Gypo's theme, which has already been heard during the opening credits music. This motif accompanies the hero and is wed to his fate throughout the film, in an expressive way more than an imitative one.

Max Steiner based his *Informer* music on a principle that would subsequently dominate nine out of ten film scores—the principle of the leitmotif. Each main character or key thematic idea of the narrative is assigned a musical theme, which characterizes the character or idea and acts as its musical guardian angel. In *The Informer* the principal themes belong to Gypo (expressively rather neutral, and rather energetic and *marcato*, evoking Irish popular song) and to Katie, the prostitute-with-a-heart-of-gold whom Gypo loves (*espressivo* and *legato*). Then there is a special motif for the symbolic character of the Blind Man (a plaintive melody evoking the formal indeterminacy of Debussy). These musical themes are heard frequently in the orchestral score as "their" characters appear; they undergo changes that reflect variations in the characters' external circumstances and internal states. The theorizing and systematizing of the leitmotif of course goes back to Wagner, but if there is one opera in particular that inspired Steiner for *The Informer*, it must be Debussy's *Pelleas et Melisande*. Despite his own caustic denunciation of leitmotif technique, Debussy used it himself, trying to make it more subtle by using more laconic, less pompous themes. Other Debussyesque touches in *The Informer* (as we shall see below) are an insistence on silences and a penchant for interruptions—when music halts and speech interposes itself into the vacuum.

The tendency here is to try to make a sound film into spoken opera, and in this endeavor we can see the hand of John Ford as well as that of the composer. *The Informer* certainly is not the kind of film that a composer was unleashed to slap a coat of musical paint onto after the shooting. The director, set designer, and composer worked together from the start, and judging from their own words, their preliminary consultations were much more extensive than is usual. Therefore, not only did Ford go along with the musical choices made for *The Informer* but he gave them his stamp of approval, perhaps even made his own suggestions.

The cornerstone of *The Informer's* style is a predilection for stylization and symbolic expression—at the heart of a cinema that had just suffered the assaults of naturalism upon the coming of sound. This stylization obviously seeks to recapture the spirit of the silent film, perhaps even to achieve what the silent film could only dream of. The orchestra's punctuation of gestures and dialogue aims to undermine their purely realist and concrete aspects in favor of making them into signifying elements in an overall mise-en-scene.

In fact, Max Steiner's score hardly ever imitates the immediate materiality of the events; at least it does so much less than the great majority of film scores past and present. It does not emphasize the sounds of the elements, doors closing or bodies falling, and when it is heard with a particular movement on screen, its contours carefully *avoid* imitating the visual movement's form. Consider for example the scene to which Jaubert alludes from memory.

Let us recall that Gypo, the brutish outcast, has just turned in his friend Frankie, an Irish independence fighter wanted by the police, and he has gotten the reward money. He soon finds out that Frankie was killed on arrest. When he enters a bar, orders a whiskey, and raises his head to drink, it's because he is trying to forget. Gypo's musical theme plays at this moment, suggesting that even in his absence-to-himself, his identity insistently hangs

on. Gypo, the grieving partner in the couple he formed with Frankie—who treated him with affectionate condescendence, as if he were the brain and Gypo the body—the incomplete Gypo, will find himself only by sacrificing himself, and the film tells the story of his coming to consciousness.

Already in Wagner's work there are themes in the orchestral fabric that embody a character's unconscious, giving voice to what the character does not know about himself. For example, in the first act of the *Walkyrie*, there is the sword motif, which (through the orchestra) works on Siegmund's unconscious, even before he finds the weapon in the hut where he has taken refuge. The barroom scene in *The Informer* also seems replete with a sense of meditation, of the stirring of dark things, and of preparation. It conveys something quite similar to Wagner's opening act, which has a very fragmented and discontinuous musical fabric that includes stops, reprises, and silences.

When the barkeeper returns the change to Gypo, four instrumental notes punctuate the coins' falling. We might be tempted to consider the synchronism silly; but we must not forget that these coins are the money of betrayal, the money of Judas. The fall of the four coins is fatal: they open a running account that will culminate in the informer's condemnation. (It is by adding up all Gypo's expenditures and realizing their total equals the amount of the reward for Frankie that the independence fighters confirm their suspicions about Gypo.) All Max Steiner does here is adopt a common technique from opera that uses music for expressive symbolization of actions. Note first that the music does not *substitute* for the sound of the falling coins—the coins are heard diegetically at the same time (more or less). Second, the melodic shape here is precisely that of one of the score's leitmotifs, the betrayal motif.

So what is there about these musical interventions that nevertheless feels like imitation? The answer is that they are punctual

and synchronous. As we shall see below (and as I wrote in *La Toile trouee*, in the chapter on "the clapboard"), synchronization is an important factor in film in how it manages to glue together entirely unlikely sounds and images.⁵ In opera the frequent synchronizing of music and action poses no problem, since it is an integral part of an overall gestural and decorative stylization. In the cinema such synchronization must be handled more discreetly, so as not to be taken as exclusively imitative or slip over into the mode of cartoon gags. And if *The Informer's* attempts at operatic stylization have aged ungracefully in using this overt synchronism, we can still recognize the film for its daring and honesty. The cinema is a realist art: but it remains that this realist art has progressed only by means of straining against its own principle, through forceful doses of unrealism.

What makes the use of music unique in certain scenes of *The Informer*, what gives it a quality that might seem tedious? It's the way the music stops. Music cues do not get eclipsed by such things as doors opening or closing, as they would learn to do later in the thirties. Instead, music in *The Informer* often gets interrupted bluntly and suddenly, in mid-phrase, producing a silence in which the subsequent dialogue resonates strangely. By stopping and passing the baton to speech, it seems as if the music is pointedly referring to itself rather than remaining unobtrusive. Henceforth the cinema would tend to avoid this abrupt and conspicuous transition, preferring a more fluid relationship wherein music and scene were mixed more thoroughly together; music became more constant and at the same time more indistinct.

Punctuation: Elements of Auditory Setting

I call *elements of auditory setting* (E.A.S.) sounds with a more or less punctual source, which appear more or less intermittently and

which help to create and define a film's space by means of specific, distinct small touches. Typical sounds of the auditory setting are the faraway barking of a dog, or the ringing of a phone in the office next door, or a police car siren. The E.A.S. inhabits and defines a space, unlike a "permanent" sound such as the continuous chirping of birds or the noise of ocean surf that is the space itself.

Beyond its narrative role (establishing or reestablishing the framework of the action), the E.A.S. can also have a punctuative role, thanks to editing. It can be used with intelligence to help create the scene's overall rhythm; this possibility can renew and transfigure its functions completely.

The multiplicity of functions of elements of auditory setting reminds us that soundtrack analysis must constantly take into account the likelihood of overdetermination: that is, a filmic element often signifies on several different levels at once.

Anticipation: Convergence / Divergence

From a horizontal perspective sounds and images are not uniform elements all lined up like fenceboards in a row. They have tendencies, they indicate directions, they follow patterns of change and repetition that create in the spectator a sense of hope, expectation, and plenitude to be broken or emptiness to be filled. This effect is best known in connection with music. Musical form leads the listener to expect cadences; the listener's anticipation of the cadence comes to subtend his/her perception. Likewise, a camera movement, a sound rhythm, or a change in an actor's behavior can put the spectator in a state of anticipation. What follows either confirms or surprises the expectations established—and thus an audiovisual sequence functions according to this dynamic of anticipation and outcome.

One of the most persistent players of this game is Godard, in

his *Letter to Freddy Buache*. Ravel's *Bolero*, the music he selected to accompany the whole video, is a vast melodic curve that prepares for but ceaselessly defers its cadence, like a delayed orgasm—while for its part Godard's voiceover commentary takes a malicious pleasure in fishing about for the right words, thus making us wait for them—meantime constant pans in the image leave us to imagine, at the end of the paths they trace through urban or bucolic landscapes, who knows what revelation.

In an audiovisual sequence the audio-viewer consciously or unconsciously recognizes the beginnings of a pattern (e.g., a crescendo or an accelerando) and then verifies whether it evolves as expected. It is often more interesting when the expectation is subverted. And, at other times, when everything ensues as anticipated, the sweetness and perfection of the rendering of the anticipation itself are sufficient to move us.

In *Children of a Lesser God*, just after William Hurt has left the dance and walks out into the night air he turns around to see Marlee Matlin, all dressed in white, coming to join him. The volume of the music from the dance gradually decreases, faded out by the mixer. Consciously the spectator expects the two characters to meet; less consciously, for the music to disappear when the lovers join—for there to be silence when they touch. This indeed is what happens, the convergence of a joining and a disappearance, but so precisely and subtly executed that we are always moved when the disco music fades into silence, and the reunited couple become still, all in one breath.

We never stop anticipating, and surprising anticipation—for this is the movement of desire itself.

Separating: Silence

In a well-known aphorism Bresson reminded us that the sound film made silence possible. This statement illuminates a paradox:

it was necessary to have sounds and voices so that the *interruption* of them could probe more deeply into this thing called silence. (In the silent cinema, everything just suggested sounds.)

However, this zero-degree (or is it?) element of the soundtrack that is silence is certainly not so simple to achieve, even on the technical level. You can't just interrupt the auditory flow and stick in a few inches of blank leader. The spectator would have the impression of a technical break (which of course Godard used to full effect, notably in *Band of Outsiders*). Every place has its own unique silence, and it is for this reason that for sound recording on exterior locations, in a studio, or in an auditorium, care is taken to record several seconds of the "silence" specific to that place. This ambient silence can be used later if needed behind dialogue, and will create the desired feeling that the space of the action is temporarily silent.

However, the impression of silence in a film scene does not simply come from an absence of noise. It can only be produced as a result of context and preparation. The simplest of cases consists in preceding it with a noise-filled sequence. So silence is never a neutral emptiness. It is the negative of sound we've heard beforehand or imagined; it is the product of a contrast.

Another way to express silence, which might or might not be associated with the procedures I have just described, consists in subjecting the listener to . . . noises. But I mean here the subtle kind of noises like the ticking of an alarm clock, naturally associated with calmness. These do not attract attention; they are not even audible unless other sounds (of traffic, conversation, the workplace) cease.

We find a good example in *Alien*, when Ridley Scott, with a closeup of the cat who is the space ship's mascot, wants to create the impression of a disconcerting silence to precede sinister developments. The shots leading up to this are sonically rich, preparing for the void that follows, but care has been taken that

the silence doesn't strike too suddenly. During the first three seconds of the shot of the cat we can hear a small unidentified sound, like a tick-tock. Its presence on the soundtrack, then its rapid fadeout, help form a bridge to total emptiness.

In *Face to Face* Bergman gets a very different effect by a reverse treatment of the same kind of ticking sound. A woman, deeply depressed, is at home getting ready for bed. The ticking of the alarm clock on the night table, which had previously gone unnoticed, becomes louder and louder. Paradoxically we end up with an anxiety-producing impression of silence, all the stronger because the only sound there is is so intense, and heightened by the lack of other sounds, bringing out this emptiness in a terrible way. (Bergman's personal touch can be noted here in the swift precision with which the sound is augmented.)

Film uses other sounds as synonyms of silence: faraway animal calls, clocks in an adjoining room, rustlings, and all the intimate noises of immediate space. Also, and somewhat strangely, a hint of reverberation added to isolated sounds (for example, footsteps in a street) can reinforce the feeling of emptiness and silence. We cannot perceive reverb like this when other sounds (e.g., daytime traffic) are heard at the same time.

POINT OF SYNCHRONIZATION AND SYNCHRESIS

A *point of synchronization*, or synch point, is a salient moment of an audiovisual sequence during which a sound event and a visual event meet in synchrony. It is a point where the effect of synchresis (see below) is particularly prominent, rather like an accented chord in music.

The phenomenon of significant synch points generally obeys laws of gestalt psychology. A synch point sometimes emerges more specially in a sequence:

- As an unexpected double break in the audiovisual flow, a synchronous cut in both sound and image track. This is characteristic of external logic, frequent in *Alien* for example.
- As a form of punctuation at the end of a sequence whose tracks seemed separate until they end up together (synch point of convergence).
- Purely by its physical character: for example when the synch point falls on a closeup that creates an effect of visual fortissimo, or when the sound itself is louder than the rest of the soundtrack.
- But also by its affective or semantic character: a word of dialogue that conveys a strong meaning and is spoken in a certain way can be the locus of an important point of synchronization with the image.

A point of synchronization can stage the meeting of elements of quite differing natures. For example, a visual cut can be coordinated with a word or group of words specially emphasized by the voiceover commentary. In Godard's *Letter to Freddy Buache* several meetings between visual cuts and ends of sentences provide the principal synch points on which the architecture of the whole film is based. The synch point is indeed the place where the audiovisual "arch" meets the ground before taking off again.

Synch points naturally signify in relation to the content of the scene and the film's overall dynamics. As such, they give the audiovisual flow its phrasing, just as chords or cadences, which are also vertical meetings of elements, can give phrasing to a sequence of music.

There is also the particular case we shall call the *false synch point*. The deceptive or false cadence in Western classical music is a cadence that a particular harmonic progression sets us up for

but that does not resolve as anticipated. In much the same way audiovisual texts have false synch points. They can be more striking than synch points that actually do occur, because the audio-spectator has been given them to fabricate mentally. The best-known example is the suicide scene of a corrupt and compromised official in John Huston's *The Asphalt Jungle*. We see him lock himself into his office, open a drawer, take out the gun; then, we hear the gunshot only, because the visual editing transports us elsewhere. For reasons of taste? Not entirely . . .

Godard scatters the seeds of false synch points at the beginning of *Hail Mary*. We hear repeated plops, and in the image we see only the surface of a lake rippled by something fallen into it; the fallen object, and the place of the fall, remain offscreen. So the cause is heard on the soundtrack while its consequences shimmer in the image. But in the spectator's mind is a synch point all the more disturbing for being postulated but not actualized (into hearing and seeing what falls); the object in question, acousmatic, may be anything one wishes: a stone someone has tossed, a meteorite, or the Holy Spirit.

Emblematic Synch Point: The Punch

In real life a punch does not necessarily make noise, even if it hurts someone. In a cinematic or televisual audio-image, the sound of the impact is well-nigh obligatory. Otherwise no one would believe the punches, even if they had really been inflicted. Accordingly they are accompanied with sound effects as a matter of course. This punctual, momentary, abrupt coincidence of a sound and a visible impact thus becomes the most direct and immediate representation of the audiovisual synch point, in its quality as the sequence's keystone, punctuation, Lacan's *point de capiton*. The punch becomes the moment around which the narra-

tion's time is constructed: beforehand, it is thought about, it is announced, it is dreaded; afterward, we feel its shock waves, we confront its reverberations. It is the audiovisual point toward which everything converges and out from which all radiates. And it is also the privileged expression of instantaneity in the audio-image.

The ultrabrief image of the punch all by itself would not become engraved into the memory, would tend to get lost. But an ultrabrief but clearly delineated *sound* has the advantage of etching its form and tone directly into consciousness, where it can repeat as an echo. Sound is the rubber stamp that marks the image with the seal of instantaneity. (On the metaphor of the stamp recall the gag of the librarian in Spielberg's *Indiana Jones and the Last Crusade*.)

What is the most important object in audiovisual representation? The human body. What can the most immediate and brief meeting between two of these objects be? The physical blow. And what is the most immediate audiovisual relationship? The synchronization between a blow heard and a blow seen—or one that we believe we have seen. For, in fact, we do not really see the punch; you can confirm this by cutting the sound out of a scene. What we hear is what we haven't had time to see.

Accented Synch Points and Temporal Elasticity

This structure is already present in much nonanimation cinema, notably in all martial arts and fight films. But Japanese animated films I can see on French television add something more: an analysis of movement (as in Muybridge and Marey's famous photos, which lie at cinema's origins), the use of slow motion and radical stylization of time. These diverse techniques get their inspiration from the slow-motion and still-frames of sports

replays, but also directly from Japanese comic books, or *mangas*. In these rudimentary animated adventures the point of synchronization constituted by the punch, this point of hooking auditory continuity to visual continuity, is what allows the time around it to swell, fold, puff up, tighten, stretch or, on the contrary, to gape or hang loosely like fabric. On either side of a characteristic synch point such as a punch the capacity for temporal elasticity can become almost infinite. During an episode of the series *Dragon Ball* the battling characters constantly freeze in mid-motion, stop in mid-air (for they make incredible leaps), and converse interminably, slowing down, speeding up and changing poses like a series of discontinuous slides, before launching flurries of swift punches and kicks to one another.

In short, the punch with sound effects is to audiovisual language as the chord is to music, mobilizing the vertical dimension. In the brutal and exhausting boxing scenes in *Raging Bull* Scorsese used punches to bestow a maximum degree of temporal elasticity on the fighting scenes; thus he could use slow motion, repeated images, and so forth.

The paradox is that in the beginning, temporal elasticity was an inherent characteristic of the silent cinema. Since the silents did not have to be dubbed point by point and second by second with synchronous sound, they could easily dilate and contract time. With the arrival of sound this elasticity began to disappear from the sound film. And though it may have succeeded in entering into a few realist films, note that it shows up most insistently in action and combat sequences, such as Sam Peckinpah's. In other words, temporal flexibility was not introduced into a sound-image relationship that is vague and asynchronous but, on the contrary, it appears in scenes that have strong points of synchronization, where blows, collisions, and explosions can serve as strong reference points.

Synchresis

Synchresis (a word I have forged by combining *synchronism* and *synthesis*) is the spontaneous and irresistible weld produced between a particular auditory phenomenon and visual phenomenon when they occur at the same time. This join results independently of any rational logic. Synchresis is responsible for our conviction that the sounds heard over the shots of the hands in the prologue of *Persona* are indeed the sounds of the hammer pounding nails into them.

Synchresis is what makes dubbing, postsynchronization, and sound-effects mixing possible, and enables such a wide array of choice in these processes. For a single body and a single face on the screen, thanks to synchresis, there are dozens of allowable voices—just as, for a shot of a hammer, any one of a hundred sounds will do.

Certain experimental videos and films demonstrate that synchresis can even work out of thin air—that is, with images and sounds that strictly speaking have nothing to do with each other, forming monstrous yet inevitable and irresistible agglomerations in our perception. The syllable *fa* is heard over a shot of a dog, the sound of a blow with the sight of a triangle. Synchresis is Pavlovian,

But it is not totally automatic. It is also a function of meaning, and is organized according to gestaltist laws and contextual determinations. Play a stream of random audio and visual events, and you will find that certain ones will come together through synchresis and other combinations will not. The sequence takes on its phrasing all on its own, getting caught up in patterns of mutual reinforcement and phenomena of "good form" that do not operate by any simple rules. Sometimes this logic is obvious. When there is a sound that is louder than the others, it coagulates with the image it is heard with more strongly than previous or

subsequent images and sounds. Meaning and rhythm can also play important roles in securing the synchresis effect.

For some situations that set up precise expectations of sound—a character walking, for example—synchresis is unstoppable, and we can therefore use just about any sound effects for these footsteps that we might desire. In *Mon Oncle* Tati drew on all kinds of noises for human footsteps, including ping-pong balls and glass objects.

The effect of synchresis is obviously capable of being influenced, reinforced, and oriented by cultural habits. But at the same time it very probably has an innate basis, as shown by the limit experiments of video artist Gary Hill.⁶ Specific reactions to synchronized aural and visual phenomena have been observed in newborn infants.

The "modest" phenomenon of synchresis—modest because it is so common—opens the floodgates of sound film. Synchresis permits filmmakers to make the most subtle and astonishing audiovisual configurations. But today, when it has become commonplace to see a moving figure on a screen as we hear movement, it is difficult to imagine the amazement inspired by the first synchronized sound films in the twenties. That sound and image were heard and seen like a couple of perfectly matched dancers was a spectacle in itself. Texts written at the time bear witness to this state of mind. So do the films, especially the musical ones, which exalted synchronism as such, showing violinists or banjo players whose every visual gesture provoked a distinct sound on the soundtrack.

Habit has led us to consider this phenomenon "natural" and devoid of cinematic interest. But let us rediscover it.

Loose and Tight Synchronisms

Synchresis does not function in an all-or-nothing way. There are degrees of synchronism and, particularly in the case of lip synch,

these "degrees" play a part in determining film style. For example, the French, who are accustomed to a tight and narrow synchronization, find fault with the postsynching of Italian films. What they are objecting to in reality is a looser and more "forgiving" synchronization that's often off by a tenth of a second or so. This difference is particularly noticeable in the case of the voice. While very tight synch holds voices to lip movements, Italian films synch more loosely, taking into consideration the totality of the speaking body, particularly gestures.

In general, loose synch gives a less naturalistic, more readily poetic effect, and a very tight synch stretches the audiovisual canvas more... this canvas, the status of whose landscape or "scene" we now shall proceed to investigate.