

Epilogue

Audition and Ergo-Audition: Then and Now

Michel Chion

The question of the relationship between making sounds and hearing them is secular: does the person who initiates a sound (the musician, the machine or hi-fi operator, the craftsman, etc.) hear it more clearly than the person who is simply present and hears it without intervening, not so clearly, or differently? The parameters of the question have shifted owing to the emergence of new types of interactivity. Which is why I find it pertinent to review the fundamental aspects of what I call “ergo-audition.”

In 1998, in a publication entitled *Le Son* (Sound),¹ I proposed to baptize the “audition particular to whoever is in some way simultaneously the producer of the sound heard or has a measure of influence over that sound” as “ergo-audition.” It was necessary to invent a new, specific term, because ergo-audition has its own laws, different from those of simple audition (i.e., hearing without intervention in a sound)—not that the latter may be deemed “passive” on the pretext that “audition” is invisible, as listening motionless to music or to a person speaking to his psychoanalyst can in fact mobilize the whole individual and the best of his capacities.

“Ergo-audition” thus describes a situation in which the hearer is simultaneously responsible—totally or partially, consciously or not—for the sound he hears when playing an instrument, operating a machine or vehicle, pouring liquid into a container, or producing noise by means of his footsteps, clothing, movements, or actions, as well as when he speaks. This can cover a broad range of situations, including the lesser-known instance of tinnitus.

A Singular Case of “Ergo-Audition”: Tinnitus

Tinnitus had long been a far-off, abstract concern for me. Until about four years ago, when during a period of stress and fatigue I started to hear constant whistling sounds in my ears. They have since never ceased and probably never will again, even if most of the time I very fortunately forget about them. In the beginning it was harrowing, but I have gradually learned to live with the condition, thanks to the help of a psychiatrist and antidepressants. During that time I also learned through exchange with other people suffering the same symptoms that there are many different kinds of tinnitus. Mine is double (Jean-Jacques Rousseau’s was triple or quadruple); in my case, it consists of a stable medium-pitched note, bordering on a G-sharp and, above all, a strident high-frequency whistle vibrating at around 11.2 kilohertz. This whistling—described as “pulsating”—beats continuously to a precise rhythm: that of my heartbeat.

As the angiologist I consulted demonstrated by having me listen to the pulse of the artery in my neck, my acoustic phenomenon is exactly synchronous with that artery’s rhythm. Its pulsation accelerates and slows down in time with my heartbeat and, consequently, with my activities and my physical and nervous state. It is also influenced by other factors such as blood pressure, atmospheric pressure, and stress.

¹ Michel Chion, *Le Son* (Paris: Nathan, 1998), 84.

Now, my heartbeat is an internal rhythm over which I have limited but real control. I am not a yogi able to slow it down at will, but with physical activity I can easily accelerate it. Part of the suffering I felt at first derived from the fact that tinnitus subjected me to an acute high-frequency awareness of my heartbeat, something one generally prefers to ignore. It was as if I were being forced to be aware of my heart, and I was always saying, “There, it’s beating too fast,” “It’s beating more slowly,” “It could stop beating,” and so forth.

I also discovered that I could make the tinnitus almost completely disappear at will if I were to adopt a rather uncomfortable position, impossible to maintain for any length of time, with my head deeply inclined upon my chest, and my chin in contact with the upper bones of my sternum. In that position, probably owing to the change of contact between my veins and certain nerves, I no longer hear the shrill whistling sounds.

So these are the sounds that I am subject to, but which I can to some extent “play with”; they are presented to me in an interactive mode, as in fact are most of the sounds created by the environment. This means that as well as being responsible for my health and my sickness (sickness that I am supposed to have the means to prevent, at least partially, by leading a “healthy life”), I become responsible for the rhythm of my tinnitus, instead of being obliged to simply submit to it. In a way, it is something of a relief to know that. On the other hand, it is a new form of alienation and a fresh source of guilt—as if to say one only has the tinnitus one deserves.

What I mean to illustrate through this anecdote is that in our relationship to sounds there are many different situations, and that the term “listening” does not suffice to describe them. In this case, one must speak of “self-listening” or “listening-to-oneself,” insofar as it comprises a measure of “ergo-audition.”

Another Singular Case of Ergo-Audition: Listening to One’s Own Voice

Let’s take another situation, more banal, universal even, yet far more complex than it seems: listening to one’s own voice. We are supposed to be able to control the vocal sound we emit constantly by hearing it: “self-listening” or “listening-to-oneself” seems to endow us with the equivalent of a permanent mirror of our vocal activity; yet, at the same time—since the proliferation of sound recording by means of cassette recorders, answering machines, handycams, and the like—everyone knows that the voice others hear coming from us is different from the one we ourselves hear (which is, incidentally, always more highly pitched).

The difference with vision is that, independent of the mirror situation, the act of “seeing oneself” (i.e., of “knowing one is seen”) amounts to perceiving oneself through the eyes of another, or even through those of the object one is contemplating—as a finite body in space, in reference of course to the mirror stage. One may ask oneself whether an equivalent exists in the act of “hearing oneself.”

In fact, “hearing oneself” is of a quite different nature. Even if one does hear the “return,” as it is called, of one’s own voice directly from the distant loudspeaker—a familiar situation for lecturers and, given the acoustics in churches, even for preachers in the past—the fact that one hears oneself at the same time from within changes the information received: a continuum is instated—between “hearing (oneself) from within” (by means of co-vibration) and “hearing (oneself) from without” (by means of one’s ears, reflections off surrounding

surfaces, an amplified loudspeaker, etc.)—that relates one to the other. In the same way, all sound constantly links the interior and the exterior realms.

Our own voice is what it is because of a certain configuration of the larynx and certain hormonal factors (such as testosterone in men), but also because we are able to hear ourselves. The voices of deaf persons who have learned to “oralize” (in parallel with the acquisition of sign language, which fortunately is no longer proscribed) only have that peculiar timbre because those who produce the sound cannot hear themselves.

In *L'Oreille et le langage* (The Ear and Language), Alfred Tomatis nevertheless defended the—somewhat excessive—thesis that one can produce vocally only those sounds that one is able to hear. The use of this “audio-phonatory feedback” led him to treat actors and singers with vocal problems by making them hear themselves differently through headphones. They heard their voice filtered in a precise manner designed to bring them to “correct” their emission.

In his short essay *La voix et le phénomène* (Speech and Phenomena), published in 1967, French philosopher Jacques Derrida was right to emphasize the importance and the specificity of “hearing oneself speak.” However, he fails to question the strangeness of this experience which, in my opinion, he too quickly reads as a “seamless” experience of one’s presence. In his enthusiasm for this discovery, Derrida seems to overlook the complex nature of “hearing oneself speak,” which creates a relationship not only between will and effect but also between external and internal perceptions which, although they interrelate, do not combine.

We do in fact hear others speak before hearing “ourselves” speak, and their voices—which we imitate and incorporate, transposing them one or two octaves higher—prompt our own. Is there not here an element of that alienation already observed at the “mirror stage,” albeit in a different mode?

Yet, while the mirror image is—according to Jacques Lacan—totalizing because it is “totalizable” (one sees oneself whole), the image of “hearing oneself” is not only “non-totalizable” but also alienated in a will to speak. Our voice remains largely unfamiliar to us both when it is heard from the outside and when it is heard from within.

The mirror actually returns our image even when we remain motionless and inactive. The vocal mirror—if at all it does exist as such, though it is a common enough metaphor for hearing oneself—implies that one speaks and consequently projects a certain intention which prevents one from listening objectively (except in the case of professional singers, who incessantly correct by ear the sounds they produce).

Hearing Oneself with the Echo

While hearing one’s own voice return from the outside and deferred is a new experience, such phenomena have been encountered in certain natural sites—mountain passes, for example, that existed already two thousand years ago—in the form of an echo both sufficiently distant in time and sufficiently precise in the restitution of tonality for one to be truly able to hear one’s own voice from the outside, from a distance, and slightly deferred: not only the word pronounced, but also the timbre in which one uttered it. It is a strangely arresting experience, even in the age of the tape recorder.

Is the echo equivalent to the mirror? Certain psychoanalysts—Didier Anzieu, for example, in *Le Moi-peau* (The Skin Ego)—have not hesitated to answer this in

the affirmative, on the basis of the myth of Echo and Narcissus that associates both. The echo described in the myth presupposes temporal deferment. But the echo that is simultaneous—or virtually so in terms of the human temporal scale—does exist: it is the reflection we always have of our own voice returned by the environment—a reflection of which we become aware only when it is absent, which is rarely the case. The problem is that, for reasons that ensue from the physical nature of sound and audition, this sound mirror mingles with the “original.”

But to “hear oneself speak” from an external source and at a delay is of course an experience that has become progressively widespread since the beginning of the twentieth century. For a long time, the difference between how one hears one’s voice from within and how it sounds to others could be attributed to flaws in the recording process. Today, we know that the difference is the result of a fundamentally different listening position. In the 1950s, those who had already had a chance to hear their own voice from “the outside” were rare—and they usually failed to recognize it the first time. Today, millions of people (and in some countries almost everyone) have had an opportunity to experience it by means of answering machines, amateur video, and tape recorders. And the discontentment experienced on hearing one’s voice is always the same. Only those who hear themselves recorded frequently owing to their profession—actors, radio and television personalities, politicians—become accustomed to it, while often also acquiring a manner of speaking “for the mike.”

Ergo-Audition, Regulation, and Sounding

At work, the noise one produces may serve to regulate the efficiency of a hammer-blow, the correct progression of a sawing operation, the handling of a tool, one’s own footsteps. Variations on the “harmonic timbre” (the specific color related to the harmonics of a sound) of the noise produced when filling a vessel is a familiar guide for anyone who runs a coffee shop or is pouring a drink. It also guides the young blind girl praised by Diderot in his “Letter on the Blind,” of whom he says: “If one served her a drink, she knew from the sound of the liquor pouring, when her glass was full.”² She had acquired this perception, which is a good illustration of ergo-auditory feedback.

Whether emptying a bucket of coal into a stove (once a very familiar sound that has become rarer since the 1960s), pouring a liquid into a glass, or refilling a sugar bowl or salt cruet, it is always, on very diverse timescales, the audible change in harmonic timbre that serves the ergo-auditor as a reference. This subtle sensation is only one of the many examples of familiar, characteristic sound profiles which practically no one identifies and designates as such, yet which are nevertheless “archetypes,” quite universal “types.”

In the same way, in all latitudes, peeling a fruit or a vegetable produces a characteristic variation on harmonic timbre, which the peeler recognizes and by which he aurally verifies the stage of peeling he has reached—but which seems also to provide a characteristic oral satisfaction. The pleasure of ergo-audition, one’s consciousness of the sound’s response to what one is doing, could rank among motivations to undertake action. In particular, all action consisting of sounding a thing to find out whether it is empty or full could be partially motivated by ergo-auditory pleasure. The overdetermination here is interesting.

² Denis Diderot, “Letter on the Blind” in *Diderot’s Early Philosophical Works*, ed. and trans. Margaret Jourdain (New York: AMS Press, 1973).

In Virgil's *The Aeneid, Book II*, Laocoon wants to warn his companions not to let the famous Trojan Horse enter the town walls. To show them it is a trick, he throws a javelin at the wooden object

which, hissing as [it]flew,
Pierc'd thro' the yielding planks of jointed wood,
And trembling in the hollow belly stood.
The sides, transpierc'd, return a rattling sound³

Similarly, we take pleasure in the sound of our footsteps:

Over the heath my footsteps resound;
Pounding beside me from out of the ground.
—Theodor Storm, “Over the Heath”⁴

In *Idyll VII* by the Greek poet Theocritus, a person addresses the narrator who is hurrying along the road: “For 'faith, every stone i' the road strikes stinging against your hastening brogues.”⁵ Here, the sound of our footsteps is an expression of vitality, of happiness. It is also a pleasure to relate to the world, to the environment. But the stones' song belongs to them—it is their response; and this generates a particular pleasure.

Inversely, not to hear one's own footsteps is to lose one's footing. Mephistopheles says to Faust as the latter is preparing to journey into the region that is the birthplace of the Mothers:

You will see nothing in the eternally empty distance
You will not hear the footstep that you make.
—Johann Wolfgang von Goethe, *Faust II*⁶

Hearing Oneself “Doing”: Ergo-Auditory Pleasure in Excretory Activities

Ergo-audition also concerns “unsuitable” sounds that even in our modern and liberated times one cannot mention without embarrassment and significant giggles. Here we shall talk of the noises generated by our natural functions. As one Japanese poet wrote:

I pee on the
dead leaves
rustling
—Hōsha⁷

Noise is sometimes primitively associated with physiological acts: namely, urination and defecation—noises that are sometimes involuntary, socially taboo, and at times all the more embarrassing in that nothing distinguishes them from noises of a very different nature. There may be no difference at all between the

³ Virgil, *The Aeneid, Book II*, trans. John Dryden, <http://classics.mit.edu/Virgil/aeneid.2.ii.html>.

⁴ Theodor Storm, “Over the Heath” in *Hans and Heinz Kirch*, trans. Denis Jackson and Anja Nauck (London: Angel, 1999), 174.

⁵ Theocritus, *The Idylls (Idyll VII: Harvest Home)*, trans. J. M. Edmonds, <http://www.theoi.com/Text/TheocritusIdylls2.html#7>.

⁶ Goethe's *Faust, Part II*, trans. Leopold J. Bernays (London: S. Low and Karlsruhe: A. Bielefeld, 1839), <http://books.google.com/books?id=2VQHAAAAQAAJ&oe=UTF-8>.

⁷ Maurice Coyaud, ed., *Fourmis sans ombres. Le livre du haiku* (Paris: Phébus, 1978).

noise of a stream of water and that of a stream of urine; the difference being created not by the nature but by the quantity and force of the stream, as well as by the place where it falls.

All children, especially the little boy who can direct his flow, experiment with the fact that urinating in various spots and vessels makes different noises—as does tapping in the same way on various objects, shaking various toys in the same way, and so on—and this procures a particular feeling of pleasure. This ergo-auditory pleasure is associated or not with a feeling of relief or with physical exertion. While the relationship to ergo-audition is different for boys and girls when urinating, it is the same for both sexes when defecating.

For children, these games include both exploratory and functional aspects. The adolescent and the adult experience this ergo-auditory pleasure when they play with the controls of a motorbike to accelerate, when they noisily slap down their playing cards (or their mah-jongg tiles in Asian countries), and so on . . .

The Shining Effect

Ergo-auditory pleasure takes on a particular form when what alters the sound comes not from changes to the action itself but from the environment.

Since the release of Stanley Kubrick's film *The Shining* (1980), I have been struck by the thrilled attention that one of the scenes in the film always elicits: the scene in which a little boy in his pedal car rides through the endless corridors of the Overlook Hotel. The camera follows him at his height. When the car passes over carpeting, the noise changes and becomes muffled; then, when it returns again to the wooden flooring, there is a change in volume and tone, and so on. It is like when a train passes over a bridge and then regains the ballast on firm ground, except that here it is the child who is driving and propelling the vehicle with his little legs.

That is why, in homage to Kubrick, although it is also to be observed in several of Robert Bresson's films, I baptized as "the Shining effect" that thrill produced when an identical action generates different audio "responses" which activate ergo-auditory feedback and invite us to endless exploration of *how the world sounds*. Even the sound of our footsteps is not completely integrated by us as being a part of ourselves; it is to a certain extent the ground's, the world's response to our actions, and consequently triggers the ergo-auditory trap: the loop that ensnares us in its ever-changing audio responses. Here, noise becomes a symbol of control, but also a narcissistic trap.

The very fact that the sound produced corresponds, even and above all only partially, to the causal action creates a specific kind of feedback, the trap we are discussing: one can never completely ignore it although one would ignore it if, in two symmetrical cases, the sound were to dutifully reproduce the intention of the causal action or, inversely, the sound were to have only a random relationship with that act.

The Ergo-Auditory Loop

And so "the Shining effect" is related to the absence of any systematic correlation of cause and sound (or of an action and its result). The spell it casts is that of a mirage of the effect: of different effects, predictable or not, produced in reaction to the same cause.

But the ergo-auditory loop concerns only a part of the sound; it leads us to feel that which in the sound's course is not the direct product of our action, and which escapes our control in prolonging the movement of our action.

One must distinguish the cases in which:

- a sound is the action's identical counterpart, or as near as can be, such as when one plays with the accelerator of a car or motorbike by pressing the pedal or controls. In this case, the slightest variations in the "audio response" to the action suffice to maintain the ergo-auditory pleasure. The response seems "logical" and consistent with the cause: an action produces a louder or higher sound according to its force.
- an isomorphism does exist, but expertly and asynchronously, in the case of certain instruments: the prolonged note of a violin played *coll' arco* has nothing to do with "push and draw."
- a sound is not isomorphic to the action: as when one briefly awakens something in its course of being; when briefly striking a resonant body such as a bell, for instance, that continues to vibrate even when we can no longer see it. In this third example, not to listen to the sound we have produced until it ceases is the most common case, as if we were liberated from the weight of listening. "Triggering" or "making ring" something that escapes our control also predominates in games such as electronic pinball and other, more recent "interactive" games.

The Ergo-Auditory Trap in Relation to Movement

As Heraclitus said, "one never bathes twice in the same river"; in other words, the water flowing in a torrent is ever renewed. However, that same water flows in banks and a bed that do not move, at least not on the human temporal scale. But the sound that the water produces is massive, uniform, and statistically identical. If we move along the river or if we approach it or move away from it the sound will change incessantly in its higher tones, and we will receive very different "illuminations" according to the position of our two ears. It is one of those sounds we call "points of sound," whose (statistical) stability demonstrates the effect upon itself of our change in distance and attention.

The quality of this rich, high-pitched sound enables acute apprehension and a subtle sense of localization. It also presents two characteristics: its harmonic timbre changes considerably according to whether we move closer or away and to how we place our head. A person with a slight auditory defect in one ear—which is not uncommon—will hear more or less clearly the higher (more directional) frequencies depending on which way he turns his head and which ear he directs towards the source. Here we find ourselves in a loop where there is a suspicion of intentionality: the sound variations seem to "respond" to our movements. We are half caught, trapped in an ergo-auditory loop—just as I am with my tinnitus, which is more or less audible to me according to the position of my head—and trapped also by the pleasure of interactivity.

Conscious and Unconscious Ergo-Audition

Sergio Leone's Western *Once Upon a Time in the West* (1968) opens with three men silently waiting at a remote, deserted station for the train that will bring in their victim: a bald black man (played by Woody Strode), an ill-shaven white man, and a man with a beard. The black man and the ill-shaven man are seated; the bearded man is standing.

As the bald black man sits there impassively, drops—authenticated by an image of the ceiling from which they drip—fall at intervals on his shaven head. Instead of moving, he puts on his broad-brimmed hat. The sound of the drops falling on the hat's fabric is muted.

The ill-shaven man is bothered in turn by a fly walking across his face or buzzing around him. He manages to capture it in the barrel of his revolver, which he plugs with his finger. Then he lifts the closed barrel to his ear, murmuring to himself the sound of the imprisoned fly while controlling it. We can hear the intermittent buzzing of the insect that we imagine flying around inside. Because it comes from inside the closed barrel, the sound is muffled.

And the third of the waiting men, the gray-bearded man, is slowly stretching and cracking his knuckles.

These three sounds can be compared in terms of their interaction with the film characters: the drops of water, the fly, the cracking of knuckles are sounds that these men “pending death”—Charles Bronson will shoot them down fifteen minutes later—are producing directly with their bodies; that is to say, they receive these sounds from external reality but more or less integrate them by allowing themselves the means to modify them.

In putting on his hat, Woody Strode changed the key of the passive sound of water dripping on his shaven head, muting it. He makes an active sound of it. Jack Elam captured the fly to dispose of it in the barrel of his gun and make it buzz at his command before releasing it. He too has to some minute degree “controlled” a sound from the environment, for the sound of the captured fly resounding from inside the gun barrel has changed timbre.

Additionally then, these sounds interrogate the mystery of the ergo-auditory relationship: Is the character aware of the noise? Is “the Shining effect” produced consciously or unconsciously? Does the man who sets his hat on his shaven head do so to protect his head? Or to accumulate water on the brim without moving and, finally, quench his thirst (an indication of patience)? Does he do it also to change the sound? Such is the mystery of ergo-audition as demonstrated by cinema.

The movies do in fact give one the power to zoom in on faces—and Sergio Leone is one of the directors who zooms in closest—thereby underlining their opacity and the fact that we cannot view intentions or thoughts from the inside. Among other things, we cannot know the nature of the person's own listening capacity, his own “ergo-audition.”

Another interesting sequence is from Jacques Tati's film *Playtime* (1967), in which Hulot (the main character) has to wait in a glass-walled room where the absence of noise makes the slightest sounds potentially embarrassing. He shares the room for a few minutes with a very agitated businessman who, while waiting, signs papers, moves around, and seems intent on producing sound with a sort of auditory narcissism.

The man does not seem to be aware of the sounds he is making unless, on the contrary, he is exhibiting them to manifest his importance, his competence. His sounds trace precise, clear, and completed activities and seem to return to him a self-image of exactitude and brisk efficiency. The functional client would claim his territory by producing clear, energetic sounds, with no response. One might say he derived pleasure from hearing himself work: pleasure in that the sound exactly fills the contours of the activity while endorsing its imperative, determined nature. But the question is, can he hear himself?

When I give lectures, I sometimes hear a student playing compulsively with a retractable ballpoint, triggering the little click, and I wonder whether he is aware of the exasperating noise.

When we are disturbed by the voices of people speaking loudly into their cell phones, we always wonder whether these people talking so loudly are conscious of the fact: everyone is supposed to be conscious and responsible in terms of clothing and physical appearance, but not in terms of their auditory appearance.

Partial or Total Scotomization of Our Own Internal and External Sound Emissions

The human being starts life by emitting powerful, far-carrying sounds in the form of cries, but he is not conscious of emitting these sounds; he does not hear himself emit these cries. (Incidentally, deaf children make the same cries as other children.) The adult continues to take no notice of a great many of the sounds he produces and of which he becomes conscious only in certain circumstances, when seeking to dissimulate his presence or to make it unfelt.

Fortunately, we cannot remain constantly aware of the sounds we make. Just as we are almost continually obliged to filter out certain visual perceptions (e.g., our nose seen from a different angle by each of our two eyes), we are obliged most of the time, but in a far more systematic and fundamental manner, to scotomize our internal noises of mastication in order not to be deafened by them. Powerful feedback reflexes are there to help us achieve this . . .

A New Kind of Ergo-Audition?

Has the modern world of technology created new ergo-auditory situations? Certainly. When I trigger a brief electronic note by pressing a button on the keyboard of an ATM, or when I press the keys of a remote control to increase the volume of the sound coming from loudspeakers placed several meters away from me, I am experiencing new ergo-auditory situations in which there seems to me to remain but a meager relationship between my intention, my physical act, and the nature of the result. In this case, the same gesture may frequently trigger different sounds. Inversely, different gestures may create the same sound.

The act of pressing a plastic surface (on the ATM) should not emit a different note depending on the point one presses, and pressing on the tiny key of a remote control or sliding the mouse of a computer over an icon on the computer screen should not multiply the sound's volume a hundredfold. Contemporary art has grasped these possibilities, and its interactive video installations often invite the visitor (or should one say the spectator? the auditor? the player?) to influence a visual and/or audio phenomenon, the aim being to re-create the pleasure we felt as children exploring the universe's reaction to our activities. Often, however, I find this interactivity more interesting in terms of image, of what is visible, than in terms of sound.

The factor common to interactive multimedia productions on video-game cassettes and CD-ROM, but also on the standard DVD, is to enable the audience to intervene in the duration and speed of audio and visual events, sometimes very actively (when it's a game), and sometimes with more limited interactivity (in the case of DVDs and certain CD-ROMs). In both cases, the difference in tem-

poral terms between a sound and an image is considerable: although it is possible to stop or “freeze” an image, it is impossible to stop or “freeze-frame” a sound, because sound requires time as one of its spatial dimensions. A piano sound, for example, is defined as a gradual evolution from a maximum to a minimum intensity, and a certain lapse of time is required for this evolution to reach its end. If one were to freeze for a twentieth of a second the sound made by a piano and to replay the fragment repeatedly, all one would obtain is a prolonged note or (if one can hear harmonic timbre) a chord which no longer has anything at all to do with a piano sound. By means of digital sampling, certain current digital systems allow musical or linguistic sounds to be compressed or expanded within a limited range, while still preserving the identity of the sounds, which comes in very handy for audiovisual editing, film dubbing, and the like, as well as for resynchronizing elements of sound and image. But as soon as one slows down or accelerates sound to any degree, even without modifying its pitch, the timbre is completely transformed.

A sound is true only if it has a precise duration and if one respects that duration. Technical progress will never change the fact that there always subsists in sound something that is impossible to dissociate from time and to “freeze.” The problem will be the same a hundred years from now, in spite of the development of ever more powerful hardware and ever more sophisticated digital software applications. Sound will always be characterized by a certain form of resistance to interactivity.

It is plain to see what happens on an ordinary DVD when one presses “pause”: the image “freezes” but the sound is interrupted or, alternatively, one hears a musical theme repeated every ten or fifteen seconds (on a DVD menu, for example). Thus, sound is either absent or replaced by the hypnotic repetition of a fragment of the soundtrack or musical theme. To make sound correspond to a fixed image, one is obliged to resort either to silence or to a fragment of time repeating itself ad infinitum—but which remains time, and within time.

The importance of noises in video games is evident; I remember the early pinball machines with electronic sounds that one used to hear in cafés in the 1970s: the sound emitted was either an invitation (“play with me”), gratification, or, when one lost, boozing; it simulated a very amusing and sometimes hypnotic dialogue.

But is it possible to create rich sound with the most recent video games? I don’t believe so; or else, it’s possible only within certain limits. Video games, visually quite basic at the start (remember Pac-Man!), have made immense progress in terms of the quality of image texture and the variety of decors and colors, thanks to substantial financial investment by designers. But in terms of sound, first, there has not been the same level of investment and, second, video sets are equipped with tiny loudspeakers unable to reproduce high-quality, comprehensive sounds. Another problem is posed by spatial cohabitation: when sound installations are located in the same space, game sounds will mingle with the other sounds produced there, including the (mainly vocal) sounds emitted by the player. Whereas an image remains neatly within the frame of a screen, a sound has no frame and must be prepared to mingle with all kinds of other sounds in a random mix.

In certain already outdated games such as pool and table soccer, sound is gratification—when one hears the ball falling into the right hole. A pinball or video-game player plays not only to win and mark up a score, or to beat an adversary, but also to activate sounds, nowadays electronic—brief beeps and musical themes symbolizing triumph or defeat. However, the themes are always the same and are not produced by the action itself. Their deferred activation is a source of fascination; it casts a specific spell.

Electronic machines have therefore completely changed the rules of the traditional ergo-auditory game because they offer sounds—whether it be the four beeps (“tonic impulsions,” in Pierre Schaeffer’s terms) that in France accompany the input of a bank-card PIN code, or the warning sounds of a digital system, computer or otherwise, to let us know we have not pressed the right button—that remain indifferent to the force and the form of expression of our gestures, merely punctuating them or responding to them with a note of “protest” (the reiterated beep when one presses the wrong key). Likewise, whether one presses hard or softly on the keys of a cheap electric piano, the sound remains the same; only on the more costly models will touch patently influence the sound—that is, produce a better quality of interactivity, which the humblest sound system would provide for next to nothing.

Certainly, as I stated earlier, the noncorrelation of the intensity of the causal gesture and that of the sound produced was already the rule, centuries ago, for the clavichordist or the organist, whose instruments do not allow the influence of the force of pressure on the intensity of sound. But these musicians represented only a small part of humanity; moreover, the mechanical nature of instruments such as the clavichord and the organ introduces with each sound subtle variations in pressure, which means that a note is never exactly the same, even if its intensity does not strictly “obey” the player’s touch.

Above all, we live in the midst of sounds that no longer are the “natural” punctuation of gestures; instead, these sounds are maintained or added so that ergo-auditory feedback may function correctly. The electronic keyboard does not make the sound itself; the beep is added, created, and therefore can be isolated and adjusted (offering us, for example on our computer, the pitiful choice between a duck’s quack and a drop of water).

The woman operating a supermarket’s cash register is informed by a resounding beep that the barcode corresponding to each product has been read correctly, but this monotonous beep is influenced neither by the position of her hand nor by the object she presents. She handles a thousand different products that she must lift and pass before the laser beam in different ways: bottles, yoghurt jars, vacuum-packed slices, cartons, newspapers, clothing. Each one of these complex, diverse, and laborious gestures receives the same sound in response.

This situation—when the contrast between the diversity of an action and the uniformity of the responding sound is extreme—disconcerts the ergo-auditory loop, in some way dispossesses us of our action, and, in compensation, engenders the fallacy whereby sounds exist for which we, their producers, are entirely responsible. Which is not true; at least, not spontaneously. Even when we speak, formulate a vocal sound, we are employing an instrument that, even though it is inside us, is unfamiliar to us. And one can appropriate an instrument only by playing it often and at length.

But I must insist on the fact that, to me, there does not seem to be a definitive break between ergo-auditory situations that (because acoustic) might be deemed “natural,” those that predate the invention of electricity, and those enabled by the latter (in terms of sound amplification), which might therefore be deemed “artificial.” The exception would seem to be the rule. If one were to compile a typology of ergo-auditory situations, one would observe that there is a vast but limited number of situations, but that these do not culminate in an opposition between natural acoustics and electro-acoustics.

Between Doing and Listening: Naming

One of the tools with which I endeavor to make my students aware of the ergo-auditory relationship is the “doing/listening/naming” exercise. I have in fact elaborated a graphic and instructive exercise that I have already tested in several schools and from which I have derived certain lessons.

First of all, I initiate students in Schaeffer’s theory of reduced listening (i.e., of listening to a sound for itself), and I have them practice this form of listening in groups with ready-made sounds while providing them with the most elementary criteria pertaining to classification and description: the distinction, for example, between sounds of a precise pitch, known as “tonic” in Schaeffer’s terms, and sounds with no precise pitch, known as sounds of “complex” mass.

In a second phase, the students “create” and record sounds that respond to a simple definition and are situated within a reasonable range of duration (less than thirty seconds). The sounds are defined according to Schaeffer’s terminology. Examples include an “X” or “complex impulsion”—that is, a very brief sound with no precise pitch; and a “Y,” an “iterative variant sound”—a sound that continually varies in pitch but is prolonged by a repetition of impulsions, a volley of brief sounds that make it seem like a dotted line.

All sound sources are allowed: acoustic or synthetic musical instruments, everyday objects, one’s own voice, and so on. I have previously showed them how, using the voice or *a fortiori* ordinary objects, one can produce the whole family of the nine basic sounds in Schaeffer’s typology: the tonic sounds, of course, but also the complex sounds. For example, with the mouth one can create a prolonged “ch,” a complex sound, maintained continually or iteratively, whose site (situation in the vocal range) and caliber (pitch) can be controlled approximately.

The student has a few hours to create the sounds requested and is required to present these to his fellow students and the group; then (after having made them heard, without comment from either him or myself) he explains how he hears them. Do they correspond to his intention? What characteristic details do they comprise in addition to the required specifications? The final grade I give a student is based equally on his oral presentation of the sound he made to order and on the sound itself, for the presentation is the only way of verifying that the concept, a notion such as “tonic sound,” has been intelligently assimilated. This circuit—listening (without creating sounds), the creation of sounds, and oral explanation—seems essential.

The obligation to name the sounds that one has “made” also helps raise awareness of the traps of ergo-audition and causality, and helps maintain one’s role of observer. It may be, for example, that one believes one has created a tonic sound by means of a string instrument (a guitar or violin), or a complex sound by using an everyday object (a metal or cardboard box), on the basis of a misapprehension that one has of a stereotypical correspondence between instruments and notes, trivial objects and noise. However, it may be that a specific action executed with a guitar string has created a complex mass sound and, inversely, that the “trivial” object (a wooden stick or a salad bowl) has rung a note. It is therefore far more important to know how to hear what the sound one has recorded really is, then to have “controlled” the production of that sound by creating an object exactly to order, in line with the initial intention.

It may happen that the sound obtained does not in fact correspond exactly to the required specifications: that doesn’t matter as long as one is aware of it. But it does matter exceedingly if the student doesn’t know how to hear what

he has produced. It would be far better for him to have produced a complex sound where a tonic sound was required and yet to hear that the sound is complex and be able to describe it as such.

Another observation—and this has proved true on several occasions—is that a student may have intended to produce an impulsion by briefly striking an object; the object, however, has unpredictably prolonged the resonance, and the sound has become a continuous percussion tone. Anyone who focuses on the question of his gestural control (i.e., sounds obey my conscious will) will refuse to hear that the sound may be something quite other than the reflection of his intent and will be inclined to minimize, or even to scotomize, the “unintended” resonance.

Likewise, it is far more important to hear than to have created the required sound, and to thereby ascertain the potential and frequent shift between gesture and sound—a further illustration that sound is not necessarily isomorphic to the gesture or the movement that produces it.

This exercise does not claim aesthetic value or to be what is today called “an educational game.” However, it is evident that one must not eliminate the element of pleasure from the making of sounds. As a composer of concrete music, when I create sound recordings, I find there inevitably exists a sort of audio-phonetic cycle between the sounds we emit and what we hear, which enables us to make sounds in a dynamic way and implies a haziness in our listening capacity. But afterwards, when we listen to those sounds again as material for a composition recording, we must listen scrupulously to what is on the tape and disconnect this from the situation of ergo-audition. At this stage, one must find or create adequate listening conditions. To listen when one is the creator of a sound and to listen to a sound in which one no longer intervenes is not the same thing, nor the same moment in time.

An instrumentalist may be able to hear himself directly. In order to do so, he will have had to acquire many years of experience of the same sound source; he has had a teacher, with whose ear he has, so to speak, identified himself. Certain instruments demand a very critical ear: the violinist, for one, must necessarily verify by ear what he is playing, the notes not being pretuned by keys. When using material sources with unknown potential, however, a new apprenticeship is always necessary, and the finest musician reverts to the beginner stage. A pianist cannot always hear what he sounds like when singing or playing percussion: his body and his brain have to learn an entirely new procedure in order to correlate sound and gesture.

The final point is that the complexity and the challenge of the exercise lie in the difficulty of listening to sounds one by one, of hearing particular sounds and not a standard model.

In reality, in situ, for certain sounds of a repetitive or cyclic nature one is obliged to train oneself to create a standard model based on grammatical data in which each sound is a phoneme. But when we observe a fixed sound by means of reduced listening, we must extract it from the global image type, from the many images that overlay it and that are all the kinds of sounds of the same form, obeying the same rules as the standard model. For example, it is not customary to pay great attention to the passage of a single car, of that car and not another; we take it for granted that they are all the same because they all stem from the same model “in delta.” But that would be like saying that all forms of summits seen from the valley are the same, whereas there are pointed ones, jagged ones, and so forth.

Music without Ergo-Audition

One must also consider the types of music that are neither interactive nor open to a variety of interpretations. My own compositions of concrete music cannot be classified as interactive or instrumental, which leaves the interpreter some creative freedom in terms of sound and tempo. My ideal, regarding my own music, would be that the auditor's intervention in my music be as limited as possible, not because I wish to impose my will on others but because I believe that what I am endeavoring to do exists only when presented in a certain fashion.

My music is made

- to be listened to without making a noise
- to be listened to without moving
- and to be listened to wholly, from beginning to end.

“Without making a noise,” the auditor must be silent and calm, as my music comprises many subtle details that one must perceive (it is absolutely unsuited to open-air productions). “Without moving,” because one must be focused, attentive, receptive, and not vary the way in which the sound reaches one's ears. “Wholly, from beginning to end,” because to my mind a work of music is like a phrase whose meaning becomes clear only when all the words from the first to the last have been spoken and heard. It cannot be taken up midway or interrupted before the end without altering its meaning (which does not exclude playing excerpts on the radio, on condition that it is clearly stated that these are only excerpts).

Are such listening conditions possible? Yes, of course. This is not passive listening—listening is an action in itself, in the same way that thinking and meditating are actions that one cannot see.

Like yoga and certain schools of meditation, motionless listening is, of course, an entirely cultural and contrived situation. It must not and obviously cannot replace ergo-auditory situations; but nor must the latter for their part become the norm, the predominant situation. In a sense, my practice as a composer is contrary to and a criticism of interactive audiovisual works; often, these give rise to sounds that I find of little interest, and to feedback between gesture and sound that seems to me to be repetitive and stereotypical.