Listening strategies in electroacoustic music

Stephen McAdams

Introduction

Consider the following couple. On one hand, a *human being* endowed with sense organs capable of transcribing air vibrations into a form that can be transmitted and processed by the neurons of its brain. We will call it "the listener". On the other hand, a structured set of air vibrations that result from the movement of an electromechanical device that is particularly omnipresent today: a loudspeaker. The structure of the vibrations emitted by this device is supposed to bear a specific æsthetic project conceived by another human being, who realizes this project by the arrangement of sound events that, in the case of synthetic or electronically processed sounds, could not have been produced mechanically by any physical object set into "natural" vibration. We will call this set of vibrations "electroacoustic music". When our listener intends to apprehend and, if possible, to understand the information coming through its sense organs, we say it "listens". And lastly, faced with complex acoustic structures like music, our listener can, more or less consciously, direct its attention to this or that aspect of the music or try to register this or that relation between successive events, or among groups of events separated in time. In this case, one would say that the listener employs one or a number of "listening strategies". What interests us more specifically in the framework of this set of papers is the nature of the strategies invoked while listening to a music, the sound events of which are, from a biological viewpoint, "unnatural".

Sound structures

Acoustic ecology and auditory processing - default listening

Biological (wo)men, humans as animals in other words, use their ears for what purpose in everyday life? One of the principal functions of audition for many animals is to alert the organism to the presence of animate objects in the environment that thereby produce noises without necessarily being perceived by the senses of vision or touch. For this, the animal needs simply to detect the acoustic events and to calculate their probable position in the environment. It can also orient itself to fix the object with its gaze or continue to analyze the

Appeared in French: McAdams, S. (1996). Stratégies de l'écoute : L'auditeur et la musique électroacoustique. In M. Borillo & A. Sauvageot (Eds.), *Les cinq sens de la création*, pp. 70-76, Editions Champ Vallon, Seyssel.

sound in order to determine, with as much precision as possible, the identity of the sound source. Therefore detection, localization, and identification are among the most important functions of *default listening*. In our current lives, filled with mechanical objects that we must manipulate with a lesser or greater degree of dexterity, one might possibly add the functions of organization of motor activity by the rhythms that are heard—typing, assembly line work—and of continuous control of the functioning state of machines—a car, an industrial machine (McAdams & Bigand, 1993).

Nonetheless, all of these functions depend on a number of types of processing performed beforehand by the auditory system, the most important of which would be what we call today the analysis of the "auditory scene", that is, the building of distinct mental representations of the behavior of sound objects through time (Bregman, 1990). The auditory system appears to be sensitive to a small number of acoustic cues that characterize the vibratory behavior of physical objects. This sensitivity, probably acquired during the evolution of our species in the physical world, allows us to segregate perceptually the sources that vibrate at the same time and to follow through time the stream of events they emit. What emerges from research on this subject is that the physical objects have a certain coherence of behavior, a coherence to which the auditory system is acutely sensitive. A priori this shouldn't be too surprising. However, what interests us is how this same system, with its bias toward processing events produced by physical objects, reacts to artifactual events created by humans. These events may no longer respect the constraints of physics in their behavior: electronically and digitally synthesized events, acoustic events recorded and transformed by these means, and even the clever arrangement of acoustic events to produce auditory illusions—what one traditionally calls "orchestration".

Cultural artifacts and mental representation

The composer assembles a sound structure in time, the aim of which is more than simply to signal the existence and identity of objects in the environment surrounding the listener—it is a cultural artifact rather than a "natural" event. The events can be considered as vehicles of a sort of "message". The composer must therefore lead the listener toward a perception of the abstract qualities of these events and of the relations among them. In the context of a given culture, certain qualities like pitch and duration are organized in *abstract systems of relations*—scales, harmonies, rhythmic patterns, meters—that serve as a framework for the mental organization of hierarchical relations realized in time (Bigand, 1993). So it is no longer the sources themselves that matter, but the message emitted by these sources.

If, however, the composer wants to organize a complex message carried by the coordinated activity of several sources, or even several coordinated messages, which occurs as soon as there are two or more instruments, the independent perception of the different components of this message may depend on the discrimination of the sources. This discrimination often takes place on the basis of their timbral qualities. The sensitivity of the auditory system to the behavior of physical sources is such that one has to make a concerted effort to "trick the ear"[1] to make it believe that a mixture of several sources is a single musical object—a practice that became more and more current with the symphony orchestra from the second half of the 19th century (Boulez, 1987). Near the end of the 19th and the beginning of the 20th centuries, timbre—the aspect of sound quality not defined by pitch, duration or intensity—began to occupy the minds of composers that were, as ever, desirous to extend their musical vocabulary toward new horizons. If one examines orchestration practice closely, one finds that timbral mixtures were often intuitively based on the use of the acoustic cues that characterize the behavior of physical sources that were alluded to above: synchronization of attacks, composition of harmonic relations among pitches, coherent fluctuations of amplitude and frequency, common position in space, and so on. On can conclude in a sense that these composers subconsciously aimed to create images or mental representations of virtual sound sources by the combination of multiple physical sources, but all the while maximizing the cues that reflect the behavior of *one* physical source. One might perhaps even hypothesize that they tried, always intuitively, to simulate a physical causality that could be picked up as such by an auditory system already biased in this direction. Obviously, the complexity of musical organization can go beyond the auditory system's capacity of analysis, discrimination, and representation of simultaneous images. The listener would thus be led toward a sort of "surface listening" where only the larger-scale elementary and textural features would be perceived (McAdams, 1987). This is the case with a number of works for large orchestra in the second half of the 20th century, such as *Atmosphères* by György Ligeti, *Threnody* by Krzysztof Penderecki, or works from the "spectral music" approach in France (Tristan Murail and Gérard Grisey, among others).

Sound synthesis and processing

With the first explorations in electronic sound production toward the end of the 1940's, this connection with the physical world was broken. And with the complexification of the

[1] In reality, this term is false, since it is not really the "ear" that is tricked, but rather that part of the brain that processes auditory information—the entire central auditory nervous system.

available technologies (first analog and then digital), the way in which a musical sound was produced moved farther and farther away from the physical process to which the auditory system is sensitive. By this statement, I do not mean to imply that sound thus fashioned has no musical, or more generally no æsthetic, value, but simply that default listening often no longer has any evident landmarks to which to adhere. And this is all the more so if the music is conceived as polyphonic or multi-voiced. Without the acoustic behavioral coherence of physical sources, in a sense the sensory overspecification of the sources inhabiting the acoustic world, the listener has greater difficulty separating and following the different voices. The result can become a kind of acoustic mud. In short, "natural" listening, oriented toward the apprehension of messages emitted by recognizable sources, is obliged to step aside to leave room for other listening strategies, without actually turning off completely—default listening will always operate in the background. It would be impossible to describe all the possible strategies that could be used to complement the mental organization of an extremely complex acoustic environment, since they would depend a great deal on the listeners' familiarity with this kind of music, as well as on their degree of explicit musical training and of implicit learning by the auditory system through exposure to these new sound worlds. Nevertheless, a few examples should be considered.

Behavioral coherence of sound objects and the problem of polyphonic tracking

A first strategy that is set in motion, when a listener no longer succeeds in organizing the acoustic information into perceptual images of sound sources, is to orient listening toward a "surface reading". The listener follows with more or less attention the changes over time in intensity, register, spectral composition, and texture. But at times landmarks are missing, the discontinuities that clearly indicate the temporal edges of events. Even more, sufficient information to track the trajectories of a "musical object" across registers is often missing. In analog electronic music, there was a lot of this kind of phenomenon, but there were also some notable exceptions. It is interesting to consider why these latter succeeded from a perceptual point of view. One of the masterpieces of electronic music is Karlheinz Stockhausen's *Kontakte*. This piece has a certain compositional complexity, but the composer knew how to organize the sound materials so that the timbres were easily distinguished when they played simultaneously, and they each generally stayed within a particular register. As such, the composer intuitively used Gestalt rules, expressed by a psychological school that describes the human tendency to organize sensory information into distinct "forms" or "objects" (Köhler, 1929). One example of a piece created with digital synthesis on a computer, where this organization into sound strata is not possible (in fact such was not the æsthetic goal) is *Antony* by David Wessel. In this piece, the intensity and frequency of each of a large number

of oscillators are controlled over time. Sound densities are created that slowly evolve as waves on a temporal beach lasting about 15 minutes. There is a minimal number of discontinuities in this work, which leads the listener toward a sort of textural reading of the piece. Similar phenomena are also obtained in orchestral music where the arrangement of a large number of instruments in relatively compact registers creates bands of sound (as in *Threnody for the Victims of Hiroshima* by Krzysztof Penderecki) or swarming textures (as in parts of *Atmosphères* by György Ligeti) that are not analyzable in detail by the auditory system.

Certain composers have adopted other compositional strategies with respect to the use of synthetic sounds. One approach consists in using electroacoustic sound to extend the physical sound world (Saariaho, 1987). A example *par excellence* is *Verblendungen* by Kaija Saariaho in which, as its name indicates, the goal is to mix the two worlds: the orchestral sound, predominant at one moment, is augmented by the synthetic taped sound, until it is transformed progressively toward the noisy and textural characteristics of the synthetic sound, which, in its turn, becomes the principal actor in the sound drama.

Another approach, introduced by Pierre Schaeffer, founder of the Groupe de Recherches Musicales (GRM) at Radio France, has been called *musique concrète*. In general, environmental sounds are recorded (the sound of a closing door, a buzzing fly, flowing water). They are then processed by electronic or digital techniques and are arranged in a sound structure. This quasi-cinematographic music, which elicits a series of very striking images, is projected through an ensemble of loudspeakers (for example, GRM's famous *Acousmonium*). One of the problems of a perceptual and cognitive nature with this approach concerns the identity of the sound objects. In fact, if the *identity* of each sound predominates over the *musical relations* among the perceptual qualities, the listener tends to adopt a listening strategy of object recognition in trying to "make sense" of the sound environment projected by these auditory images, to the detriment of a perception of the *musical discourse* that is supposed to be borne by these sounds. The references to the external world direct the listener's attention beyond the music itself. Aside from program music, or ritual music that incorporates a particular symbolism, music is generally an autoreferential system. An external reference runs the risk of breaking the grammatical coherence of the musical structure. It is not impossible to overcome this default listening tendency in the listener, but the composer must pay specific attention to attract the listener toward the relations among the sounds. Two particularly successful examples of this genre are *Gesang der Jünglinge* by Karlheinz Stockhausen and *Vox 5* by Trevor Wishart.

Perceptual control of digital synthesis and processing

In the realm of sound synthesis, research on the perception and mental representation of musical timbre have a primary importance. This auditory attribute is in fact multiple (McAdams, 1993). In the classical definition adopted by the psychoacoustics community, it includes everything that is not pitch, intensity, duration, and spatial position. This leaves a lot of possibilities and the problem for a composer interested in digital synthesis and timbral composition is the intelligent control of this attribute in sound synthesis and processing. The study of timbre in hearing research has a lot of similarities to the study of odors in smell research. Odors are multiple, have very complex origins in the chemical world, and researchers rack their brains trying to understand all of their perceptual and semantic components.

At IRCAM, we have been working for over a decade on developing tools for the perceptual control of timbre. Our approach consists of a multidimensional analysis of the "perceptual space" of timbre in trying to model it with geometric structures (McAdams, Winsberg, Donnadieu, De Soete & Krimphoff, 1995; Wessel, 1979). We then try to find a correlation between the axes of this space and psychoacoustic parameters. This relation, appropriately called "psychophysical" since it establishes a quantitative link between the worlds of perception and physics, orients us toward the control parameters that are likely to be the most intuitive for musicians, since they are perceptual and thus potentially musical. Without this more or less rational control of synthetic sound, many composers have the impression they are groping in the dark, since the "traditional" acoustic parameters (amplitude, frequency, phase, spectral envelope, temporal envelope, ...) often have a relation with the dimension of perception that is often quite indirect. They are thus difficult to fathom for very complex sounds. And if the search for a link with hearing is vague for the composer, what must it be for the listener?

The apprehension of artistic form

By way of conclusion and in playing, not the devil's, but the listener's advocate, at least for the listener with an open mind and a will to feel a certain depth of experience when listening to music, I would like to pose a few questions.

Didn't the modernist project presume that there were no limits to the comprehension of structure and form, whatever their complexity and rules of organization.

Mightn't we think today that the human brain cannot necessarily evolve ontogenetically at the level of its inter-neuronal connections to process absolutely *any* structure in time?

In short, mightn't there be certain forms that the evolution of our species predisposes us to perceive and to understand? Don't there exist biological limits to the apprehension of artistic form? The other, slightly more optimistic, side of this question is the following: what is the space of possible forms that can give us an æsthetic and affective experience that has a certain psychological depth? This space is probably constrained, but the possibilities within it are nevertheless very certainly infinite.

References

Bigand, E. (1993). Contributions of music to research on human auditory cognition. In S. McAdams & E. Bigand (Eds.), Thinking in sound: The cognitive psychology of human audition, (pp. 231-277). Oxford: Oxford University Press.

Boulez, P. (1987). Timbre and composition—timbre and language. Contemporary Music Review, 2(1), 161-172.

Bregman, A. S. (1990). Auditory Scene Analysis. Cambridge, MA: MIT. .

Köhler, W. (1929). Gestalt Psychology. New York: Horace Liveright. .

McAdams, S. (1987). Music: A science of the mind? Contemporary Music Review, 2(1), 1-61.

McAdams, S. (1993). Recognition of sound sources and events. In S. McAdams & E. Bigand (Eds.), Thinking in sound: The cognitive psychology of human audition, (pp. 146-198). Oxford: Oxford University Press.

McAdams, S., & Bigand, E. (1993). Introduction to auditory cognition. In S. McAdams & E. Bigand (Eds.), Thinking in sound: The cognitive psychology of human audition, (pp. 1-9). Oxford: Oxford University Press.

McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. Psychological Research, 58.

Saariaho, K. (1987). Timbre and harmony: Interpolations of timbral structures. Contemporary Music Review, 2((1)), 93-134.

Wessel, D. L. (1979). Timbre space as a musical control structure. Computer Music Journal, 3(2), 45-52.