Contemporary Music Review
1999, Vol. 18, Part 3, pp. 41–77
Reprints available directly from the publisher
Photocopying permitted by license only

Live electronics or...live music? Towards a critique of interaction

Marco Stroppa

Interactive systems are throughly scrutinized in the lengthy article. Of particular interest to the author are the question of space and sound projection, interpretation of electronic sounds and the relationship between performer and technology.

KEY WORDS: Sound projection, Synthesis control, Interpretation, Performance, Radio opera, Orchestration, Virtual space, Virtual interpreter.

1. *Prélude*

Interactive systems are truly popular nowadays: virtually each computer-music studio around the world is equipped with commercial real-time machinery, if it has not even developed its own one. New musical applications are no longer described in terms of their expressive power, richness or flexibility of the environment, or adequacy to tackle a given problem, but rather in terms of their real-time[1] features (such as number of simultaneous voices, oscillators, harmonizers, filters), as if they were the only and unique criterion of validity.

It is sufficient to take a look at the summary of the ICMC Proceedings to realize how dramatic this change of mentality has been over the last fifteen years. Aside from poorly concealing a delusion of grandeur and a certain dose of pretentiousness — is the rest of the world the domain of "fake time", therefore devoid of any aesthetic value? — it is astonishing to discover that the musical debate has rarely delved into the truly essential issues raised by the emergence of interactive systems, and has been often hidden behind purely technological concerns.

It is surely too simplistic to state that such a world-wide success is the consequence of the tremendous growth in computing speed during the last years, because one could use it for other purposes as well. But what is the musical significance? Is "faster" better? This is not just a pun: real time has somehow become a new dogma, unquestioned and unquestionable, a "conditio sine qua non" in order to reach the path of true music. How many composers pretend that what is not real time is merely not at all music, without inquiring, first, what music is, second, where the problems lie, and considering then whether or not this technology provides an adequate solution? I am not ignoring the technological advance: after so much waiting in the early years of computer music, such an immediate and efficient relation with synthetic sounds is extraordinary. But every development has a price: if much is gained, some is lost, for ever.

So, once the "demo" effect has subsided — it is indeed much fancier to show some action going on a screen than to play the result of a complex computation, no matter how banal the former and sophisticated the latter! — the question is still appropriate: what are the repercussions on composition and musical aesthetics?

In this paper I will first critique interactive systems starting with the most frequent objections against non-interactive music. It is not my intention to plead for any particular system, because any solution may fit, provided that it is suited to the composer's demands. I will then shortly examine the issue of performance and interpretation applied to computer-generated material, and its musical and technological constraints. My aim is to show that the problem is elsewhere, and that, when closely perused, these harsh objections, although not useless, are largely devoid of fundament.

2. *Allemande*

A few years ago the Parisian Ensemble Intercontemporain invited me to conceive and perform the sound projection of Stockhausen's *Kontakte*[2]. Although I very well knew the piece from my student time, I had never played it in concert. It was quite an astounding experience: the synthetic sounds were so "real", energetic, lively, that I hardly believed they were recorded on a tape produced a third of a century earlier; that music was much more effective than any real-time pieces I ever heard. Which mysterious forces were responsible for that? How would this energy spring out from such an obsolete medium as a quadraphonic tape? Clearly, the essence of the piece, directly linked to the composer's creative power, seemed to transcend the medium[3]. Or, perhaps, when used appropriately, won't it still have something to offer to the composer?

The performance was as free and fresh as any interpreted music. In no moment did the instrumentalists or I feel constrained by the presence of a tape running uninterruptedly from the beginning to the end. Yet, one of the first and most imperious criticisms toward tapes is precisely their temporal rigidity. Though there is no doubt concerning the crucial importance of time in music, this is a far subtler problem than dealing only with temporal fluctuations. A performer at ease with a click track will find other ways to express his or her interpretive choices. If the composition is done in a certain way, nobody in the audience will perceive any temporal awkwardness and the performance will be judged as free as usual. But supposing that a rigid medium is indeed a major constraint, how would interactive systems solve the problem? And at which costs?

There are usually three levels of sophistication: immediate reaction, score following and tempo tracking. The first level designates the ability to start something (a sequence, a sound, an automatic process, a recording and so on) at will and have the system react with no perceivable delay. I have already observed that this characteristic is now feasible with direct-to-disk technology; therefore it should no longer be a discriminating issue. The second level refers to systems provided with an event detector[4] and some internal information, usually a coded score, about the music that is being played. Their task is to know where the performance is with respect to the internal score. The third level adds to the second some information about time, that is either the absolute speed of a sequence of events or the current metronome, so that the tempo of the machine is adapted to the one used in the performance. This is a very captivating feature, since it gives the mirage of a musical behavior. Unfortunately, if this barely works with baroque music or with less demanding styles from the standpoint of temporal accuracy, the classical XXth century repertoire has such a sophisticated, multifarious and delicate relation with time that nothing of what I heard, despite imaginative technical proposals, ever displayed the faintest gleam of a musical attitude.

Hence interactive systems, if they do provide more ductile methods than a tape to deal with time, are far from constituting "the" definite answer to the need of making lively music. Moreover, they come with their own burden of constraints, some of which are elaborated further in this text. This is not just a technological problem, but primarily a conceptual task: we do not yet have any models understanding the essence of musical life well enough. If we examine the relationship between the temporal adjustments of an interpreter while playing in a concert and what is notated in the score, we will realize that it is an extremely complex, refined, mutable and "multimedia" phenomenon, in any case far more intricate than a machine's elementary skills. The performer's fluctuations depend upon information arriving to several senses at once, the hearing, of course, but also the sight, the touch, sometimes even the smell and the taste. Feeling an audience, experiencing the physical vibrations of an instrument, receiving the energetic swells from other performers, mastering one's own bodily tensions are as important as hearing correctly. Furthermore, these flows of communication are constantly open and interacting within the performer and are under the additional influence of other psychological, circumstantial or intellectual factors. This — and indeed much more! — is what turns each concert into a unique event. How many times a piece, which sounded uninteresting when heard in a recording or in a mediocre performance, was so metamorphosed in a concert that it would be hardly recognized as being the same music?

Assimilating a real-time device to a performer simply because it starts its own sequence alone is degrading the performer's competence, asserting that a rubato or a slightly different tempo is the heart of true music is intellectually questionable and musically weak, if not a purely ideological claim[5].

If the inner liveliness of electronic music is fortunately not the property of any specific technological system, it is nevertheless influenced by a system's inherent drawbacks one must eventually always cope with. Devising the best compromise is sometimes an arduous task. The deficiencies of tapes are too well known to be here discussed in further detail. Yet, they are not absolute. If expressive music with a perpetually oscillating tempo poorly fits in with tapes, music that is either built on a naturally steady pulsation or notated proportionally is rather well suited[6]. On the other hand, tightly intertwined rhythmic mixtures between a performer and a machine are all the time cumbersome, and virtually unfeasible with tapes. Provided that a piece does not insist too much on its very limitations, the association between a performer and a tape can be as natural and musical as any other.

Exploring the musical space between a piano and a world of synthetic material was one of my primary goals when I started to work on *Traiettoria* for piano and computer in 1981. On the side of the instrument, I had just discovered and understood the gestural control of that marvelous secret universe which is the resonance of the piano. I started to investigate the relations between the resonance and personal harmonies and was experimenting with multiple combinations of instrumental gestures as structured sources of excitation. In addition, I wanted to see how this approach would meet with synthesized material. Since the computer never simulates or processes piano sounds, the synthesized material bears no direct acoustic relationship to familiar instrumental or concrete sounds. How will then a piano and a computer manage to influence each other in such a intimate fashion that sometimes they originate a single, hybrid "instrument"? By interacting in symbiosis at the level of musical language, that is by inventing synthetic sounds that are "organic morphologies" and having them intersect with the figures played on the piano (for some more details about the relationship between piano and computer see the program notes included in the recording of *Traiettoria* published by Wergo, [Aimard, Stroppa, 1992]). This required that I adopted an unusual approach to computer music and developed special software suited to it: all the synthetic morphologies were classified into families sharing similar phenomenological characteristics. Each family was thus subdivided into elementary "morphemes" (or morphological units) whose characteristics could be directly coded into the computer thanks to my special software. Each morpheme corresponded to an individual sound file and was stored on disk. Several morphemes were then digitally mixed to form an instance of a morphological family; distinct instances were assembled in successive stages so as to gradually form more complex structures[7].

Each stage required specifically fine-tuned operations to generate a satisfactory result: the dynamics and spectral contents of each file had to be ceaselessly updated, their temporal placement was all the time revised by few milliseconds. Empirically done by ear, this operation was nonetheless crucial: minute modifications of time or dynamics were sufficient to make a bunch of morphemes fuse together into an organic morphology, or, on the contrary, to tear it apart. I did not discover any general model of behavior: each morphological instance had to be individually crafted, since the perceptual constraints were too context-sensitive. Now and then the temporal links had to be broken down and redefined from scratch, for the delicate equilibrium of energies that worked properly within a musical flow would no longer function when merging several flows together. However, the final outcome fitted in with the piano very naturally, in spite of the high density of the synthetic material (from a few hundreds to several million simultaneous oscillators!).

Because of this compositional approach to computer synthesis, the best technological option was then — and would still be today![8] — to perform all the mixing digitally and to record the final result on a tape. The desire to work with various types of relationship between the figures played on the piano and the morphologies synthesized by the computer and the wise to approach moments of so complete fusion and exquisite interlink that the instruments become indissociable, required a musically perfect synchronization between piano and tape. This perfection "while difficult to achieve, is made possible by the organic nature of the relationship established between piano and computer as well as by the specific compositional techniques adopted by the composer. Thus, (...) the piano part is not metric. Rather, it usually "hinges" on "temporal pivots"[9], focal points of structure and articulation around which other material is organized.

Such material possesses relatively "flexible" properties and its rate of execution and momentum are determined by internal "temporal dynamics"

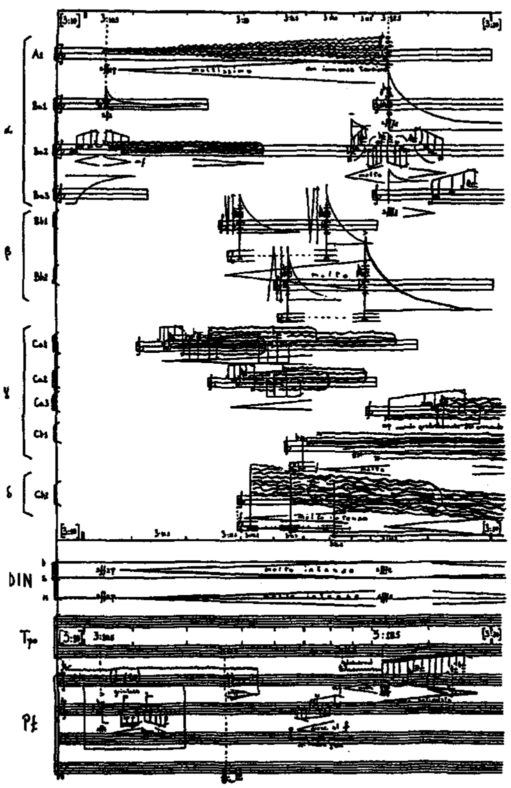

Figure 1 one page from the score of *Traiettoria...deviata* (© G. Ricordi and Co., Milan; reproduced with permission). The upper part is the notation of the computer material; each alphanumeric code (A1, Ba3, Cb2) identifies different instances of the same sound morphology; Greek letters group codes into super-families constituting a single musical flow (for a detailed analysis of this notation see the introductory notes to the score and Stroppa, 1991). The dynamic signs on the "DIN" staff refer to the sound projection in concert. For convenience, time is in minutes and seconds, although the performance of the piece never requires a stop watch. In fact, since the tape must be tuned with the piano prior to each performance, a change of 1 Hz of the reference "A" will produce a difference of 1 second after approximately 7' 30". The time indications on the lower staff are not for synchronization purposes; they just give the pianist a general framework. The figures in the box are a pianistic morphology with its own independent internal speed. "3:10,5" and "3:18,5" are the two main "sync" points, or "temporal pivots" of the page. They respectively start and conclude an intense crescendo. Due to the perceptual salience of the synthetic material, with its organic, quasiorchestral unfolding, and to the kind of piano writing I adopted, it is impossible to miss the pivots, while the rest of the figures will still find their natural place in between.

intrinsic to the nature of the material itself (for example, a glissando cannot be played below a certain speed) as well as by decisions stemming from personal interpretation"[10] (fig. 1)

All this caution, however, was still not satisfactory to attain the superior experience of a real performance. My conception of time solved the synchronization problems, but the issue of a good musical balance between piano and computer was still unsettled. Unfortunately, this is so dependent on the hall and the equipment available for the sound projection, that no real alternatives are left: when mixing the master tape, only the local dynamics were accurately shaped, that is the relationships of level and timbre between the sound files constituting a morphology, then between morphologies belonging to the same family, and so on, until the major musical flows. The master itself was recorded as "hot" as possible, for reasons of signal-to-noise ratio, by doing a manual dynamic compression. All the global dynamic shadings are the task of the performer in charge of the sound projection: loud and soft passages, crescendi and diminuendi, the computer's imperceptible fading in and out from the piano, the subtle interplay between figures yearning for total fusion are not recorded on the tape, but notated in the score and interpreted during the concert[11].

The decision of not producing a definite master and to rely on the sound projection to "finish the work" was in fact quite extreme, risky and very unusual on the computer music scene. Nevertheless, it revealed that performing a tape could be a greatly musical role, as lively and varied as any interpretation. If the speed of a tape is unchangeable, the psychological feeling of the perceived time is however always different and depends on the sound projection: when a rest separates two sound events, the identical physical duration will not be perceived as being an equal rest if the sounds come from the same loudspeaker or from two speakers far apart[12]. The acousmatic experience of the Groupe de Recherche Musicale in Paris comfortably demonstrates how much diversity is achievable through a ingeniously planned setup of loudspeakers[13]. Sound example 16 is a short excerpt from *Dialoghi* that applies some of the principles expounded above.

Another property that is often regarded as particularly attractive in interactive systems is to use microphones and pick up sounds directly during the concert, so as to grasp the unsurpassed thrill of a performer's inspired gesture. Whether they are immediately processed or stored in a sampler's memory or on disk for a later retrieval, these sounds are thought to retain the soul of the performer and the seeds of true music, something a machine will never be capable of! I will not object that real-time processing might yield extraordinary results and deserves all our attention. But, once more, this has nothing to do with the life of a musical performance and it has its own inherent problems. Such performances may be as "deadly" as the worst tape music, if the problems are not properly coped with![14]

Recording sounds is a very delicate process, extremely fragile and subject to many constraints. The position of microphones and the dynamics and timbral qualities of a signal depend on which processing algorithms are applied to it and severely affect the final outcome. How is it imaginable to fulfill all these constraints during a concert, where the controls are so flaky and isolating an instrument for sampling virtually unrealizable? To get a cleaner signal, one might try to "bury" a microphone into the body of the instrument. This would yield a better isolation, but then, the sound quality is utterly deteriorated and the foremost attributes of interpretation lost forever.

Furthermore, an instrument played in a "natural" way[15] will engender not only an acoustical phenomenon, but also a strong and persistent perceptual "archetype", that is often badly suited to signal processing and will tend to be heard as a poorly deformed timbre rather than as a new sound family[16]. To create sounds whose processing would be more effective requires special conditions and, chiefly, a different approach to instrumental playing and to recording but this approach is not automatically very meaningful during a performance. This is not a catch, though; it simply shows that there are never universal, problem-free solutions and that any decision must always be carefully evaluated and "tuned" with the basic compositional needs.

For technological reasons, the available algorithms for real-time processing are not so many and sometimes yield quite predictable, stereotyped timbres[17]. By definition, non-causal algorithms are excluded, as are those which first look ahead at a certain arrival state and then work their way backwards. The usage of purely synthesized sounds suffers even more from the same drawbacks, since substantial processing capabilities and musically more relevant control models are needed to achieve satisfactory results. An alternative has sometimes been to pragmatically reduce the complexity of a synthesis algorithm and to extract control parameters directly from a live instrumental sound[18]. The secret hope is that what is musical on one side would still remain musical once transferred to the other side. But why? By virtue of which mysterious alchemy will it work? The interaction between a performer and his or her instrument is so tightly coupled to the physical behavior of the instrument itself, that I do not see why, for instance, extracting some phrasing characteristics from a clarinet and applying them to a cello should sound anything but clumsy. It is even harder with synthesized sounds, since then the acoustical instrument, being usually superposed to the processed sounds, cruelly pinpoints the difference, or one would have to store the analysis parameters and use them later. But then the relationship with the originating gesture will probably be unhearable. Were the machines more powerful and (this is much more important) had we better models to understand a performer's playing and to control a signal-processing network, this approach would be greatly enthralling. For the time being, however, I am sure that such experiments will certainly constitute very fancy demonstrations, but am afraid they are still too problematic to be an integral part of a compositional process.

In the meanwhile, we have to be pragmatic, cope with the current state of the art and strive for the most musical compromise. But genuine real-time "aficionados" would probably have another reaction: let us deal with what we have, no matter how limited it is, provided that it is real time; if the current sound quality is not so great, never mind, the technological progress will automatically improve it in the future. If the dogmatism of similar claims is too flagrant to be discussed, its potential repercussions are perhaps even more misleading; in fact, this suggests that sound quality is not at stake, since it is sacrificed on the altar of interaction. It is definitely not a "fault" as such, although one might conjecture why to use an electronic technology that cannot match the sonority of instrumental sounds. In this case, either the electronics is obliged to remain segregated in the background, or the dichotomy between the natural qualities of instrumental sounds and the stiffness of the electronics will simply be too blatant. But when sound quality is not just an optional accessory of a piece, but is deeply rooted into the composition itself, no compromises are possible without altering the very essence of the piece.

In most circumstances, interactive music is tightly connected to the MIDI protocol of communication. The advantages and drawbacks of MIDI are too widely known to be reconsidered here; fortunately, new more powerful protocols are currently under way and might substitute it in the near future. But the advent of MIDI, under the pretext of a universal communicator, has also caused a major leap backwards, since it enforced again a commercial keyboard-biased view of music, with its notes (or keys) and velocities, while musical research had already gone beyond the concept of "note" as an elementary component to come up with other, more powerful concepts, better adapted to today's musical concerns.

By substituting the common units of measure[19] with very idiosyncratic ones, MIDI acted like a huge distorting glass: those who already knew what the sound world "looked" like, could adapt their knowledge to the new "Weltanschauung", with some efforts. But those who only grew up with MIDI — and for whom pitch is a key number, amplitude a velocity between 0 and 127 and attack time a slope — are running the same risk as

Plato's prisoners chained in a cave who watch on its inner wall the shadows of statues carried along a wall behind them and take these figures for the reality itself.

Although an interactive system is not the "panacea" to infuse electronic music with the vital energy that is often missing, it can surely represent the best solution in a given circumstance. Suppose that the technological constraints are accurately dealt with and that an adequate system is built and ready to perform, another substantial problem will then instantly arise, especially if our system is to be integrated with an ensemble: rehearsal time! It is today infrequent to get enough rehearsal time to go through the musical aspects of a piece, even if the instrumental writing is strictly conventional. But if an electronic device is to behave as a real performer, one would need further rehearsals to adjust the levels and fine tune all its delicate control parameters. Since they are inherently dependent on the performance context, there is not much one can simulate in advance. Combining acoustical instruments with sounds coming from loudspeakers is already not at all easy, even without any interaction at all, and demands a creative attitude from the composer and the sound engineers, a scrupulous plan and a lot of time[20]. This is today almost unfeasible. In the end, real-time setups must often be improvised during musical rehearsals, when not directly during the concert. Thus, in spite of the best possible will and proficiency, the behavior of the system and the sound projection risk having that sour taste of an unavoidable amateurism.

Courante

In the previous chapter, I pointed to still unanswered questions concerning the usage of electronics in a performance and clarified why the only viable approach is first to analyze the salient compositional concerns and then to work out the most effective technological strategy. However, I have not yet delved into a more essential issue: what is "live music"? Is life the same as interpretation? Does it imperatively require some living "appearance" on stage? As a matter of fact, we all wish our music to be as lively as possible; I have never met a composer who enjoyed writing a "dead electronics" piece! By which mechanisms is this vital energy going to be transmitted? And who should perceive it? When, and where?

Since this topic is too thorny to be undertaken straightforwardly, I will first begin with some obvious facts: in a concert hall, the performer's presence and attitude can do wonders and is often necessary and sufficient to infuse life into a piece, at least for all but very few experts. Two components seem to be indispensable: the visible presence of a performer and his or her playing an instrument that is accepted as such by the musical community. Hence, a live piece whose player is hidden backstage and amplified through some loudspeakers around the hall will probably lack a few fundamental "life ingredients". Similarly, a person operating a mixing board on stage and managing to accomplish the most inspired sound projection will nevertheless fail to give the music enough vital sparkles, if nobody else is on stage, since the mixing board is not yet accepted as a real instrument.

Conversely, a piece in which a performer would "fake" to interact with a tape, will be instinctively respected as having its own life, independently on the value of the music. A rough measure of whether or not a performance is real is provided by the audience's reaction at the end of a piece: people simply do not clap at tapes and barely never at "sound projectionists"! But they always react to a performer, even if the performance is particularly lousy.

The visual aspects are so preponderant in a concert, that any situation that does not comply with its ritual is "kindly requested" to search for other spaces and less conditioned audiences! However, when listening to a compact disc at home, music can be as lively as in concerts, although the experience is entirely dissimilar. So, life can inhabit recorded music if we get out of a concert hall and accept that there exist multiple, equally fascinating manners to experience music, and not only an imperial path, the concert, and a cluster of third-class palliatives for when one needs music, but can't afford to go to a concert!

Pop musicians and audiences have already understood it for a long time; why are classical musicians always so timorous? Were we still at the time of piano rolls, all right! But in an era in which technological advances[21] have given us new supports for music of excellent quality, this unilateral fetishism for concerts frankly sounds a little démodé. Why not taking a recording as an autonomous first-class artistic product, neither more nor less enticing than a concert, but different from it? Both are unique occasions to interpret a piece. When will it be normal that a disc is not a sort of concert without a stage, but a new perspective over a piece, one that will never be heard in concerts[22]?

If people accept the recorded images of films as a new form of art, but not the recorded sounds of discs, it is certainly for complex personal, social and psychological reasons, not just to disdain this technology. Getting out of home to go to another place is sometimes sufficient to turn on people's "life receptors". But, sometimes, the refusal to give recorded music the same credit as a concert is just the symptom of a mind encumbered with prejudices.[23]

In 1990, I participated in a concert at the Gnessin Institute in Moscow.[24] Three pieces for soloist and electronics were played; all of them were radically different in their usage of the technology: Pierre Boulez's *Dialogue de l'Ombre Double*, for live clarinet and the same clarinet prerecorded on tape and spatialized over six loudspeakers around the hall, Philippe Manoury's *Jupiter* for flute and live electronics (the 4X signal-processing machine) and my *Traiettoria* for piano and computer-synthesized tape. The performers were remarkable and no technical problems spoiled the concert. The audience received the music enthusiastically, but none of the people paid attention to the technological differences. They all just heard lively music and excellent interpretations. If I had said that Boulez's recorded clarinet was a synthetic sound or that the computer material of *Traiettoria* was performed by me in real time, nobody would have objected, because their primary concerns were exquisitely musical. All the people with whom I spoke were either musicians or musicologists.

As extreme as it may be, this experience has far-reaching consequences: of course, to those who knew about machines, a perceivable difference did exist among the three works. But there is no correlation between a piece using interactive technology and the perception of authentic interpretation in music: interaction is not interpretation, since the latter, if it ideally implies the former, is a much subtler and complex phenomenon, as we have already seen. Life in music sprouts suddenly, when some kind of communication is going on between the performer and the surrounding electronic material, between both of them and the audience and, naturally, when the piece is worth it. Whether the electronics is a modest tape or the "cleverest" interactive system, that's a completely different kettle of fish!

There is also another intrinsic peculiarity of live electronics that is rarely taken into account: the instrumental music is so predominant over the electronics, that the latter is frequently relegated to the role of a faint "halo" of the instrumental material. Unfortunately, there are not so many alternatives: highly interactive systems that play together with a performer via MIDI — and use commercial MIDI-controlled synthesizers or samplers to make themselves heard — sometimes do display remarkably enterprising behaviors, particularly in quasi-improvised situations. But the difference in quality between the acoustical instruments and the performer's articulation, and a "MIDI" sound is so startling that it is like comparing gourmet food with inexpensive can food: no matter how hungry I am, the latter is always problematic to me, and most of the times indigestible. Too bad, for such a sound poverty often turns the listener away from the interesting sides of the system as well.

We have also already seen that real-time synthesis algorithms are usually problematic from the standpoint of sound quality and mostly require special-purpose hardware. So, processing instrumental sounds is often the most reasonable procedure to quickly obtain electronic material of a certain complexity. But this means that, since the instrumental sounds must be played first, in order to be processed, they will probably be heard by the audience. It is very hard to escape from this pattern, as is proved by most of the existing live electronic pieces. In the end, if we also add in the limitations of real-time algorithms we already discussed, the whole process will risk becoming tedious and trite. This is hardly a guaranty of life!

A typical example of the phenomenon of the instrumental preponderance is a piece that is sometimes taken for a milestone of interactive music, Pierre Boulez's *...explosante/fixe...* for solo flute, ensemble and live electronics. There are no doubts that the work is lively and very ingeniously written, with a special color given by a super-dense, yet transparent and mobile orchestration. But where is this energy sparkling from? My feeling is that it comes out in spite, and not because of the live electronics.[25] Replacing all live electronics with orchestral writing might certainly change the piece a little, but it would not at all strip off its vital energy, nor alter its orchestral subtleties, and it would greatly reduce its performance costs! Instead, with the exception of two solo cadences, whose sound quality is rather dull when compared with the instrumental writing, the electronics is almost always present, mostly in the background, often with balance problems with the live ensemble, who is also playing all the time. Aside from a few sampled flute slaps and some successful, albeit conventional, effects of infinite reverb, the electronic material sounds a bit like a thick soup whose taste changes depending on the listener's seat.

Sarabande

In the domain of instrumental music it is not possible to analyze a certain interpretation out of the context of a written score and of a historical repertoire, since it is often simply defined as the performer's own contribution to what is notated in the score. Because score notation and the understanding of musical styles change in history,[26] any performer is to have a good knowledge of the principles in force at a given epoch and of the prevalent interpretive trends.

Considering that, by definition, there is no repertoire if a piece has just been composed, it is not always clear whether contemporary music can be interpreted or is merely "executed". Moreover, in the case of improvised pieces, it is even harder to separate the role of the composer from the contribution of the improviser, at least as one would do it, say, for a performance of Beethoven's Fourth piano concerto.

But with synthetic material[27] not only does a score hardly exist,[28] but a repertoire is just impossible, for the sound result is not known in advance. The fundamental question then is: will synthetic material ever be "interpretable", that is will it ever become more that what it is? Or are we to accept its inherent immanence ineluctably? Actually, before even discussing about its life, what is its genuine nature? If, as pretended in the 50s, composition has finally extended its domain over sound itself, there is not much space left over for interpretation. Sound has to be determined once for all together with the other structures of the piece. In this case, working with a real-time system is highly undesirable, since the composer would have to come up with gestural controls that are not at all required at this stage of the creative process.[29]

But if we seek to infuse "interpretation" into synthetic sounds,[30] we must change both their nature and our view on them. While every traditional instrument is a mechanical structure that does not communicate with a performer only via a sound, but also via its own physical vibrations,[31] none of this is possible with electronic material. On the contrary, what the performer hears through monitor speakers, has often nothing to do with what the audience hears, dynamically shaped sounds in space.

So, interpreting an electronic sound entails a different approach to both sound as such and the meaning of interpretation.

If we do not want to give up our compositional control over synthetic material, because computer music is our only opportunity to compose sounds directly, but we also wish to give it the vital energy of performances, we must adopt a double and complementary perspective on it: synthetic material possesses two "states", a "**composed state**" and an "**interpreted state**". Generated by more or less complex processes, the former is determined at the moment of the composition and is often represented by a sound file or a patch with control parameters attached to it. It acts like a fixed reference and takes the place the score had in instrumental music. During the interpreted state, the reference is modified through real-time processing under the control of an instrument player at the moment of the concert. The set of all the possible changes a given sound can undergo constitutes its "**interpretive potential**", usually specified in advance as a function of the composer's creative needs and of the nature of the material itself. The performer is then able to freely explore this realm of possibilities. Appendix 1 and figures 8 to 10 introduces some more technical aspects of this idea with the proposal of a "virtual interpreter", and shows a simple application of the interpreted state.

That an acoustical instrument or a special-purpose interface be utilized, the performer, by finely tuning his or her choices and motor reactions, is in the situation of an ideal interpreter. Certain subtle mixing effects, like the timbral modifications due to very short delays between the sounds, will be notated with a given musical figure belonging to his or her gestural repertoire. The realization of this figure will yield a unique sound event that is both under the control of the composer (through the written score) and of the performer (through the actual gesture).

In spite of its unavoidable constraints, I imagine that a similar system might generate a repertoire of original gestures and that a new generation of performers will learn to deal with a sound's interpretive potential as naturally as with the traditional repertoire. I hope that eventually the composer's approach will change and that a whole set of gestural typologies specific to each sound family will be discovered and serve as a platform for analyzing computer music.

There is often an implicit correlation between an instrumental gesture and the perception of organicity in sound material, and between the latter and the feeling of a good performance. For example, once a composer has written "crescendo" or "accelerando" the performer still has infinite possibilities to realize it, some of which are musical, while others are clumsy. It is the performer's task to "liberate" a score's implicit naturalness[32]."

When I wrote the radio opera *Proemio*[33] I tried to transfer this feeling to abstract synthetic material bearing no direct reference to any given physical behavior.[34] At first, my attempt seems a paradox: why producing organic and natural sounds with material that is not born from a human gesture? My need stemmed from the very conception of the piece: the three actors' voices recited in an intelligible and very expressive manner, with just a little reverb added from time to time, and are combined with exclusively synthetic computer material. I intended to find out whether a strong connection between the two, apparently distant and unrelated sound realms might nonetheless show up not only at the timbral, rhythmical and articulatory levels. If each realm were able to "emanate" a sort of peculiar, expressive "aura", I would then explore these potential interactions and their psychological resonance.

It was not difficult to ask the actors to recite in a musically inspiring way. I was actually surprised to realize how musical the infinite nuances of the same word can be.[35] But a strong "aura" had to exist for the synthetic sounds as well, and the easiest way for me was to give them an organic shape. How to do it without a physical gesture? Adopting a phenomenological approach, I examined not the gesture itself, but the sound that is produced by it, in order to seize its characteristic acoustical patterns. It is not the place, here, to analyze them thoroughly, especially because they involve both a combination of explicit and intuitive knowledge and a fertile aural feedback that are hard to convey on paper. My purpose was to achieve a perceptually plausible unfolding of the sound energy over time, where "energy" is to be understood in a musical context.

I had no wish to "humanize" the computer; on the contrary, by creating sounds that would also appeal to our "perceptual receptors" as being rich, multidimensional, lively and expressive, I was fully availing myself of the computer's best potential. At first, I expanded the view of sound synthesis that I put forward in *Traiettoria*, by introducing a separation between the perception of a **sound chunk**[36] and its computational reality, the patch or the sound file. So, what our system perceives as "one" sound event is actually a **composite image**, obtained by assembling several short sound files together, each of which contributes to determining some specific features (fig. 2).

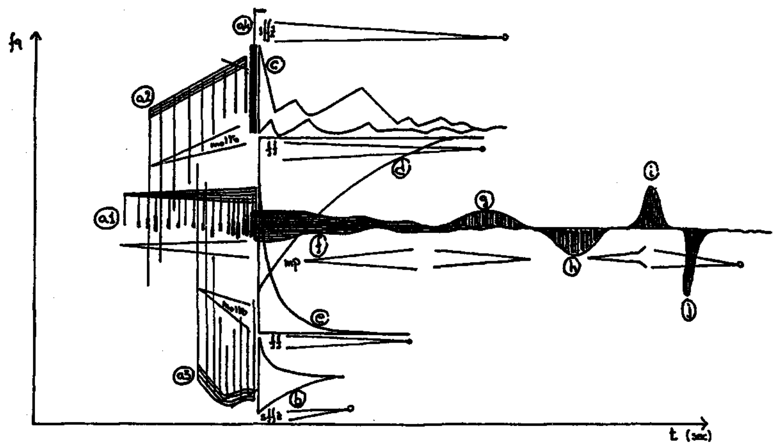

Figure 2 composite sound image, hinged on a unique temporal pivot (**sffz**) and composed of several sound classes. The main accent comprises 5 short sounds, "**d**" and "**e**" constitute the main "armor"; their role is to give the sound its basic spectral structure, its "body", and to sustain the overall loudness. The details of their spectral contents are not important at this stage and are therefore not shown here. Their volume is rapidly reduced and they end up respectively in their higher and lower regions. "**c**" and "**b**" bring a high and low resonance, that help increasing the brilliance of the attack and reinforcing the image's low region. "**f**" to "**j**" represent the compositional part of the composite image: a pitch (it is not essential here to say which one) emerges from an undulating narrow-band noise that gets progressively shorter and thicker. At the end, the pitch is alone. Sounds "**a1**" to "**a4**" constitute the energetic "uplift" of the accent; their role is to prepare the impact of the other events: "**a1**", the longest, anticipates the narrow-band noise, "**a2**" and "**a3**" start from a common spectral area, and aim at a region respectively near "**c**" and "**b**"; "**a4**" is a short "appoggiatura" whose task is to add some timbral brilliance to the accent. Although this is an abstract model, it may be used to analyze most of the sounds of *Traiettoria* (especially of *Dialoghi* and *Contrasti*) and of *Proemio*. Even greater compositional salience can be added, for instance, by having the low resonance "**b**" point to musically significant pitch, while the high resonance "**c**" is built around a harmonic pivot, while following a known rhythmical profile, and so on.

Through carefully controlling by ear, during the mixing phase, the relative dynamics, spectral contents and fine temporal locations of each component, I could generate sound events that were both fully natural and completely synthetic. This kind of process is, in fact, a sort of computer-based micro-orchestration applied to a single sound, half-way between composition and interpretation. I then divided the composite sound images into several **classes**, according to their perceptual salience. Each class had its own "aura" and was developed in an episode of the piece.[37] When an instance of a given class reappeared elsewhere, it was easily recognized and associated to the expressive mood of the episode it belonged to. The composite sound images of *Proemio* have also several other singular properties that are partially described in appendix 2 (fig. 11 and 12 and sound example 17).

At each level of the process, I devised some principles of "orchestration" appropriate to it: each composite image was composed thanks to micro-orchestration techniques; composite images were assembled together to form organic gestural events. When needed, events were juxtaposed to yield a musical flow. Some flows could then be superposed to generate polyphonic frameworks.

This process spurred me to take up a more flexible view of composition, that I also exploited when writing instrumental music: the succession of sound events and their approximate location within a section are specified in advance according to my structural plans. However, the exact placement of each composite image may move around its theoretical insertion point as a function of the degree of organicity I wish. After some experiments, this view was very efficient in discovering the correct temporal position by ear.

Sound example 18 illustrates an application of this process in an overall "adagio" atmosphere. A sequence of three composite images builds up an "organic" sound event in such a way that a given image comes in only when the previous one had the time to completely "manifest" itself. The meaning of "completely manifest" is context-dependent and always refers to the image's phenomenological specificity. The overall context should be a slow sequence and not a superposition of images.

Using the classification introduced in appendix 2, the first image is a relatively bright timbral chord, with five slow undulations. The sound class is of the same family as the one discussed in sound example 17,[38] but with no temporal beatings. Each instance is in its fused or semi-fused state (fig. 3 and sound example 18a).

The second image belongs to another class of the same family, with a more "velvety" timbre and very soft beatings. Only the first swell is used in building up the sequence (fig. 4 and sound example 18b). The third

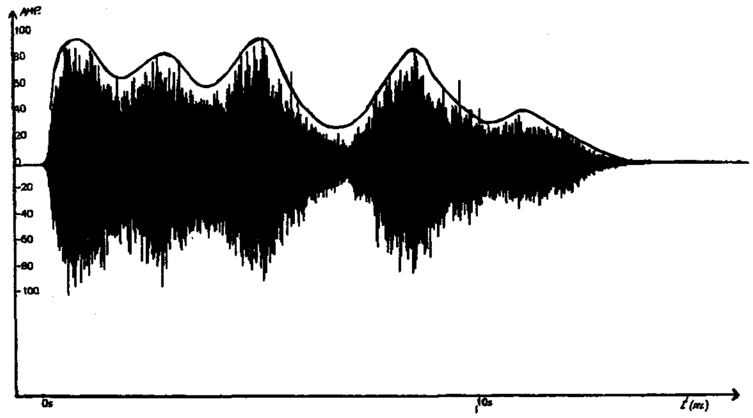

Figure 3 time and amplitude representation of the composite image played in sound example 3a. The five undulations are very visible, as is the gap between the third and the fourth in which the following image can be inserted.

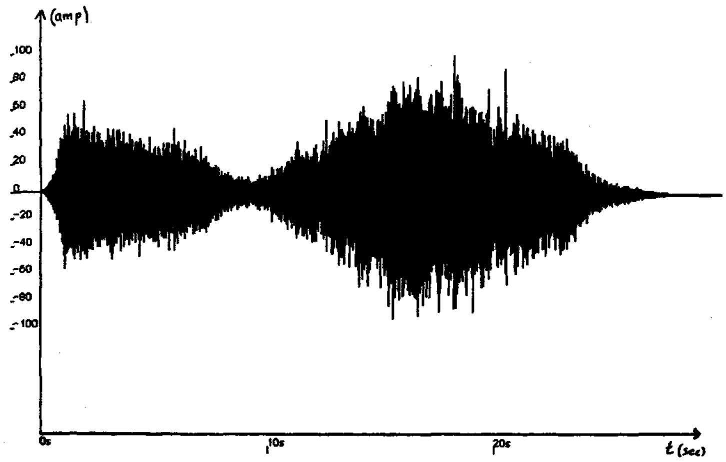

Figure 4 time and amplitude representation of the composite sound image played in sound example 3b.

image is the same instance analyzed in sound example 17b (in its rhythmical state) (fig. 5).

Therefore, the three composite sound images belong to distinct classes of the same family. This assures a good homogeneity. From a timbral standpoint, the brightest image is followed by the most "mellow", while the last is mid-way between the first two. Their inner temporal activity, on the other hand, increases steadily. Sound example 18c plays the final sequence (fig. 6).

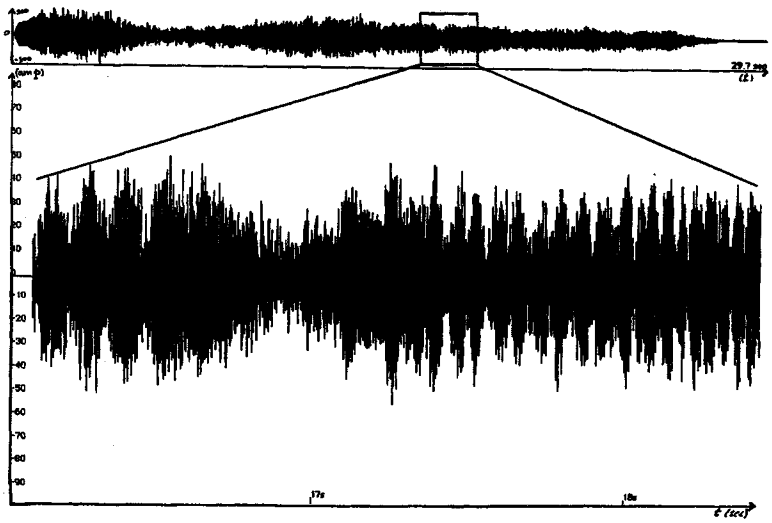

Figure 5 time and amplitude representation of the composite sound image constituting the sound example 17b and the end of sound example 18c. The upper part is the graph of the whole image. The lower part zooms into a 2.5-second segment of the sound, thus showing the interferences between beatings of different speeds.

In order to make the perceptual segmentation easier, each image is panned differently: the first is center-to-left, the second center-to-right and the third starts in the center and gradually "occupies" the whole stereo width.

When I composed *in cielo in terra in mare* (1992), I pushed my research of a musical "life" in radio operas much further. For this piece, my sound palette was extremely varied: eight actors (four women and four men), an eight-voice vocal group singing without "bel canto", a few samples of double bass, concrete sound effects and computer-synthesized material. With so manifold sound realms, my first concern was to elaborate efficient strategies for synchronizing them with each other, without losing their specific naturalness of expression. The way time is dealt with by recitation is very different from the instrinsic temporal properties of a vocal group or a sound effect. In a concert situation, common synchronization references (like a beat or a series of signs) must exist in order to hold everything together. This implies that either something is to give up its natural temporal flow and to comply with the common reference,[39] or that the relation is extremely weak.[40] Since the dramatic requirements of my radio opera called for a very tight, almost virtuoso coupling between all the sound realms, none of the alternatives were exploitable[41].

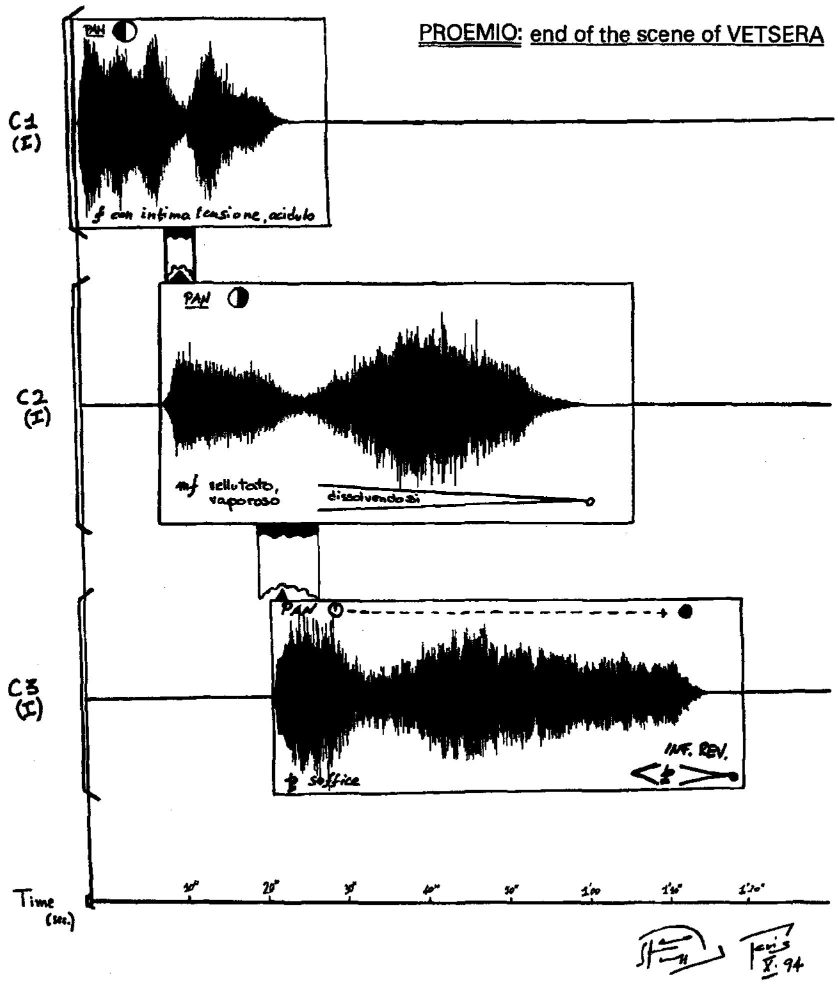

Figure 6 mixing plan for the sequence of the three images showed in fig. 3, 4 and 5. The representation is explained in [Stroppa, Duthen, 1990]. Each source image (the sound that is to be mixed in) is assigned a "temporal pivot" which is connected to an "anchor point" of the target image (the reference sound context). The curled trait under an anchor point indicates the amount of flexibility in the placement of the source as a function of the degree of organicity that is to be achieved.

To synchronize the eight actors' voices, I used "relative positioning", a technique derived from my theory of "temporal pivots[42]": I had the

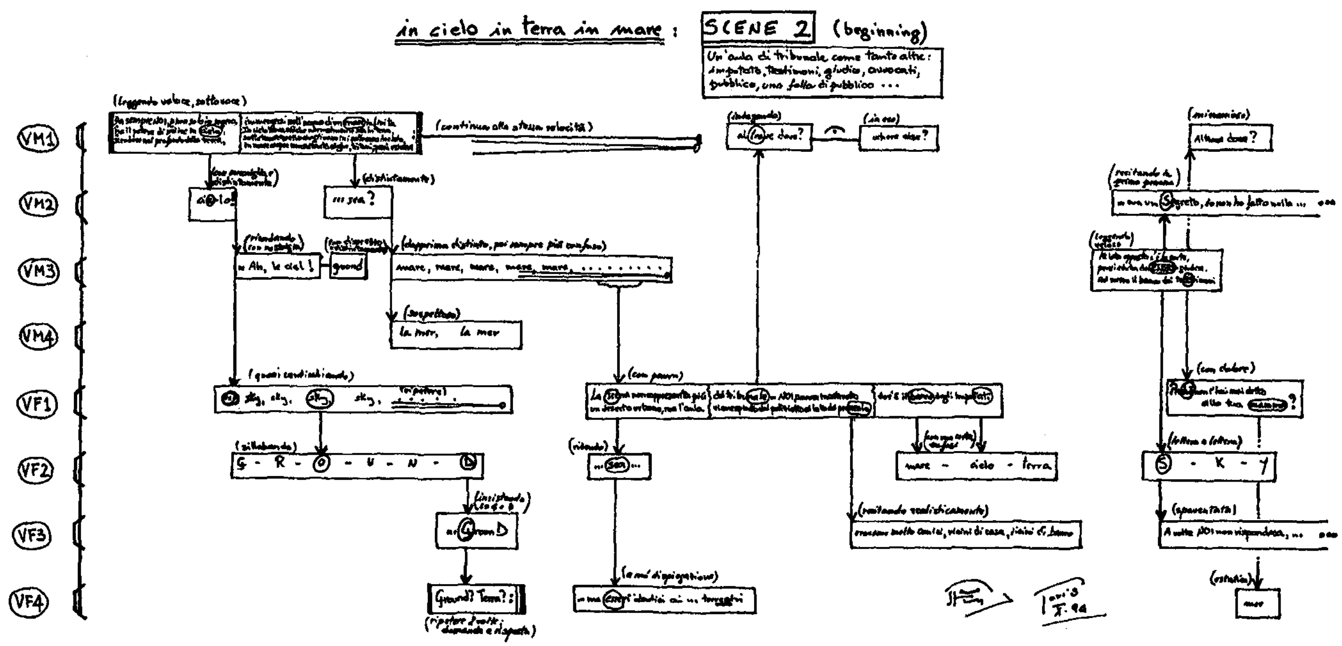

Figure 7 Relative positioning of the eight actors' voices from the beginning of the second scene of *in cielo in terra in mare*. The original script is already in eight columns, one per actor, so that what is on the same row is pronounced at the same time. This score slightly modifies the original text and adds in the exact cues for the synchronization. The speed of the recitation is not shown here for simplicity's sake. VM1-4: male voices 1 to 4, VF1-4: female voices 1 to 4.

leading voice of a section start reciting at his or her own expressive pace. A second voice was then asked to come in when a certain word, accent or inflection was pronounced in the active voice and to continue independently. Each voice had its own external cue, until the scene was complete (fig. 7). Following cues was very natural for the actors, since it did not require any musical training. The same strategy was used with the rest of the material. The excerpt from the end of the third scene (sound example 20) illustrates such a virtuoso-like synchronization.

Dramatic music, however, often requires a stage: in radiophonic productions, this is usually stimulated by adding more or less reverb, so as to yield the feeling of a volume in which the action is being developed. In *in cielo in terra in mare* I intended to transfigure it and compose a whole realm of "virtual spaces". Here, "space" refers to the perception of location and distance within a stereo field, whereas "virtual" means that no efforts were made to produce realistic results. On the contrary, the piece exploits uncommon and illusory spaces that bring in a sort of "cinematographic" relief to the music, with dramatic purposes.

To achieve a better independence, each sound was "spatially" processed before being mixed down onto a 24-track master and had its own unique "volume" attached to it. Only excellent commercial machines were used, but their parameters were "diverted" so as to generate unusual effects. It was a real compositional endeavour: several "spaces" were superposed or embedded into each other, before reaching the final shape. Panning and delays were also used to dynamically modify a sound's spatial image.

Since this is used throughout the piece, sound example 20 will only give a faint idea of the work on illusory spaces in *in cielo in terra in mare*.

Gigue

Starting from some simple reflections about real-time performance, this text dug in the complex issue of "life" and "energy" in music, especially in machine-made music. It is quite normal and reassuring that no simplistic technological solutions exist, for this issue has its own roots in the creative act itself, whose deep essence is both autonomous and original.

I also showed how "life" necessitates some kind of interpretation, but it is not restrained to the presence of musicians on stage. If the circumstances are propitious and properly thought out in advance, "life" will be found outside of the concert situation; tapes, radiophonic music, compact disc recordings[43] as well as all the other available media deserve respect and as highly an imaginative approach as possible.

I do not intend to substitute to "real-time" another dogmatic view, or to oppose concerts to other ways of producing musical art, but to enrich the potential of musical experience and to incorporate today's technological advances. This text is a "critique" of certain aspects of real time in the Greek meaning of the word, that is of revealing what something really is, rather than what it appears to be. I myself often use and love the immediacy of interactive systems, when this is what I need, as I appreciate the excitment of working in a studio and the passionate thrill of composing sound material.

There is still a huge amount of work to be done before computer music can really match the richness of instrumental music, which is normal due its very young age and small repertoire. This should spur us to explore the immense potential at our avail and to crave for fresh experience, struggle for new models, invent expressive musical interfaces and produce creative work. It is my impression that in the last period too many efforts were devoted to purely technological concerns, and that fundamental musical research was somewhat forgotten, to the profit of immediate, but somewhat ephemeral results.

For several reasons, both strategical and technical, these efforts have so much clustered around building real-time systems, that it is still hard to glimpse at any significant aesthetic or musical outcome. If it is the goal of every composer working with technology to get the most lively music he or she is capable of, by no means is it automatic that any piece of music, just because it is real time, by virtue of a silent wand is also... real music!

Gavotte (Appendix 1): overview of a Virtual Interpreter

A Virtual Interpreter (VINT) provides a way to govern the composed and interpreted states of synthetic material. From the perspective of its relation to the overall tempo of the piece, any composed sound material may be either "**punctual**" or "**articulated**".[44] Punctual sounds are single events, without or with a negligible inner temporal life; percussive sounds or homogeneous layer-like textures are most frequently encountered. This material is neutral with respect to the reference tempo. Articulated sounds are usually the product of more complex compositional processes and tend to display a temporal or even a rhythmical structure of their own, which might interfere with the global tempo. These sounds are less neutral and require special-purpose manipulations in order to be smoothly harmonized with the temporal context of a piece.

A VINT must therefore know about and cope with these distinct aspects of composed sounds and adopt the best strategy. A simple functional diagram (fig. 8) should include the following features:

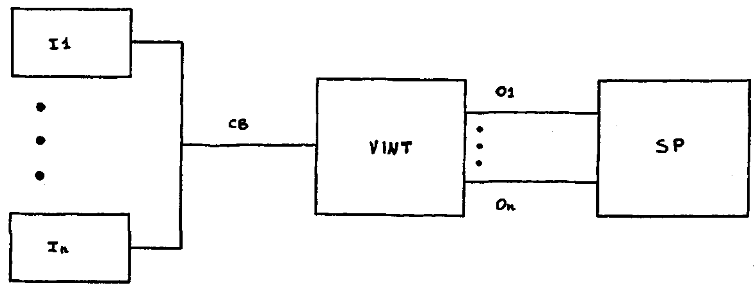

Figure 8 basic functional diagram of a virtual interpreter. I1-n: input interfaces; CB: communication buffer; VINT: main core of the system; O1-n: output channels; SP: sound projection system.

An **interface**, that is any sort of tool able to send grammatically correct instructions to the **communication buffer**. An interface is divided into two components: a **trigger** (from the simplest foot pedal to elaborated tracking systems connected to a performer's playing attitudes) which starts the process and a **lexical "enveloper"** that gives the event the correct syntax by providing supplementary information.

A **communication buffer**, dealing with information coming from several interfaces. Each instruction must contain the following data: a starting keyword, signaling the beginning of a new event and its type, and ID of the interface that is sending it, the address of the sound to be played, an optional code with selected control mode (see below) and the control parameters needed by the event.

A virtual interpreter, the core of the system (see fig.9).

Controls for a multi-channel **sound projection.**

The core features of a VINT are the following (fig. 9):

A score follower, including a tempo tracking mechanism, if needed (see control programs below).

Control programs, dealing with the temporal flow and deciding which processing mode is to be applied. Mode **"direct-to-disk"**: the usual one, playing a sound file or a sample as it is. If tempo tracking is needed, the pitch will vary accordingly. **"Blind"** mode: a sound file is automatically cut into several sub-files that are triggered more or less rapidly according to the tempo of the piece; to avoid pitch shifting, each sub-file is read at its original speed, but will be triggered slightly earlier or later as a function of the tempo tracking; this will produce phasing effects due to the unavoidable crossfades. Although the quality is not ideal, this mode is computationally economic and perfectly works with certain sound families. **"Witty"** mode: the same as above, but the sound file is divided into musically salient segments (the "wit" comes from the user, of course!). **"Stretching"**

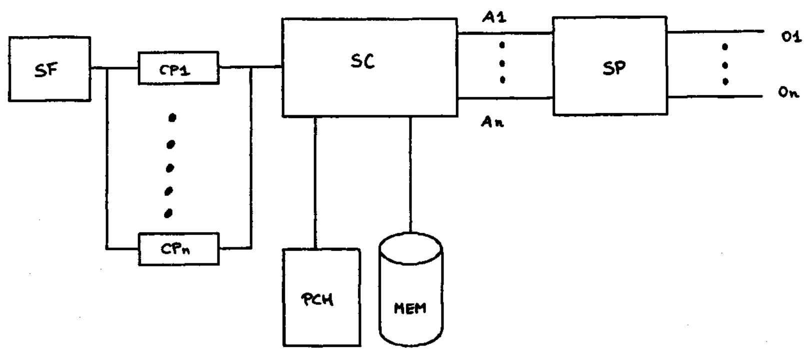

Figure 9 core diagram of a virtual interpreter. SF: score follower; CP1–n: control programs; SC: main sound controller; MEM: hard disk or sampler's memory; PCH: patch; A1-n: audio channels; SP: signal-processing device; O1-n: output channels.

mode: the sound is time-stretched using a phase vocoder like algorithm, which gives the best quality but is computationally very demanding.

A sound processing device that will deal with the interpreted state of the sound (see also fig. 10).

The interpretive potential of a sound is specified while conceiving the sound and controlled by a performer thanks to a signal-processing device. Figure 10 is an example of a simple, perceptually pertinent usage of the interpreted state of a sound.

Let's suppose that a synthetic sound already exists in its composed state. The signal-processing patch available to the performer is made of a low-pass filter applied to the sound's amplitude envelope and of a filter bank applied to the audio signal. The control parameters may come from two MIDI controllers, respectively connected to the amplitude envelope and to the audio signal. For technical reasons, a set of scaling functions is needed (fig. 10). Even within the realm of this modest patch, changing the control parameters and transfer functions will modify an interpretive potential at will: all sorts of experiments are allowed, from the most macroscopic changes to fine tuning a given sound family so as to adapt it to a performer's gesture.

It is not my purpose here to delve into further technical details concerning a VINT, but to show its feasibility and its relatively simple hardware demands. A lot of direct experience by both composers and performers will be needed to explore its possibilities. Isolated features of a VINT have already been used for some time in concerts, although there is neither a general environment that deals with the interpreted state of synthetic sounds, nor a system that efficiently integrates them. Since a VINT always necessitates some control information coming along the

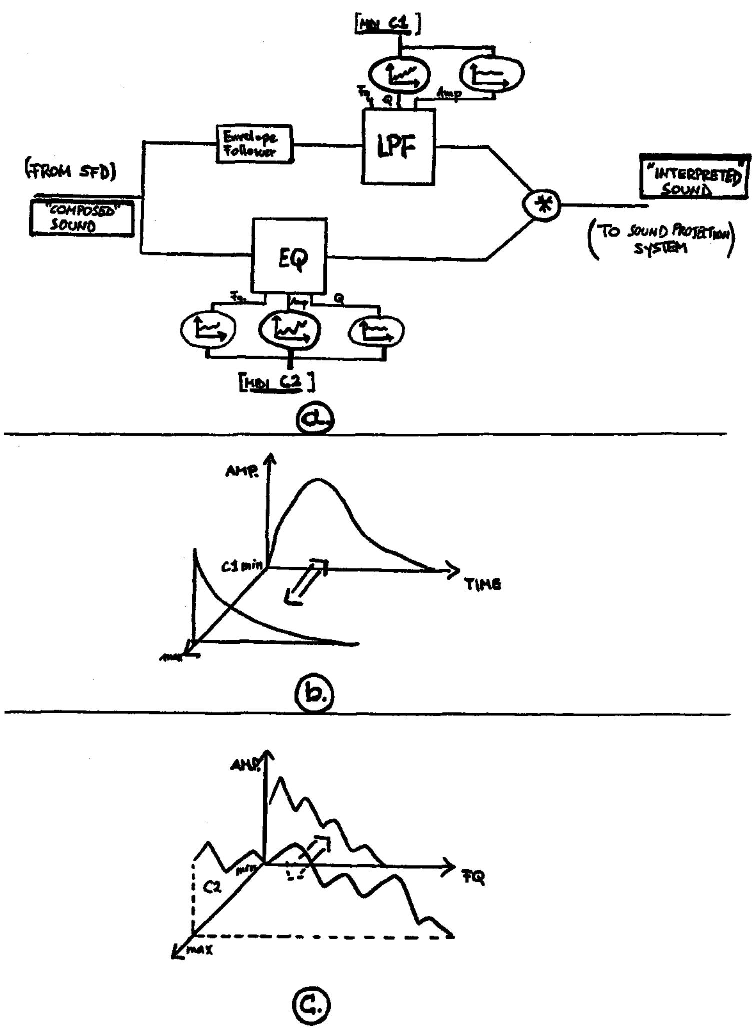

Figure 10 using a sound's interpretive potential. The composed sound is fetched from a VINT. a) LPF: low-pass filter for the envelope follower; EQ: equalizer bank for the audio signal; MIDI C1-2: MIDI controllers 1 and 2. b-c) scaling transfer functions for LPF and EQ.

communication buffer, it is this information, in the end, which shapes the interpretation, that is something activated by a human performer. Hence, no automatic models of interpretation are needed at this stage. Efficiently modelling, coding and controlling the interpretive potential of a sound still requires a lot of musical research, which inevitably lies at the heart of the aesthetics of live electronic music.

Musette (Appendix 2): the main structure of Proemio's synthetic material

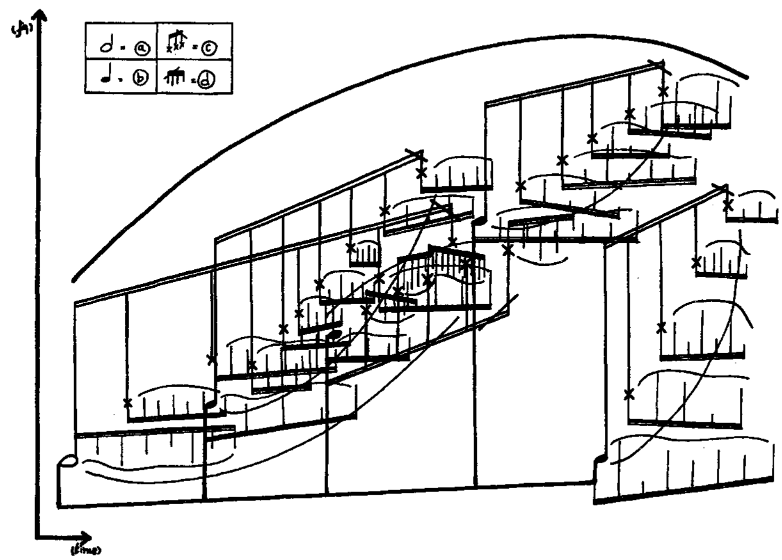

Figure 11 organization of a typical sound class in *Proemio*: **a)** main pitch; **b)** secondary pitches; **c)** pseudo-harmonic spectral reinforcements (shifted spectra); **d)** beatings pivoting around the main and secondary pitches.

Most of the sounds in *Proemio* are "twilight entities" ceaselessly on the threshold between unit and multiplicity: they are too rich to be perceived as a "simple" event (like, for instance, an ordinary note played on a piano), yet not articulated enough to sound as a full structure made of elementary components (like, to pursue the same analogy, a chord or a phrase played on the piano). Depending on the context and register, these sounds appear either as composite or as unique elements. From my poetical point of view, they are the synthetic counterpart of the "Musical Information Organisms" (see Stroppa, 1989).

The sound class presented here is composed of four embedded layers (fig. 11): a predominant pitch (main pitch), some weak pitches (secondary pitches), an "ad-hoc" spectral support, usually composed of shifted partials and a set of regular beats of varying speed.

In the high register, the class is in its **fused state**: the instances have one predominant color corresponding to the main pitch. The secondary pitches do not stand out well enough and the beatings, falling beyond our perceptual system's threshold of temporal discrimination sound like brilliant spectral sparkles (sound example 17a).

Figure 12 fused state of the sound class used in examples 17 and 18: **a**) melodic motif (each note is played by a different instance); **b**) chord sweeping from a fused state (higher pitches) to a totally defused state (lowest pitch). The structure of the chord is based on the melodic pattern of the sound's pitch references (**c**).

When transposed down, however, a curious effect of structural zooming shows up: since the sparkles are now within our capabilities of temporal resolution, they will sound like regular beats. The instance breaks down into a very complex world of pulsating units around some pivot pitches (sound example 17b). It is the class's **decomposed state.** Micro-orchestration techniques help correcting the undesirable effects of plain transposition (both longer and less brilliant sounds), through context-dependent filtering or by cutting away the unneeded portions of the sound.

Taking advantage of the fused state, melodic contours can be created, like the one of sound example 17c (fig. 12a). By combining instances in different registers, all the states will be present at the same time. In a chord like the one shown in fig. 12b (sound examples 17d and 17e), the ensuing acoustical relief is often unexpected: while maintaining at first its original harmonic color, this chord unfolds itself very organically and is much richer than if it were played on an acoustical instrument.

Similar phenomena, of course, are never accidental and must be carefully worked out. This faculty to juggle with our perception is very typical of and quite unique to computer music. There is a whole panoply of new dimensions to explore, that will confer a fresh vitality to harmonic, rhythmic and melodic textures.

Sound examples

Ex. 16 (2′ 43″): excerpt from *Dialoghi* (from 7′ 57″ to 10′ 40″ of the Wergo recording). Pierre-Laurent Aimard, piano. Reproduced with permission.

Ex. 17 (1′ 55″): instances of one sound class from *Proemio*:

a: "fused state" with temporal sparkles (6″ 22). The main pitch is B flat 4 (slightly higher).

b: "decomposed state" (same sound as before, transposed almost three octaves down): the temporal sparkles turn into regular pulses of varying speed. The first low pitch (C2) corresponds to the main pitch, but other pitches are also clearly perceivable, as are most of the partials (29″ 24).

c: melodic motif made of the same instance in its fused state, in a tenor-like register (fig. 8a). The instance's main pitch is transposed on each of the pitches of the motif. While the motif is unfolded, the "inertia" of the other components (secondary pitches, partials and beatings) gradually builds up, making the whole process more predictable (31″ 09).

d: seven-note chord containing the same instance in its different states: the chord is here slowly arpeggiated so as to single out the main pitches of each instance (12″ 11).

e: the same chord played as a whole. Its timbral and articulatory richness is impressive. At first, the harmonic color of the seven pitches stands out clearly; then, the sound melts down into a kaleidoscope of pulsating spectra that get progressively darker and slower (21″ 04).

Ex. 18 (1′ 21″): building up a sequence of three composite images:

a: first image: slowly undulating timbral chord (16″ 05).

b: second image: two velvety swells, with soft temporal beatings. Only the first swell will be used in the final example (30″).

c: complete sequence, including modified panning: center-to-left (first image), center-to right (second image) center to all over (third image) (24″ 17).

Ex. 19 (3′ 49″): excerpt from *Proemio*: scene of Medea (from 18′ 16″ to 22′ 05″).

Ex. 20 (3′ 33″): excerpt from *in cielo in terra in mare* (from 21′ 30″ to 25′ 03): end of the third scene ("scena buffa" or comic scene) and beginning of the fourth ("the trial").

Acknowledgements

This text was inspired by a series of courses I gave at the 1994 Bartók Seminar in Szombathely, Hungary.

Notes for the CD

Traiettoria (1982–84), three movements
(*Traiettoria . . . deviata, Dialoghi, Contrasti*) for piano and computer. Commission of the Venice Biennale for the European Year of the Music.

Computer music produced at the "Centro di Sonologia Computazionale" of the University of Padua. Ricordi Editions.

Compact disc recorded in co-production with IRCAM, Paris: Wergo, Computer Music Currents n. 10, WER 2030-2, Pierre-Laurent Aimard, piano, Marco Stroppa, sound projection.

Proemio (1990), a radiophonic opera for 3 actors' voices and computer music on a script by Adolfo Moriconi.

Commission of the Italian RAI, Radio 3.

Computer music produced at IRCAM, Paris. Actors' voices: Alfredo Bianchini, Paolo Poli, Pamela Villoresi. They were recorded in the studios of the RAI in Florence.

The plot: in the form of a musical drama, where words and music are mutually and closely connected but never meet each other in the form of a singing voice, two characters (the "Author" who is two people to indicate contradiction, collaboration and mediation of the inner self) are about to write their preface ("proemio" in Italian): a preface that is a beginning and an opening, but also a reflection, research, impact, intuition, synthesis.

They start by telling each other exactly what Giovanni Boccaccio (the greatest Italian writer of the XIVth century) said at the beginning of his "Decameron", "it is only human to have compassion for the afflicted" and then, little by little, they discover the reason why there are so many women among the afflicted who need consolation. The words, their ideas are not yet clear: they reflect, argue, discuss, repeating end even throwing in each other's face some of those words. During the tensest moments, three childlike women appear, one after the other (Maria Goretti, Maria Vetsera and Anna Frank); they are historical characters whom the authors see and converse with. The women are in turn transformed into other feminine characters, fictitious this time (Dante's Beatrice, Eurypides' Medea and Ibsen's Nora). Only after these meetings do the authors know how to express their preface clearly and properly. Finally, though, these words no longer belong to the author but to the computer. As a new owner, the computer decides to make some words clearer, some more confused and to manipulate others; therefore those words will have another reality. Those words, rewritten in a different order, will convey other meanings and other reasons. The authors wonder, is it the present or the future?

Text of the excerpt:

Giovanni:	Non ha potuto resistere all'amore . . . come . . . Medea . . .	He could not resist love . . . like . . . Medea . . .
Boccaccio:	. . . o Medea . . .	. . . or like Medea . . .
Vetsera:	(*Tragica e dolce al tempo stesso*) Fra gli esseri del creato che vivono e pensano . . . noi donne siamo i più infelici . . .	(*Tragic but sweet and soft at the same time*) Among the human beings who live and think . . . we women are the unhappiest . . .
Male choir:	(*lontano*) δνσινχης...δνσινχης	(*far away*) δνσινχης...δνσινχης

Vetsera:	... prima, per comprare uno sposo, forti somme ci sono necessarie e solamente con la dote possiamo ... e questo è un male assai più grande dell'altro ... avere al fianco un padrone del nostro corpo ... ma come si mostrerà, dopo, il padrone?	... first, to buy a husband, a lot of money is necessary and only with the dowry can we ... and this is an even greater evil than the other... live side by side with our body's possessor... but how will the possessor show himself, afterwards?
Male choir:	(*lontano*) αφενιης...αφενιης	(*far away*) αφενιης...αφενιης
Vetsera:	Buono o malvagio?	Will he be kind or cruel?

in cielo in terra in mare (in the sky on earth in the sea) (1992), a radiophonic opera for vocal group, 8 actor's voices and computer music on a script by Adolfo Moriconi.

Commission of the Italian RAI, Radio 3. Computer music produced at IRCAM, Paris. Vocal group: Electric Phoenix of London (8 solo voices), recorded in London.

The actors were recorded in the studios of the RAI in Turin.

The plot: NOI (we, us) has always been endowed with the power — reaching well beyond the realm of his dreams — to rise up into the sky, descend into the depths of the earth and submerge himself in the waters of a limitless sea. What is normally to be seen on land he encounters in the sky, what men place in heaven he finds beneath the ground, while in the sea he discovers not only seaweed, tritons, fishes and mermaids, but also beings that are identical to land creatures.

After many years — NOI is now prevalently a schoolboy in the third grade who is not always attentive — he reveals his secret to a companion of his own age, who for a long time remains the exclusive sharer of that extraordinary knowledge.

One day at the hospital however — NOI is now prevalently a lovely young nurse — the ward sister is bewildered and irritated by NOI's sudden and protracted episodes of absent-mindedness, and so punishes her very harshly. In despair, NOI confides the true reasons for those lapses to her fiancé, who rushes to explain the situation to the ward sister.

As from that moment, NOI knows no peace: he is arrested, locked up in a cage and put on trial. The case is on everyone's lips by now, including television and the newspapers. ESSI (they, them) are stricken not so much by the multiple identities that NOI is able to assume — a child-person or an elderly one, a woman, a man — as by the inconceivable absurdity of NOI's dreams, little mattering whether these occur while NOI is asleep or awake. With the sole exception of friends and family members, ESSI have few doubts: whereas to alter one's age, gender and physical appearance is, in the end, but one of nature's straightforward little tricks, to mistake the earth for the sky, on the other hand, and to imagine the existence of sea-living beings who are in all respects similar to their terrestrial counterparts, is a crime which must be stopped at once and punished as severely as possible.

In his cage as the days pass by, NOI is losing contact with the outside world. He is left with his visions, but these too become more rarefied and less and less distinct... so that between one round of questioning and the next he patiently awaits that great purifying event at which ESSO-judge (HE-judge) so frequently hints in whispers. The fateful day has arrived. As NOI prepares for the approaching end, ESSI who represent the authority and the law make their entry one by one, each wearing an identical mask over his features. They enter the cage as a single body and silently lead NOI toward the altar. To the utter amazement of all presents, the altar rises... rises up into the sky, then plunges deep into the earth and finally vanishes into the abyss of the unending sea.

Aural guide to the excerpt (the original text is in eight columns; what is clearly perceivable in the music is underlined and followed by an English translation within parentheses; the indications of time are from the beginning of the exerpt):

SCENE III: the change of age and sex (comic scene, end)

–B–Coda:

–a–Choir [start of excerpt], *"fotoscultura"* bombastically sung in the foreground, as if it were the most important thing in the world, followed by *"bagolamento fotoscultura"*, recited in Italian with a strong British accent, as do certain tourists when reading an Italian conversation booklet.

–b–Scene IIIb (follows immediately [0:06]): (ESSO-<various professions>, running after each other) " . . . *spettatrice TV . . . asettico . . . tutti i giorni a tutte le ore, sempre, sempre . . . ESSO ESSO ESSO . . . seguono inseguono, seguono inseguono . . .* " [TV watcher . . . aseptic . . . every day, at every hour, all the time, all the time . . . HE HE HE . . . they track him down and follow him, track him down and follow him . . .]; in the end some "perky" noises, few shy kisses, always with several voices at once.

–c–Choir and actors: *"bagolamento"*, bombastic, as usual, while the actors continue running after each other with words: " . . . *ESSI che lo seguono ed inseguono . . . spettatrice TV tutti i giorni, a tutte le ore, sempre, sempre . . . ESSO-giornalista . . . ESSO-famiglia . . . ESSO-amico d'infanzia . . . che lo seguono e l'inseguono . . . (risate) . . . ESSO-medico con l'ostentazione di un rigore scientifico . . . rigore, rigore, rigore . . . fantasia del cronista, giornalismo e fantasia . . . ESSO-famiglia con la sciocca trepidazione di chi ama e non capisce . . .* "[. . . THEM track him down and follow him . . . TV watcher, every day, at every hour, all the time, all the time . . . HIM-journalist . . . HIM-family . . . HIM-childhood friend . . . who track him down and follow him . . . (*laughter*) HIM-doctor with the stubbornness of a scientific rigor . . . rigor, rigor, rigor . . . a reporter's imaginativeness, journalism and imaginativeness . . . HIM-family with the silly trepidation of those who love but do not understand . . .]. A bashful kiss and a gentle smack end this episode.

–d–Finale [0:44]: in a whirling and restless motion, mutterings of amazement and a tongue twister on the word

"bagolamento" (recalling some kids that play at the Indians), while THEM are always harping on the same string: " . . . *che è stato il primo a saperlo . . . ESSO-lettore di*

giornali . . . straordinario . . . " [. . . who always knew it first . . . HIM-newspaper reader . . . extraordinary . . .]. The quips keep going . . . until a last, sudden, unusually amplified wince [1:16] interrupts them with its metallic squeak. It's a strange noise, as an ominous rising blade: it triggers a final frightened, fleeting run (*"uomo, donna, vecchio, bambino . . . man, woman, child, öreg, kind, . . . gunaika"*).

SCENE IV: the trial (adagio religioso).

Inside the prison cell, a cell like any other: the prisoner awaits the end. Over a low menacing computer background, a sudden "stroke" introduces this scene. The whole atmosphere is "slowed down", totally opposite to the gay frenzy that preceded it. Simulacra of previous characters walk by in a fatalistic pace. A tragic ineluctable doom seems to have invaded everything.

-A-Computer: A sudden shock starts the computer's background. At the same time:

–a– 2 actors (a man and a woman) [1:34]: a cry in the foreground, albeit extremely faint; the man is trying to hold himself back, whereas the woman's crying is more manifest.

–b–computer [2:17]: the same stroke begins a descending scale made of increasingly lower double bass pizzicati, until they are turned into a cavernous grunt.

–c–actors [3:10]: THEM repeat the words "death" and "capital punishment" in several languages (*Todesurteil, morte, halálra ítélés, pena capitale, death, halálra ítélés*), in a hopeless mood.

–d–choir (split into four groups of two singers each, starting with two basses): the excerpt fades out right after the beginning of this episode.

Notes

1. When closely examined, this term is, in fact, rather ambiguous and can relate to a great deal of applications in quite heterogeneous contexts. In this article, however, for simplicity's sake, I intend any special-purpose machines, mainly used in concerts, such as signal-processing cards or effect processors, excluding the usual sound-projection equipment. My remarks are concerned with certain musical views that are often associated with this technology, and not with any specific applications. Likewise, the usage of interactive systems as a quick test in studio productions, or of real-time hardware as a means to speed up computation is not taken into account here. As a matter of fact, the terms "interactive" and "real-time" are not always interchangeable: the first refers to a sufficiently rapid response of the machine permitting an ordinary communication to take place, the latter requires so fast a reaction, that no delay is perceivable between the command and the result and usually implies a direct control over some of the algorithm's parameters.
2. Pierre-Laurent Aimard, piano, Michel Cerutti, percussion, Etienne Bultingaire, sound engineer. Théâtre du Châtelet in Paris.
3. In this text I will always write "tape" to indicate a support where the temporal flow is somewhat fixed and cannot be continuously varied during the performance. However, nowadays the difference between absolutely rigid media and totally flexible systems tends to be increasingly smaller. Tapes are often replaced by multi-track direct-to-disk or by samplers with a lot of memory. It is not difficult to imagine that each single event might be an independent sound file played back more or less automatically according to the global tempo. While this is a purely technological solution, my concerns tackle a more general problem, independent on any physical choice.
4. For the detection of isolated events a simple envelope follower will work properly. But for more complex situations a pitch detector is needed coupled with a pattern-recognition mechanism, whose role is at least to "understand" the errors of either the performer or the pitch detector. This is usually quite reliable for monophonic parts with little noise. However, tracking polyphonic music (from a keyboard to a group of instruments, even if they have separate inputs to the machine) or performing more complex recognition algorithms (such as the detection of a whole chord or of a melodic pattern) is still very time consuming and extremely hard to carry out reliably.
5. A serious analysis of score following and tempo tracking would be fairly long and complex; not only are the current results amazingly crude from a musical perspective, but also often absent from the concert sites. It is not rare to listen to so-called live-electronics pieces in which the tempo and the other control parameters are completely fixed, in a sort of real-time simulation of a direct-to-disk playback system! Evidently, finding salient controls during the concert is a very difficult undertaking. There even exist some pieces where the interaction between a performer and a MIDI device such as a synthesizer, an effects processor or a sampler does not go beyond a sequencer with immovable tempo and constant dynamics!

   It is also very instructive to compare a real-time vs. a direct-to-disk version of the same piece (such as Kaija Saariaho's *Noa Noa* for flute and electronics). Most frequently the best musical interpretation is the latter! Trapped by invisible eyes and flawless electronic ears, chased by a fragile tracking mechanism, uncertain about the behavior of the system in case of errors, the performer of a real-time piece will inevitably give the priority to playing as correctly as possible; as a result, the interpretation becomes stiff, aseptic and often mechanical. On the other side, fearing for potential failures, the musician in charge of the sound projection will be monitoring the state of the system rather than concentrating on the musical aspects of his or her task.

6. This might explain why works for conducted ensemble, in which a reference tempo is almost always mandatory, are more problematic and yield inferior results than solo pieces when a tape is used.
7. See the analysis of a composite sound image further down and fig. 2.
8. Nowadays, the substance of my choices would not change, although I could partially work at home and use more advanced hardware, like, for instance, a direct-to-disk system — where each musical flow is recorded separately as a long stereo file — and a slightly better strategy to start the sequences.
9. For a more extended analysis of this concept see [Stroppa, Duthen, 1990].
10. Aimard, Stroppa, cit., p. 20–21. As a matter of fact, the rigid unfolding of the computer sounds turned out to be even beneficial for the pianist, since it provided a safe reference around which to turn freely. Actually, it is often more problematic to come up with convincing temporal strategies to combine the different musical figures in the sections for piano solo than in those for piano and computer.
11. I always marked the musical effect that is to be obtained instead of the concrete actions on a mixing board, because they depend on local factors. For instance, most of the crescendi and diminuendi will be produced through filtering, by displacing the spectral energetic balance towards the high register of a sound flow, rather than simply increasing the sound pressure level.
12. For a usage of this phenomenon, see the beginning of *Contrasti*.
13. Of course, a performance of music for tape solo requires totally different spaces from concert halls that are imperiously calling for something to happen on stage. There is nothing more depressing than an empty stage! Unfortunately, the great enthusiasm for alternative musical experiences that enflamed the research of composers in the 60s has now left its place to a very traditional view of concerts which, "de facto" is highly discriminational against anything that doesn't require a performer on stage. I hope the next century will find again the excitement of fresh adventures!
14. Some of the late music by Luigi Nono is an excellent example of a very simple, yet cogent usage of real-time processing algorithms that are particularly well adapted to the composer's language and instrumental writing.
15. Mixing a piano's left foot, mapping a violin's tailpiece, blowing a lonely reed or playing a tuba with a triangle bearer might certainly produce interesting sound effects, but it won't highlight the performer's skills. These effects are hard to faithfully reproduce, refinedly shape and incorporate in a compositional plan. Conversely, using the most efficient modes of vibrations of an instrument and benefitting from a performer's gestural repertoire will probably generate more traditional sound figures, but will also allow for a much finer articulation.
16. Since the beginning of computer music, a very popular procedure has been to make a continuous timbral transition from an instrument to another (for instance from a trumpet to a voice), a kind of acoustical correspondence to the wide-spread "morphing" in computer graphics. However, my reaction has always been very mild: the timbral quality of the first instrument would get more and more mediocre until it reached a kind of "no man's land", where it was frankly boring, to suddenly jump to an increasingly more specified image of the second instrument. I couldn't hear any real continuity, or the transition was so fast that I couldn't tell the difference with a banal cross-fade between two audio channels. The more archetypal the instrumental sounds or the articulation, the stronger was this effect.

17. How many times have we heard the most trite infinite reverb or harmonized sound? There is a lot to explore in this area, but this always requires a high degree of musical imagination and a painstaking hands on experience to feel the potential of each algorithm in relation to every possible sound.
18. This approach combines an analysis and a synthesis phase, both computationally demanding, like, for example, an FFT analysis providing amplitude and frequency values to the corresponding channels of a resynthesis bank.
19. that is either the units learned during musical studies or those associated to signal analysis.
20. This is one of the major issues of all electronic music, yet one of the least worked out. There is a huge mismatch between the richness and complexity with which an instrument radiates sound in space and the way a loudspeaker behaves. Hence, material coming from loudspeakers tends to sound dull and lacks depth of image and presence. To ameliorate it, one would have to create, for each sound family, a virtual source with several loudspeakers and a signal-processing machine changing radiation patterns as a function of frequency and amplitude. So, a rich electronic texture might require a few hundreds audio channels to be accurately diffused in space. But this appeal is still today purely utopic.
21. Among them, the advent of digital recordings, of home-based sound processing and editing systems, and now of the video disk, of computer graphic software and interactive videos.
22. This remark applies, of course, to all musical styles, and not only to electronic pieces! One would need, though, a much more intimate collaboration between composers, interpreters, sound engineers and production companies, another schedule for the recording sessions — but not necessarily more sessions — and, chiefly, a creative approach to recordings and imaginative solutions. Today, this seems in thorough conflict with the show-business attitude of most recording companies and of too many complaisant performers. But various courageous attempts are fortunately still under way. One should also not forget the remarkable example of Friedrich Gulda.
23. There is also a social aspect to going to a concert or to a movie: the feeling of "communing" with other people and sharing the same event in the same place. This aspect will never be reproducible elsewhere and I have no wish to replace it with a home-based virtual-reality environment! But this should not prevent us from exploring other opportunities to disseminate art as well.
24. It was on February 25, in main hall of the Gnessin Institute, part of an IRCAM / Ensemble Intercontemporain tour to Moscow and Saint Petersburg.
25. I do not question, here, the intellectual justifications for using live electronics, nor its strategic importance. I am also confident in the composer's seriousness in controlling the effect of the signal processing and in the professionalism of the technological setup.
26. As an example, in baroque music a performer has the right to abundantly flourish a written melodic contour according to a set of codified conventions. In its principles, this somewhat recalls today's jazz modules.
27. For "synthetic material" I mean what comes out of speakers, whether they are concrete, processed or synthetic sounds, that is what has an electrical, rather than a mechanical source. The special case of an amplified instrument is to be included in my definition if the amplification significantly transforms the acoustical image of the instrument, but not if it is just a sound reinforcement, for balance purposes.
28. For an articulated account of the difference between a musical score, an aural score and a computer score see Stroppa, 1991.
29. It is probably not an accident that the composers at that time did not use any of the live-electronic instruments at their disposal, and perfectly accepted tapes.

30. It ought to be clear, by now, that using some random processing or varying the amount of reverb added to a sound so as to change it during a performance is not interpreting it, since there is no relationship to a style or a repertoire.
31. One notable exception is the organ with an electrical console, but this is considered much inferior to mechanical organs. Other electrical instruments, such as the Hammond organ or the onde Martenot, only have a very peripheric role in our repertoire.
32. When taken from the intellectual side, this issue will lead us to endless discussions about what is "natural", "organic" or "artificial". But any correctly trained musician has such an infallible feeling for what is musically natural and efficient that no definitions are necessary.
33. For a more complete discussion of my two radio operas in the context of the Italian tradition of radio dramas, see Ferrari, 1995.
34. More precisely, "abstract" means that I did not attempt to purposedly simulate any properties of physical instruments or to use any data derived from analysis models of acoustical sounds. However, an instrumentally-biased hearing would still recognize organ-like harmonic structures that are due to my avoiding the typical attack characteristics of mechanical instruments. Although this conditioning has psychological roots that consitute a real problem in appreciating abstract sounds, it will not be analyzed further here.
35. The secret musicality of a recited word is lost when singing, because other connotations are more apparent. The "ouverture" of *Proemio* grows out of the multiple timbral, rhythmical and expressive resonances of a single Italian word: "proemio" (preface).
36. taken in its psychological meaning of one single piece of information storable in short-term memory.
37. See the plot in the notes for the compact disc. The global structure of *Proemio* has a kind of "recitativo-like" backbone (corresponding to the intervention of the male voices) around which turn six lyrical episodes (assigned to the female voice). The whole is preceded by an "ouverture" and concluded by a double "finale".
38. A sound class is a set of (potentially infinite) sound instances sharing certain phenomenological specificities. A family is rather a collection of classes with similar characteristics (in object-oriented jargon a family would be a "metaclass").
39. This is the case, for example, of rhythmical recitation. Although it can produce impressive results, as in Stravinsky's *Histoire du soldat*, its severe expressive limitations would have been unfit to the dramatic framework of *in cielo in terra in mare.*
40. as in Prokovief's *Peter and the Wolf*, where the synchronization is reduced to starting a recited passage at specific moments of the piece.
41. As a matter of fact, musicals adopt some interesting strategies to give the recitation a richer expressiveness, while still controlling the synchronization with the instrumental music.
42. see Stroppa, Duthen, 1992.
43. "The recording of *Traiettoria* was designed to be an original studio product, and in no way reconstitutes a live performance with its spatial projection. The computer's depth of sound therefore appears flattened, since it no longer enjoys the spatialization of a concert hall. Yet it acquires a highly pronounced stereo quality which affords excellent fusion between computer and piano. The high-dynamic recording of the piano was done with six microphones placed at three different distances. The nearest pair of microphones underscored resonance, while the two other pairs furnished a close auditory image (main microphones) and a distant image. This system allowed for a nearing or distancing of the instrumental image depending on the context (. . .). The two main microphones were directed toward the piano so as to give the instrument a "stretched" spatial image along an axis with the bass to the left and the treble to the right. This

approach gives listeners the impression of being within the piano itself. This decision not only clarifies relationships between the auditory images created by the tape, but above all permits a spatial organization of the various musical figures. Passages involving abrupt changes in register are therefore very striking from this standpoint." (Aimard, Stroppa, cit., p. 23).

44. These are not independent states, but two poles of a continuum. The composite sound images of *Proemio* are always on the threshold between the two states. Computers are perfectly suited to explore these ambiguous and somewhat illusory worlds.

Personal Bibliography

Ferrari, G. (1995) Vers un scénario imaginaire: l'opéra radiophonique italien. In *Les Cahiers de l'IRCAM, n. 6, Musique: Texte,* p. 71–80. IRCAM Edition, Paris.

Stroppa, M. (1989) Musical Information Organisms: An Approach to Composition. In *Music and the Cognitive sciences,* S. Mc Adams and I. Deliège eds, *Contemporary Music Revue,* Vol. 4, Harwood Academic Publishers, London, 1989.

Stroppa, M., Duthen, J. (1990) Une Représentation de Structures Temporelles par Synchronisation de Pivots. Proceedings of the Symposium: *Musique et Assistance Informatique,* Marseille, 1990).

Stroppa, M. (1991) Un Orchestre Synthétique: Remarques sur une Notation Personnelle. In *Le timbre: Métaphores pour la composition,* J.B. Barrière ed., Ed. C. Bourgois, Paris, 1991, p. 485–538.

Aimard, P-L., Stroppa, M. (1992) Program notes to the CD published by Wergo, pp. 20–21.