SOUND DESIGN FOR NAVIGATION IN TOPOPHONIES

Roland Cahen

ENSCI – Ecole Nationale Supérieure de Création Industrielle
48, rue Saint Sabin 75011 PARIS
cahen@ensci.com

1. ABSTRACT

This paper attempts to develop concepts of multimodal objects made out of the interaction between space and sound though motion. The present paper is mostly descriptive and tries to gather experiences in order to delimit a new field of applications, concepts and method. My purpose is to describe some specificities of sound navigation design, starting from everyday life experience and cultural backgrounds towards virtual reality. My point of view will be that of a sound designer. I will examine different methods of implementation related to several sound navigation concepts and metaphors related to this particular approach. For this purpose, I will take different examples mainly in my works to show how some kind of scenarios propose different ways to link motion to sound, what kind of sound behaviours they produce, how they can be created and controlled, what kind of human and artistic experience they offer.

2. INTRODUCTION

Sound navigation is an everyday life experience. The word *topophony* [1] chosen for the title of this paper has already been used [2] for different matters related to sound and space. It is appropriate to designate structured sound fields, spatial sound architecture, graphic or spatial playing sound interface, navigable scores, etc. More precisely, I would like to focus on the meaning and the rendering of movement or navigation, of a person / subject / avatar, between different situated sound sources (fixed or in motion) and on the interaction with sound situated objects or beings. This figure, which we call sound navigation, is significant in real life as well as in sound virtual realities or in audiovisual interactive displays. For virtual worlds, sound navigation may be considered as a realistic kind of evidence, unless the aim of the subject's movement is to produce sound and music. But considering sound navigation also appeals to some specific questions such as why and how movement produces sounds and how they can be conceived and designed.
To a certain extent, sound navigation can be considered as a way to compose music or create sounding trajectories through motion in space. It differs from *sound browsing,* another related expression which usually refers to "an exploratory, information seeking strategy that depends upon serendipity ... especially appropriate for ill-defined problems and for exploring new task domains" [3]. In an artistic or leisure context, where the purpose is less functional, more aesthetic and expressive, the concept of sound navigation seems more appropriate. In addition, *sound browsing* architectures are often abstract and do not use a spatial metaphor.
There are many situations and artistic traditions in which spatial motion and sound or music production are linked together, such as walking sounds (footsteps, etc) and walking songs, aboriginal *songlines*, dancing with musical instruments, music instrument playing, music games like Rez, Frequency, PHASE, Guitar Hero [4] etc.

Figure 1. *The Art of Dancing, explained by Reading and Figures Kellom Tomlinson (London 1735), book II, Plate XIV*

In such examples, the sound-producing actions and modalities may be different but all have coordinating movement in space and sound production in common.
The visual aspect of these musical expressions is not compulsory, but when present, it has an essential function of locating in space. Therefore the visual and the sound are complementary, helping one another [4]. I am not interested here in making correlations or relations between sound and vision as in a synesthetic way. Visualization here appears to be a representative reduction of the mutlimodal real space. Space is linked to time, because moving through space simultaneously composes a sound sequence and a visual pattern linked to one another by the interaction of space and time in a multimodal approach [5].

3. SPACE, TIME AND MOVEMENT

In *Musique Architecture* Iannis Xenakis [7] distinguishes between *out of time* "hors temps", *in time "*en temps*"* and *time/temporal "*temporel*"* object categories or classes. To simplify, let me say that music material, such as scales, is *out of time*, score is *in time* as a way to organize structures between *out of time* and *time* objects, whereas music and sounds themselves are *temporal*. This is certainly true when the time axis is represented on the score, but what happens on a kind of score where there is no representation of any time direction. Would it even still be a score? Does a score necessarily needs a time axis? Is the time axis the necessary link between out of time and temporal objects? What about if time was simply defined by the spatial trajectory between objects or places? An attempt is made here to reply to these questions though different examples and application.

4. SONGLINES

In *The Songlines,* Bruce Chatwin [8] describes how Aboriginals mythology creates an oral music map of the earth. The ancestors sang the world and in the present, a person who *walks about* and sings the song of one path, brings to life its myth. Therefore, each path / song seems to be composed of mythological stories and spirits (dreaming) and some descriptive elements of the physical path. Aboriginal people can know hundreds of songs. Thus, the whole territorial map is a musical corpus of songs all related as in a 2d web, or if this makes sense, the brain itself. The geographical representation completes the musical threads' architecture, in a multimodal mental figure. The songs are mapped to their paths as a music corpus on a map.

"Richard Lee calculated that a Bushman child will be carried a distance of 4,900 miles before he begins to walk on his own. Since, during this rhythmic phase, he will be forever naming the content of his territory, it is impossible he will not become a poet." Traditional cultures are often linked with singing epics about travelling, such as the Exodus or the Odyssey. Singing and music are strongly rooted to travelling and motion. Walking is a fundamental reference for music as well as a spatial experience.

5. SUBJECT CENTERED EXPIERIENCE

In the VR sound installation *Icare* [9] the user could virtually fly over a local map up to the burning sun. All objects were sonified and spatialized. When we presented the installation at Ircam in 1998, we localized more precisely the most important sources with the *spat*, especially the moving ones that produce continuous sound. These kinds of actors (such as another user who moves independently in the same scene) have a movement of their own, different from the player's and the world. We wanted to be able to move the position of the subject, relatively to the world and the moving actors, which was not yet implemented in the spat at that time. The subject reference then became predominant and the sound rendering with spat was outstandingly enhanced by the Ircam spatializer compared to a simple volume panoramic. But in order to have a good sound spatialization, it is necessary to simulate a room acoustics, and it is not possible to achieve the same quality result for an open-air sound field. The audition device was a VR headset and the experience was limited to one (user) point of view. As referring to subjective everyday life experience of sound, the primacy of the subject in sound navigation is very important. The subjective point of view allows the user to act as in the real world. This does not mean that fidelity to reality is the only measure, because symbolism and convention can replace realism, but incoherencies are awkward. The aim is achieved when the experience become real (rich) even if the world is only a simulation. Some sound navigation metaphors are not subject-centered, they can be multi-user, the points of view, of listening and the origins of actions can be dissociated.

Figure 2. *Icare [7]. The user can fly moving his arms, represented in the world by the avatar of the two blue wings in the foreground all objects and movements having autonomous and interactive sound behaviour.*

6. NAVIGABLE SCORES

We have tried to figure this metaphor in the concept of navigable scores as defined in the ENIGMES project [10]. Traditionally, scores are separated from instruments for practical, physical and autonomy reasons. But also because of the limits of mechanical instruments, and probably because executing music in following a score, may be considered as a technical process. The possibilities offered by digital interactive graphic animated objects allow interactive scores, which are also instruments, to be created. Interactive score is an interesting metaphor for sound navigation. It questions the relation between indeterminate real time movement of the player on the score and a predetermined music time shape.

In order to be a navigable score (or a *score–instrument*) such a device should allow real time sound navigation (including the control of time objects) and possibly time reshaping :

1) Moving a cursor or an avatar on the *topophony* and interacting with the notation itself produce the sounds and the music.
2) It is possible to program evolutions of interactions with the notation. As if the instrument would transforms itself according to the score, the action etc.

The instrument and the notation then somehow merge. How can we play, how can we compose, what kind of music can and cannot be made this way?

7. BROWSING THROUGH SPACE AND MUSIC INTERPRETATION

Music *composing while playing* experience includes a direct encounter with sound-producing objects though movement. What can be done in real space, can also be implemented in virtual reality and be used to produce games sounds and music control, data *sonification* (giving significant sound representation to data) or artistic creation. As a field of artistic expression, sound navigation proposes free or constraint exploration with music. In the PHASE game and installation [11] (Fig. 6) this exploration is at the same time contemplative and competitive allowing the player to move continuously from artistic expression to game action. The most surprising discovery was that music occurred both in the very sensitive slow exploration and in the fastest race where virtuosity was required.

8. ENCOUNTERING OBJECTS

The subject/player influences the objects sound behaviour depending on the possible interactions.

Type	Sound production cause
Self producing sound object	Autonomous
Interactive sound object	User action on object
User action sound without object	Movement
Off sound without object	External reason

Table 1: Main types of sound objects and sound behaviour

As in the real word, the situated origin of the sound causal action can be distant, close or in contact. Sound interaction can happen with or without contact with the objects. Interaction with distant objects may happen by using a medium like throwing objects, controlling a flow, etc or without a medium or with a metaphoric one, such as pointing to an object to control its sound behaviour. In sound navigation the gesture of selection can be the action at the same time.

Touching or hitting objects can be performed in different ways and with different instruments: the most usual are a cursor or a mouse and triggering a related sound, but the cursor can take different shapes: point, line, 2d surface or 3d objects. Cursors can have their own behaviour, for example using the metaphor of electrons turning around a nucleus as in Plumage [12] [VIDEO PlumageCuts-s.mp4], particle projections, etc. In Plumage, the goal is to trigger many objects, corresponding to organized sound grains, in a structured way. Three reading heads performed this, each of them surrounded by 3 rolling triggers (Fig.3). The cursor can be attached to a physical model and controlled with a specific tangible interface to emphasize the physical experience, for example force feedback or another *enactive* device [13]. But the simplest interactions: approach and contact already raise many interesting questions.

The distance between two sound objects determines the time that separates the two interaction events in their visual and sound expression. Thus, the speed of these sound events occurrences is dependant on the proximity of the object and the facility to move from one object to another. The trajectory can be lengthened by taking a longer trajectory or shortened one, or by going faster from one point to another.

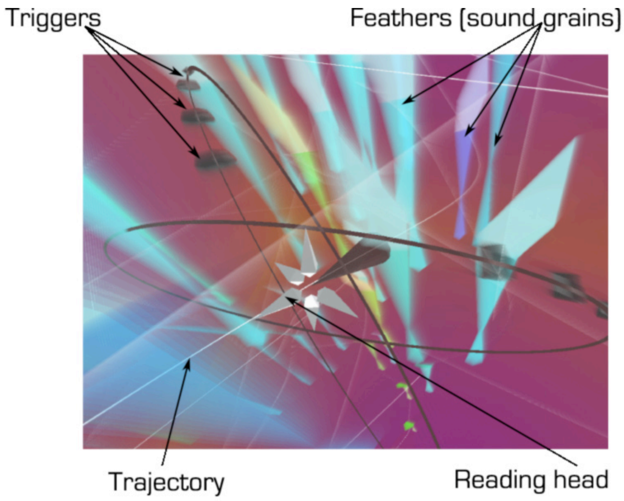

Figure 3. *Plumage real time corpus based synthesis design Yoan Ollivier, development Christian Jacquemin, Diemo Schwarz.*

9. INDETERMINATION

This freedom can be used either to improvise or to interpret, remix, play or execute etc. The degree of freedom, the existence and the symbolisation level of a music model can determine the difference. Improvisation may signify moving freely through a corpus of objects as in “Piège à Rêve” (Fig 4) [10] [VIDEO PiegeAReve-s.mp4]. Interpretation would rather mean a limited freedom in replaying a pre-composed trajectory. This freedom limitation can be realized by: creating constraint paths, succession rules, putting distance between events or by changing the topophony along the time. The interpretation generally evolves from one performance to another, which can be done totally by memory and invention but also by modifying the score itself from one interpretation to another. For example comparing in real time one’s trajectory to all other users’ interpretation, using for example statistical interactive representations.

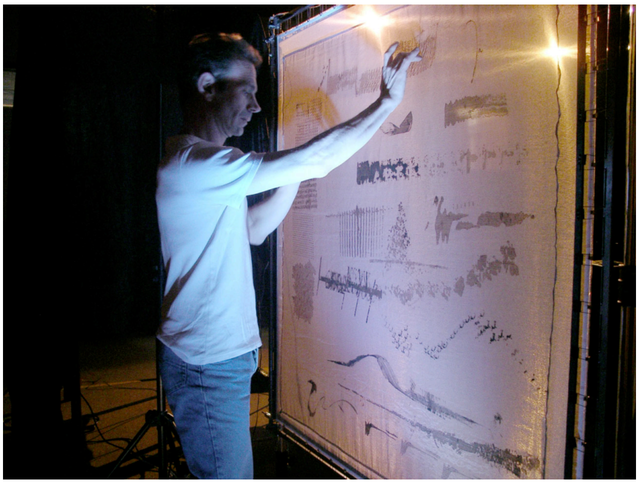

Figure 4. *Piège à Rêve : textile interactive score device by Maurin Donneaud, music Marco Marini.*

Moving hands on the sounds representations on the material plays the sounds.

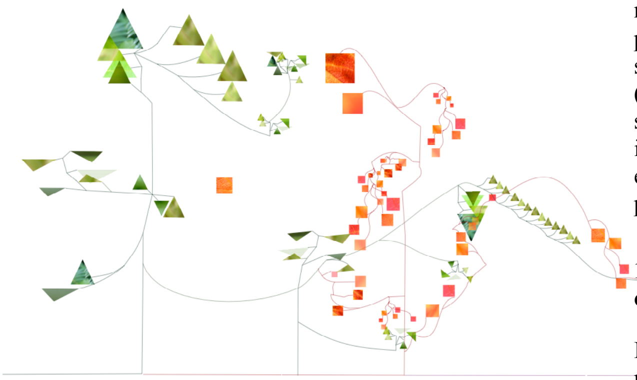

Figure 5. *Feuillage fragment of an interactive score: simple 2d graphic implementation by Katalina Quijano, music Eric Broitmann*

10. COMPOSING AND DESIGNING TOPOPHONIES

The activity consisting of composing topohonies and sound interaction can be called topophony design and/or sound navigation design. It requires music composition skills as well as visual and spatial design ability. But it also means controlling the interaction modalities in order to produce the expected musical result. Ideally, the composer should be able to create the whole process by himself or at least to conceive it: music conception, gesture design and related graphic design. In the ENIGMES project [10] we privileged the use of musicals models. The composers Eric Broitmann and Marco Marini both composed electroacoustic music pieces, which we disassembled, to reconstruct them in an interactive way on navigable scores. In this second phase, we designed the topography (graphics) and the real time sound interaction devices in order to make the scores playable. To start the interpretation process, the capability to reproduce the original piece as closely as possible demonstrates that the interaction device is accurate enough. This is only significant if there is a musical model, if the interaction is more sophisticated than a simple player and allows interpretation features. If there is a model, its reproduction may be easy, like following a visible trajectory. Interpretation variations may consist of shifting from the model line in time, distance, taking shortcuts, passing repeatedly, approaching the objects differently, modifying them, etc.

11. REAL TIME INTERACTION AND ELECTRO ACOUSTIC SOUND LIMITATIONS

11.1.Polyphony

We have two hands to play percussions, 10 fingers and two feet to play keyboards and pedals, but digital interactive devices are still mainly limited to one or a very few interactive cursors. It is important to distinguish between cursor polyphony, action polyphony, and sound objects polyphony. The polyphony may therefore be often limited to one action voice (one action at a time), because augmenting the number of simultaneous controllers or parameters makes the use more complex and longer to master. In that case, the instrument controls must be particularly sharp or the sound interesting enough to play solo. The solution is often to mix slow interactive sounds (long sounds with little interaction) with a fast interaction solo part, producing a subject on a background in an interactive way. Reintroducing real time interaction to electroacoustic music reduces the polyphony to action possibilities and sound object time behaviours. I

11.2.Real time sound interaction, redundancy and calculation limits

If one programs a sound to figure a reaction to a percussion movement, the usual choice would be between a synthesizer or a sample, but the synthesizer will generally sound very artificial and the sample will be repeated without changes. The next step is to control envelopes and filters applied to samples, which may be more satisfactory but quite processor-costly, especially when variable polyphony is required and hardly corresponds to the gestures. It is quite difficult to produce a coherent and rich sound behaviour this way, for example to perform good transitions between different kinds of reactions. It is not a question of realism, but rather a question of design and credibility. The tendency nowadays is to create physical models of acoustic objects, but at the moment only a few models of musical instruments can be used and tools such as *modalys* [14] are processing-costly and not yet powerful enough to be used easily in complex real-time devices, which often include hundreds of objects and sound behaviours.

11.3.Spatialization and timbre

Physical modelling is certainly a good approach today, but electroacoustic sound design and restitution do not allow as many different sounds as would seem to be possible. The problem may be that we often assimilate a real physical sound with its simulacra coming out of any loudspeaker. As a matter of fact, most real life sounds are not only spatialized but also attached to physical objects. Dynamic sound radiations are transferred to objects, surfaces, resonator, etc. This aspect of timbre is fundamental to produce sounds we could qualify as kinetic sounds. There are lots of them, such as sounds of friction. The model for friction is very close to an iterative movement of hitting successive objects distributed in space, like rubbing a computer keyboard. In this kind of sound spatialization, the sources do not move, but a particular movement triggers a great number of events in place. Then, introducing motion and spatialization of both the subject and the sound sources, the number of required spatialized objects may blow up the processor capacity. This is why many electroacoustic interactive sounds are often metallic, glass-like or poor instrumental imitations. For sound navigation, the physical body, tangible or simulated, of sound objects is decisive for the localization and precision of the actions.

12. SOUND NAVIGATION METAPHORS FROM REAL TO VIRTUAL WORLD

The model of sound navigation movement is the body movement reduced to an avatar (cursor). Virtual interactions with objects are generally much poorer than in real life, even if more imaginative. Making proper music encountering nearby real objects requires skill. The more realistic the topophony is, the more skill is needed from the user. Complex gestures like grasping throwing, rolling, extending, tearing, etc are not used very often to produce music, probably because they are not very easy to use as triggers or controllers.

The sensory-motor daily life experience is used here as a reference model. But interaction design simplifies the gesture and allows many expressive variations the user would not have been able to produce with an instrument [13]. This skill is limited by the device affordance, the user's comprehension and his practice.

Referring to the metaphor of the earth and the sky, especially in 3d is convenient, because it allows at least two very different kinds of motion: i) on the ground, interacting with it and the elements on it and ii) flying in the air. Flying in the air sound ambiances are not that easy to produce. In 3d animation for architectural projects one of the principal motion to illustrate is flying over the roofs of buildings. In real life, unless you are a bird, that may only happen with a noisy machine. Often, you would really like to recreate the audition experience of a bird or a person on a flying carpet and this is what the sound designer is supposed to do. In order to create certain realism, this effect can be achieved by simulating the movement of a large number of sound sources, individually synchronized to all possible events as you pass by each of them. A correct implementation may require to drive simultaneously and in real time, time behavior and space control of each source.

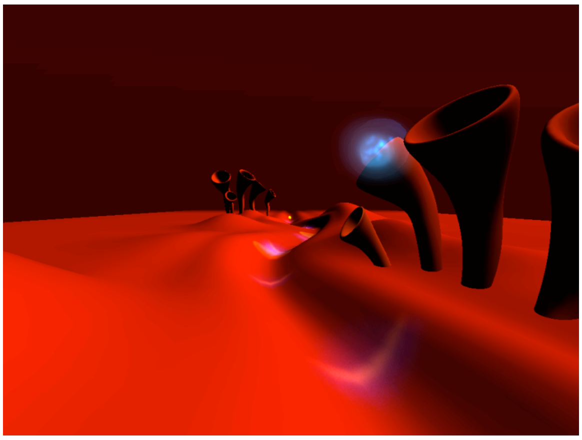

Figure 6 *PHASE A writing head, (faraway light spot) prints the music notes in a groove (clear spots in the groove) while the player holds the reading head (blue light). The player can interact and play music with the objects. When the reading head is positioned in the groove, the ground starts moving and a music race between the writing and the reading head happens.*

12.1.The reading head metaphor

The cursor can be compared to the reading head of a tape player, which produces the sound attached to a space item. In PHASE installation, the music is generated by a music automat and carved in a 3d groove by a writing head entity. The player is holding the reading head, with 6 degrees in and 3 degrees out haptic feedback arm [11], and is pursuing the writing head like a car race duel. [VIDEO PHASE-Demo-m.mp4]

12.2.Cameras and microphones

Distant objects may not be heard when they are too far away. An audibility zone for each object and a volume function, comparable to the level of detail in the image is defined in order to map the volume to the distance of the audition point. The audition point is transformed in a virtual world into a microphone.

The point of view (camera), the point of audition (microphone) and their link are very important. In games the camera most often show either the view of the player or the view of the avatar. If a camera is placed at the user's point of view, and the microphone is placed on the avatar itself, it may be satisfactory for a subjective audition in certain conditions, but the sound effect may seem disconnected from the global sound mix.

It can be interesting to install several virtual cameras and microphones for different functions. For example, you may be able either to watch and hear the whole scene or to be looking and hearing from the point of view and audition of a vehicle like a moving object. The point of view and the point of audition may be different if the listening function must be distinct from the view one. For example if the microphone is used to isolate and listen manually to different sources to arrange them, it is preferable to see them all.

12.3.Spatial mapping

Mapping sound behaviours on objects consists of giving sound to actions. Actions are temporal. There are several kind of sound interactions: from simple triggering to more sophisticated behaviours as flow or distribution control, conditional conducting, time browsing, etc.

The three categories of Xenakis: *out of time*, *in time* and *temporal* are interesting to articulate mapping strategies. Mapping links these three paradigms to gesture. Gesture is temporal and can be mapped to *temporal* objects in a real time control manner, may conduct *in time* structures and manipulate to *out of time* objects. For sound navigation, the topophony is *out of time*, but can be at the same time *in time,* when the time orientation is determinate and *temporal* when the topophony itself evolves.

When mapping a sound behaviour on a spatial pattern and playing with it for some time, a *privileged sound mapping shape* generally emerges. It can be defined as the shape of the usual pattern of gesture produced to obtain the more interesting or chosen effective result. Some mapping tuning is necessary to optimize the shape in order to improve the sound result and the movement, for example expand the movement resolution in detail-required areas. This shape depends on the shape detection, the mapping and the sound generation. This shape can also be designed by adapting the shape detection and the mapping to designed topophonies, as in *Les îles,* in order to associate gesture and sound behaviour in the appropriate spatial representation [VIDEO LesIles-Marine.mp4].

Figure 7 *Les îles Marine Rouit time mapping surfaces*

12.4. Spatial sequence and time tracking

Time objects such as music phrases can be mapped in space, either on a line or a curve, as in (Fig 7) or round a circle, etc. Curves can separate or merge, proposing variations, or folding the sequence moments into different threads of space. Different moments can also be mixed together or shifted in space echoes. Gesture activity can also be mapped to progression in a sequence giving freedom of movement.

12.5.Spatial and performance scenarios

In sound navigation the route the user chooses determines the course of the scenario. There are several ways to construct a scenario. It can be continuous, progressive, divided in regions, levelled, etc. Topophony design influences musical development. When performed in public, the progression of the gesture and the sound creates theatralization. The way it is composed and designed determines the performance.

In instrumental music and in electronic music, the visual aspects are pre-determined by the mechanics but generally not very closely linked to the music structure or the sound. It is often laborious to identify the relationship of cause and effect between the sound heard and the visual action which produces it in an orchestral concert. The visual action of performing electronic and electroacoustic music is even less significant. In interactive audiovisual performance, the visual and the actions can be built up the way the artist wants. Thus it is possible to choose what appears and what does not. A topophony can be used as a kind of program map and action substratum shared between the artists and the audience.

13. CONCLUSION

It is possible that sound navigation in topophonies may sound like a vast or loose domain that should be restricted to be more consistent. But it makes sense, because its situated specificities diverge from other interaction concepts. If it should be restricted, it may be according to accurate metaphors or applications such as the ones described in this paper. Artistic realisations and scientific works are welcome to develop the field.

14. REFERENCES

[1] *Topophonies and score navigation in ENIGMES* (ENACTIVE/07)

[2] Thierry Madiot, Pascal Battus, Eric La Casa and others artistic action: topophonie.free.fr

[3] Eoin Brazil, Mikael Fernström ic *Where's that sound? Exploring arbitrary user classifications of sounds for audio collection management.* (ICAD 2003)

[4] Matti Gröhn *Comparison of auditory, visual and audio-visual navigation in a 3d space* CSC - Scientific Computing Ltd

[5] Y. Bellik, S. Ferrari, F. Néel, D. Teil, E. Pierre, V. Tachoires *Interaction Multimodale Concepts et Architecture* (publication Loria Pedagogie)
Y. Bellik *Interfaces Multimodales: Concepts, Modèles et Architectures.* Thèse de Doctorat, Université Paris XI, Orsay, 1995

[6] Roland Cahen, Xavier Rodet and Jean-Philippe Lambert, Springer Verlag *Sound Navigation in PHASE Installation: Producing Music as Performing a Game Using Haptic Feedback.*

[7] Iannis Xenakis *Musique Architecture* (Casterman 1971) *p. 42 - Musiques formelles (ed. Richard Masse)*

[8] Bruce Chatwin *The Songlines* 1987.

[9] *Icare* a virtual reality artistic installation by Ivan Chabannaud, music by Roland Cahen, engineer Bruno Herebelin, production CICV Pierre Schaeffer Montbéliard 1996

[10] *ENIGMES* http://projetenigmes.free.fr
ENIGMES: la partition navigable (IHM 07)
ENIGMES: Compte rendu de projet
http://theatrophone.free.fr/ENIGMES/CRProjet-ENIGMES-25.pdf

[11] *PHASE project* : Ircam - CEA-List – Haption – Ondim – Graphic design Félicie d'Estienne d'Orves and Xavier Boissarie – Sound design Roland Cahen

[12] Diemo Schwarz *Musical applications of real-time corpus based synthesis* (ICMC 2007)
http://mediatheque.ircam.fr/articles/textes/Schwarz07b
Christian Jacquemin *Plumage: Design d'une interface 3D pour le parcours d'échantillons sonores granularisés.* (IHM 07)
Christian Jacquemin *Visualisation 3D interactive: contrôle et intégration de l'audio*
http://sic.sp2mi.univ-poitiers.fr/GT-rendu/articles/Jacquemin.pdf

[13] Enactive devices are devices that propose a body experience of knowledge. The concept of enaction, was created by Francisco J Varela : T*he Embodied Mind: Cognitive Science and Human Experience. MIT Press 1991*

[14] Ircam software for modelling physical interactions for programming instrument models.