

Acousmatic discourse and sound projection under the new multichannel surround formats.

Past, current and future.

Daniel Schachter

National University of Lanús, Argentina (UNLa), Department of Humanities and Arts,
Center for Studies in Sonic Arts and Audiovisual Production (CEPSA)

Audiovision Degree career.

dans.ds@gmail.com

1. Introduction

This is a part of a greater research project, in which currently I'm working in at the National University of Lanús (Argentina), referred to the influence that the affirmation of surround formats as a standard for the concert and for the global commercial distribution should have in the way of thinking the sound discourse. After the EMS'09 Conference this paper was also presented at the EMUFEST'09 in Rome. The complete research does not refer exclusively to electroacoustic music. It also looks for their sway on different musical genres. In fact we can affirm that the surround sound formats have a strong influence in products thought for massive distribution where the better or more convenient way to obtain a high quality output has to do with the market success probabilities, but even in such cases there is a lack of uniform criterion about the format in which a discographic production may see the light.

2. Is this a valid matter of discussion for electroacoustic music?

The evolution of recording media technology (and the great advances in physical data storage formats) have had a strong impact on the electroacoustic music discourse since the early times of the analog magnetic audio tape to the present digital domain. As time went by, each step forward in the technological evolution brought important changes in the way of listening, thinking, composing and performing being therefore a decisive factor to define the final form of a sonic art composition. During the second half of the XXth century we may recognize some key milestones which had such an important role, as for example the transition from stereo to the early experiences in discrete multichannel sound, trespassing the borders first to quadraphony and later to octophony.

In fact, the confirmation of four and eight channels as new standards for the concert, was undoubtedly a big step forward for the composers of acousmatic music. The CD standard introduced (the possibility of having) an acceptable media for distribution... with clear sound and a larger dynamic range, and this added the need to obtain the final output of each new multichannel piece not only in the format thought for the concert, but also on a necessary stereo reduction which should be available on audio CD for distribution as well as for radio diffusion.

The evidence that it's not the same to think on a two channel universe than in multiple discrete loudspeakers does not only have an aesthetic side, but also has to do with the ability to work with a multi channel environment in mind, distinguishing among different models of these, and developing the ability to identify the most suitable format for each project. This paper tries to focus on the idea that the recent most widely and generally accepted physical and logical standard data formats designed for multi channel surround sound may assume an almost paradigmatic role in acousmatic composition, just as digital sound on CD has been in the 80's and octophony in the 90's.

To understand this idea, we should compare those influences with the new ones, related to more recent technologies, and of course, we should include some possible alternatives for converting already existing compositions from stereo or 8 channels to surround formats and vice versa. But, at the same time we should not forget the massive distribution of music on the internet as well as the huge production of low quality sound devices and the development of software tools designed for online sound compression, streaming, uploading and downloading as a weak point. The blooming of social online networks as well as Internet radios and other sites that include fragments or even complete electroacoustic pieces in lo-fi show that this aspect, which is apparently out of the scope of this paper, has also become a valid matter of discussion in our field. So we may guess that the future may show at the same time the evolution of hardware and software towards higher quality devices on one hand, together with lossy compression formats and players on the other.

3. Differences between the acousmatic and other sound discourses

The acousmatic discourse is in fact a very specific experience. So, when thinking about the influence of these new technological resources in the acousmatic field, we need to take into account its main differences with other musics. A piece of acousmatic sonic art, works on the idea of a deeper perception not related with the sound sources: The perception of the unseen. This is fundamental, as the surround sound formats were thought as a solution for the movies' soundtracks, where the visual references are almost permanent. So, there is undoubtedly a big difference between the sound for the cinema and the "cinema for the ears" which is one of the main definitions of acousmatic music. More, in surround sound for the cinema the front central channel is thought for dialogs, with a clear motionless reference there, where the central channel reinforces the front stereo image leaving a less significant role to the rear channels. So, how to manage the front central channel, will be one of the challenges when adapting a 5.1 for the acousmatic discourse.

Acousmatic music into a large loudspeaker system should provide a wider dynamic range. Quadraphony and Octophony introduced the spatial placement as a variable and any stereo reduction for CD distribution ends on a strong limitation of the spatial deployment. If we think on an original acousmatic multichannel piece transcribed into a surround environment, our transcription should retain or recreate as much of the original spatial placement as possible, and more if we think in surround sound as an alternative from the starting point.

So, thinking for instance on 5.1 as the original loudspeaker distribution, the composer should consider a particular or innovative strategy in the management of sound Gesture, including even some inventive layouts related with the use of Texture and also a satisfactory and complex use of Trajectory (terms introduced by Dennis Smalley in his article "*Spectromorphology and Structuring Processes*"), considering that he would not have to reduce it for the version to be distributed on a surround sound media.

4. Physical Media and Logical Formats

When doing a research work about the impact of the new surround sound systems, we also have to consider the Physical Media and the different available Logical Formats. This was not a main point before, because the CD Standard at 44 (.1) KHz 16 bit without compression has been also one of the most usual formats for Digital Audio, and also the only possible logical format for the available physical medium. There were no options in that field. Now instead, there are many different available logical formats for sound on DVD. So, we need to do a right choice to ensure that we are actually taking advantage of this multichannel structure.

The reduction to CD of a multichannel piece, or just the CD version of a stereo composition always had a great importance -or it was even an obsession- for the acousmatic composer, who normally used to make some revisions for that specific edition, different to the one thought for the concert, due to the qualities of the media, and the evident differences on these two listening experiences.

From the composers' point of view, the possibilities to unify the commercial distribution and concert formats into one compatible for both, making possible to recreate that concert experience in different venues, is very appealing. Then, the choice among the available for the most suitable format is a central, and may also help to construct a new way of thinking the sonic arts discourse. But we also should not forget that, as shown by the wide acceptance of DVD-video, the physical media format and its related logical formats usually depend more on commercial strategies than on aesthetic evolution.

5. New reference models – new paradigms

Just like it happened with the Audio CD in the 80's we are now on a possible new point of inflection, if actually the media for distribution changes and becomes universal, or at least widely accepted by the consumer. In that case it may become a new standard or reference model.

But we must be aware that the technological development doesn't stop, and each new technological step has a shorter obsolescence time cycle. After all, the CD era began in the early eighties so we may think for sure in less than twenty years, and perhaps the next challenge is the complete replacement of the whole tradition of music distribution on physical media by new technologies based on high quality digital sound streaming on the internet. So, reading the past we may work in the present, preserving our sonic thoughts for the future.

6. Thinking the sonic discourse on multiple channels

Same as what happened with the transition from analog media to CD, but perhaps in a much deeper sense, a multichannel environment makes possible to think in Dynamic Range as a relevant attribute. Many of the new tools for analysis and sound processing, work in the same sense and confirm that the evolution of technology usually brings forth a constant evolution of the language of Electroacoustic Music. As the number of loudspeakers is increased, as far as we do not reduce the resolution in a higher proportion, we may think on the use of wider dynamic contrasts, considering that we handle a larger and more satisfactory Signal to Noise Ratio. Then, we may elaborate the sonic discourse using intensity as a variable, and this helps to get an Improved textural perception from the listener. So, one of our main challenges will be no to loose those essentials of a four or eight channel composition when thinking our music into a new formats standard.

The global acceptance of the surround formats for massive distribution, is a plus, and thanks to this, we'll have multichannel dispositions available out of the concert venues. But on the other hand we get a minus: The surround formats were originally thought for the cinema, not for music, and they allow many different sound file standards, some of them with a lesser resolution. So, which format should we use?

As said, there is no relation between the quality and the market's rules. Same as happened with the Beta-VHS war in video, the SACD and DVD Audio formats were not adopted for global distribution, spite of their much better quality audio specifications, and finally the DVD-Video won the struggle.

7. Some ideas and references put into perspective

I would like to introduce some ideas and developments made in this field and put them in perspective. First, Dominique Bassal's article "*The practice of Mastering in electroacoustics*" published by the Canadian Electroacoustic Community in 2002 and written before the war between multichannel formats get to a result. This is a very useful reference and guide. About the questions and possible solutions thought by Bassal in 2002, he writes about the possibility of placing

four channels of LPCM uncompressed audio in 48KHz 20 bit or 96KHz 16bit inside a DVD-Video structure, obtaining a quadrasonic output that should be recognized by any standard DVD-players as a valid disc. This theoretical solution was not considered by any DVD authoring software so far.

Another interesting work is Felipe Otondo's paper "*Some considerations for spatial design and concert projection with surround 5.1*" presented at the Digital Music Research Network Conference, Glasgow. Focusing on the composition of one of his pieces, he describes the strong and the weak sides of a surround system when used for an electroacoustic discourse, and very specially the differences in perception between monitoring and concert diffusion.

Also in 2005, Jean-Marc Lyzwa published in France the article "Prise de son et restitution multicanal en 5.1. Problematique d'une oeuvre spatialisée: Réponses de Pierre Boulez" an article about the questions related to the sound take in Surround system used for the recording of the piece *Repons* by Pierre Boulez. The paper refers to the recording sessions held on March 15th – 16th 2003 at the *Cité de la Musique* with the *Ensemble Intercontemporain* conducted by the composer.

The piece asks for an ensemble of 24 players placed at the middle of the venue, divided into three groups (strings, woodwinds, brass), with the public sitting around the musicians. Six of them are placed in a square outside the public, they were recorded individually and their sound was treated in real time. The others remained in the center and were recorded together. Lyzwa explains that the normal stereo recording would have made impossible the perception of the sonic scenario thought by the composer. The sound take was done with the standard 5.1 channel disposition including the front central channel. The author justifies this procedure due to the massive distribution of the DVD format.

We should not forget of course all the research and development on the Ambisonics System by Michael Gerzon and Peter Fellgett, at the Oxford and Reading universities in the UK. Any multichannel system has a weakness in what is called the "sweet point" or the area in which the listener retains the spatial perception. This critical area is much wider in stereo systems than in a standard surround in 5.1 channels, and this goes worst due to the limitations of space at the normal consumers' houses, not only due to the size but also because the sound system placement is usually not such a main thing to care about for many people. But here is where the Ambisonics system should win the war, providing an extended "sweet point". And not only this, it just needs four instead of six loudspeakers. Ambisonics records a 360-degree soundfield using a single-point microphone with four or eight capsules. The sound take is processed with the UHJ encoding getting a stereo compatible file. The weak side is that Ambisonics needs a dedicated equipment to return to surround in four channels, or eight including speakers on the ceiling and the floor. This system is effective and much more related with the sonic arts universe than 5.1 surround, but it did not stick in the massive market, so it is not easy to find an Ambisonics decoder for home theater purpose and it remains instead as a valid alternative for the concert. After the arrival of DVD, the English label Nimbus that has a catalogue of recordings made with Ambisonics re-issued some of them as compatible with the DVD-Audio with DTS audio format. Without any doubt Ambisonics deserves a better future and will for sure be also one of the following steps in this research.

8. Alternatives to incorporate surround sound formats into acousmatic production.

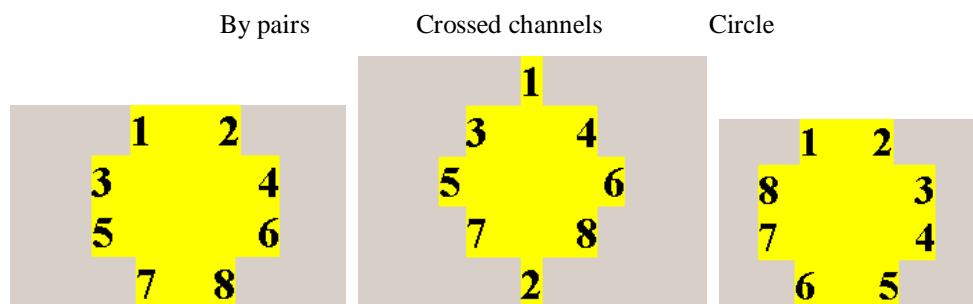
Considering that the 5.1 format on DVD appears to consolidate as a standard, and taking into account that the new BluRay format provides a much bigger physical container more than new logical formats or changes in the channels' distribution, introducing the new lossless Dolby TrueHD audio data format that allows the use of uncompressed audio data to replace the old lossy Dolby AC3 used on DVD, keeping DTS as an alternative format, it makes sense to expect that in the near future we'll be able to enjoy high quality 5.1 audio as an alternative available in mass market players. We may of course expect even bigger surround configurations on standard DVD but the five channels plus sub-low seem to consolidate as the standard for discographic distribution. With this in mind, we may define various strategies to incorporate the 5.1 to acousmatic sonic discourse, such as:

- To think the piece for the 5.1 disposition from the very outset, unifying then the concert and distribution formats. In this case we'll have to define criteria about skipping or not the Central Front Channel (C), that appears as the discordant element in the balance of powers, and also about the inclusion or not of a discrete sub-low ".1" channel in our output, that may come out of the sum of the other channels passed through a crossover filter.
- To keep on thinking in stereo as a starting point, constructing strategies to expand it into five channels, always with or without the central and sub-low.
- To keep on thinking octophony, designing alternatives to reduce the eight channel space into five. In this case we'll have to define a role for the Front Central Channel and the sub-low ".1" as well.

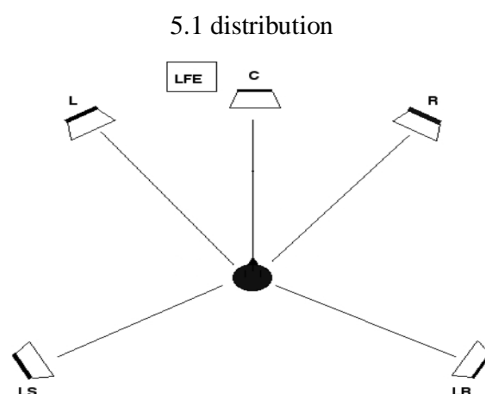
My point of view is close to the first alternative. The quadrasonic or pentaphonic channel disposition (I prefer 'pentaphonic channel disposition' rather than 5.1 when referring to the use of a set of discrete channels for acousmatic music), gives us a bigger choice for the placement of the sound in the space than stereo. In spite of its cinematographic origin, we'll still be able to use it as an interesting alternative for sound gestures and trajectories, provided that we set the five channels as discrete forces and as far as we are able to solve the role of the central front. Regarding the sub-low, today is usual to find it in many concert venues equipped with eight channels, configured as an internal sub-mixing plus filtering of the whole. Same happens with most of the consumer 5.1 equipment available in the market. So we may just think in five channels and leave the sub-low reinforcement to the installed hardware. But we have to admit that the price to pay for this is the loss of the flexibility and versatility of the different possible eight channel dispositions when we change that for the massively distributed schematic strictness of the surround setting.

9. Differences between the 5.1 channel distribution and some of the usual octophonic loudspeaker setups.

The octophonic disposition is always discrete, where every channel may be thought autonomous from the others. There are different distributions of eight channels. Some of the more usual octophonic setups are:



In contrast to the many possible alternatives in eight channels, the 5.1 distribution is fixed and comes out from the Surround Matrix according to the specifications of Dolby Laboratories Inc.



In an octophonic setup the role of the channels may change during the course of each piece, going from a disposition in circle to crossed channels, etc. Compared with this degree of freedom, the surround sound setting is a regression. But if we focus on the possible unification of the concert and distribution formats, allowing spatial distribution on the disc that will reach the hands of the potential listeners, it appears as an advance regarding the stereo CD format.

More, there are a lot of acousmatic works in the repertoire, edited in CD, that are stereo reductions of original eight channel pieces, and same as happened with the restoration and remastering of recordings from the analogical age, the consolidation of the surround sound formats and supports as a new standards will make the design of transcription strategies absolutely necessary.

10. Alternatives for the transcription from stereo to 5.1 channels.

First of all, to my view, the use of the term *transcription* is more adequate than *conversion* since *conversion* refers to a more technological than aesthetic question, being this a case where the aesthetic is at least as important as the technological aspect. In other words, here to transcribe will be equivalent to orchestrate a piece for piano or to reduce a symphonic score to the piano.

To carry out this work there are different hardware or software tools based on the application of the *Dolby Surround Matrix* (a complete references to this can be found at www.dolby.com), which perform the job doing the sum or subtraction of channels and phase shifting. This procedure, which is not at all related to acousmatic music, allows to obtain a sonic image similar to an extended stereo field, but it is fragile in relation with the perception of trajectory.

Instead of using one of this tools, we can develop an original criterion coherent with the acousmatic discourse, thinking about how to preserve the perception of the sound objects and the peculiarity of their energy development in the time axis, and their spatial displacement. An original criterion based on these premises will work on the manipulation of the sound gesture, which is an absolutely appropriate and typical resource of the acousmatic language.

11. Manipulation of the Sound Gesture – Gestural Saliencies

In a paper presented as an address to the Sonic Arts Network Conference in 2004 in Leicester (“Towards new models for the construction of interactive electroacoustic music discourse” Organised Sound Magazine, Vol 12.1), I introduced some qualities of the acousmatic discourse as well as different questions referred to the difference in sound reception from the public and its perception by the participants of an interactive performance between acoustic instruments and electroacoustics. The considerations about the audience’s perception are completely applicable to the case of the transcription from an original acousmatic stereo piece into a surround sound setting.

To talk about *Gestuality*, first of all we need to mention the *Gestalt Theory*. Although the Gestalt takes the visual experience as a starting point, it is also very useful for the analysis of the comprehension of the musical discourse because it is strongly related to the human perception. According to the *Gestalt*, we perceive naturally the whole as

greater than the sum of its parts. This way, we reorganize the information and perceive a *form* or *figure* and can discern the global idea of a thing, so that the entirety and unity remain in our conscience, beyond which we could not retain all the details. Departing from this idea, in an acousmatic discourse we can recognize different *Gestural Saliencies*, or recognizable elements that we perceive in accordance with our focal aptitude to comprehend the texture of that discourse. The perception of these *Gestural Saliencies* will depend on the distribution of the sound energy along the timeline as well as on its position and spatial displacement.

12. The Spectro-Morphology of Sound

Starting from Pierre Schaeffer's Typo-Morphology, Dennis Smalley installed the concept of Spectromorphology (1986) extending those concepts and deepening the aural analysis in his already quoted article *Spectromorphology and Structuring Processes*. Later other authors continued this idea. Among them, Lasse Thoresen extends some concepts introduced by Smalley in his *Spectromorphological Analysis of sound Objects*.

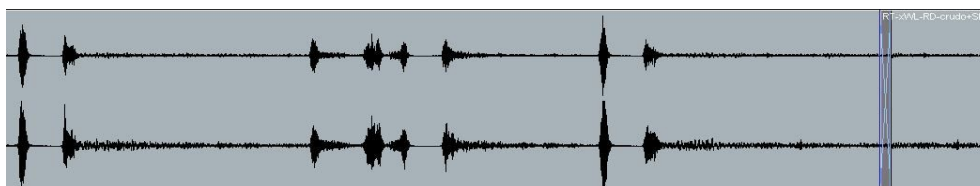
Focal Depth is one of the main ideas introduced by Smalley, mentioned inside *Structuring Processes* in the *Level and Focus* section. The author writes that "... we feel the need to change our perceptive focus happening for diverse levels during the process of listening ...". Smalley describes as fundamental the relation between the *Gesture* and the *Texture* of the sound. He writes that "... *Gesture has to see with trajectory, with the application of energy and it is joined to the causality ...*". The textural internal lines of the electroacoustic discourse favor the relations of "causality". This way, the perception of sound gesture is tied with the perception of the texture and so with the comprehension of the whole discourse.

13. Gestural Emphatization

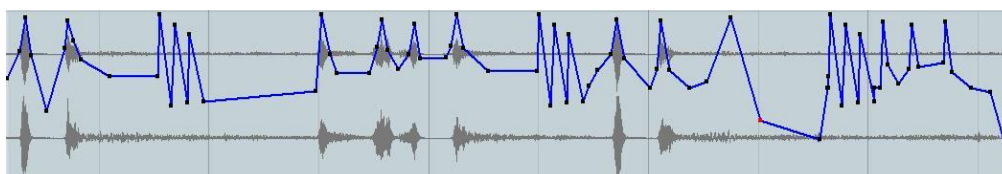
We may try to apply this ideas when doing the transcription of a given texture, originally conceived for a certain spatial distribution, in order to maintain its essence into another different distribution. In other words, we may start considering the *Gesture* of the sound discourse as contained in the perception of its *Gestural Saliencies*, and from there we may emphasize them, altering their relations of intensity and thus giving them the ability to impact on our perception in a different way as they should at their original placement and with their initial energy distribution. By doing this we'll be able to obtain a more vivid sonic image than the one embedded in the stereo spatial allocation. We may call *Gestural Emphatization* to this procedure that will remark what in fact is already there, and it's application may allow us to work around the listener's perception on a surround sound system, locating for example some saliencies in the front channels and others in the central or the rear. Of course, we'll also be able to apply the same criteria to the trajectory of the sound, modifying the panoramic distribution between front and rear channels. The use of this emphatization to dynamics as well as the trajectory will replace the sums, substractions and phase inversion/shifting of the standard Dolby Surround Matrix procedure.

14. Gestural Elaboration

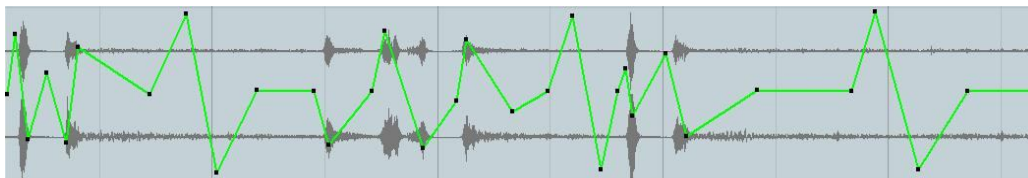
A different and very effective strategy, is to create saliencies not present in the original. This can be very resourceful since it will generate new saliencies that may be placed in some loudspeakers while not in others. Thus, we can impact the listener with different textures for the same speech. It is not a question of contradicting the original idea, but of making use of the multiplicity of channels to obtain a major textural richness. Let's imagine for instance the recording of a symphonic work, where thanks to the surround sound disposition there is possible a more detailed listening of certain instruments, by routing them to different loudspeakers. The following graph shows the wave form of the original stereo mix corresponding to a 45 seconds excerpt of an acousmatic piece:



If for a moment we imagine this fragment placed in two of the five channels of a standard surround setup, we'll be able understand the next graph. Here the same music is assigned to another two channels on the same surround distribution. Following the timeline we can see first the application of Gestural Emphatization, just after that a Gestural Elaboration. After this, the first procedure appears two more times and so does the second one.



The following graph shows the same idea applied to the spatial position, so that we maintain the original on one pair of channels and we show different trajectories on the other.



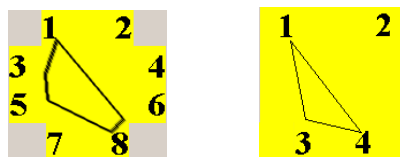
As said, we need to define the role of the Central Front. We may of course exclude it or we can assign it some specific sonic materials instead. For that purpose we may do the sum/subtraction of the original channels, or also apply a different Gestural Emphatization or Elaboration, or any other possible composer's strategy.

15. Transcription of an original octophonic mixing to 5.1 channels

To reduce to five channels an original octophonic idea is undoubtedly a big challenge, more if we bear in mind that there are many different distributions for an eight channel setup. Considering this diversity, we'll have to build a particular strategy for each case, and of course the transcription strategy will up to some extent defy the spatial ideas thought by the composer. Thus, in all cases our main concern will be not to contradict the author's intentions. Each one of the different eight channel distributions present a particular confront. Those that include front and rear central channels will need a particular treatment to achieve a satisfactory transcription, as they may appear more close to the idea of leaving the front central into the five channel output. Nevertheless, this is more imaginary than real, given the static roll of the above mentioned channel in the 5.1 disposition, opposed to the natural flow of forces of all channels in any of the octophonic distributions.

On the other hand, between the more usual octophonic setups, the disposition by pairs of channels is the least controversial as it does not include central channels. This original setup allows to think in discrete quadraphony as a possible intermediate step between eight and five. To do this, a possible strategy may be to embed four channels into the other four, intending in all cases to keep close of the original trajectories. Of course, in almost all cases, some of the spatialization design will suffer, but this strategy may be satisfactory and the reduction of eight into four helps us to limit this loss, but as soon as we consider the inclusion of the central front channel, we run the risk of going completely away from the original.

The following images illustrate this procedure. On the left side appears the original octophonic distribution showing an example of the possible trajectories that we'll try to preserve. The graph on the right illustrates the result we get after embedding four channels into the others.



To get this result, taking advantage of the pair distribution, we may consider the left channels on one side and the right on the other. Following this idea we can proceed as follows: $1=1+3$, $3=5+7$, $2=2+4$ and $4=6+8$.

All these procedures are arbitrary and show the rigidity of the 5.1 scheme that may denature an original octophonic spatial idea. The obvious conclusion is that the channel disposition of a surround sound system allows innovative solutions for the expansion of the stereo field, but on the other hand it is not completely satisfactory when having to carry an octophonic discourse into five channels. This circumstance should be a motivation for the composers who work directly in eight channels, to originally think in a surround sound setup when their aim is to unify the concert and commercial distribution formats.

Bibliography

SCHAEFFER, Pierre. *Traité des objets musicaux* - Ed. du Seuil Paris, (1966)
 SMALLEY, Denis: "Spectromorphology and Structuring Processes", in *"The Language of Electroacoustic Music"*, Ed. by Emmerson, S. - Macmillan, London (1985) pp. 61-93
 BASSAL, Dominique: *"The practice of Mastering in electroacoustics"* Ed. by CEC -Canadian Electroacoustic Community (2002)
 OTONDO, Felipe: *"Some considerations for spatial design and concert projection with surround 5.1"* presented at the Digital Music Research Network Conference, Glasgow (2005).
 LYZWA, Jean-Mark: "Prise de son et restitution multicanal en 5.1. Problematique d'une oeuvre spatialisée: Répons de Pierre Boulez". Service Audiovisuel, Conservatoire National Supérieur de Musique et de Danse de Paris (2005)
 FELGETT, Peter: "Ambisonics. Part one: general system description", *Studio Sound*, (1975) pp. 20-22
 GERZON, Michael: "Ambisonics. Part two: Studio techniques", *Studio Sound*, (1975) pp. 24-26 and 28
 GERZON, Michael: Multi-system ambisonic decoder (2-Main decoder circuits), *Wireless World*, vol. 83 (1977) pp. 69-73.
 GERZON, Michael: "Multi-system ambisonic decoder (1-Basic design philosophy)", *Wireless World*, vol. 83 (1977) pp. 43-47.
 GERZON, M. "Psychoacoustic Decoders for Multispeaker Stereo and Surround Sound", AES-An Audio Engineering Society Preprint, Presented at the 93rd Convention (1992) pp. 1-25
 SCHACHTER, D. "Towards new models for the construction of interactive electroacoustic music discourse" Organised

Sound Magazine, Vol 12.1 Cambridge University Press (2007) pp. 67-78

WERTHEIMER, Max: "*Laws of Organization in Perceptual Forms*". German Ed. "Untersuchungen zur Lehre von der Gestalt II", in *Psychologische Forschung* Nr.4, (1923) 301-350

THORESEN, Lasse: "*Spectromorphological Analysis of Sound Objects - An adaptation of Pierre Schaeffer's Typomorphology*" Paper presented at EMS Conference (2006)