Copyright statement

LISTENING PATTERNS

From Music to Perception and Cognition

by

MASSIMILIANO VIEL

A thesis submitted to Plymouth University
in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

February 2017

Abstract

Massimiliano Viel

Listening Patterns.

From music to perception and cognition.

The research aims to propose a narrative of the experience of listening and to provide some first examples of its possible application. This is done in three parts.

Part One, "Words", aims to methodologically frame the narrative by discussing the limits and requirements of a theory of listening. After discussing the difficulties of building an objective characterization of the listening experience, the research proposes that any theorization on listening can only express a point of view that is implied by descriptions of listening both in linguistic terms and in the data they involve. The analysis of theories about listening is therefore conducted through a grammatical path that unfolds by following the syntactic roles of the words involved in theoretical claims about listening. Starting from the problem of synonymy, the analysis moves around the subject, the object, adjectives and adverbs to finally discuss the status of the references of the discourses on listening.

The Part One ends by claiming the need to reintroduce the subject in theories about listening and proposes to attribute the epistemological status of the narrative to any discourse about the listening experience. This implies that any

proposed narrative must substitute its truth-value with the instrumental value that is expressed by the idea of "viability".

The Part Two, "Patterns", is devoted to introducing a narrative of listening. This is first informally introduced in terms of the experience of a distinction within the sonic flow. After an intermission dedicated to connecting the idea of distinction to Gaston Bachelard's metaphysics of time, the narrative is finally presented as a dialectics among three ways of organizing perceptive distinctions. Three perceptive modes of distinctions are presented as a basic mechanism that is responsible for articulating the sonic continuum in a complex structure of expectations and reactions, in terms of patterns, that is constantly renewed under the direction of statistical learning.

The final chapter of the Part Two aims to briefly apply the narrative of pattern structures to dealing with the experience of noise.

Part Three aims to show the "viability" of the proposed narrative of listening. First, a method for analysing music by listening is discussed. Then, a second chapter puts the idea of pattern structures in contact with music composition, as a framework that can be applied to data sonification, installations, music production and to the didactics of composition.

Finally, the last chapter is devoted to the discussion of the idea of "soundscape" and "identity formation", in order to show the potential of applying the proposed narrative to the context of cultural and social studies.

List of contents

List of illustrations

Acknowledgments

Author's declaration

Thesis submitted for the degree of Doctor of Philosophy at Plymouth University has been composed by the author, and has not been submitted for any other degree or professional qualification. The research conducted herein is the work of the author and has not formed part of any other research degree programme either at Plymouth University or at any other establishment. At no time during the registration for the degree of Doctor of Philosophy at Plymouth University has the author registered the research for any other University award without prior agreement of the Graduate Committee. The research was partially financed as a "doctoral leave" granted by the Ministry of Education, University and Research of the Republic of Italy.

The research works have been presented at the following conferences: Logic Lane: Giornate di Studi Antonio Caronia (Milano, 2015); Soundscape and Sound Identities - FKL (Castelbeseno Rovereto, 2015); Sound Studies: Mapping the Field 2014 (Copenhagen, 2014); Pattern: Forma, Formação, Informação. Encontro Internacional de Pesquisadores Em Design (San Paulo, 2013); #12.ART, 12 Encontro Internacional de Arte e Tecnologia (Brasilia, 2013); Tracce Sonore (Milano, 2013); Technoetic Telos: Art Myth and Media Consciousness Reframed International Conference Series (Kefalonia, April 2012); iStream eText: Words in Motion, International Research Conference (Milano, 2012); Presence in the Mindfield: Art, Identity and the Technology of Transformation, 12th Consciousness Reframed International Conference (Lisbon, 2011); Always Already New: Thinking Media, Subversing Feeling, Scaffolding Knowledge: Art and Education in the Praxis of Transformation (Milano, 2010).

The research works has been published in the following publications:
Viel, M., 2016. Quello che riconosciamo. In Bianchi, A. And Leghissa, G., eds, 2016. *Mondi Altri. Processi di soggettivazione nell'era postumana a partire dal pensiero di Antonio Caronia.* Milano: Mimesis.
Viel, M., 2015. Nell'Alto dei Giorni Immobili: une Analyse Esthésique. In Arbo, A., ed., 2015. *Anamorphoses. Études sur l'Œuvre de Fausto Romitelli*. Paris: Hermann.
Viel, M., 2014. In the Name of the Pattern. *Noema - Essay, Ideas 22 June 2014* [online] Available at: <http://noemalab.eu/ideas/in-the-name-of-the-pattern/>
Viel, M., 2014. Nell'alto dei giorni immobili. Un'analisi estesica. In Santarcangelo, V., ed., 2014. *Have your Trip. La musica di Fausto Romitelli*, Milano: Auditorium.
Viel, M., 2012. Sounds and Objects. *Technoetic Arts*, 9(2/3), pp.233-238.
Viel, M., 2012 On and on. A Survey of the role of Repetition in Music. *Limina,* 2012(2)
Viel, M., 2011. Ricercare l'ascolto. *D'Ars Magazine*, 207, 2011.
Viel, M., 2011. Living in music. In Ascott, R. and Girao, L.M. , eds, 2012. *Consciousness Reframed 12.* Universitade de Aveiro, Portogallo.

Word count of main body of thesis: **86.069**

Date: 20th February 2017

Signed:

1. INTRODUCTION

After having been left for centuries to be hidden within discourses about the Word of God or music aesthetics, we can now say that "listening" has finally reached the status of a respectable topic of research, although it still hasn't received the honour of being directly involved in the name of a discipline. If we consider the statistics of the use of the morpheme "listen-" in book titles (I refer here to a search among the 11 million books listed in the Internet Archive website), we see a gradual, slightly irregular—yet constantly growing—application of the term over the course of the last century, bringing the initial percentage of 0,002 from the first decade up to the 0,0356 in the 80s. But it is only in the 90s that the presence of "listen" in book titles has reached an average plateau of 0,041 % and a degree of appeal of "listening" in written texts have been finally ratified. These statistics are not really meaningful because on one hand they do not distinguish, for example, novels from essays, nor do they consider texts that are strongly related to listening yet do not include the word in their title, a situation that, as we will see, is rather common. Nevertheless, "listening" today has become a highly discussed topic in the academy because it is situated at the crossroads of myriad disciplines related to music, (neuro-) physiology, cultural studies and computation, and consequently involved in an intense activity of publication.

The reasons for this emerging "appeal of listening", have to be found in the complex articulation of a variety of factors that are related to the individual disciplines and involve technological and conceptual advancement, but also,

as Kuhn (1970) has shown us, in complications of the institutional side of knowledge, such as the possibility of obtaining funds or the necessity of attracting students. However, I think that another possible reason for this increased appeal of listening might also be related to an emergent sensibility towards a mode of sensing, which lets us reach the world in a different way with respect to vision and its sensorimotor metaphors.

Marshall MacLuhan's idea of "auditory space" (1960) as an experiential region that is not involved in dealing with boundaries or perspective seems to connected to the experience of polyphony, of the articulated presence of a multiplicity of sources in a spherical non-directional space, to which we respond emotionally. This concept immediately evokes the "internet sphere" and especially the virtual sphere of social networks that is instantiated by the technologies of tele-presence, as a general term. It is therefore possible to understand the current turn of research towards "listening", as a search for a paradigm that can account for the peculiarities of meaning making, and for the implications of media control, in the age of global hyper-textualization.

This premise is quite general, yet it might stimulate the researcher to turn to listening in order to look for a novel conceptual framework to address the peculiarities of contemporary global society. However, I must also add a further motivation that moved me to engage in this research.

In fact, as a musician, a composer and a music teacher, it often occurred to me to be involved in discussions about techniques of composition, for example those related to serialism, that as being based on the manipulation of the graphic symbols of notation, don't seem to be sufficiently rooted in the sounding aspects of music, both as listened and realized in performance.

Moreover, when dealing with didactic methods, it is easy to stumble in explicative apparatuses and patterns, such as the very idea of harmony, the sensorimotor metaphor of a tension between dominant and tonic chords and so on, which often claim an objective existence without any discussion of their ontological status.

This prompted me to look at the experience of listening for a weight that could anchor discourses and practices related to music, but also simultaneously engage a discussion of the conditions that are peculiar to the topic of listening and for being able to share those discourses and practices.

In fact, the cognitive and cultural instruments that we use to deal with listening cannot help but be shaped by the paradigms that emerged from linguistic practices in the first place and have been organised and developed by the academy. As a consequence, the concepts and the words we use in our discourses about listening seem to mostly not be the best suited for dealing with the evanescent auditory world, especially if we aim for a new perspective.

This is why discourses about listening often appear to serve the purpose of promoting paradigms and knowledge that are based on assumptions, such as the paradigm of the object as it will be shown in Sections 3.6 and 8.8, that we need to get rid of if we want to foster an understanding of sounding practices as a general term that includes music, which is able to deal with its cultural situation, but also to possibly acquire an innovative perspective about cognition in general.

The problem at the core of this text can be therefore summarized in two points:

1) How can we deal with the variety of discourses about listening and possibly expose the paradigms concealed within them? This problem is discussed in Part 1.

2) How is it possible to propose a discourse about listening that is able to account for the variety of the experiences of listening while also fostering new practices? This problem is discussed in Part 2.

Throughout the first two parts, the research therefore starts by calling into question the inherited paradigms that are involved in discussing "listening", particularly in terms of linguistic patterns, as sources of understatements that are possibly absorbed within the academic research on listening. It then proposes a narrative of listening that, starting from the very simple aural experience of distinction, follows a path of articulation and complication up to a discussion of meaning-forming that has the potential to show how patterns emerge in linguistic practices themselves (Section 8.8).

Finally, the third part is devoted to a discussion of some of the applications of the proposed narrative, which are especially directed to topics related to music and music teaching. These applications are inevitably related to my practice as a composer and as a music teacher. Therefore, the first application is directed towards music analysis; followed by composition, the didactics of composition and a general discussion of the idea of soundscape, which is a concept involved both in (music) teaching and in composition.

The idea of narrative is paramount to this research as the expression of a methodology that is built around the contingency of the narrator and the

reader.

Specifically, by using the word "narrative", I put my research in an explicit epistemological context by stressing the textuality of academic research and therefore inscribing it within the hermeneutic circle.

In Chapter 4, a narrative is characterized by authoriality, contingency and negotiability. The first character implies that, as being narrated by a musician and specifically by a composer and music teacher, the narrative involves primarily examples, listening experiences (in terms of mental experiments) and applications in the field of music. Nevertheless, the simplicity and openness of a starting point (the experience of distinction), that is situated before the cognitive constitution of music listening maintains the possibility to apply the narrative to linguistic practices and to other sensorial modes. That is why "listening" in its very general sense is always in the background, even when not explicitly stated, while "music listening" is involved only as an expression of the authoriality of the narrative and for the sake of example.

The character of contingency is related to the individuality of the author and it therefore implies the incompleteness and temporary quality of the content of narration. Negotiability, on the other hand, is related to the relational character of a narrative and is the basis of the determination of its "viability", that is the ability to "survive" in the world of engaged readers as a source of new knowledge and new practices.

And indeed, the third part of this text is meant to underscore the viability of this research, involving it in musicological, compositive and pedagogical practices.

Both the characters of authoriality and negotiability have a resounding impact on the linguistic style of this research. The intensive use of the first person throughout the text intends to narrate the experience of listening in terms of a "first-person" knowledge, while also seeking to engage the reader in a dialogue with the author. In fact, the first person is an explicit way to avoid the suppression of the narrating subject (see Section 3.8), as a typical aspect of the standard academic style, which, in its attempt to promote objectivity, contradicts the epistemological claims of the research (Section 2.6). It simultaneously explicitly evokes personal experiences that can appeal to the reader, who might share or diverge from the narrative, while also remaining consistent with an epistemology of the contingent. A brief summary of the main arguments of the research follows, which maps the three-part path to propose and explore a narrative of listening.

PART 1 – WORDS. This section discusses the constraints and requirements of a discourse about listening.

Chapter 2 - This chapter intends to establish the context of the whole research by touching upon elements that will be dealt with and developed in the rest of the text.
In general, this chapter aims to place the research within an epistemological framework that denies the possibility of being objective and fosters its inter-subjective status by stressing its textual character as the expression of a point of view, a narrative. For this reason, this chapter will be granted a more detailed examination than the following one.

Section 2.1 starts by characterizing listening in terms of subjective experiences (qualia) and considers Thomas Nagel's proposal (1974) of building an objective discourse about them. The chapter aims to show the problems involved in Nagel's project by presenting three attempts to build objectivity by connecting ontological subjectivity (which is dependent upon the subject) to epistemic objectivity in terms of measurements (Searle, 2008).

In the first attempt (Section 2.2) subjective experiences are connected to measurements related to the listening subject (that are in charge of expressing objectivity and therefore are also called "objective references") and are immediately labelled, in order to be able to deal with them, in the most inexpressive, schematic way. A problem arises when trying to distinguish within the infinite number of possible measurements (the "overall reference") what is meaningful in order to relate with the listening experience measurements from what is not. In the end it is not only impossible to define a finite set of meaningful measurements to reference a listening experience, but even if it were possible, it would still be useless because it would merely replicate the taxonomy of measurements (which is not objective in itself) without saying anything about the experience of listening.

As a second attempt, a private language is invoked in order to provide a meaningful taxonomy of labels, so that we can attach objective references to listening experiences without the burden of the measurements' taxonomy. But private languages, if even they exist, cannot be communicated because they are possibly linked to hidden variables and unconscious experiences (Wittgenstein, 1953).

As a third attempt (Section 2.3), private languages are substituted by inter-subjective languages, which can be seen as a system of "common" labels with rules of composition. These labels are organized according to a meaningful taxonomy, both on a syntactic and a semantic level, and therefore can be connected both to experiences and to sets of measurements (what I propose calling "phenomenic dyads") in such a way that the researcher can say something that can be shared about the experience of listening. But of course, even the taxonomy of "common" labels cannot but follow an articulation that is far from objective. Therefore, the universe of "phenomenic dyads" appears as organized in a complex game of relevances and contexts that mirror the way natural languages have evolved in time. In the end, even this attempt cannot be followed to build an objective discourse of the experience of listening and Nagel's proposal has to be abandoned.

From this point on, attention is drawn to the very idea of objectivity by addressing three topics emerging from the previous attempts.

The first one (Section 2.4) is the disruptive role of time in building objectivity. Not only are experiencing, measuring and labelling asynchronous activities, but the necessary introduction of time in a discourse about listening implies a passage to existence, according which experiencing, labelling and measuring are contingent actions that happen in time and have to be contextualized in order to be properly understood. For this reason, the ideas of "biology of cognition" and "epistemology of the observer" (Maturana and Varela, 1980) are introduced as a (constructivist) frame that is able to account for the failure of building objectivity and will be further addressed in the second part of the text.

The second topic (Section 2.5) is a discussion about the objectivity of measurements. As science itself is engaged in social practices, measurements also emerge in these social practices. But such measurements cannot avoid the "paradox of categorization" (Scheffler, 1982), according to which I cannot base the objectivity of a reductionist representation on a cultural representation and at the same time base the objectivity of a cultural representation on a reductionist representation, without being laden by theory (Heidelberger, 2013). The third topic (Section 2.6), which serves as the chapter's conclusion and introduces the rest of the first part, is the role of natural language. Ultimately, an objective discourse about the experience of listening is impossible because experiences are possibly still hidden in private language descriptions that cannot be connected to a common language. We only have intersubjective (mis-)understandings based on the common labels of natural languages, which completely override the organization of measurements and also the possibility of talking about experiences. But labelling is an experience itself and using labels is an experience that "logocentrically" (Derrida 2010) substitutes the aural experiences with its experiential presence.

Finally: objectification is the attempt to suppress the material conditions of existence (and knowledge) by organizing (in the categorization circle) measurements and experiences according to the structure of "logocentric" public language. It is only by reintroducing the subject that we can give listening back to its experience in terms of the point of view that the listener expresses through their description of listening.

Chapter 3 – Once the primary role of how we linguistically deal with the topic of listening through fostering points of view has been determined, the text turns to discussing the paradigms of listening that emerge in language. This is done by examining a series of topics that are fostered by specific uses of parts of speech and are representative of the work of notable researchers and musicologists.

Section 3.1 is dedicated to verbs and to the relationship of synonymy as it is involved in the "four listenings" of Pierre Schaeffer (1966).

Section 3.2 is dedicated to adjectives and specifically to the topic of "modes of listening" in Schaeffer (1966), Michel Chion (1994), Kai Tuuri and Tuomas Eerola (2012) and others.

The next five sections are dedicated to the fundamental topic of the object of listening. Section 3.3 introduces the opposition between subject and object as a relationship that emerges in language that is capable of driving discourses about listening. The idea of cognition as the act of "distinguishing", as taken from the epistemology of Humberto Maturana (1988), is introduced as a basis for the "paradigm of the object" that underlies our descriptions. Section 3.4 develops an analysis of descriptions in terms of their constitutive linguistic parts, as the premise for the introduction of Harold Fiske's idea of a "copy paradigm" (1996), according to which listening is described in terms its objects. Section 3.5 discusses the object of listening as the source of three kind of references, which are typically used in claiming an objective status of discourses about listening, namely: (western) music grammar, acoustical or (neuro-) physiological measurements and linguistic /graphic/ gestural reports of listening. Finally, Section 3.6 explicitly presents the "paradigm of the object" as

a product of our language that charges our discourses on listening with cognitive metaphors that, as coming from other senses, risks obstructing an understanding of listening as an aural experience.

Sections 3.7 and 3.8 are dedicated to the subject. First, the philosophical topic of "the subject" is introduced as it is presented by the work of Michel Foucault (1972) and linguistically expressed by verbs of action and the related adjuncts. The idea that the erasure of the listening subject in discourses about listening or its substitution with other elements of description can give rise to practices of subjections is developed in Section 3.8. Three ways of erasing the subject are shown: the use of a generalized impersonal subject (he, she, they), the use of an "actorized" verb, as in using "listening" as the subject of a sentence, and the substitution of the act of listening with "music" and "sound" as the focus of discourses that implicitly involve listening.

Finally, it has been shown that relying on parts of linguistic descriptions of listening, such as adjectives, synonyms and objects as legitimate sources of objective knowledge can risk fostering paradigms might not prove useful in accounting for the peculiarities of listening. Section 3.9 proposes the idea of "intransitive listening" as an approach that avoids relying on the structure of the objects of listening, in order to focus on the diversity of subjects in the attempt of giving back the experience of listening its cognitive strength, which operates before the constitution of its objects and its linguistic descriptions.

Chapter 4 – In this chapter, the epistemology of the narrative is presented in order to respond to the impossibility of building an objective discourse about listening on one hand (Chapter 2) and the proposal of giving back discourses

to the subject in spite of a language loaded with paradigms on the other hand (Chapter 3). As noted previously, this is a way to deal with authoriality, contingency and negotiability, which are three ways of accounting for the involvement of a diversity of subjects in the attempt to build "viability" as the possibility for a narrative to survive the never-ending emergence of the context of applications.

Finally, Fiske's requirements for a theory of listening (1996) are presented and discussed, by following the idea of narrative and its reintroduction of the subject.

PART 2 – PATTERNS. This part is dedicated to the presentation of a narrative of listening. This is done in three chapters (5, 7 and 8), with a brief Intermission to discuss Gaston Bachelard's idea of time (Chapter 6) and a Coda, which intends to apply the narrative of listening to the concept of "noise" (Chapter 9).

Chapter 5 – This Chapter offers an intuitive approach to the proposed narrative of listening as a top-down process that takes the complications of the world that we live in for granted. If the first part dealt with listening in the most general way, even if many quoted researchers were specifically involved with music, this chapter more explicitly relates to music in the proposed examples and mind experiments as it expresses a narrative told by a musician. The chapter therefore attempts to illustrate how the way we deal with and describe our listening experiences can be rooted in a constructive process of distinguishing and assembling that is simultaneously cognitive and culturally driven.

Section 5.1 aims to introduce a characterization of the experience of listening in terms of the distinction of relevant sounding moments within the sound flow.

Section 5.2 is devoted to a fundamental and complex aural distinction, which is usually expressed as the opposition between music and what we consider the sounds and noises of the world. In this section, the idea of "noise" is introduced in an informal way, which will be characterized in Chapter 9 as its "semantic" meaning, in order to deal with the informal description of listening experiences. This use of the word "noise" will be maintained throughout the whole chapter.

Section 5.3 connects the act of distinguishing to the cognitive ability of segmenting. Specifically, the idea that different contiguous segments might share a constant character, an "identity", as the background against which they are distinguished because of their "difference" in some other character, is introduced for the first time.

Section 5.4 discusses the listening cues for understanding that a music piece is finished as the occasion to introduce the ideas of "pattern" and "scheme".

Section 5.5 aims to develop the idea of "difference", as the cue for segmenting, in terms of change in sonic quality. Examples from music that explicitly avoid providing such cues, as a way to promote a subversion of the very concept of "music", are presented in Section 5.6.

Finally, in Section 5.7, the idea of "quality" is exposed as a narrative tool that needs to be discussed in order to address the constitutive aspect of listening in dealing with the paradigm of a world of objects.

Chapter 6 – This chapter is meant to be an Interlude, which is not necessary for the proposed narrative of listening, but by making connections to the philosophical discussion about time, and especially to the work of Gaston Bachelard, nonetheless opens the narrative of listening to the wider topic of cognition and existence.

Chapter 7 – This chapter, together with the next one, constitutes the core of this text, and introduces the narrative of listening as a constructive path that starts from what is posed as the simplest experience of listening up to the complexity of our world of music, sound and noise.

Specifically, this chapter is dedicated to exploring the idea of "distinction" and places it within the articulation of cognition as a process of expectation.

If the experiences and examples in Chapter 5 were typically drawn from music works, in this chapter and the following one, they are described as sonic events that are not explicitly related to music, but rather belong to the experimental world of psychoacoustics. This is because the constructive path of these chapters are directed towards listening experiences that are constituted before the distinction between music and speech, or sounds that can be related to a cause.

Section 7.1 introduces the idea of "distinction" both as a fundamental concept of Maturana's epistemology and as a basic experience in psychoacoustics.

In Section 7.2, qualities are finally defined as perceptive dimensions. These are formally defined as parameters, so that on one hand they are consistent with psychoacoustic experimentations and on the other they are open to the

possible complications of their order structure, allowing a variety of types of dimensions, such as binary, partially ordered and nested ones.

Section 7.3 presents the fundamental concept of "perceptive regions". These are three fundamental perceptive experiences that are connected to the extent of changes in a perceptive dimension and to the speed with which they occur. These experiences are described in terms of the impossibility of distinguishing changes (the "chunking region" or CR), the ability to detect the order of occurring changes (the "serial integration region" or SIR) and the impossibility of grouping changes as part of the same series of events that result in splitting the series into more "streams" (the "parallel representation region" or PR). These regions constitute a key concept in characterizing the cognitive process in general as the act of distinguishing a foreground from a background.

Finally, Section 7.4 puts the ideas of "distinction" and "perceptive regions" in the context of a narrative of cognition that, even in presenting very basic and constructive concepts, is already involved in the complications of existence. For that reason, the cognitive frame of a dialectics between memory and attention is here introduced together with a brief history of the involvement of this theoretical framework in the study of music cognition, from Leonard Meyer, to the "resonance model". This section is meant as a bridge to the presentation of new concepts in the next chapter.

Chapter 8 – This chapter aims to develop the basic ideas presented in the previous chapter, in order to be able to put them within the context of the complications of an operating cognition. While the introduction of the concept

of "pattern" gives the whole narrative a leap in scope, this chapter is indeed just a natural continuation of the previous one.

Section 8.1 is dedicated to introducing and defining the idea of "pattern" as a fundamental apparatus for a narrative of listening. First of all, it presents a discussion of the technical meaning of the word "pattern". It follows an informal definition of "pattern" as "something that we recognize in terms of its components".

Section 8.2 introduces the idea of recognition in terms of the detection of a resemblance. Consequently, the philosophical problems of the concept of resemblance are presented and framed within a process of anticipation and reaction both in the context of "intention" (Husserl, 1970) and "structural coupling" (Varela, 1992).

The next three sections aim to develop the idea of pattern in a more formal manner. This is realized first by introducing the possibility of having hierarchical layers of patterns with the concept of "pattern structure" (Section 8.3), then by discussing the role of "perceptive regions" at different levels of "pattern structure" (Section 8.4) and finally by dealing with sonic events that while giving rise to experiences that are consistent with perceptive regions, cannot be easily framed in terms of patterns, such as the detection of global changes in a series of random events. This allows the introduction of "statistical dimensions" (Section 8.5) as cues for observing attention in operation.

The next three sections are devoted to bringing the narrative to its most complex consequences, such as the emergence of extrinsic listening with the idea of sound as the property of an object. For this reason, the narrative now turns from the experience of listening towards how we describe it.

First of all, in Section 8.6, "statistical dimensions" are connected to the process of learning, in terms of the openness of an attentional window, and consequently the idea of "statistical learning" is introduced (Huron, 2006). Then, in Section 8.7, the idea that the different instances of a pattern can be distant in time from one another is introduced, so that a "dimension" too, in the end, seems to be the result of recognition and therefore can be framed as a pattern in itself. At this point, the "misattribution effect" (Huron, 2006) is presented and given a role in attributing properties to expected recognitions, so that from being the result of the cognitive act of a subject, they are ready to become the objects, i.e. the instances of a pattern, that we are used to referring to in our descriptions. As a conclusion, Section 8.8 introduces the idea of "bounded patterns" as patterns that are marked by a recognition both at their beginning and at a subsequent point that is marked as their end. These kinds of patterns, together with the effects of "misattribution", are given the capability of fitting a sensorimotor metaphor that is finally able to concede to them the attribute of object and can insert them in indexical relationships, with physical events, in a process of building meaning.

The last section of Chapter 8, Section 8.9, is a sort of a Coda. In fact, after having developed a narrative of listening from the simplest experience of distinction to meaning formation, the text now turns towards the unavoidable topic of meaning in music as a way to circularly link the ending of Chapter 8 to the beginning of Chapter 7, when the issue was the act of recognizing the beginning of the performance of a music work. After briefly summarizing the history of the debate on this subject, it is proposed that the lack of relevance of the topic as a foundational issue, once the process of meaning formation is

framed is a consequence of a process of objectification of recognitions, within the effects of "misattribution".

Chapter 9 – This chapter aims to apply the proposed narrative of listening to the topic of "noise". After having discussed some common definitions of "noise", the text focuses on the fundamental contradiction between the characterization of "white noise" as the most unpredictable signal and the apparently simple and very predicable experience of listening to it. This contradiction provides the occasion to introduce the concept of "the indistinct" as a perceptual category that is consistent with the idea of "statistical dimension" and addresses, when drawn to the higher levels of the pattern structure, the organization of music in the subversive terms of life's unpredictability.

PART 3 – MUSIC. This part is intended to show the "viability" of the proposed narrative, so that it can possibly offer the reader a framework for applying it in analysing, composing and teaching music.

Chapter 10 – This chapter aims to propose a method for music analysis by listening that is shareable, and therefore open to discussion, provided that it positively presents a connection between listening experiences and references on the "neutral level" (the timeline of the audio track and, if present, the score).

In Section 10.1, a discussion about esthesic (listening) analysis and its epistemological status is discussed, while in the successive sections, two examples of esthesic analysis are presented, namely: an analysis of the first 56

bars of György Ligeti's "Lontano" for orchestra (Section 10.2) and an analysis of Fausto Romitelli's "Nell'Alto dei Giorni Immobili" for ensemble (Section 10.3).

Chapter 11 – This chapter is dedicated to the application of the proposed narrative of listening to music composition and teaching.

Section 11.1 presents a sonic installation, "Ordo Coelestis", that is based on creating a connection between astronomical parameters and perceptive dimension, in terms of the structuring of sounding dimensions that are able to account for the structure of astronomical data.

Section 11.2 focuses on the didactics of pattern composition as a part of the wider didactics of compositive "linear functions". The logic of attending is here paramount for driving a technique of composition that is based on the discontinuities of recognition and is open to arbitrarily complex developments.

Section 11.3 introduces a versatile software that has been realized for live performance and is directed to perform "pattern structures" in which the single levels are expressed by the attacks of sonic events.

Finally, Section 11.4 is dedicated to the analysis of the composition of "Cluster", a complex work for fixed media, that is based on a strategy of setting perceptive dimensions and pattern structures, which can be generalized as a technique of composition driven by listening.

Chapter 12 – The last chapter of this text is the application of the proposed narrative to the discussion of the very general topic of "soundscape", as relevant both in musicology and in the context of contemporary musical practices. The chapter starts from a survey of the most important definitions of "soundscape" and then discusses it in the context of the work of Jacob von

Uexküll (1934) and of Maturana and Varela's concept of "autopoiesis". Finally, the idea of "soundscape" is interpreted within the frameworks of perceptive regions and "misattribution", so that its relevance in terms of identity forming can be addressed. The chapter and the research text end with an appeal on the active role of the subject in contrasting the processes of subjection and identity forming that are engaged in contemporary global society.

In conclusion, the research moved from the need to address the problems of dealing with listening, in order to bring the extremely diverse discourses that are related to listening under a common analytical framework and to reveal the hidden paradigms they conduct. For this purpose, the text proposes an examination of theories of listening in terms of the relationship between grammatical elements as the vehicles of assumptions and metaphors that are at the basis of theories. A first achievement is therefore the proposal of founding the epistemology of listening on the idea of narrative, in order to be able to bring the assumption of paradigms out of possible claims of existence and within the observable reaches of "viability".

However, the most important result of the present research is the proposed narrative of "distinctions" / "patterns" as:

1) it is potentially able to be directed towards the study of the whole spectrum of the experiences of listening, from music to speech, because it is situated before the constitution, in terms of patterns, of the target of listening, and even before the constitution of the difference among senses (Section 8.8);

2) it is able to constitute the basis of shareable discourses about listening experiences in the way it connects experiences and terms of distinctions to singular points on the "neutral level". For example, in the form of a sonogram or a score (Section 10.1);

3) it is able to foster its viability in a wide range of practices, from the analysis of (musical or not) sonic events to music production/composition to the didactics of music (Sections 10.2, 10.3 and Chapter 11);

4) it is finally able to build broader discourses that extend their perceptual/cognitive scope to issues that involve higher cognitive aspects of the mind, such as the formation of identity and bring the proposed narrative of listening within contexts related to society and culture (Chapter 12).

It is not yet possible to ascertain whether the viability of my work will extend beyond my personal artistic and analytical practices. Its involvement within the didactics of composition has certainly already brought it in contact with the artistic urges of students who, in some cases, were also able to incorporate it in their work outside the academy.

This research indeed poses a lot of possible directions for further research, for example with the possibility of the algorithmic implementation of the proposed model of pattern forming or with a more formalized description of perceptive dimensions, that, by projecting the narrative of listening in new contexts and environments, will provide new challenges for its viability.

PART 1
WORDS

2. IN SEARCH FOR OBJECTIVITY

2.1 What is it like to be a listener?

"If no one asks me, I know what it is. If I wish to explain it to him whom asks, I know not" (Augustine, *Confessionum,* 11.14).[1] This very famous statement about time could well also be made about listening. Most of us are used to the experience of listening, to such an extent that we consider it part of our general experience of living. Listening appears to be a way to access the external world by getting some clues, the sounding clues, of what is happening within it. And yet, if a person who was born deaf would ask us "what is it like to hear?" we wouldn't know how to answer. Of course, we could explain the experience of hearing by linking it to the physics of sound waves or to the physiology of the auditory system. We could also try to take advantage of some similarities with the experience of seeing or touching by looking for metaphors and comparisons. Nevertheless, we would still fail to describe the experience of hearing or listening in its peculiarities. On the other hand, even we would not have any more luck even if the person who posed the question was sound of hearing. In fact, if we try to go beyond the characterization of sonic experiences in terms of the objects that provoke them, of some very general effects on emotion and taste, or the possibly corresponding notions derived from a specific music grammar, we are mostly unable to communicate the

---

[1] "Si nemo ex me quaerat, scio; si quaerenti explicare velim, nescio". The translation follows Outler's version (Augustine, 1955).

quality of the sensations we experience while listening with any significant degree of detail.

What is special about the sensation of an octave interval in respect to the sensation of any other interval? How is the sensation of a rising melody different from the descending one? How do we distinguish the sensation of timbre from, let us say, pitch?

While the first question exemplifies the difficulties in introducing uneducated listeners to "exotic" distinctions related to music grammars, the second one is a typical problem met when starting to practice melodic dictation, at the very beginning of the music education path. On the other hand, the third question is, so to speak, the black sheep of basic music theory textbooks: as timbre, which is sometimes also referred to as "colour", cannot be expressed in explicit quantitative terms within the western music grammar, there is mostly no other way to characterize it besides defining it as a vague "quality" of sound or by simply mentioning the name of the playing instrument. For example, in Hewitt "tone quality—also called tone colour or timbre—is the property that enables the ear to distinguish between the sound of, say, a flute and a violin playing the same note" (2008, p. 8). We don't find more precision in Wright: "colour in music is the tone quality of any sound produced by a voice or an instrument. Timbre is another term for the tone quality of musical sound" (2000, p. 39).

In the end, the embarrassment we experience when trying to describe the qualities of our listening sensations is not very different from the one we

experience when we recount a dream that has intensely moved us and no body seems to share the same emotion. And when we move from the simplicity of a single tone to the complexity of a music work, we are still engages in a useless struggle of trying to convince someone to appreciate a music piece that we love by describing how we feel while listening to it.

The problem is the same whenever we have to deal with subjective states that we lack the proper words for—because words are formed as the result of collective interactions in the first place. Thomas Nagel concluded his seminal article on consciousness by admitting that "at present we are completely unequipped to think about the subjective character of experience without relying on the imagination—without taking up the point of view of the experiential subject" (Nagel, 1974). Almost half a century later we still don't have a common agreement on how to deal with subjective states if we do not want to surrender to denial as the cost to pay for a "serene reductionism" (Dennett, 1991) or to the necessity of posing their existence as a fundamental issue, which needs no explanation (Chalmers, 2010).

Dealing with our subjective experiences, which are sometimes also addressed with the words "qualia", "phenomenal consciousness" or simply "experiences", as I prefer by following Chalmers (2010, p.5), implies first of all the ability to distinguish among them, so that we are able to reference them in our discourses with the least ambiguity. In other words: we need a shared vocabulary that is capable of directly connecting a satisfying number of labels

to an objective phenomenology of experiences.[2] The requirement of objectivity is paramount to allow a discourse on subjective experience that, by holding a scientific epistemological status, can be publicly shared, as otherwise it seems we don't have an reliable way to compare the experiences of different listeners. We will therefore try to follow Nagel's request of setting the conditions of an objective discourse on qualia, in terms of the experience of listening, even if it will take the form of a *reductio ad absurdum*. In fact, Chapter 2 is entirely devoted to showing the problems that emerge from trying to build an objective discourse about listening, in order to set the context for the proposition, in Chapter 4, of *narrative*, as the necessary epistemological status of a research on listening.

We will begin by linking the character of our subjective experiences to elements that are able to grant each subjective experience an objective reference, such as a description that relies on some reductionist analysis, as some physical relata (see for example Chalmers's informational principle for a theory of consciousness in: Chalmers, 2010).

The idea, following Searle's distinction between ontological or epistemic subjectivity/objectivity (Searle, 2008), is to link an ontological subjectivity (an experience whose existence depends on the subject) to an epistemic objectivity (a proposition whose truth value is independent from the subject). In order to attain this, we need to assume that an objective set of

---

[2]See again Nagel, 1974.

descriptions/observations is possible if we follow some stipulated conventions within a community, which may possibly be the entire world.

2.2 From data to experience and back

We could say that, as soon as we are conscious, our experiences are something that characterize our life in each one of its moments. Moreover, as we are immersed in a never-ending sonic flow, which does not cease even if we are standing in a completely silent space, we always have some experience of hearing.

This claim is related to John Cage's well-known experience inside an anechoic chamber in 1951 at the Harvard University. This is a room that was primarily built for research purposes following two requirements: it has to be as acoustically isolated as possible from the external world and its internal surfaces have to be the least reflexive as possible for the sounds emitted within it. This means that when no sound is made and everybody is standing still and noiseless, no sound should be perceivable.

Yet, as Cage pointed out in his "Experimental Music" text:

> [I] heard two sounds, one high and one low. When I described them to the engineer in charge, he informed me that the high one was my nervous system in operation, the low one my blood in circulation. Until I die there will be sounds. (Cage, 1961)

Even if we focus on the experience of listening, we are always open to it, even if, apparently, nothing happens acoustically, because *we are happening*: our

cognition is inevitably embodied in our noisy organism. Concerning our experimental purposes, this means that we always have the possibility of experiencing something. And even if this possibility is modulated by the ability of consciousness to report experiences, we might also consider the process of experiencing as independent from these reports, as unconscious, forgotten experiences or even extra-body experiences (or OBE) and locked-in syndrome seem to suggest.

We could therefore make an attempt to reach objectivity in our discourses about listening by first taking the measurements related to the experiencing subject at a given time, in order to objectively specify the hearing experience the subject must necessarily have at that very moment. Measurements as a whole constitute the references of that experience, but in order to easily use them in our discourses we may label them with a simple arbitrary text, such as an inexpressive, so to speak, catalog style text string, formed by letters and numbers, for example: "X458T".

A problem arises immediately in the way we set the borders of the context influencing our experiences: when do we stop measuring in order to reach a minimum set of parameters that will account for an experience? In fact, in order to maintain their subjective character, experiences must constitute themselves independently from any public language, such as the analytical language of the objective descriptions or the natural language we use to express labels. The experience of, let us say, a note is indeed different if we

change the acoustics of the room, the environment, the position of our body, the time of the day, the ability of the performer, the way we are dressed, the way we stand, our state of health. Moreover, two experiences can never be equal because each experience is also influenced by past experiences, which constantly accumulate as a reservoir that is peculiar to each subject.

Accordingly, there are no pure experiences in hearing, as in some experiences of "pitch" or "space position", because the way we extract a single aspect of an experience is linked to the way we describe it in our public language, which, in order to preserve the originality of the experience, must intervene only after we have had our experience. For the same reason, there is not even a pure experience of hearing that can be separated from other sensory modes without losing its original character. In this sense, experiences have a holistic character.

As a consequence, if we want to account for the very individuality of experiences, it seems that we need to include within each objective reference the entire world that is somehow contingent with the subject at the moment of the single experience, provided that a definition of "contingent" is stipulated. Each overall objective reference will therefore be, so to speak, the snapshot of the world surrounding the subject at the time of an experience.

Unfortunately, a set of labels that has been compiled from objective references will not tell us anything about experiences, as it restricts itself to replicating the syntax of objective descriptions. There is no need for experiences to rely on such a description of the world. The label does not express subjectivity: it

simply replaced it. Therefore, if we want to establish a connection between experiences and references, we have to engage another strategy, which starts by establishing the character of an experience and only later by attaching an objective reference to it.

Of course we have access to the character of an experience while having it, but we still have to rely on some labels in order to publicly communicate it. As we had to give up the strategy of moving from references, we can't lean on the objective metrics of measurements to compare experiences. This means that if we want to go beyond the simple determination of an infinite number of labels, pointing to an infinite number of unique experiences, we need to propose a way to articulate the whole range of our experiences, relying on the discriminatory ability of the experiencing subject. We therefore start self-observation of the subject, who will label their experiences in order to express a possible metrics of experience that is subjectively determined and can be put in relation with the objective metrics of references. In other words, we ask the subject to connect the context measurements to the words of a private language he/she must set up.

This problem has been widely debated, especially with regards to its presentation within the paragraphs from 244 to 271 of Ludwig Wittgenstein's *Philosophical Investigations*. The philosopher's claim that a private language is

simply not possible has stimulated an intense and complex discussion over the last fifty years.[3]

In this context, it is not important to determine whether it is possible to engage in a solitary linguistic practice that, in order to be a deeply private one, needs to emerge from a conscience that does not have a prior knowledge of any linguistic practice. Nor to verify if the simple knowledge of a shared language is already able, with its communicative dialectics and its morphological structures, to deteriorate, so to speak, the privateness of a new language.

What matters here is that building a private language in fact consists of the act of connecting labels to subjective experiences, which leaves out not only the possibility for other subjects to validate the proposed subjective metrics, because both the choice of labels and the labelled experience are subjective, but also the possibility for the subject itself to verify his belief in a world of experiences that is constant and stable:

> 'Well, I *believe* that this is the sensation S again.'—Perhaps you *believe* that you believe it! (Wittgenstein, 1953, §259).

According to Wittgenstein, the act of labelling experiences starting from the subject's impressions is just a matter of appearance: these labels are just "sounds which no one else understands but which I 'appear to understand'" (Wittgenstein, 1953, §269). A private language is really useless to anybody but its inventor, who is lost in a hall of mirrors. If moving from references to experiences renders experiences in the image and likeness of the structure of

---

[3]For a general presentation of this debate see: Candlish and Wrisley, 2014.

parameters, the opposite path lets the references depend on the arbitrary topology of a subjective metrics, which obscures a possible connection to parameters that the subject is unable to detect. Even if this is enough for the subject to build a repertoire of practices, which will prove their efficacy within the context of their realization, it seems to be impossible to use a set of private labels in order to trespass the boundaries of the subject to build anything but a soliloquy, which might be artistically interesting, but proves useless in terms of objectivity. We tried to build an objective phenomenology of subjective experiences by giving our private intuition the responsibility to choose the relevant measurements that we can link to the experience of listening. But we ultimately must abandon this project because we cannot rely on private language in order to find criteria of relevance that can be shared.

2.3 Here comes the natural language

If we can rely neither on measurements nor on the character of subjective experiences in our project to build an objective phenomenology of subjective experiences, we are just left with labels. We have already ruled out the "inexpressive" labels of the catalogue of references, as a useless replica of references, therefore all we can do is to look into the public discourses we make in our natural language whilst describing our experiences. In fact, we already have at our disposal a whole world of labels we normally use to address the variety of listening experiences. What we are looking for are the indexes that we use to address the quality of experiences, both in simple

"atomic" form, as in "red", and in compound "molecular" locutions, such as "a fortissimo central A". In both cases, we are dealing with the members of an organized and stipulated system of labels, which includes the rules for their correct composition in order to properly connect them to experiences. They might consist of the words of the ordinary vocabulary of our natural language or in an additional set of words that enriches our natural language for more sophisticated needs, as is the case of music grammar.

Let's take as an example the experience of *consonance.*

This experience is invisible to objective measurements, in the sense that it has to be revealed by "enriched" ordinary language in order for the process of objectification to be oriented. The octave interval, besides the mere concordance of notes and pitches in unison, has always been considered in the western musical tradition as the highest example of the maximum possible consonance between different pitches, the *consonantia perfecta.*[4]

Over the centuries and precisely as a consequence of its peculiar character, this interval has been always granted a special role in western music practices, both, so to speak, as a positive one, as it happens for tonal harmony[5] and

---

[4]For a general discussion of the idea of consonance and the role of the octave interval see, among the vast literature on the subject: Kolinski, 1962 and: Di Stefano, 2016.

[5]Without the assumption of an identity between two pitches at an octave interval that in fact share the same note name, a definition of tonal chord would be impossible. The assumption is so strong that one of the most common analytic systems for non-tonal music, the so-called Pitch-Class Set Theory, considers the interval between two notes independently from their register (i.e. their octave transposition) as identified by the same pitch class (Forte, 1964). This is despite research that claims that the octave identity might be culturally shaped (Keuler, 1999).

traditional orchestration,[6] and as a negative one, in dodecaphony, as an interval to be avoided (Eimert, 1950).

And yet, when objective measures are left alone, the octave interval, numerically expressed by a ratio of 2:1 between the two pitch frequencies, is, as an objective reference, more similar to 15:8, which is the very dissonant major seventh interval than to 3:2, the consonant major fifth interval, which is, in terms of consonance, much closer to the octave than the major seventh. This happens because the possible metrics of consonance differ from the metrics of frequency and it is not possible to apply the latter instead of the former when labelling our experiences without losing the adherence to the way we are used to describing them.

In other terms, the opposition between consonance and dissonance emerges at the same time in the way we express our judgments on pitch intervals as more or less pleasant or "usable" in our compositions and in the way we measure the ratio between the frequencies in a pitch interval. A causal relation in which the judgement is a consequence of the interval ratio has often connected these two different descriptions of the same phenomenon. This causal relation has been extended in the last 150 years to physiology, first in terms of the ear's anatomy and lately in terms of the brain's structure. Yet, it seems that what we consider the objective causes of consonance are rather the effects of the experience of consonance, which drives mathematics and

---

[6]The octave doubling of instrumental parts is the most used device in orchestration. See, for example: Rimsky-Korsakov et al., 1964.

physiology in search of a reason for the experience. This search has to confront the incompatibility between the metrics of the listening experience, which is an ever-changing one, and the metrics of rational numbers and physiological models. Descartes already realized it when he pointed out in a letter to Marsenne:

> I need […] to point out that all these calculations are only made to show which consonances are the simplest ones or - if you want - the sweetest and more perfect ones, but non necessarily the most pleasant ones. (translated from: Di Stefano, 2016, 65)

Nevertheless, once we decide to accept this causal relation we might discover that some judgments don't fit our expectations, so that we might be tempted to normatively distinguish the correct judgments from the incorrect ones.[7] Moreover, the power of the continuum, which characterizes the mathematics of objective measurements, is blind to the discontinuous categorization of words of the natural language. This is true, for example, when experiencing "red": the objective light spectrum in fact shows a continuity between infrared and ultraviolet, which is left to our natural language, resulting from our social practices, to be broken into culturally defined segments that we call *colours*, which eventually appear to us as crossfading with each other in the spectrum. The same is true for pitch intervals, which, while being a very clear set of discontinuities within the octave, differ in the exact numerical ratio according to

---

[7]See, for example, an attempt to scientifically demonstrate the superior value of classical music over contemporary music, represented by the dodecaphonic school, in terms of physical and physiological analysis in: Frova, 2006.

tuning practices so that a precise numerical definition of intervals is impossible without defining a context.

On one hand, our experiences are "experienced" before we can express them in language (as with children before they can speak, as well as in the moment before we can focus on an experience to describe it). Therefore, they are neither shareable—because we use language to share them—nor classifiable, as classes emerge through language.

On the other hand, the labels of our natural language are necessarily finite, even if we include the compound locutions. This means that labels must be in an under-determined relationship with the reference they represent. Moreover, this under-determination is already expressing a metrics in terms of the set of references addressed by each single label. Let's take, for example, the possible diversity of accounts in labelling an identical reference and assume that one subject labelled that same reference as "a tone". The reference is clearly under-determined by that label, because a "tone", in order to be experienced by listening, needs to be emitted by an instrument that produces a complex event, which also exhibits at least a timbre and a degree of intensity. The measuring apparatus will therefore provide the complete information on the acoustic context attached to the "tone" during the determination of the related objective reference. Nevertheless, if we ask the subject to attribute a label to what someone might have labelled as a "tone", the choice of that label will depend on the subject's sensitivity and expertise in using a specialized

vocabulary. For many people the simple "tone" label will be enough, while skilled musicians with an absolute ear might prefer to label their experience as "a central c sharp played fortissimo on a clarinet with the duration of a quarter note at about 100 beats for minute", which can be translated as well, not without some work, in objective parameters.

As a consequence, we can say that this kind of label (which we could confer the name of a "common" label to remind us of its stipulated character) is connected by means of metric thresholds to domains of experiences/references, which are different for each label. For example, the scope of the "tone" label's domain is larger than the scope of the domain related to the "tone played fortissimo" compound label, because the extension of the former includes the extension of the latter.

We have started from the idea of holistic experiences connected to overall references, but as soon as we introduced natural language, possibly extended by specific vocabularies, we met the necessity to deal with the possibility of distinguishing within different experiences: a possibility that is the consequence of the rules, which form compound locutions from "atomic" labels. Not only we are able to label with "a tone" what other subjects might prefer to label as "a tone played fortissimo", but there might be even someone who labels such experience as merely "a loud sound", without even mentioning if it is a tone or a noise. Or someone might even label that same reference by

addressing only the reverberation of the hall in which the sound is emitted, thus avoiding expressing the whole experience of the sound character.

If we want to account for this wide range of possibilities, we need to set up a mechanics of relevance that is able to align the uniformity of objective parameters towards the subjective distinctions we express in our descriptions. We may, for example, distinguish within the overall reference a primary set of parameters as the focus, which will be expressed by the label against the background set of the context. In this sense, we may regard the overall reference as a structure of parameters that can host a virtually infinite number of focus and context subsets, so that we can finally account for the common practice of differently labelling the same experience.

The experience of "red" connects with other experiences, such as "transparency" or "saturation", which, while being independent from it, nevertheless also contributes to the whole experience: in the same way the experience of a "tone" heard during a concert may also be lived as the experience of the timbre of the performing instrument or as the cough that disturbs it, or as any other character that is connected to the infinite variety of the possible focus and context sets.

We can therefore consider the correspondence between experiences and references as an infinite number of what can be called "phenomenic dyads", namely connections between a single experience and its related overall reference. Moreover, each of the "dyads" are linked to sets of "common"

labels, which are articulated according to their proper language syntax, within a system of a complex articulation of constraints and liberties expressing the way references can bear different focus and context sets.

2.4 Observing the cognitive actor.

The attempts to foster objectivity in discourses about the experience of listening have evolved in a very complicated direction in order to confer objectivity to our experiences. To meet this challenge we need to analyse the relationship between the simple and compound indexical locutions that we use in our natural language to describe experiences, which ultimately constitute the taxonomical structure of a world of connected labels that our language describes.

And this abstract taxonomy of labels, once it has passed through objectification by being linked to references, seems to suggest an ontology that mirrors the way we talk about it. Experiences receive in this way their ticket to a common reality in which two individuals can share the same focus reference in the same way two individuals can listen to a same note that is performed by the same instrumentalist. Finally, the world is now populated by phenomena, the experience of which can be described with different taxonomical chains of labels submitted to the system of constraints and liberties, which rule not only the possibility of labelling experiences differently, but also the very possibility of having different experiences of the same reference. We could now assume that the experience of a tone might well be

described by the same listener as the experience, for example, of a squeaky sound or of a car horn because that phenomenon was consisting in all the three experiences at the same time.

But we have ruled out the possibility for one subject to have alternative labels for the same phenomenon that is connected to an objective reference. In fact, I have said that experiences exhibit a holistic character that, as it constitutes itself before language, can hardly be split in components or be compared publicly to other experiences, because this would involve a private language. I have also doubted that two successive labelling acts could have the same reference because the subject has meanwhile changed and therefore the objective context of the experience has changed as well. On the other hand, references that are related to two simultaneous experiences cannot be identical because a subject cannot have two experiences at the same time and therefore two different subjects must make the two experiences simultaneously, each one bringing his own repertoire of previous experiences, which again changes the context of experiences.

Moreover, the act of labelling an experience necessarily happens at a time that is different from the time of the experience, because the labelled experience is always already in the past with respect to the linguistic identification of the label: the experience connected to the label is always a memory.

This is true even if the experience extends to the labelling act, because the experience of labelling something must necessarily be different from the experience of what is labelled.

With the introduction of time we made a passage to existence, which involves the presence of an operator, a cognitive actor, which may be defined in terms of a cognitive faculty such as attention or intentionality and operates organizing objective data, for some necessary or contingent reason.

Remembering is nevertheless related to an experience, which is the experience of remembering an experience. Thus we could say that within a "phenomenic dyad", references are not connected to experiences, but rather to experiences of memories—or better to the class of all the experiences of remembering that experience.

At this point, the work of Humberto Maturana and Francisco Varela bears mentioning, as it is very important for this research and it will be constantly referenced throughout the text. The two Chilean neurophysiologists, who established the basis for the constructivist theory of mind starting from a definition of living being in terms of an *autopoietic* system,[8] are mainly relevant here for their discussion on the epistemological status of science. This can be derived from the fundamental idea of "biology of cognition" as the "study of cognition as a legitimate biological problem" (Maturana and Varela, 1980, p.xv). Indeed, by calling upon a mental faculty such as memory, objectivity is projected in a "biology of cognition", which constrains the necessary character of objectivity to depend on the contingent materiality of our body and inserts the process of knowledge in a loop between observer and observed. However, this necessarily mines the idea of an objective knowledge:

---

[8]For a fundamental presentation of *autopoiesis* see Maturana and Varela, 1980. This concept will also be very shortly presented in Section 8.2.

> The basic claim of science is objectivity: it attempts, through the application of a well-defined methodology, to make statements about the universe. At the very root of this claim, however, lies its weakness: the *a priori* assumption that objective knowledge constitutes a description of that which is known (Maturana and Varela, 1980, p.5).

If we think that the measurements we use as references are necessary entities that populate a world that is not contaminated by history, then the knowledge of our experiences is only limited to the parts of that idealistic world that we have not yet accessed. On the other hand, we need to trust our capability to access sooner or later that world of immortal ideas in its entirety, in order for us to reach objectivity. This capability puts us, our species, our scientific community, in a special place of the universe as the guardians, so to speak, of objectivity. But when we realize that we ourselves are the object of the same measurements we apply in studying any other biological phenomena and that we cannot find, at the moment at least, that unique element that puts us in such a special position in the universe, we lose our faith in the objectivity of our methods of investigations, of measuring. It is not simply a matter of having precise instruments, of course, but rather of finding the right parameters of measurements, of discovering the topology of structures that involve hidden variables. Maturana's and Varela's "biology of cognition" is therefore the attempt to put that world of ideas, our knowledge, which we choose as the warrantor for objectivity, in contact with the methods of investigation that we apply to the contingent world in an attempt to understand its structure. Knowledge therefore becomes the unnecessary, contingent result of our living

46

in the world, of our both phylo- and onto-genetic history caught in a *mise en abîme* of contingency, resulting from applying contingent methods in studying contingent subjects. And the first result of this loop is the dialectics between the observer, which is not Science or Knowledge, but a human being in flesh and blood, and another observer:

> everything said is said by an observer to another observer that could be him or herself. (Maturana, 1988, 27)

2.5 The language of science.

It seems that the project of building objective discourses about subjective experience is undermined in its foundations because we cannot ignore the materiality that is at the core of objective knowledge, a materiality that is typically repressed and substituted by a timeless, immaterial geometry of ideas. We have given to the process of measuring the role of connecting this immaterial world of ideas to the earthly world of experiences, in spite of the uncertainty of the possibilities for the abstract system of parameters we rely on, to account for the contingency of matter and ultimately of cultural phenomena such as language. In fact, we only resort to language because the physical parameters are not able to address the metrics of our experiences without being driven by our descriptions. And only at the end of objectification is the data able to trace back the experiences that engaged their recording in the first place. Such experiences are now only a consumed memory of the original one—or better the promise of their future occurrences, in case of a

repetition of their objective conditions, in the form of a topology that articulates the reference parameters.[9]

But the relationship between science and natural language is much more strict than we would like to admit. In most aspects of research a scientific language cannot avoid interacting with natural languages: not only in "transporting" data to the receiver of scientific communication as a "passive vehicle which merely conveys information" (Ford and Peat, 1988), but also in a range of practices from the public release of scientific information on media to the actual linguistic framing in the development of new theories (Bohm, 1981).

Science is in fact engaged in social practices, in which natural language has a primary role in accompanying its development adrift in time. Here we deal again with the temptation of repressing temporality, which comes around in the personal genesis of scientific ideas, in the collective development and overtaking of paradigms (Kuhn, 1970) and, in the end, in the life of scientists and all the people around them. It is therefore no wonder if the process of connecting the experience of listening to a tone with the label "central A", so that we can look for an objective reference as an instance of parameters' metrics, is not too different from the process of connecting the experience of watching an indicator that stops at a certain point of a scale with the label that results from certain operations of framing and visually organizing that scale, for example "440 Hz".

---

[9]For a classic example of a complex topology driven by experience reports, see Shepard, 1982.

We can call these labels "categories", which do not have to be intended in the ontological sense, but rather as descriptive devices that are able to distinguish equivalence classes within the infinity of experiences that can be memorized. But if these categories are built from the cognitive act of an observer, even if it is a community of observers, even if all of humanity has decided to unanimously adhere to those categories, they cannot be the warrantor of a description of reality that is independent from the observers.

As Scheffler states in his paradox of categorization:

> if my categories of thought determine what I observe, then what I observe provides no independent control over my thought. On the other hand, if my categories of thought do not determine what I observe, than what I observe must be uncategorized, that is formless and nondescript—hence again incapable of providing any test of my thought. (Scheffler, 1982, 13)

We are therefore forced to face the theory-ladenness of our references as an important and unavoidable part of the process of objectification, which we can now downgrade to a less ambitious process of parameterization. The term "theory-ladennes" was introduced in the late 1950s in order to characterize scientific observation as driven by the prejudices and misconceptions that are capable to of leading scientist to erroneous results. The work of Kuhn, Hanson and Duhem, the "founding fathers of theory-ladennes", has distinguished this vice of scientific observation in two forms:

> according to its psychological form, perceptions of scientists, as perceptions of humans generally, are guided by prior beliefs and expectations, and perception has a peculiar holist character. In its

conceptual form it maintains that scientists' observations rest on the theories they accept and that the meaning of the observational terms involved depends upon the theoretical context in which they occur. (Heidelberger, 2013, p.138)

A theory-ladennes of measurement is not necessarily a bad thing. After all, it is our certainties that drive our quest for knowledge, when we are not able to find a convincing way to talk about our experiences with language. For Kuhn (1970) scientific knowledge advances precisely because theory-ladenness forms a wall, so to speak, that unaligned observations come up against until to the point that the wall itself falls down. In other words, even if our certainties do not have an epistemic value per se, they might still realize a systemic role (Wittgenstein, 1969), which serves the advancement, maybe adrift, of knowledge.

2.6 Points of view.

Constructing an objective phenomenology of the subjective experience of listening ultimately seems like an unattainable task because there can be no certain knowledge in dealing with shareable labels that have been arbitrarily connected to our subjective experiences. On the other hand, we can still go on taking measurements of sounding phenomena and listening (physiological) acts, but we must resign ourselves to an objectivity that cannot rid itself of subjectivity, because the process of objectification constitutes itself as a cognitive process in the contingency of the organic materiality, in a temporal

stretch that we might call memory and attention, and takes place within the individual existence. And the labels we use in driving objectification belong to a public language, possibly enriched, of somehow stipulated words, which we use even if we do not have access to each other's experiences. As long as our linguistic practices do not go too far in violating the consensual rules, so that they stay within the possibility of being recognized and usable, we can interact with each other because everyone has experiences to connect with the labels we use in our interactions. In this was, we exercise "subjective understanding" by using a language in which everyone is "thinking he understands, attaching some meaning to the word, but not the right one" (Wittgenstein, 1953, §269), because both the speaking and the listening subjects cannot validate their beliefs against an "independent control".

This description of experience is in the end shareable, yet it is based on misunderstanding. It offers the illusion of objectivity, but it is ultimately founded on subjectivities, on points of view, that emerge from our subjective experience that we use to fill the void (which are our experiences of the other's subjectivity). If we must abandon the proposal to build an objective description of the experience of listening, we are nonetheless left with something: we still have the labels included in the "enriched" public language we use to describe our experiences of listening.

These are the "common" labels that, in a simple or compound form, express the qualitative characters of experiences as contingent choices of adjectives and nouns within the linguistic system of possibilities. The connection they

enact with experiences may be a stipulated one, such as the terms of music grammar, or it may be synesthetically inherited from other experiences, as is the case of "height" as an attribute of pitch. It might also come from a relationship—for example, a causal or metonymic one—that the experience holds with the abstract or concrete objects that populate our world. In fact, we do not always want to talk about our experiences by describing them exactly as they present themselves to us, because it might not suit the context well. For example, we might prefer to say that we heard an "alarm sound" rather than specifying that we heard a "repeated high pitched tone with a harsh sound-colour". It's a way of labelling listening experiences that is sometime referred as to "extrinsic meaning" (Nattiez, 2002) and is realized by relying on our belief in a world of objects and relations constrained by their material ontologies, which are mirrored by the articulation of our natural language. In this way, our labels are, so to say, naturalized, so that our system of labels opens to the infinities of "naturalized" descriptions.

Our "enriched" natural language has finally reached a supervenience over experience up to the point that it is hard to discern the boundaries between the idea that we use to label our experiences and the idea that we experience what we can label. This casts doubts upon the status of the possible experiences we cannot label, such as the unconscious experiences that are possibly due to momentary or permanent failure of memory related to altered states and on the

status of labels that blur the "naturalistic" separation between the subject and the world experienced.

This latter case is related to the problem of the "naturalized" interpretation of simultaneous experiences between two subjects related to references which, beyond certain thresholds, can be considered as being the same one, or stated otherwise: the experiences that two subjects have of a same sounding object.

I have already pointed out that two subjects necessarily bring along different reservoirs of past experiences, which conditions how they experience the present. This implies that a complete reference set, such as the physiological measurements of the subject's body, should not lack the parameters that are related to the memories of past experiences. This is particularly clear when we deal with labels that address taste or emotions.

Let us take the example of a "sad melody". The attribute of "sad" is related to the melody as if it were an object in the world and for this reason we can describe the label as a "naturalized" one. But what is sad about the melody? Where do we find the parameters that define the "sadness" of this melody? It is very difficult to find these parameters in measurements related to the melody because a happy melody is, as concerns the raw parameters, indistinguishable from a sad melody. We could drive our research in the same way we did with consonance and dissonance, but it seems that even if we might be able to detect some complex parameters that are related to "sadness", it might be difficult to distinguish a proper set of focus references because "sadness" is

the property of an entire melody, rather than the quality of a single note or chord. In this respect, I need to mention the Italian composer Giuseppe Chiari, a member of the Fluxus artistic group, who in 1967 composed "Triste", which translates as "Sad", for solo piano. In this philosophical work the performer is supposed to only play a single note, in order to engage the paradoxical effort of expressing sadness with just one musical element that is inexpressive per se, a short note. The pianist goes on playing it until he reaches an agreement with the audience, often caught humorously by his pretension, on the satisfying level of sadness that he was able to express.

Of course, Chiari's work also addresses the problem of the subjective character of sadness, as it is almost impossible that an entire audience might agree on the level of sadness of a single work, not to mention of a single note. We have therefore to look into the physiological measurements on the subject's side, in its internal reference, not only to detect its degree of sadness, but also in order to look for those parameters that can be connected to the "sadness" of that particular melody. Even so, how do we find the proper parameters to put into the focus set? And even if we are able to find them, this would only mean that the experience is independent from the melody, whose reference parameters would lay in the context set. Thus, it ultimately seems that the subject, so to say, is experiencing itself under the context of a melody. This actually holds true for all of our experiences, as any perception is first of all a modification of our own body and the focus reference must therefore be

sought within it. If the focus reference is external to our body, we can resort to our world ontology as expressed by our language to detect it. On the other hand, it is quite difficult to discern the pertinent parameters that can be connected to our labels in our physiology: there is no physiological parameter or neuron of the clarinet sound.

Looking for the focus that is connected to a label in the internal reference means transferring the ontology that is expressed by natural language to the analysis of objective data related to our physiological states, so what I observe provides no independent control over my thought, therefore engaging the paradox of categorizations.

Of course, we need to put our cognitive actor in a temporal process that somehow builds the experience of sadness while listening to that melody. But again, I want to not that "self pointing" labels, as we could call them, are the models from which "common" labels are built. In fact, objectivity requires that all the factors involved in experience must be accounted for and the possibility of sharing labels, which descends from the possibility of having two simultaneous experiences of the same reference, is just a "naturalistic" appeal to the belief in a shared world. Natural language results de facto from the co-presence of private languages that share their words, so that the mere presence of a label, for the simple reason it is shared, brings the possibility of having a related subjective experience. Every word is a "floating signifier", a term first proposed by Claude Levi-Strauss (1950) to account for words/signs

that do not have a specific referent and are applicable to a variety of discourses. Therefore every word, when inserted in a system of rules of use, is able to allow interactions between subjects and to provide the illusion of understanding each other's experiences.

But labelling is an experience in itself. When we deal with a label, for example by pronouncing it, we renew its connection with the memories of labelled experiences, together with the memories of labelling those experiences and with the constraints and liberties that rule the uses of that label within the system of labels, within the enriched natural language: actions we take as the consequence of having pronounced that label. This is why the label is on one hand experienced with the stigma of the presence, which we are used to attribute to objects of the world, just because it is an experience itself, and on the other hand it steals the place of the experience that it is connected to, whose presence, whose actual experience, becomes the promise enacted by the system of constraints and liberties that realize the context of our actions (Derrida, 2010).

This is also why it is so difficult to discuss subject experiences, why we need to appeal to zombies and hallucinations (Kirk, 2005) to possibly separate experiences from the words that describe and label them, in order to possibly allow us to grasp what the experience attached to its label is—minus the label.

In conclusion, the process of objectification can only be framed within the context of a discourse that, because it rooted in the contingency of the subject, cannot express anything besides the context comprised by its repertoire of experiences and therefore necessarily expresses a point of view, even when it is stipulated by a community, such as the scientific one.

This is ultimately not surprising. Most of the time, we engage the public practices of our language, a practice concealed within the projections of the naturality it fosters. Following Derrida, we could call this practice "writing" and accuse it of suppressing its material conditions of existence in order to elevate itself as a replica of the world, lost in the charm of *logocentrism*. Nevertheless, we have characterized the arrangement of the labels of our vocabulary of subjectivities as an experiment of shareability. We do not need to reach the thing-itself, because subjective experiences, as traces of transcendence, are autonomously completed by the speaking subjects. We do not even need to pretend that naturalized labels hide themselves as natural objects in the world, as they are placeholders for something very different from the objectual structure of the world that is described by objective references: they are indexes for the aural experiences that are related, mostly but not necessarily, by causal relation, to the world's logocentric objects.

So, when we say that whilst talking about listening we merely deal with our public language, we mean that we do not have to deal with objectivity (as we cannot) but that our interest rather goes in the direction of forming and

detecting points of view expressed by a language based on shared conventions. We do not deal at all with what listening is, but rather with how we describe listening, in search of a description that is most usable.

3. WHAT DO WE TALK ABOUT WHEN WE TALK ABOUT LISTENING

3.1 Hearing, listening, and other dangerous things.

Up to this point I have interchangeably used the words "listening" and "hearing"; it is now time to make a distinction between the terms. In his "Per una teoria dell'ascolto musicale" (2015), the Italian musicologist Marco De Natale aptly described the intricate game of mirrors at play in defining the referent of these words. In fact, on one hand, the domain of listening emerges in relation to the activity of the auditive system, by a remotion of its role within the synesthetic continuum, which involves its haptic aspects, for example. [10] On the other hand, the word listening is inserted in a process that moves from the vegetative threshold of, in De Natale's terms, the *aural* to the *audial*. In other words, discourses about listening seem to oscillate between considering it as a fact of the body, the *aural*, be it in biological, (neuro-) physiological or behavioural, and considering it as a fact of the mind related to sound and music, the audial. This happens in a process of re-contextualization both in terms of its relationship to an object, or better to the process of constituting that object, and in terms of the social receptive environment that interacts with the process itself. There is no wonder then that it seems so difficult to talk about listening without referring to a grammar of musical objects if listening is often addressed within the context of other related topics.

---

[10]See: Mazzeo, 2000.

59

Through a swirl of linguistic inventions, Marco De Natale ultimately proposes

the idea of *auditio*, as a dynamic layer that drives the transition between the

aural and *audial* aspects of listening and engages the "acquisition of

experiential referents, that (…) are brought back to symbolic meanings in the

context of music" (translated from: De Natale, 2015, p.44).

This is the plan that concerns this text. Not because music is its primary

concern, but because De Natale's *auditio* is outside the continuum between

aural and *audial,* like a narrative that claims its objectivity and is realized as

such in the onto- and phylo-genetic drift of the social narrative of the mind. It is

therefore within our shared discourses, in the relationships between the words

we use, that a referent of "listening" and "hearing" can be found. "Listening" is

therefore the mark of a region (in Husserlian terms) of experience with blurred

boundaries, relating to a constellation of other words, in the process of

realizing an ontology. This is how I use the word "listening" in this text, in an

attempt to expose the unconscious narratives that listening is involved in.

A set of words of some natural language, even when meant to exhaustively

represent the different aspects of a category, is not a very good ontology (Hirst,

2009) because the subcategories it enacts are rarely disjointed.

For example synonymy, in spite of its definition as an identity of meaning, is

hardly realized as a clear and clean relationship among words. In fact, it has

been proposed (Cruse, 1989) that degrees of synonymy could be established,

as ranging from *absolute synonymy*, the very rare case that two lexemes can be exchanged in all the possible propositions of a language, to *cognitive synonymy*, which links two lexemes to equal truth-conditions when exchanged in the same proposition, and to *plesionymy* or *near-synonymy*, in which the truth-conditions are unequal.

It is doubtful that the majority of what we are accustomed to calling *synonyms* fall into the category of *cognitive synonymy* or *near-synonymy*. What is certain is that on one hand a number of semantic relations, such as *hyponymy*, *meronymy*, *troponymy* and many others, can modulate the lexical context that differentiates pairs of synonyms, while, on the other hand, it is the syntagmatic context (Geeraerts, 2010) that is able to articulate the use of words.

If it is nearly impossible to avoid the inherent *polysemy* of natural language words, so that we can set distinct borders among the *lexical fields* of the words (I am borrowing here, with some freedom, the term introduced in 1931 by Jost Trier), detailing the same wider field of meaning, it is nonetheless possible to organize them in a functional relationship with each other.

In the context of the analysis of lexemes related to perception, for example, it has been proposed (Viberg, 2001) that the lexical field of perception could be organized according to the five traditional sensory modes, together with the functional relationships of: *activity*, the degree of control by the perceiver; *experience*, the degree of automation; and *being source-based*, as in syntagms where the perceiver is omitted. The research concludes with the claim that the

perceptual lexicon is organized by following one direction, from vision to listening, followed by touch and smell/taste, in which the lexemes of one sensory mode are extended from the lexemes of the preceding one within the line, which behave as sensory prototypes.

Sweetser (1990) enforces the idea of a hierarchical structure of sensory modes by claiming that further extensions related to more abstract concepts such as "knowledge", for vision, "obedience", for hearing, and "internal self" for taste, are the realization of what she calls the "mind-as-body" metaphor (see also: Ibarretxe-Antuñano, 2002).

If we turn to the specific analysis of lexemes related to the regional experience of listening, we can still try to determine a series of functional relationships, which we can organize in a structure that is possibly consistent with their use. Koprytko, for example (1990), proposes seven coordinates that he uses by characterizing a number of verbs of sensory cognition in terms of specificity, result/success, volition, state of affairs, exhibited attitude, direct, intensive, furtive or active manner of perception, and finally duration. As a result, he gets to reduce the number of relevant parameters to four and is able to select three categories of *resultative*, *intentional* and *existential* verbs, according to the respective predominance of the resulting degree, the combination of volition and activity degrees, with the "state of affairs" exhibited by the object of perception.

The three categories are related to English verbs as follows: "to hear" is a *resultative* verb, "to harken", a second use of "to hear", and "to listen" are *intentional* verbs, and "to ring", together with "to sound, related to the category of *existential* verbs. Other missing verbs, such as "to eavesdrop" and "to overhear" need to be accounted for by the secondary parameters that have been left out.

The coordinates proposed by Pierre Schaeffer in distinguishing his "four listenings" (Schaeffer, 1966, p.112 and following) should now be mentioned.

It should be noted that, on one hand, the universal validity of both a parametrical/structuralist analysis and the setting of (cognitive) metaphors in interpreting the meaning of lexemes are debatable: primarily, because they are made by lexemes and therefore a circle is implicated that must be accounted for. It is not the validity that is questioned per se, but its presumed universal character.

On the other hand, the claim of universality seems to rely on phenomena that come from the consequences of colonialism. While it has been found and confirmed that verbs related to vision are typically the most used in languages across the globe (San Roque and others, 2015), there is no proof of a hierarchically preferred role. Even Sweetser's claim of a "knowledge" universally driven by vision has been contradicted by the hypothesis, for example in Australian languages, of the prevalence of the ear as the metaphorical organ of cognition (Evans and Wilkins, 1998). It seems

unsurprising that the universal claims of the primacy of "vision" are a Eurocentric cliché (Aikhenvald and Storch, 2013).

Similar arguments can be made for the analysis of a lexicon in terms of the subcategories of a single sensory mode. In fact, it is easy to detect semantic gaps or non-corresponding widths of lexical fields between lexica in different languages. For example, in Italian there is no equivalent for "overhear" while there is "origliare" for "eavesdrop". In Spanish, "sentir" is used for "hear", but also for "feel" and "feeling sorry", amongst other meanings. There are many more examples that cannot be accounted for here, such as the Dutch "horen", which besides "hear" is also connected to interrogating and fitting; or the Greek "αφουγκράζομαι", which relates to sensing (for example somebody's needs).

Consequently, if we want to avoid the risk of projecting the categories that emerge in individual languages, even when shared within a wider international community, the analysis of a lexicon can only refer to the language it belongs to, together with the syntagmatic context that defines its use.

With these observations in mind, we can now consider Schaeffer's articulation of the most important French verbs related to the lexical field of listening. In spite of Claude Lévi-Strauss' Structuralist critique of *musique concrète* (1969, pp.22-23) (see also: Goldman, 2010) and of the programmatic adherence of the "Traité des objets musicaux" to the themes of *phenomenology*, Schaeffer's scheme of the "four listenings" is definitely in debt to linguistic *Structuralism,* at

least for the application of *componential analysis* (Geeraerts, 2010). This is

conducted along two dimensions that are based on oppositions: the first one

between subjective and objective, the second one between abstract and

concrete. The scheme entailed by the combination of the two coordinates,

which in the "Treatise" is expressed by the "Tableau des fonctions de l'ecoute"

(Chart of listening functions) lists four possible categories (Schaeffer, 1966,

p.116): the abstract and objective, the abstract and subjective, the concrete

and objective and the concrete and subjective.

4. TO COMPREHEND (comprendre) - for me: signs - in front of me: values (meaning-language) Emergence of a sound content and *reference to*, *encounters with*, extra-sonorous concepts.	1. TO LISTEN (écouter) - for me: indicators - in front of me: external events (agent-instrument) Sound *production*.	1 & 4: objective
3. TO HEAR (entendre) - for me: qualified perceptions - in front of me: qualified sound object *Selection* of certain specific aspects of the sound.	3. TO PERCEIVE (ouïr) - for me: raw perceptions, vague idea of the object - in front of me: raw sound object *Reception* of the sound.	2 & 3: subjective
3 & 4: abstract	1 & 2: concrete	

Figure 1: table of the listening's functions (Schaeffer, 1966, p.116)*.*

These are intended to articulate listening according to the possibility of turning,

during the perceptive process, towards the object of perception (objective) or

to the activity of the perceiving subject (subjectivity) on one hand, and to the

focus of listening as meaning (abstract) or as the immediate given (concrete).

We therefore have four verbs that point to the four combinations that result

from this analysis, inscribed in a "circuit" that numbers them circularly starting from the objective and concrete perception.

The first one is "listening" ("écouter" in French), which is described in terms of the identification of the source of sound, by treating the sound as an index, a trace of this source. The second one is "perceiving" ("ouïr"), the instantiation of a subjective and concrete perception. This is "the crudest, most elementary level of perception" (Chion, 2009, p.20); a passive process directed to the raw sonic data that we do not try to listen to or to understand. The third one is "hearing" ("entendre"): a subjective perception that also abstracts sonic events. We have here the intention to listen, and therefore a choice regarding what we listen to in order to make a description. Finally, we have "comprehending" ("comprendre"), an objective and abstract perception, which means "grasping a *meaning* (…) by treating the sound as a sign, referring to (…) a code." (Ibid.)

It is not relevant here to verify whether we are in the presence of lexical gaps or in the presence of a meaning overlapping, in comparing Schaeffer's use of verbs and their ordinary use in French; all the more so that Schaeffer himself advises the reader that his table might appear too synthetic, because it has just "the methodological goal to describe targets that correspond to the specific functions of listening" (translated from: Schaeffer, 1966, p.113). In this respect, the "four listenings" each realize a targeted act of perception that the composer interprets as a communication: it is therefore a semiotic process, which can be described in the terms of Charles Sanders Pierce's semiotics.

It is not easy to summarize Piercian sign structure without overly simplifying, as the author constantly developed both theory and terminology.[11] Yet a basic triadic structure can always be found, in which a sign, or *representamen*, is related to an *object* by means of an *interpretant*. If *representamen* and *object* are in a relationship that is close to the well-known one between *signifier* and *signified*, the idea of *interpretant* acts as a sort of mediation between the two so that it can be considered as the representation of the relation between *representamen* and *object.*

On one hand, the *interpretant* as being a representation in itself, that is a *sign*, can be caught as another *representamen* in a further triadic relationship that gives rise to a infinite set of triads, an *infinite semiosis*. On the other hand, it is possible to find recurrent schemes in the relationship between *representamen* and *object* so that it is possible to list them in terms of types of *representamen*: namely, *icons* (*signs* that resemble their *object*), *indexes* (*signs* that have a correspondence in fact, such as a causal one, with their *object*) and *icons* (*signs* that represent an "imputed character" of their *object*).

Schaeffer himself seems to refer to this semiotic model by posing an opposition between *indexes* and *signs*, as the semiotic actualization of the opposition between abstract and concrete listenings (Schaeffer, 1966, p.119).

In Peircian terms, the sonic phenomena are dealt with intentionality by selecting a *representamen*, such as the brute qualities of sound in "perceiving", or the abstract qualities that we have learned, for example, from a music

---

[11]For a concise account of Perice's semiotics also from the point of view of its development see Atkin, 2013.

grammar in "hearing". The chosen *representamen* can therefore be expressed in terms of *objects*, along the *infinite semiosis*, so that they can be referenced by further *representamen*, such as the source in "listening" or the musical content or even the linguistic meaning in "comprehending".

The passage between the two *representamen*, or better between "perceiving" and "listening", needs to be conducted as a process of abstraction driven by another semiotic instantiation so that we can connect an abstract element as a *representamen* to the *object* of "brute sound", because this is how we connect the linguistic class of a word or a note to the tokens of the possibly endless variety of their instantiations. In this sense, the passage between "perceiving" and "hearing" is mediated by extending the original domain of the *object* and selecting an *interpretant*. In the end, the original *object* in the chain of semiotic processes is the same for all the four "listenings" and consists of the experience of a sounding phenomenon, which can be described in terms of measurements resulting from the engagement of a certain kind of semiotic processes in the experiential domain that we call science. Each listening is therefore the enaction of a semiotic process that, starting from a same *object*, engages the shaping of a semiotic chain. In other words, each "listening" needs to be expressed by a linguistic description and therefore it cannot avoid the engagement of a semiotic process: there is no "listening" in terms of sound objects, as is the case of "perceiving" (see Schaeffer's "reduction to the object" in: Schaeffer, 1966, p.261 and following), without a verbal description of the listening experience according to a grammar, or better a *solfege*, of

sound objects. So, even when dealing with "brute sound", a "listening" enacts a description of experience that refers to "brute sound" as the label of its *interpretant.* In the end, there is no "concreteness" as opposed to a degree of "abstraction" in listening, unless we consider them a narrative at the service of a discourse about music production.

In fact, Schaeffer appeals to Husserl in order to address objectivity, in terms of the "transcendence of the object" (Schaeffer, 1966, pp.262-264) by quoting the widely known example regarding the perceptive constitution of the idea of "table" (Husserl, 1983, p.86). On the other hand, he prefers to rely on some objectivity of the (sonic) world without dealing on one hand with a *passive synthesis*, whose role is directed to show how objects are constituted in the conscience (Husserl, 2001) and, on the other hand, with the social practice of a public language as the requisite for a theory of consciousness.

Moreover, his idea of a sound object shows a tendency to rely on a hyletic adherence to sensorymotor, and therefore spatial-oriented, cognition, which by hastily relying on the Husserlian example of the "table", seems to avoid the discussion of a *regional* specificity of listening. Schaeffer ultimately cannot avoid a naturalistic approach to listening, in which the paradigm of a world of objects is still not entailed by sound sources, by sounding objects as physical objects in a 3-dimensional world, at least until the semiotic process of "listening" is engaged. And yet listening still relates to the distinction of sound

69

objects within the sonic flow, a distinction that follows a grammar based on the reception of emerging technological practices, a "solfège des objets sonores".

Finally, the discontinuity of lexemes brings us to a discontinuity of "listenings". Even if the "four listening" are intended to possibly overlap with each other, according to the composer, it is not clear what it means to engage different semiotic processes simultaneously, especially if the lexical fields of verbs do not entirely overlap. And yet, how is it possible to think about listening without relying on the verbs we use when talking about listening? Can we imagine a way of sensitively dealing with sound as it relates to the sensory mode labelled by "listening" without the use of "listen", "hear", "harken", "overhear", "eavesdrop", but also "audition", "auscultate", and without relying on metaphors or articulated locutions? For example, is "sound massage" still listening? And what about dealing with environmental music or with the unconscious recognition of a song amidst the din of a (cocktail) party? Is being habituated to some sound still a form of listening?

Of course, if we connect subjective experiences with verbal descriptions, we always face some kind of a semiotic process, as this is a hermeneutic scheme that can be involved whenever we have any process of connection. Moreover, as we have seen, we cannot deal, with subjective experiences without connecting them to labels that we can use in our public language.

It ultimately seems that, from the semiotic point of view, our aural experiences need to depend on an *interpretant* that should be pragmatically selected as

appropriate in relation to the context of that experience in order to be meaningful.

On one hand the range and articulation of the possible contexts suggest that the "listenings" might lie on a continuous and multidimensional spectrum, which is modulated by the possible pragmatics.

On the other hand, the process of selecting the appropriate *interpretant* for a context suggests that a more usable analysis of listening might be better conducted by starting from the way we deal with those contexts.

But if "listenings" are connected to pragmatics, it might be possible to deal with listening as a core process that can be modulated and articulated by pragmatics. This means that we should step back from a semiotic perspective and turn towards some elementary aspects of listening that do not rely on cultural taxonomies and rules such as language or music grammars, but are nonetheless still able to show how listening emerges in discourses as a part of the "praxis of living".

3.2 Modes of listening.

If using near-synonym verbs to articulate the category of the aural mode of perception is the main way that we deal with the different aspects of listening in our daily conversation, at least since the work of Schaeffer, the use of qualifying "listening" as a technical term, by associating it with adjectives, has been followed in this text.

I have to mention that, if we look at a criterion of distinction between "listen" and the other audition verbs in the ordinary use of English, a search in lexical databases for English language, such as WordNet or Dante, gives us back what at least a native English speaker already knows: "listen" is related to "hear" with the qualifications of attention, intention or care. It seems therefore that "hear" is often dismissed as a homogeneous cognitive activity that does not need any further specification. Even the prepositions and adverbs that occasionally go with this verb are typically addressed to the relationship between the subject and the object of hearing, often in terms of its source or its content, rather than qualifying the process of hearing in itself.

On the other hand, if "listen" is related to a certain increase of attention towards a target, it seems consequent that it might be connected to the strategies engaged in order to appropriately address that target. As a consequence, "listen" becomes "listening", a substantivized verb that is specified by adjectives.

This is an ordinary process in the English language, but in the cases I am presenting, it is a relevant aspect of creating technical terms within the research on listening.

In fact, in several cases the association between the adjectives and "listening" would be unusual within the ordinary use of English and is therefore directed towards naming particular ways of listening that emerge from specific cultural contexts, such as with the use of particular technologies or with social environments.

As early as his "Treatise", Schaeffer proceeds with the analysis of the

distinctive traits of listening by connecting the four "listenings" to a couple of

oppositions: the first one between *natural* and *cultural listening* and the second

one between *ordinary* and *specialist listening* (Schaeffer, 1966, pp.120 and

following).[12] It is not important here to go into the details of the single couples

of adjectives and substantivized nouns, because it would involve the

discussion of individual theoretical contexts. What is important here is the

differently pursued urge to render the continuum of listening in discrete slices.

The result is a collection of ways of listening, which are often called "modes of

listening" (Chion, 1994), "types of listening" (a locution especially used in the

mottled literature about counselling and therapy) and sometimes also "styles of

listening" (Clarke, 2005, although Tuuri and Eerola, 2012, distinguish *modes*

from *styles of listening*) or "listening strategies" (Huron, 2002). These emerge as

a constellation of objects that fragment the continuum of listening and leave it

with gaps. And this happens because we articulate listening by fostering some

instances of listening within the ontology that emerges from our natural

language: those instances that are useful for our theoretical context.

We have therefore *reduced listening* (Schaeffer, 1966), to

> the listening attitude which consists in listening to the sound *for its own
> sake*, as a *sound object*, by removing its real or supposed source and the
> meaning it may convey (Chion, 2009, p.30)

---

[12]I have to mention some affinity between *ordinary* and *specialist listenings* and Theodor
Adorno's notorious analysis of types of listeners in his *Sociology of Music* (1976). In this
respect, the passage from Adorno to Schaeffer is well marked as a passage from a person
whom is characterized in social-cultural terms and the cognitive process of listening.

to which Chion adds, as part of a "trinity" of listening: *causal listening*, that is

related to Schaeffer's "listening" directed to indexes, and *semantic listening*,

that is related to Schaeffer's "hearing" and "comprehending" (Chion, 1994).

On the other hand, Denis Smalley, when introducing Schaeffer's research to

the English speaking world, diverts it, so to speak, by introducing within the

listening "perceiving" process a reference to concepts that come from a

technological approach to sound, such as the idea of spectrum and the use of

sonograms, whose uses involve other sensory modes. And here we have

technological listening (Smalley, 1986).

We also have, in open order, *active listening* (Roden, 2005), *deep listening*

(Oliveros, 2005), *cumulative listening* together with the opposition between

linear and *non-linear listening* (Kramer, 1988), *autonomous* or *heteronomous*

listening (Clarke, 2005) and finally *structural listening* (Subotnick, 1988). These

are only a few examples drawn from the literature about listening, but there is

virtually no limit to the number of possible listening types, as is also witnessed

by the proliferation of locutions used in the world of counselling, such as

discriminative, *comprehension*, *critical*, *biased*, *evaluative*, *appreciative*,

sympathetic, *empathetic*, *therapeutic*, *dialogic* and *relationship listening*. I have

taken this list from one of the many websites about counselling

(ChangingMinds), only to show the extent of possible variations in inventing

new types of listening.

There must be therefore some humorous intent in David Huron's proposal of a list of 21 modes of listening: *distracted listening*, *tangential listening*, *metaphysical listening*, *signal listening*, *sing-along listening*, *lyric listening*, *programmatic listening*, *allusive listening*, *reminiscent listening*, *identity listening*, *retentive listening*, *fault listening*, *feature listening*, *innovation listening*, *memory scan listening*, *directed listening*, *distance listening*, *ecstatic listening*, *emotional listening*, *kinaesthetic listening* and *performance listening* (Huron, 2002a). And of course, the author concludes, "this list is not intended to be exhaustive"!

What we can learn from this proliferation of listenings is perhaps that we should not describe listening in terms of descriptions of the targets of listening acts, but we should rather pursue an explanation of some kind that is able to show how these targets are constituted within the act of listening in the first place.

In this respect, the approach of basing an articulation of listening modes on a model of sensory cognition might help to resist the temptation to formulate a theory of listening that simply mirrors our aesthetic or technologic assumptions. Kai Tuuri and Tuomas Eerola's proposal (Tuuri and Eerola, 2012) is dedicated to find a characterization of listening that adheres to the dynamic articulation of emotional responses to sound, as caught in a process of interpretative meaning-creation.

The proposal is built upon an older one (Tuuri, Mustonen and Pirhonen 2005) by bringing it from the domain of emotional states to one of embodied cognition (Lakoff and Johnson, 1980.). The original account of Tuuri and Eerola's proposal is based on David Huron's scheme of *activating systems* (Huron, 2002a) as mechanisms that evoke emotional states from sound, which is at the root of his ITPRA model of expectation (Huron, 2006). These consist of a set of six sources that are connected with the rise of emotions as well as with the emergence of other meaningful experiences. They are:

1) the *reflexive system*, which is related to fast and automatic physiological responses;

2) the *denotative system*, which is related to the identification of sound sources;

3) the *connotative system*, which is related to the identification of physical properties and passively learned associations connected with the sound;

4) the *associative system*, which is related to arbitrary learned or conditioned associations;

5) the *empathetic system*, which is related to the identification of cues regarding some state of mind connected with the sound;

6) the *critical system*, which is related to the self-awareness of the perceptive process and to the verification of the appropriateness of the listener's response.

In Tuuri's 2005 article, a set of eight listening modes is presented, which integrates Chion's three modes with the instances of Huron's *activating systems*.

While the scope is well defined in the article as instrumental to the sound design of user interfaces, in the 2012 paper the authors feel the need to express the listening modes as strategies within the context of action-based cognition directed by meaning-forming.

In order to set the stage of the experience by listening, Tuuri and Eerola start from the idea of an embodied cognition as resulting in an experiential domain of meanings that is essentially sensorimotoric and imaginative. With their use of the term "imaginative", they refer to Lakoff's idea of the ability to

> project from certain well-structured aspects of bodily and interactional experience to abstract conceptual structures. (1988, p.121)

This projection can be otherwise described as a *resonance* among *action-relevant mental images*, which are the abstract structures, and two elements of experience, namely *patterns of sensation* and *well-structured patterns of recurrent experiences*.

Another important referent of the Tuuri and Eeroal proposal is the concept of *enactive perception*, as the idea that perception emerges first of all from our sensorimotoric, actions:

> [it's] an activity of exploring the environment drawing on one's understanding of the ways in which one's movements affect one's sensory states. (Nöe, 2008, p.663)

This means that listening at its most basic constitution is experienced in terms of recurrent schemes that are action-dependent in a motor-specific way and work as enactive efforts in *imitative enactions* (in which the ideomotoric processes follow the sound event itself) and *responsive enactions* (in which the ideomotoric processes counteract to the sound event). These enactive efforts may be seen as

> synonymous to [the] imaginative effort of the embodied resonator, resulting in ecologically relevant mental images of doing. (Tuuri and Eerola, 2012, p.28)

In the end, Tuuri and Eerola's revision of listening modes organizes them in three classes as a progressive abstraction from the *experiential modes* to the *denotative modes* and finally to the highly abstracted *reflective modes*. The three *experiential modes* follow three main levels of structured *action-sound* couplings, which are

> structured experiences able to project meaningful action-relevant mental images relating both to our body […] and the environment. (Tuuri and Eerola, 2012, p.28)

They consist in *reflexive couplings*, *kinaesthetic couplings* and *connotative couplings*. Three sub-types of couplings distinguish the last one: *action-sound-object*, *action-sound-intersubjectivity* and *action-sound-habit.*

The four *denotative modes* are ordered from more source-oriented modes to more context-oriented ones: *causal listening*, *empathetic listening*, *functional listening* and *semantic listening*.

Finally, there are only two *reflective modes*: *reduced listening* and *critical listening*.

I will now present a very short description of all the resulting eleven listening modes, with the words of Tuuri and Eerola (2012, p.29 and following).

Reflexive listening is based on automated schemas as innate action-sound-reaction affordances.

Kinaesthetic listening is related to the perception of movement and in general to the sensitivity of "coping with the physical word". It is based on primitive structures of "kinaesthetic schemata concerning bodily movements, coordination and postures".

Connotative listening class deals with schemata based on interactions with both natural and cultural constraints, which are therefore acquired by learning.

The first *connotative listening*, *action-sound-object*, refers to "sonic experiences that are about actions of encountering and manipulating objects in the environment". The second one, *action-sound-intersubjectivity*, refers to

"sonic experiences of interpersonal encounters". Finally, *action-sound-habit* refers to "various habituated aspects of cultural ecology that are involved in actions".

Denotative modes of listening deal are used to conceptualize their perceptual content and therefore are oriented towards sound sources or to the context of sound. The first listening mode is the *causal* one, which we have already met in Schaeffer and Chion and is directed to the source of sound; *emphatic listening* follows as the detection of someone's state of mind or intentions; *functional listening* is to detect the function related to the context and finally *semantic listening*, another term derived from Chion, is related to the use of sound as an abstract symbol.

Reflective modes are related to the highest abstraction of considering the perceptive process in itself. *Reduced listening* aims to capture the listening experience in itself and results in describing sound qualities, while *critical listening* puts the sonic event within a social context so that it is possible to evaluate its appropriateness.

There is much more in Tuuri and Eerola's article that cannot be mentioned here, such as the illuminating description of the eleven modes as reactions to a cell phone ringing. The article in itself is quite relevant for this text, as it touches upon several paramount tenets of this research. The idea of enactive cognition

and the connected paradigm of embodiment, the importance of repetition and context in shaping perception and the pursuit of the significant diversity of listening, together with the importance of putting listening in the wider context of the "praxis of living", are all fundamental issues in letting experiences break into the loops of semiosis.

Nevertheless, it seems that from the 2007 article to the 2012 article, the proposal's character has shifted from presenting a hermeneutic ontology for sound design to expressing a phenomenology of some degree. So, in this respect, what the theory seems to miss is the constitution of listening, so that again the discrete character of listenings matches the discreteness of the words we use to describe the ontology of our world.

It is therefore quite surprising that the problem of ontology forming in language is not fully addressed, especially as it directly involves listening.

In summation, it is indeed interesting to explore the different ways we listen to a sound and it is definitely didactically important to deal with a structure when building a degree of meaning that, upon further study, will possibly be broken in order to build a deeper meaning. Yet, in case we might be tempted to express some objectivity with our theory, we should at least have to deal with the very likely possibility that the way we enact listening differs significantly from the way we describe it in our daily life. Moreover, in spite of the great number of modes and related schemes, in spite of the complexity we build, which are all elements that seem to emerge whenever we want establish a

foundation for our linguistic descriptions, we always end up referring to some sort of subsumed ontology that will leave out an unexplained phenomenon. Schaeffer's description of "listenings" was ultimately an achievement that was necessary to handle the temporary ontology emerging from the technological possibility of repetitions, in the same way Chion's modes were necessary to deal with the "audiovision" of sound film or Smalley's *technological listening* dealt with the public availability of software applications for sound analysis. We cannot escape our subsumed ontologies, creating a taxonomy that aims to describe an ontology, but we can use that taxonomy to explore our subsumed ontology so that we can reveal what is left out, by looking for neglected experiences, without presuming that someday, we will arrive at a *finis terrae* of listening.

3.3 The Ontology of Listening.

The symmetry break that characterizes the human body since its banishment from the hyperuranic world of sphere-shaped Platonic beings (I am referring here to Plato's myth of the progenitor of modern humans in shape of a double-bodied creature, which is presented by the character of Aristophanes in the *Symposium*) has provided us with a preferred direction for bodily actions. As far as listening is concerned, the shape and position of *pinnae* already orient us in such a way that we can turn most of our sensing accesses towards the optimal area for distinction by listening. We can then integrate our distinctions

in the sound flow with the distinctions in the other sensory modes, so that we can best take action against the possible source of sound.

No wonder then, that the opposition between a *subject* and an *object* via a *verb* has been used as one of the most important bases for classifying languages. In the attempt to determine the variables capable of capturing cross-linguistic similarities and differences, the typological branch of linguistics (Bickel, 2007) makes a basic distinction between *verbs* and *nouns*. According to a widely recognized connection between the syntactic and semantic roles of these clause constituents, *nouns* are generally defined as "a class of words referring to entities" (Laudanna and Voghera, 2002), while *verbs* constitute "a class of words referring to processes". The distinction between *subject* and *object* is therefore conducted according to a same "informal semantic sense" (The World Atlas of Language Structures Online, 2014) as the opposition respectively between a "more agent-like and a more patient-like elements". All the world languages are therefore grouped according to the order of the three clause constituents so that a majority of languages, 565 according to the World Atlas, uses a *subject-object-verb* order (SOV), a similar amount of languages, 488, exhibit a SVO order, then we have 95 VSO languages, 25 VOS languages, 11 OVS languages and only 4 OSV languages. 189 languages do not show any order preference. English is a SVO language.

In the relationship between *subject* and *object* by means of verbs of perception, it is not really clear which is the agent and which is the patient.

However, I will adopt, for the moment, an "informal semantic sense", as is used in linguistic typology, in order to address discourses about listening.

We can usually describe the object of our aural perceptions, which relates, as we have seen, to modes of listening and can be expressed semiotically in terms of the relationship between an *object* and a *representamen.* However, what I am interested in here is how relevant the object of listening is to developing discourses about listening.

In fact, as a side effect of the idea that there are different ways of listening connected to those aspects of the sonic flow that are selected by the act of listening, it is quite consequent to link a single way of listening to the connected aspects that have been selected and that we could call the *object*. A class of objects seems therefore a feasible way to express a listening mode. While this is doubtful, the possibility of transferring the discourse from the impalpable domain of mental acts to the material domain of presences in the world is certainly convenient. But these presences are constituted by perception in the first place, so they are unable to account for it: at most they can be taken as the interference pattern between our ineffable experiences on one hand, and our narrative of how experiences arise, our medical science on the other hand, so that they can be used to reveal what is missing from our description of perception. And the impossibility of relying on an object, with its well-known structure, that is well-expressed in our words, in order to found a description about sensorially relating with that object, in the end constrains us in a hall of mirrors that we mistake as knowledge.

Jean-Luc Nancy effectively describes this sort of *cul-de-sac* of knowledge when he comments on Gérard Granel's critique of Husserlian analysis of time conducted on the widely known example of a melody:

Husserl, according to Granel, perpetuates the "forgetting of being" in the Heideggerian sense, and this occurs to the very extent that he does not concentrate his ear on musical resonance but rather converts it ahead of time into the object of an intention that configures it (Nancy, 2007, p.20)

> and consequently: "Husserl persists in 'seeing' the melody instead of listening to it" (Nancy, 2007, p.21).

But what I am doing by writing this text is expressing relationships among lexemes by instantiating sentences in English that are intended to represent, or at least to be consistently connected, with non-linguistic practices, which I cannot avoid accounting for in linguistic terms. Here, we need to again appeal to the epistemology of Maturana, who unequivocally states: "the physical domain of existence arises in language as a cognitive domain" (1988, §1). This strong claim is essentially meant to express that within a text everything is just text, even when it seems to evoke something that is beyond the written page, that is, so to speak in the pen of the writer. In the end it is just a disposition of words that enacts the magician-like trick of letting appear, for a moment, the invisible beyond words. But reading, writing or even holding a book is already part of our existence, a part that seems to be too easily forgotten. And the trick

suddenly transforms into an opposite one: it is the text that disappears and disguises in a world of objects. Even though Maturana is willing to express the circularity of showing within the boundaries of the closed system of language that which is outside of the system—which is inevitably expressed in terms of elements that are inside the system—he still cannot abandon the idea of relying on experience, constructing a "biology of knowledge", in order to anchor knowledge in its drift (Maturana and Varela, 1992). First of all, his gnoseology is strongly connected with the reductionism of scientific measurements, which in the end is based on trusting the senses, although caught in an onto- and phylo-genesis. Moreover, there is a fundamental dualistic tension in Maturana's thought, which results in a certain diffidence to words and in the need to clear their ambiguities with an epistemology. We are therefore far from Lakoff's *cognitive metaphors* in natural languages, which thanks to their inherent ambiguity allow the knowledge to advance. Conversely, while necessarily dealing with language Maturana fosters a logicistic approach that, in spite of being connected to the existential aspect of the actions of an observer, is more similar to a scientific language than to an ordinary one. For Maturana, a scientific explanation therefore consists in "an *ad hoc* mechanism that generates the phenomenon explained as a phenomenon to be witnessed by the observer in his or her praxis of living" (§4.1): "witnessing" is at the root of cognition as the optical-eidetical *operation*, which is in fact defined as "distinction" and is enacted by an observer.

Distinguishing is here the act of separating a foreground from a background, as the result of operations within the observer's "praxis of living". While the background is often interpreted as an environment, which is very similar to Uexküll's *Umwelt* (2010), the foreground is typically addressed as a *unit*, a whole, or better: an *object*. As thrown in an existence-by-language, our basic cognitive operations cannot avoid adhering to the structure of our "object language because this is the only language that we have (and can have)" (Maturana, 1988, §5.2). As we will see in the second part of this text, a more appropriate idea of distinction, which emerges in dealing with sound, mainly consists of the disregard of an expectation, so that the very detection of objects needs to be shown as constituted by, so to say, a series of expected distinctions instead of being a primitive concept.

However, Maturana stages the constructive narrative of cognition in a sort of proof by induction in terms of the dialectics between two constructive moments, the first one being the very basic distinction of a unit, which is a simple one, from its background and the second one being a sort of cognitive meiosis, which gives birth to an intermediate layer, so that the simple unit has now become a composite unit, resulting from the mutual interactions of other simple composing units. The constructive attitude of distinction is suggested by the definition of a simple unit in terms of a collection of properties, which implies that a language to describe these properties already needs to be there somewhere. The act of distinction is therefore caught by an observer as a relationship between the complex unit of a subject and the complex unit he or

she distinguishes as a simple one by means of an interaction with his or her environment, which the observer interprets in its language as the setting of certain properties. There is indeed a difference of scope among Maturana's constructivism and *cognitive linguistic*s' claim that "a grammar is ultimately a neural system" (Fauconnier, 2003). Therefore, it is not surprising if the former is forced to face the inherent circularity of language by aiming to show its emergence from interactions within a system of units and therefore needs to relate any such system to a meta-language that describes it (see for example as an attempt to formalize Maturana's epistemology: Möller, 1991). The latter, by taking language for granted, does not need to face circularity and therefore is unable to deal with the ontologies it fosters.

3.4 Describing.

I will attempt an intuitive approach to Maturana's epistemology here, in order to deal with the way we talk in our natural language about the object of listening without having to rely on a long introduction to its terminology or an explicit meta-language for the observer's ontology.

A proposition is here defined as a simple clause realized by a subject, a verb, an object and a series of adjuncts/complements that specify the verb and/or the object. In order to simplify the argumentation, I will need to follow some serious constraints, the first one being the restriction of the subject to the first singular "I", so that the active role of the speaking subject is explicit. A second restriction relates to the logicistic characterization of adjuncts/complements

within a taxonomy of objects and relations among them that we must take for granted. This restriction is due to the ambiguity of natural languages in connecting a complement or an adjective to the related element and in expressing the proper predicate/property in a standardized way. This means, for example, that when saying "I play a note on a piano" I need to be explicit about the meaning of "on" as related to the position of the "note", we have already excluded that it could be related to the subject, or as the specification of "play" in terms of the instrument played.

A description related to a noun is therefore a proposition, which includes that noun as an object and is considered as appropriate by a community of natural speakers in the language in which the description is formulated.

A description identifies a group of people who accept that description as meaningful for some reason, which might be its docility to intersubjectively connect to the promised experiences of the community members that have been occasionally verified without excessive inconsistencies; in short, because they share the same semantics.

Consequently, there are as many different kinds of descriptions "as there are different criteria of acceptability of reformulation of the happening of living of the observers that the observers specify" (§4.0).

Let's define a *domain of description* related to a noun as the class of all the meaningful propositions that are identified by all the possible verbs that can share the same adjuncts and all the possible adjuncts that can share same verbs compatible with that noun and the subject "I". For example: "I write a

89

note on a music sheet" and "I erase a note on the sand" share the same domain of description regardless of whether "note" references a musical element or a textual one, because if we exchange the two adjuncts while keeping one of the two verbs or exchange the verbs while keeping one of the two adjuncts, the resulting clauses are still correct. On the other hand, the two propositions "I hear a note at a concert" or "I play a note on my piano" do not express the same domain, neither with one another nor with the previous two sentences, unless certain unusual requirements are satisfied. For example in "I hear a note on a music sheet" the requirement is that a note on a music sheet is able to directly produce something that I can hear; or in "I erase a note on my piano" I need that a note is drawn (or in case of a textual note is laying) on the surface of my piano and I can erase it.

In Maturana's terms, the operation of distinction is made by the observer, "I", as a practice, the verb plus the verb's adjuncts, so that a unit is distinguished from my environment by all the possible adjuncts, which are the unit's properties, to which it can be connected.

On the other hand, if different verbs can be connected to the same noun as their object and to the same set of object's adjuncts, then it means that they are expressing different interactions with different properties of the same unit and therefore belong to a same domain of possible descriptions of that noun, as a unit distinguished by the "operational coherences" of the subject "I".

A simple unit is therefore the object, which is expressed by a noun, of a simple proposition that connects an operation of the subject to properties expressed

by the object's adjuncts. A noun can belong to different domains of description, which can be totally different or may overlap to some extent. As we have seen, a note, as a simple unit, can be connected to the different domains of description according to its connection to verbs such as "write" or "play". A noun expresses different units when it is connected to different domains of description: the "note" that is written is not the same "note" that is heard, in spite of the use of the same label "note".

Let's turn to composite units. These are expressed by one or more propositions that connect a verb and its adjuncts with different nouns and their adjuncts.

For example: "I've heard the noise of a car immediately followed by thunder" instantiate in language two units as components of a composite unit, realized by their succession in time. "I play the second note of the melody" instantiate two units, the first one, the "note", is simple and is a composing unit by means of the order relation of the second one, the "melody", which is a composite one. I would like to stop here and leave the development of this "conceptual translation" between the two "conceptual domains" of Maturana's gnoseology and of the domains of description, to another occasion. What is important here is that a simple unit is realized in a different domain of description than the one that involves the same unit, as a composite one. On the other hand, the composing units of a composite one realize it in a domain of description that is not the same as the domain of description of the composite unit.

In other words, the complete set of properties that define a written melody as a whole differs from the set of properties that define it in terms of the relations among its possible components, which is also different, although partially coincident, from the set of properties that define the written notes as the simple units realizing a melody.

This complex interplay of conceptual domains of description necessarily comes up against the redundancy of language. A metaphor, for example, while being considered as the main source of knowledge by *cognitive linguistics,* might be nevertheless considered as a possible source of fallacies in Maturana. In fact, only if a second composite unit mirrors the organization of a composite unit, which is the set of the relationships among the components' properties of the unit defining it, will we have an analogical relationship between the two units. Of course this is not a metaphor, but rather a correspondence between sets.

On the other hand, it is the cognitive friction between the domains of description of the two units that are connected with each other in a more or less analogical relation. This is at the root of the process of knowledge formation, both in proposing more or less legitimate explanatory models and in engaging new practices that are more or less serendipitous. I have already presented another important effect of analogical thinking in shaping the redundancy of language, that is, the possibility for an object/noun to be determined by different domains of description. This issue is particularly

relevant in dealing with the object of listening because it is at the core of the misattribution of domains that is provoked by indexicality (Hofstadter and Sander, 2010, p.33 and following).

Indeed, a written musical note is connected to a performed musical note and a listened musical note. I write a note on paper, then I can immediately play the corresponding key on a piano, which has the same name attribute, and when I am playing it I can also hear it, although only experienced musicians will be able to engage a practice, such as in using a tuning fork, in order to recognize the note without having seen that note on a paper or having seen the performers gesture.

Nevertheless, the range of operations that are expressed in descriptions by adjuncts defining the respective domains of description are totally different: we can use notes to compose a melody, but we cannot do the same by listening; we can be moved to dance by listening to a song, but we cannot do it by simply writing it or even reading it; we can vibrate a note on a cello, but we cannot do the same on paper or by listening.

Each experiential-enactive region is connected to a domain of description realizing a completely different set of practices, which are open to mutual translation and interaction, but are fundamentally different from each other.

As John Cage has pointed out:

> composing's one thing, performing's another, listening's a third. What can they have to do with one another? (Cage, 1961, p.15)

Yet, the word "note" is successfully used as an index of the different enactive experiences of writing, listening and performing because it is seized in a net of metonymic (in the general sense) relationships among words.

On one hand, these relationships form equivalence classes (determined by the different use of the same noun) that overlap with the different domains of description related to that word and semantically modulate them (like between a musical note and a textual note).

On the other hand, they serve the purpose of building an ontology (a note is an object that can be written, listened and performed), which nevertheless still cannot avoid being blurred by temporary (emerging from collective practices) linguistic games.

In the end, when we listen to "a music piece", a "sequence of sine tones", "what other people have to say" or "the Word of God", which are typically referenced objects in the publications about listening, we enact a cognitive process that consists in distinguishing from the cognitive background a constellation of units that is (linguistically) defined by a domain of description. But this is different from the domains of description that are possibly related to non listening practices, such as measuring, writing or playing, although a meta-domain can be set which includes properties of the involved domains of description as units.

When we build a discourse about listening we should therefore avoid the adjuncts that are part of domains of description related to different uses of the same object-name.

This is not a simple task, first of all because it is not easy to disentangle ourselves from the connected regional ontologies that we are used to consider as different aspects of the objects that populate our world. Secondly, and as a consequence of this, it is not immediately clear which objects and adjuncts can be considered as the proper ones that can be used in characterizing descriptions based on the "verbs of listening", when our linguistic practices are formed by embodied metaphors and are therefore embedded in a flow of onto- and phylo-genetic cultural transformations.

Therefore, it comes as no surprise that discourses about listening are often hidden in texts that are explicitly devoted to other topics, which may be connected to more tangible and therefore more accountable ontologies.

In fact, if we take the titles related to the most relevant theories about listening, we rarely find an explicit reference to "listening" or related terms. Starting from Leonard Meyer's *Emotion and Meaning in Music* (1956) to Schaeffer's *Traité des Objets Musicaux* (1966), Abraham Moles' *Information Theory and Esthetic Perception* (1966)*,* Mary Louise Seraphine's *Music as Cognition* (1988), Delalande's *Le condotte musicali* (1993), Huron's *Sweet Anticipation: Music and the Psychology of Expectation* (2006) and Aniruddh Patel's *Music, Language, and the Brain* (2008), it seems that there is some modesty in dealing directly with the topic of listening. And "listening" suddenly becomes an elusive

topic, when it is not possible to connect it to a materiality, to a *neutral level* (I will define better this concept when I address Molino's semiotics and the idea of esthesic analysis in Section 10.1) in order to rely on some positive basis.

Harold Fiske refers to this attitude as typical of the *copy paradigm*, which is engaged from the assumption that "music, its structure, content, and maybe even its meaning [...] [is] located 'out there' as an objective acoustical object" (1996, p.2) and consequently becomes a direction of research according to which "anything heard, experienced, understood, or felt is directly attributable to a real physical object, even one that is passing through time" (1996, pp.1-2). This is indeed is connected with the *misattribution effect,*[13] which states that

> whenever we experience a strong emotion, that brain has a tendency to associate the emotional state with whatever salient stimuli exist in the environment (Huron, 2006, p.136)

or better: whenever we get used to a stimulus, this becomes a quality of a related element of the environment.

As a consequence, in this process of ontification, listening is studied as if we were "standing back and objectively observing a musical object much in the same way as we would observe a piece of sculpture" (Fiske, 1996, p.1). Theories that follow the *copy paradigm* are usually based on an idea of listening as a decoding process, where cognition by listening

---

[13]For a more detailed discussion see the Section 8.7 in Part II.

serves only as a translator or an information receptor, as a system for information storage and retrieval, This leads to the conclusion that, give a performance of a Chopin etude, all listeners 'hear' the same phrasing, same pattern structure and pattern interrelationship, same compositional scheme, and so on (Fiske, 1996, p.41)

and potential differences between the listeners' reception would be related to different translations/representations of the same source. But how do we deal when we face such differences among listeners?

Fiske presents two typical reactions to such a problem.

According to the first one, it is assumed that "the 'experienced' listener has the correct internal representation while the 'naive' listener's version is faulty, incomplete, misperceived and so on" (Fiske, 1996, p.42).

This is, for example, the solution to diverse listening proposed by Hugo Riemann (2010). Fiske claims that relying on the sole internal representation of a music source necessarily comes against the continuity of such a representation so that it is always possible to find a difference between the representations of two listeners.

But in case one of the listeners is an experienced one and the other is the composer himself, the possible difference in the internal representation could be attributed to a misperception of the former as well as to the latter's misjudgement of the way a listener should listen. Therefore, such a position would lead to a music tradition of works that are always misperceived because it is impossible to verify the occurrence of a proper representation, especially if the composer is dead.

An alternative solution is "to point to an objective measure of the performance such as […] analytical reductions and hierarchical schemes, computer sound-generated graphs, or whatever" (Fiske, 1996, p.42).

But, as we have seen in the first chapter of this part, it is not possible to found the validity of the objective representation on the possible isomorphic internal representation of it, while at the same time founding the validity of the internal representation on the objectivity of the "external" representation.

A representation bases its objectivity on the "operational coherences" (in Maturana's terms) that are theory-laden and therefore depend on some internal representation, whose objectivity cannot be proven (and that is why Maturana pursues the strategy of putting "objectivity in parentheses" (1988, §5).

Secondarily, objective representations are subject to technological and interpretative limitations and constraints: there is no way to demonstrate that these limitations and constraints are not leaving out what is necessary to build a proper representation of music.

Fiske ultimately concludes that we must accept the validity of all the different internal representations of a same musical source. Therefore, the *copy paradigm*

> is an inadequate description of the relationship between sound objects and perceiver [because] it cannot support the idea that musical organization is contained in sound objects while also supporting conflicting internal representations of a musical performance. (Fiske, 1996, p.43)

Nevertheless, Fiske's criticism addresses two key assumptions in current theories of listening: the unreliability of objective references and the uncritical adoption of ontological paradigms that are not directly related to listening. The first assumption will be discussed in Section 3.5, while the second one will be dealt with in Section 3.7.

3.5 References.

Regarding the first point, there are basically three kinds of references, which are used with the roles of Maturana's "operational coherences" in order to anchor the discourses about listening to a certain objectivity. The first one is related to the elements of western music grammar we use when talking about listening. Mary Louise Serafine offers a historically important presentation of this criticism, claiming that when using the concepts taken from traditional music grammar to deal with music listening,

> the very derivation of tones, intervals, and chords depends on the conscious usually intellectually motivated *analysis* of compositions [and in the end] although the analysis is useful and necessary for intellectual purposes, there is no reason to believe that the units of analysis are also the units of perception.
> (Serafine, 1988, p.25)

This is a consequence of a mismatch between domains of description and holds when talking about subjects related to cognition through listening, both when referring to music score and to analytical schemes derived from

99

abstraction of notational items such as, among many others, the analytical schemes of Heinrich Schenker (Schenker and Salzer, 1969) or Allen Forte (1964).

The second kind of reference comes from the data obtained by measurements of phenomena related to the listening experience. This is the case of acoustical data, such as the sound wave or the spectrogram/sonogram (Smalley, 1997), but also the case of data taken from neural activity.

Here we meet the two problems mentioned by Fiske. The first one is related to the technological status of the data retrieval techniques, such as Positron Emission Tomography (PET) or "functional Magnetic Resonance Imaging (fMRI)" (Noë, 2009) both in terms of their resolution in time or in the data domain and for relating data to their pertinence in relation with the referenced experience. As concerns the latter, the comparative method is questioned in its ability to deal with a neural activity that is not working in just one direction but "is rather characterized by loops and two-directionalities (Noë, 2009, p.22 and following) and context-depending.[14]

The second problem is related to the degree of translatability between the domain of description of the composite units in the data set and the domain of description of the composite units in the referenced experiences that are connected to the data set.

_____

[14]See the analysis of *place cells-remapping* in: Datteri, 2012, p.109 and following

In order to have a scientific explanation, in Maturana's terms, the organization of the units in the data set, which corresponds to "operational coherences" of the observer, must in fact be isomorphic to the organization of the units within the *explanandum* composite system of units in such a way that some units in the domain of description of the data set, which constitute the *explanans*, are causally connected to the isomorphic composite system of units, which is also in the data set. In sum, the *explanans* must be in the same domain of description as the collected data. Otherwise we fall into the *paradox of categories* that I mentioned in chapter 1.1. In fact, the translating connection between the referenced experience, which is presented in terms related to musical grammar and the internal symbolic representation that should be isomorphic to it, is hardly realized, as "the generalizations expressed by the symbolic hypothesis are extremely difficult to 'translate' into neural terms" (translated from: Datteri, 2012, p.116). This happens because

> the risk, which is related to many procedures of direct investigations, consists in allowing the experimental data to surpass any interpretative hypothesis, by abstracting them from the very methodological condition of their detection: if a theoretical frame that is able to host and interpret data is lacking, the final result of research reduces itself to a collection of more or less homogeneous data. (translated from: Di Stefano, 2016, p.142)

In the end, research should not start from a description of the experience in terms of a culturally evolved system of labels such as a music grammar, but it should rather aim to show how this culturally evolved system of labels is

possible in the first place. In particular, a discourse about listening to music or sound in general cannot be independent from a discourse about listening to words, especially if the research is based on an existing vocabulary.

Finally, the third kind of reference comprises human reports of listening experiences. We can roughly distinguish two different subcategories: verbal reports in a natural language and non-linguistic reports, such as drawings or movements. The problems that are possibly encountered in the first subkind are again connected to the use of the "enriched" vocabulary that comes from a conventional taxonomical description, resulting in an ontology, such as the music grammar. An investigation of listening should show how listening is able to account for a specific music grammar, so that it can also account for grammars that are different from our own, and therefore it should not be based on an analysis that starts from a specific music grammar if it wants to avoid the circular *paradox of categories*. This of course does not mean that language should be totally avoided, but rather that an apparatus of listening, be it a narrative or a scientific explanation, should face the circularity of language and therefore account for language, which cannot be done by starting from music grammars.

The second subkind has the advantage of avoiding language, at least in the constitution of reports, although any theory that has to be communicated to a community cannot really get rid of language. Delalande, for example, directs

the analysis of listening by driving it with the apparatus of listening conducts.

These are strategies

> of focusing on particular elements, which not only contribute to shape a perceptive image of a music piece, with its symbolizations, its sense, but also provoke sensations, possibly emotions, which in turn reinforce or redirect attention. [They are therefore] acts in which finality, strategy, perceptive construction, symbolizations, emotions are in relationship of dependence with each other and progressive adaptation to an object. (translated from: Delalande, 2013, p.42)

These perceptive acts may therefore result in references that witness this complex, holistic, sensorial approach related to listening with verbal and non-verbal reports. These reports are a consequence of the functional role of listening within the "praxis of living" and are therefore the symptom of a possible basic root of cognition that is working before the distinction among senses and possibly constituting them before we account them in our descriptions. In this context, I can only refer to the experiments of sense substitution as a possible clue regarding the emergence of senses as patterns of embodied behaviour (O'Regan, 2011).

3.6 The Object

The second point of Fiske's criticism gestures towards the wide spread adoption of the *paradigm of object* in dealing with listening. I am referring here mainly to my "Sounds and Objects" (2012) when addressing a paradigm that, while having been introduced for the first time in relation to listening by

Schaeffer in his treatise (1966), it is only from about 1990 that the word "object" has reached such a status of common use in dealing with listening that it is taken for granted.

We find therefore

> an increasing use of the word 'object' defined as 'sound object', 'sonic object', 'music object' or 'auditory object' in research mainly related to acoustics, neurophysiology, psychoacoustics, philosophy, semiology, sound design, musicology, music composition and music informatics. […] [All these disciplines] try to define the boundaries of phenomenic contexts that are ultimately related to the experience of sound, be they subjective sound sensations, verbal descriptions of sound sensations, visual representations of physical or physiological data, visual signs related to musical intentions or others. Each one of those disciplines uses language to gather a set of descriptions of the phenomenic context they refer to and at the same time establishes the connections with another set of descriptions. (Viel, 2012, p.234)

In the adoption of the *paradigm of the object*, I distinguish three main aspects that ultimately connect the different descriptions of listening to domains of description that include *object* as the label for composing and possible composite units.

First of all, *objects* are distinguished from a background in the same way a visual figure detaches itself from a plane that loses the focus. In this way, the *object* inherits the perceptive patterns, such as Gestaltic laws, that were connected to vision in the first place. This aspect is also connected to the use of referring to the graphic objects of notation, as we have already dealt with, which in the end is related to indexical or causal associations of the listening

experience to a *sounding object* (Rocchesso and Fontana, 2003) in terms of "naturalized labels".

The second aspect is related to the *container* image schema (Lakoff, 1987) in order to build a taxonomy in which bigger *objects* include, as a lower hierarchical level of the taxonomy, smaller *objects*.

Sometimes the schema involves levels that are not described as *objects*, so that only a specific layer of the taxonomy entails the *paradigm of the object*, as it happens in Schaeffer's opposition between *object* and *structure* (1966, pp.277-278).

Finally, it is the logic-linguistic use of words as categorical labels that permits a use of this paradigm in order to adhere to the standard of *Object Oriented Paradigm*. The *object* therefore becomes a *class* that needs to be *instantiated* in order to connect with experience and, in the perfectly reversible process, listening is directed to identify *instances* of a *class* in the sound flow. This aspect does not simply point us to a domain of description, which entails units and relationships among them as a "theoryladenness" for our analysis of listening, but rather, and more dramatically, ascribes our search for such domains of description to the very structure of our language, as based on noun, verbs and objects.

This is not a new topic, as David Bohm (1981) has already addressed it in his attempt to build a proper domain of description for quantum mechanics. The paradigm according to which linguistic typology classifies world languages by

determining their word order takes for granted that the structure of nouns, verbs and objects is universal.

On the other hand,

> this is a pervasive structure, leading in the whole of life to a function of thought tending to divide things into separate entities, such entities being conceived of as essentially fixed and static in their nature. When this view is carried to its limit, one arrives at the prevailing scientific world view, in which every-thing is regarded as ultimately constituted out of a set of basic particles of fixed nature. (Bohm, 1981, p.37)

Therefore, the interdependence between our set of labels and our experiences seems to set us in a pervasive system of meaning which does not allow addressing what is outside of this system, or at least makes it a rather difficult task. Language has become an apparatus, something that has "the capacity to capture, orient, determine, intercept, model, control, or secure the gestures, behaviours, opinions, or discourses of living beings" (Agamben, 2009, p.14) and is inscribed in relationships drives by games of power (Foucault, 1995).

As a consequence, it is not easy to deal with languages that do not adhere to the noun-based structure of our language, because we might not even realize what that means. In fact, even in dealing with possible different language structures, not only the Indo-European language speakers, but also linguistics itself as a discipline born in a culture based on Indo-European language, tend to reduce languages that are different from ours, such as Iroquoian Cayuga or Polynesian Tongan, within the schemes of the noun-verb-object structures. The appeal "to look at the distinction languages really make, rather than to look

for distinctions only because we have inherited them from the study of Indo-European" (Broschart, 1997, p.160) can indeed be extended to the study of listening. Maturana's epistemology of simple and composite classes is ultimately not so different from the taxonomy of *objects* with its spatial metaphor, which comes from the sensorimotoric's and visual's domains of description.

Again, in dealing with listening we cannot take the *paradigm of the object* as a primitive model for describing our experiences, which has to be taken for granted, but we should instead try to show how using this paradigm can describe listening without resorting to "naturalized labels".

Perhaps, adopting an idea of distinction that does not imply a differentiation between a foreground and a background as an elementary cognitive act can help to construct a narrative that ultimately might make available to our descriptions Schaeffer's objects as well as notational ones as the result of more basic cognitive processes.

3.7 The Subject

When dealing with descriptions, I have set the constraint to only use the subject "I". A description related to a given object has been accepted only if the association among the subject "I" with a verb, and the adjuncts, both related to the verb and the object, have been considered as meaningful from a community of native speakers in the natural language of the description. "Meaningful" here means that the description has been judged as

grammatically correct and can be connected to a semantics, which is related to the "operational coherences" of the community of speakers in such a way that, when expressed by the subject itself and under certain conditions, the description can be considered appropriate.

Conditions are related to the range of domains of description that are entailed by the subject "I" and the object. On the other hand, domains of description are defined by ranges of compatible verbs and adjuncts, which are mutually implied and are connected to domains of "operational coherences". We can therefore overturn the point of view, which was initially based on the object, in such a way that, once we have rendered explicit the possible homonimic uses of verbs by having distinguished homonimies as occurrences of different domains of description, we can identify the set of possible verbs together with their adjuncts, as the spaces of actions of the subject "I" that are distinguished by the different domains of the discourse they belong to.

In sum, a subject selects, within all the possible descriptions related to it, the actions it realizes together with the range of possible target-objects, on which the actions are directed, as is implied by their domain of description. We can call the set of selected actions *a domain of action*.

A subject is therefore connected with a domain of action, that, when we move from the subject "I" to a different "third person" subject-unit, mirrors Maturana's idea of a "domain of interactions" (Maturana, 1988, §6.13). Moreover, with the possibility to change the subject, a specific object can also

be connected to a variety of subjects, so that it is possible to relate an object to a domain of subjects.

We can develop our discourse on the subject in several different directions if we start from the idea of a domain of action and a domain of subjects.

First of all, Maturana's epistemology can been connected to Foucault's analyses of discursive practices (Foucault, 1972) so that we can insert the practice of listening in a process of subject-formation by means of the interaction between the experience of listening and the ways of describing them.

In this regards, a discourse on listening is always related to a subject, even when it is, as it happens in most cases, tacit.

Secondly, it is possible to imply the subject by dealing with an object of listening, whose domain of subjects defines a more or less precise type of listener or opens to a broader generalization.

The idea of the subject developed by Foucault, is, as is well known, born from the Nietzschian questioning of the self as expressed by an uncritical use of the pronoun "I" (Nietzsche, 1996). This places the discussion of the subject within the realm of a search for identity, "become who you are!" (Nietzsche, 2006, p.192), which is painfully stretched between practices of subjectivation and apparatuses of subjection and is therefore necessarily driven by ethics (Beard,2009).

Foucault (2006), in his investigation on the idea of subject in ancient Greek philosophy as an alternative vision from the modern one, interprets the process of building the self, which we could translate as the cleaning process of one's domain of action, in terms of discursive practices, as the process of occupying the subject's position in descriptions that are true. This process therefore enables us to

> become the subject of […] true discourses, to become the subject who tells the truth and who is transfigured by its enunciation of the truth, by this enunciation itself, precisely by the fact of telling the truth. (Foucault, 2006, p.332)

In opposition, Christian tradition fosters self-renunciation, a discipline that leads to confession, which Foucault describes as the transformation of the self in an object of discourses, as the "objectification of the self in true discourses" (Foucault, 2006, p.333). In both pagan and Christian perspectives, truth plays an important role in the process of inner transformation. However, if the Christian truth lies in the Word of God as is expressed by the Holy Scriptures, the pagan truth lies, in this context, in the adherence to the "operational coherences", in the recognition of experience as the guide to disentangle oneself from the linguistic intricacies that come from using a finite vocabulary to defer infinite experiences. And a disciplined subjectivation is therefore the process of cleaning up the lies of an involuted lexicon by counterposing to it the truth of practices.

In his analysis of source texts and especially of Seneca and Epictetus, Foucault analyses the discipline of listening, which is essentially and necessarily addressed to the expressions of *logos*, as a three-part process that involves the practice of silence, the set of a correct posture and commitment. Nevertheless, the question of counterposing the practices of subjectivation against the apparatuses of subjection can and must also be posed for a listening practice directed to other aural phenomena than verbal discourses and ultimately should be treated as a general issue, which is related to the formation of the personal identity within a multi-cultural, media-directed, mass society.

But this cannot be done without narrating the construction of the subject, because any discourse about a third-person subject cannot avoid being shaped by a first-person subject that might be trapped in the net of subjected knowledge. As we have seen, a subject's domain of action is bounded by the domains of description it can use to describe its actions, but even if "the limits of my language mean the limits of my world" (Wittgenstein, 1961, §5.6), a subjected experience has the power to limit the possibility of instantiating my linguistic world by transferring the active potential of verbs into the passivity of objects. In the second part of this text I will propose an interpretation of the process of self-building, which by starting from the very basic act of distinguishing through listening is able to show an openness to both a discipline of subjectivation and to practices of subjection.

3.8 Erasing the subject

Subjection is realized by changing a domain of action, as a consequence of erasing the identity of the actor, namely the subject, from the truth conditions of a discourse. A domain of action is changed by erasing or substituting its domains of discourse and, as a consequence, by modifying within the domains of the discourse the verbs and the related proper objects that compose them, together with their adjuncts. There are several ways to put a subject actor in the background; for example it is possible to use generalized subjects such as "a person", a "man", or impersonal subjects. Other ways that are particularly important in dealing with listening are the substitution of the subject actor with other parts of the descriptive clause that have been "actorized", like substantivized verbs, such as gerunds, or anthropomorphized objects, which are now capable of an action of their own. As the result of these linguistic practices, the identity of the acting subject is no longer involved in the "operational coherences", which confer a truth-value to descriptions and therefore its identity is implicitly inferred from descriptions, when it does not disappear from the discourse entirely.

The use of "listening" is an example of an "actorized" verb, which is used throughout this text. Indeed, the simple "listening" subject seems to suggest that there can be a "listening" that is independent from a listener and therefore it can be shared, or better it has to be shared, by all listeners. Being unable to

refuse an idea and ultimately a practice of "listening" is therefore, for a listening subject, the highest possible attempt of subjection.

As I mentioned before, the experiments of sense substitution seem to suggest that senses are more connected with each other than we are used to. If we experience some evocation of vision when hearing for some time, while being blind, the sounding translation of the image that comes from a camera, which is attached to our body (O'Regan, 2011), then what we call vision is not simply what reaches our eyes from an external word, but rather the result of some regularities in the interplay between an activity, which we call "our body", and a passivity, which we call "the world". Our distinction among senses is already part of an apparatus of subjection.

Therefore, if a narrative, which starts before the process of subjectivation, cannot avoid moving from the highly subjecting condition of a non-subject, in its constructive path, it nevertheless needs to propose a way according to which the subject as an individual self emerges from the starting condition and in the end is able to propose a perceptive mechanism which is subject forming.

The "actorized" object is another device for letting the subject actor disappear from discourses. We have seen how the actor-subject is a sort of uncomfortable presence in discourses because on one hand it limits the discourses to a context, which has to be explicitly defined in order to be able to verify the truth-value of descriptions, on the other hand a subject is involved in subjective experiences, which, as we have seen, are very difficult to deal

with. That is why most research on listening is instead explicitly devoted to music or to perception as caught in actions that are engaged by them. Even "sound studies" (see for example: Sterne, 2012), the newest academic field which emerged in the last decades as a consequence of the increasing attention directed towards listening practices, is named after a possible object, in the grammatical sense, of listening.

But dealing with music or perception without putting the identity of the subject-actor into play means that music and perception are independent from a specific subject so that in the end, as we have seen with "listening", the subject is implied from the domain of subjects connected to the domain of actions.

When dealing with perception, the subject is embedded in the ontology expressed by a reference, so that in case of reductionist measurements, the subject is implied by the topology of parameters that are involved in the scientist's operations. But the data emerging from experimentation are related to specific forms of acoustic presentations that, by wearing the mask of objectivity, do not account for their esthetic/esthesic specificity and therefore are not able to detect those perceptive abilities that are entailed by the identity of the subject. Even if the experimental results are constrained within the proper conditions for an epistemology, so that the domains of description involved are properly distinguished from each other, the absence of the subject among the conditions for truth evaluation might simply let them miss the perceptive specificity of the subject.

Consequently, an average abstract subject is created, as defined by the measurements averages, a simulacrum, which is endlessly instantiable and by its replicas expresses a rhetoric of objectivity and necessity, while in the end it is the mere expression of a (stylistically) situated apparatus of subjection.

Turning to music in substituting the subject-actor, the observations to be made cannot be so different with respect to perception. The claim of objectivity is not at work here, as it is replaced by the claim of universality. This is a sensitive topic, which articulates between the academic debate—according to which the universality of music as a common cognitive approach to sounding practices is mostly recognized, while the existence of universals in music is mostly doubted —and the general acceptance of the mass culture apparatus, that is essentially related to pop music repertory, according to which music is a universal language (Wilson, C., 2007). It is not within the scope of this text to discuss the reasons for accepting or refusing to consider music as a universal phenomenon that is present in all the human cultures.

What is important here is to at least point out that some considerations should be made about this topic, before taking for granted that dealing with music can be safely done without accounting for the subjects involved. Nattiez, in an age that still was not entangled with the construction of the neurocentric golem (Lilienfeld and others, 2015), discusses in his *Music and Discourse* (1990) whether we have a stable definition of music in our culture. Moreover, he asks whether using the term "music" to describe sounding phenomena in other

cultures that do not have such a notion, or do not distinguish between what we call music and non-music, is a colonialist way to project our categories onto other cultures. This last issue is clearly quite similar to the problem we have already encountered in attributing word categories that come from our language to understand languages that are very different from our own. First, Nattiez concludes that even inside our tradition, the notion of music is far from homogeneous. Secondly, he presents several examples, which are intended to show that our notion of music can be very far from the categories of other cultures.

An important example is the one of *zoomusicology* (Ullrich, 2014), an academic discipline that was born from pushing the "music" paradigm to deal with the sounding practices of non-human animals. As a consequence of the idea that non-human animals may also develop cultures that emerge from reciprocal repeated interactions within a community, the notion of animal culture seems to foster a humanized version of non-human animals instead of a post-human recognition of human musics (plural intended) as specific sounding practices that inside some cultures are labelled as "music".

Concerning the erasure of the subject in discourses about listening, relying on music as the unique source for a theory of listening, or a theory of cognition (Serafine, 1988), could be misleading if a theory is not driven by the realization that music is not merely sonic content that somebody listens to, but it is a practice of listening, performing, composing, organizing, dancing and many other activities, that emerge in specific symbolizations from the cultural, in the

general sense, drift of a community. Therefore, its aspects are always situated and specifically related to a group, no matter how large it is. Ultimately, a discourse about listening always and necessarily expresses a point of view, as we have seen, but it also needs to leave an open door for diversity, as an essential part of the discourse, so that it can show how listening can deal with sounding practices that are fundamentally different from what we are used to. Only in this way can we possibly be protected from the temptation of subjection and, for that matter, also able to deal with the sounding practices of non-human animals without projecting assumptions that come from the subjections that entail us, so that we can finally engage a true post-human perspective—regarding humans as well.

3.9 Intransitive listening.

Dealing with listening means dealing with verbs, like any verb on Schaeffer's list or any related verb in English (such as how the label of the experience of listening is seen as an action connecting a subject with a possible object). We have also seen that in relating a verb with a truth-value by the "operative coherences" of an observer a subject is needed in terms of an identity that puts the subject-actor in a specific context. If a subject is needed, we should also wonder whether an object is needed.

We have seen that the presence of an object can be problematic in two ways: first by entailing a sensorimotor metaphor which might drive our considerations away from properly outlining the character of listening and secondly, by

conceding a specific object the supervenience on the subject. On the other hand, erasing the object from a description would indeed impoverish the domain of actions of the details expressed by the objects and their adjuncts, but at the same time it would result in expressing the most general and possibly deeper character of the verb, as independent from objects. It would express the intransitive aspect of action. How could we therefore possibly rule out the object of listening from the truth constraints operated by practices?

Peter Szendy proposes the idea of a *sur-écoute* (2017) as a condition of the subject between overhearing, which is the literal translation of the French term, and eavesdropping. In this case, the listener is always striving towards objects, which s/he expects to occasionally show himself or herself in a sort of proliferating polyphony. The listener appears to be in a never-ending cocktail party-like situation in which he is always directed to the detection of some hidden conversation, while some other unexpected conversation is leading its way to him. The object is nevertheless always there, caught in a process of being decoded by the subject. Therefore, it is no wonder that Szendy shapes his analysis of listening in terms of the activity of spying.

Nancy, on the other hand, in posing a fundamental philosophical distinction between the Schaefferian's verbs of *entendre* and *écoute,* cannot reduce the idea of listening to sensorimotor and visual paradigms-metaphors, which he connects to the tension towards meaning:

> perhaps we never *listen* to anything but the non-coded, what is not yet framed in a system of signifying references, and we never *hear* [entend] anything but the already coded, which we decode. (Nancy, 2007, p.39)

If vision is entangled in an ontology of shapes and contours, in a material discontinuity that constitutes itself in surfaces and objects that can occupy a place and can be surrounded, listening on the other hand, when detached from vision, becomes a sense that is open to the modulation of space-time and that reacts to this modulation by means of a resonance.

> The sonorous present is the result of space-time: it spreads through space, or rather it opens a space that is its own, the very spreading out of its resonance, its expansion and its reverberation. (Nancy, 2007, p.13)

Nancy's idea of an object of listening gives back its blurred and metamorphic character. Listening is the way to deal with a specific sensorial flow and its constitution in terms of space and time, so that being immersed in an ocean of sonic waves, which move in space and reflects in time, is resolved by our language of objects. But these linguistic objects are secondary with respect to structure of our sensing, a structure that cannot be thought in terms of objects, in the first place.

This is why listening (*écouter)* before, or independently from, hearing (*entendre*) is an encounter that is realized by the tuning between the capacity of the subject to engage a wave-like sensing action and the waves of the sounding flow: it is an act of resonance.

Sense is first of all the rebound of sound, a rebound that is coextensive with the whole folding/unfolding of presence and of the present that makes or opens the perceptible as such, and that opens in it the sonorous exponent: the vibrant spacing-out of a *sense* in whatever sense one understands or hears it. (Nancy, 2007, p.30).

In conclusion, this grammatical path has led us to the idea of listening as an intransitive verb.

But "listening" is here intended, so to say, as an act of perception that is before musical grammars, before the "naturalized" use of labels with their ontology of objects and therefore before language; but it is also before the modes of listening. Only the subject is present, in a sort of cognitive isolation tank, an ideal starting point for a constructivist path of narrative. And the subject, which is not yet a human subject, with an identity that is related to its place in the world, enacts and reacts to the goad of resonances. I will not use the word "resonance" any more in this text, but it well expresses the sounding friction that will give rise, in consciousness, to the separation of the self, which will develop in a subject, from the world, by sensing.

4. NARRATIVES

In this final chapter of the first part, I would like to summarize some requirements, which have been mostly met already in the previous sections that should be followed in discourses about listening.

Here, I will follow and comment upon several of the 21 principles for a theory of music listening that are listed at the end of Harold Fiske's *Selected Theories of Music Perception* (1996), as one of the very rare examples of a discussion about the requirements for a theory of listening. However, I neither propose a theory of music perception, nor a theory of perception here. I am not even proposing a theory.

My intention is rather to tell a story, to build a narrative. In fact, taking seriously the inevitability of expressing a point of view means that it is not possible to talk about a reality, or better it means abandoning the idea of words, such as "substance" or "matter", with which we can express an absolute independent from our discourses. We can therefore only implement a fiction, possibly able to give a direction to our actions, so that we can happily expand that fiction in the world we inhabit, by creating something that we are used to calling "reality".

A narrative cannot avoid the character of authoriality because an individual narrates it. Consequently, a narrative is contingent in that it is connected to the contingency of the author as opposed to the necessity of an impersonal and absolute truth. Therefore, a narrative cannot claim a foundational (and

therefore objective) status, as it is necessarily incomplete and temporary.

Finally, a narrative is entangled in a process of negotiation, an open dialogue with the reader, who must be able to link the narrative to his own experiential references and possibly might want to embrace it, or part of it, in his practices. The concept of negotiability is connected with "viability", a term that was introduced by Ernst von Glasersfeld (1981) as an epistemological translation of the biological idea of evolutionary adaptation, in order to deal with the idea of truth in a radical constructivism.

In fact,

> we construct ideas, hypotheses, theories, and models, and as long they survive, which is to say, as long as our experience can be successfully fitted into them, they are *viable.* (In Piagetian terms we might say that our constructs are viable as long as our experience can be assimilated to them). (Glasersfeld, 1981)

In this context, viability is related to the possibility of a narrative to survive the "environment" made of its readers, without appealing to values of truth. As long as readers are capable of interacting with a narrative, by using it in their life as a tool to positively foster new reality, in its very general meaning, that narrative has to be conceded the right to be narrated.

In Part 2, I will therefore propose a narrative that will possibly show, in the third part, its viability, as a method for music analysis, and music composition, and in dealing with more general ideas, that are related to discourses about our relationship with the aural phenomena that surrounds us. In developing the

narrative, I will focus on concepts, such as perceptive dimension, distinction, pattern and region of distinctions, in order to let them interact with each other and to show the potentiality of their system of interactions in building an explanatory, and possibly implementable, mechanism; a task that I leave for further research.

But first, it is necessary to discuss what we should expect from a narrative that, by starting before the objects of listening and even before a properly formed subject, promises to include in its development the language we use in the narrative and the cornucopia of listening experiences, the experience of music included.

Fiske lists 21 principles that conclude his analysis of some well-known theories of music perception: "talk about music and you are talking about the musical mind: descriptors of music are descriptors of cognitive processing" (Fiske, 1996, p.138). As being the proposer of the "copy paradigm", he quite surprisingly adopts it in a chain of objects, "music" as "musical mind" as "cognitive processing", that apparently never touch the subject of listening, which is, in cognitive terms, the mind as a whole system—no wonder then if in the first principle, the subject is already substituted with the brain.

I will try now to list Fiske's principles (1996, p.153-154) from number 1 to number 15 of the total 21, because the last six are too involved as consequences of Fiske's theory to be reported here as basic assumptions.

Each quote from Fiske is followed by a brief comment contextualizes the principles within my research.

1) "the brain seeks perceptual economy" – It is the narrative that should minimize its resources in order to optimize its viability;

2) "perception in a search for element interrelationship" – it is actually a domain of discourses that relies on the description of a unit in terms of its components;

3) "music perception is a search for […] any type of sonic patterns" – the distinction of patterns is indeed an important part of my narrative, not only in relation to music, but mainly in relation to any aural experience and, in the end, in relation with any sensorial experience overall;

4) "perception requires time and effort" – an embodied enacted listening needs to deal with the finite aspect of biological organisms: sensing is directed to survival, as is narrative;

5) "perception is a construction process, not a copy process" – the explanatory mechanism is indeed made in terms of a causal description;

6) "music cognition is unique to human brains" – "music" is a word that is used within a community to label a set of practices that are born within that community. Sounding practices on the other hand, which include spoken language are diffused among human as well as non-human animals;

7) "the function of music cognition is the realization of tonal-rhythmic (pitch durational) patterns" – the listening process directed to sounding practices are focused on the elements that emerge as peculiar (and therefore are referenced

by other practices of the same community, such as talking or teaching) of those practices;

8) "music cognition is limited exclusively to the realization of the tonal-rhythmic (pitch durational) pattern structure and interrelationship" – that is simply not true as it is witnessed even in our tradition by certain compositional (and listening) practices of contemporary music or contemporary society music styles, such as drone ambient or noise;

9) "music cognition requires time and effort" – it is the consequence of §4 and of the embodied enactive character of sensing;

10) "music pattern comparison procedures represent a semantically closed, self-referenced, modular system" – this principle is too conclusive to be commented upon without having yet presented the narrative of listening and therefore shall not be;

11) "music communication is the result of a shared social-cultural contract" – while the association between the words "music", "communication", and "language" is somewhat problematic, it is relatively uncontentious to claim that music has emerged within a community as the encounter of some of its practices;

12) "music listening requires active, rather than passive, participation on the part of the listener" – or in other words: listening, as a way of sensing, is enaction;

13) "music listening is a hypothesis-testing process" – this is acceptable only if we remember that "hypothesis-testing" is here an anthropomorphic metaphor. "Hypothesis-testing" is what mirrors a basic mechanism of sensing;

14) "perception is dependent upon a cognitive context created by the listener" - in the narrative I adopt, listening is context-related as the result of its intrinsic process of abstraction;

15) "a musical language is defined by the set of syntactic rules describing the permissible tonal and durational relationships which represent that language" – this is, in my opinion, a highly simplistic way to represent music even in the western tradition. Moreover, it is completely incorrect if it intends to address the broader sounding practices of the world.

These 15 principles seem to run into the fallacy of the subject substitution at all levels, so that "listening" appears explicitly only in numbers 12 and 13, and as a gerund, while the listening subject is involved only in numbers 12 and 14. The actant seems therefore to vary according to need, sometimes referring to the brain, sometimes music cognition, perception and so on. And as a consequence, it is rarely recognized, at least in the formulation of these principles, that listening exists before music, before the idea of perception and, in my narrative, before the language itself.

Conversely, the narrative that this text focuses on, starts as an experience "always already" in our daily experiences (see the first chapter of the second part), while its starting point of narration originates from the very constitution of

the enacted life, in the Maturanian "praxis of living", of the organism as is described by an observer.

Therefore music, not even the word "music" as applied to the world sounding practices, cannot be the possible incipit of a discourse on listening, because music cannot but work as a part of an apparatus of subjection that removes the listening subjects. A discourse on listening cannot even be founded in language, because language too is a result of listening practices. We therefore need to find a generative experience that, while being necessarily expressed by language, is "always already" actualized by our experiences and entangled in the "praxis of living" so that it involves time before the idea of time and effort before the idea of effort.

Contemporarily it should be possible to also include the basic skills of listening that we are used to attributing to our ontologies with their references, such as Gestalt laws or habituation effects, as far as we are able to describe them with our language. It should be possible to account, for example, the experience of perceiving a sine wave whose frequency is randomly modulated with an increasing range as a cross fade between a sine wave and a white noise.

This brings us to the last and most important point: a discourse about listening that seeks to capture the experience of listening should involve, as a practice of truth, the ways we talk about listening as an activity of the subject, with narratives that do not erase the subject. It is therefore essential for a discourse about listening to be able to account for, at some point along the path of its construction, how the diversities of listening (and consequently sounding)

practices emerge from the basic mechanism-without-subject. A goal that is not possible to reach if the starting point, the music or a universal music, the brain, music cognition and so on, is already a subjected one.

Let us now move to the first informal narrative about listening as expressed by the simplest experience of it: the distinction of a change.

PART 2
PATTERNS

5. IN THE ISOLATION TANK

5.1 Through our headphones

I would like to begin this second part with a narration of what we could call "the discoveries of listening", which intends to intuitively introduce some of the elements that are at the core of my research in the form of a mental experiment. I shall lead this narrative through the experience of listening to music, but only as a sort of tribute to the tradition of dealing implicitly with listening through music. Nevertheless, the same path of "discoveries" could also be explored through everyday aural experiences.

Let's imagine that we are comfortably sitting in the dark, listening to the noise and sounds that come from our headphones. It is something akin to floating inside an isolation tank, an environment that was invented by John Lilly in his research on the isolated brain (Lilly, 1988), whose purpose is to suppress all the sensory inputs as much as possible, in order to simulate the condition of a "brain in a bowl". Let's imagine we cannot see, smell or taste, we cannot even have the sensation of touch or the other sensations that constitute our most updated inventory of senses: we can only hear. This is likely an impossible condition, but I will nonetheless try to set the starting point of my narrative on this fantasy, so that I can experimentally pretend to talk about listening without involving the other senses, at least for now: we are sitting in the dark and we

are listening through our headphones to the noises and sounds that reach us from a (far? near?) music-hall, just before the concert begins.

We are immersed in a continuous stream made of noise and sounds that we mostly recognize: the voices of people talking, coughing, moving and shuffling their clothes on the seats. After some time, we can also distinguish a chaotic mixture of melodic fragments, inside of which we can recognize the sound of musical instruments, which are more or less familiar to us. Suddenly, the noise gets very soft, almost silent. We can only hear a single sustained sound that is soon followed by an intense mixture of slow and fast melodic fragments. Then silence again. After a while, a thunderous noisy texture bursts in: applause. Apparently the conductor has entered the stage. And again silence. All at once, a completely new sound experience bursts in and dominates our experience of listening with a strong intensity and highly predictable sound changes, to the point that we can also tap our foot in synchrony with the sound of the orchestra and, presumably, with the gestures of the conductor. The concert has definitely begun.

We have been able to understand and follow what was happening in the hall, probably because we have already been in similar situations. We are therefore generally able to associate the changes in the sound quality of the sonic stream with specific events, which follow a quite precise order that corresponds to what usually happens before a concert begins.

This capability is not banal at all. In fact, if we were asked to listen, within the same sensory isolation, to the sonic stream coming from a less "formal" situation, let us say some experimental music performed at a vernissage in an art gallery (for example "It's Gonna Rain" by Steve Reich, a composer whose music has been intensively performed in art galleries, during the late 60s), we most probably would not be able to detect the beginning of the performance without some serious effort.

In sum, I have characterized the experience of listening in terms of the distinction of relevant sounding moments, marking the succession of sound events, or, in other words, in terms of the auditory segmentation of a sounding stream, which, as we know from Cage's experience in the anechoic chamber, is a never-ceasing continuum.

And actually, any discourse on listening cannot but start from the existence of a sounding continuum, which constitutes the horizon of listening. But immediately, we need to turn our attention to the unfolding of the segmenting process, which articulates the sounding continuum in time.

5.2 - The music of noise

Even in the case of an absence of "signals", such as the audience's applause and the tuning of the orchestra, the cognitive act of separating the beginning of a music piece from the environmental noise that precedes it in the sounding continuum, is usually an easy operation. The distinction between the "before"

and the "after" of a musical beginning is commonly addressed as the distinction between "noise" (or "environment", "background noise" or simply "background") and "music" (or "sound"). I have to postpone the discussion on the different uses of the word "noise". For now, it is enough to say that "noise" is here intended as in the sentence "this is not music, it is noise!", which we could call a "semantic" acceptation, according to which noise is something that "doesn't speak to me", as is often said about some music we don't like.[15] Therefore, the distinction between "noise" and "music" is here a distinction of meaning, between two attitudes of the listener, which emerge within the complex articulation of social practices. In the mode of listening related to the word "noise", our attention is directed to the source of the sonic phenomenon and any quality that we detect in the sound/noise we hear is related to a property of the physical phenomenon that originated it. This is what we have called the "extrinsic meaning" of a sound phenomenon, which we have connected to "naturalized" labels. On the other hand, when we use the word "music", as opposed to "noise", listening is directed to the inner qualities of the sonic phenomenon, or better to a structured selection of them, in order to follow how these sounds change in time. This is what we can call "intrinsic meaning" and it is usually referred to as the typical music listening attitude.

I will use the terms "music" and "noise", in opposition with each other, with this meaning throughout Chapter 5 and I will distinguish them from other possible uses by putting them between quotation marks.

---

[15]A use of the word that most contemporary music composers have met at least once in their lifetime when dealing with people annoyed by their compositions!

It is nonetheless obvious that the process of segmentation is here connected with a distinction, the one between "noise" and "music", that goes beyond the simple discrimination of the sonic event (although all the perceptive mechanisms that are investigated by psychoacoustics are inevitably at work) and moves towards a more general articulation of the subject and the world of social and cultural practices it is immersed in.

The problem of the threshold between "music" and "noise", that is the problem of detecting the conditions according to which what we usually consider as "noise" starts to be considered as a part of "music" and vice versa, has especially intrigued composers over the course of the last century.

I need to mention at least:

- the birth of *musique concrète* in the 40s and the application of the idea of *objet trouvé* in music composition, as a consequence of the diffusion of tape recorders in radio stations, with the related techniques for manipulating sound events that were not previously considered as proper elements for music (Schaeffer, 1966);

- John Cage's idea that "everything we do is music" (Konstelanetz, 2003, p.69) as it was first expressed in 1952 through his seminal work for any instrument combination "4'33"", in which the performers do not actually even play their instruments, but rather sit or stand quiet letting the world sound;

- the idea of "soundscape", which Murray Schafer first introduced (Schafer, 1969) in order to grant the acoustic environment we live in, an attention that is

usually devoted to music, up to the point that it stimulated the development of a specific musical practice/genre with the name "soundscape music".[16]

Managing the threshold between "noise" and "music" in the previous examples is mostly a matter of letting the categories of our subjected ontology calibrate the psychoacoustic process of segregation, which is the ability to distinguish simultaneous layers in the continuum. Some composers on the other hand want to soften the passage between "noise" and "music", with the result of blurring, at least for a while, the boundaries between background (the "noise") and foreground (the "music", which emerges from the environment).

For example, Gerard Grisey's "Dérives" for orchestra, the masterpiece of the spectralist movement, begins with almost no perceptible difference with the preceding moment in which the musicians of the orchestra tune up their instruments. The i(m) "clapping moment" from Karlheinz Stockhausen's "Momente" for solo soprano, choir and orchestra, which in all the actual versions of this open work immediately follows the audience's applause, consists of the choir clapping back to the audience in a rather humorous mirroring, which slowly transforms into a musically articulated texture. During this more than two hour long work, the choir inserts several "signals" which are derived from the typical noisy reactions that the audience of that time used to express for its disapproval of avant-garde music, such as hisses, shouts or

---

[16]I will come back to the topic of "soundscape" in Chapter 12 of the Third Part, where I'll discuss the idea of soundscape more in detail.

clapping noises. In this way the boundaries between the foreground "music" and the background "noise" were so blurred that the audience was no longer sure whether its protests were going to become part of the "music" or not.

An historical and therefore less literal example of fading from "music" to "noise" is the ending of Franz Joseph Haydn's Symphony No. 45, the "Farewell" Symphony, dated 1772, in which the musicians of the orchestra are supposed to leave their place gradually one after the other, so in the end only two violins are left to conclude the piece. In this way the musicians, under the composer's guide, could express to the orchestra's patron the Prince Nikolaus Esterházy, as politely and cautiously as possible but in an unequivocal fashion, their desire for a vacation from work. Not unlike Haydn's work, in the theatrical version of Stockhausen's "Luzifers Tanz" ("Lucifer's dance" from the opera "Saturday from Light", which premiered in Milano in 1984), the performance is interrupted by a sudden, this time violent, strike of the orchestra which protests against overwork and therefore leaves the last nineteen bars of the work unperformed. In this case it seems that the reason has to be found as a "satanic" mockery of the orchestra's strike that, during the staging of the previous composer's opera "Donnerstag aus Licht" ("Thursday from Light") at Teatro Alla Scala in Milan in 1981, truly impeded the realization of the third act for several days. Nevertheless, the examples of fading in and out between "music" and "noise" are rather infrequent in works that are not involved with theatrical narratives. On the contrary, as soon as the theatrical part of a work starts to play a more important role, the interaction between music and the

137

world in which that music is played, which usually fades in the background during the performance, starts to intervene and a complex mirroring of *mise-en-scènes* may take place, often with an entangled articulation between narrative expedients and reality. The concept of *diegetic* music exemplifies the relationship between "noise" and "music" well, which is a functional relationship, not unlike the relationship between sound and silence that was rendered explicit by Cage's experience in the anechoic chamber (Cage, 1961). We call music that takes place as an event within a narration *diegetic*. It is therefore opposed to the music that accompanies the narration or is part of the way the story is narrated (the so called *non-diegetic* music) and especially emerges as a narrative resource in cinema (Chion, Gorbman & Murch, 1994), even if we can already find it in operas such as Mozart's "Don Giovanni" (in the scene of Don Giovanni's ballroom during the first act).

The relationship between "noise" and "music", when translated in the relationship with a narrative plan, is caught in the *mise en abîme* of representation, in which the "music" of one level of narration may become the "noise" of a deeper level. For example at the end of the third act of Stockhausen's "Donnerstag aus Licht", a section called "Vision", the triplicated role of Michael assists to a shadow play in which seven excerpts of the very same opera are presented as memories. Nevertheless the *non-diegetic* music is still present so that the excerpts, which previously were part of the *non-diegetic* music are now overlapped with the actual *non-diegetic* music as a different layer. But like the interplay between the possible plans of narration,

which is driven by apparatuses that, so to speak, lock the range of our shared narratives (Foucault, 1971) which are now open to iterated or combinatorial articulations (I will just mention Raymond Roussel's "Locus Solus", which is a progressive accumulation of nested narrative layers and ends with seven different nested stories), music too is forced to follow the complications of narrative paths in possible *meta-diegetic* or *cross-diegetic* roles, that still need to deal with the mechanisms of perceptive segmentation. This is why music is most of the time only partially involved in the multiplication of layers and, whilst we can easily find the presence of *diegetic* music, often involved in a transition from a *diegetic* and a *non-diegetic* role, it is very difficult to find music that reaches us from a second level of *diegesis*. It is worth mentioning as a rare example of second-level diegetic music, Richard Strauss' "Ariadne auf Naxos", an opera that suffered, not without reason, some difficulties in being premiered due to the excessive length of its realization and the demands in number and versatility of the cast: problems that are almost inevitable when producing a work based on a meta-narrative. In the second part of the work (which is entitled "Opera") Zerbinetta and his group of burlesque actors play a little show for Ariadne to cheer her up for having being abandoned by Theseus. This little show has a *diegetic* role in Ariadne's plan of narrative, which in turn is the plot of an opera that is premiered, and therefore has a *diegetic* role, in front of the richest man in Vienna, who commissioned it, and his guests. And of course the whole work is performed in front of the audience, such as the one who attended the premiere in 1912, at the Hoftheather in Stuttgart.

As I have said, no matter how many the nested levels of narration there are, our perception has still to cope with the distinction between the "composed" overlapping events in the arbitrary complex narration and the "non composed" events, in which we are immersed.

What is interesting for my purposes is that even if the passage from "noise" to "music", for example in Grisey's work, is gradual or imperceptible, there is always a precise point, perhaps different for each listener, in which what was previously considered as "noise" is now considered as "music". In any case, we face a discontinuity in the listeners' attitude and it is this very moment that distinguishes the sounding continuum in a "before" and an "after". Of course, at the beginning of a music work a discontinuity for some experienced qualities in the sound is often present, which are certainly useful in helping to establish the beginning (or the end) of the work. For example a sudden chord in the brass section might be an unequivocal signal that marks the beginning of a work, as it happens for "Dérives".

In any case, it is the change of the listener's perspective in considering the sound events (we could call it the intentionality that drives listening), which establishes what is "music" and what is "noise": the discontinuity in sound has only the role of reinforcing the change of perspective.

Therefore, if a possible reinforced discontinuity is not recognized by a listener as the proper signal to mark his change of perspective, it will be useless for him.

5.3 Segments

Of course, even if we do not want to consider the concert event as a "music form" in itself, many music works are composed by a series of shorter pieces, which are separated from each other by "noise" inserts.

For example, an opera is usually segmented in acts, a sonata and all the forms that have been derived from it are subdivided in movements, the suites or the collections of similar works such as the preludes, studies, variations on a same theme and so on, when performed, are mostly separated by brief moments, which are not written and are not considered part of the music.

Even if the effect of these breaks into "noise" results in loosening the "perceptive tension" of the audience, so that people might cough or even get up and move from their seats, in case of a longer break, their direct purpose is not easy to decipher because is lost in the details of the historicized practices that they singularly emerged from. Therefore, a more detailed discussion of the typical relationship between "music" sections and "noise" in a larger work cannot be unrelated to a more general consideration on the cultural practices that generated it.

In any case, along with the development of western tradition, we can find both the tendency for bigger segments of "music" to be more and more separated by moments of "noise" until they can hardly be perceived as unified under a

same larger segment of "music" and the tendency of segments of "music" that are originally separated by "noise" to be connected with each other in a uninterrupted segment of "music".

Let's take the case of the Christian mass as an institutionalized sounding practice. At the origin of the rite, the mostly uninterrupted reading of the liturgical text primarily depended upon an unordinary "musical" use of the voice, which served as a contrast with the mundane world (Corbin, 1987, 47) by transfiguring the spoken voice in cantillation up to the ecstatic part of the Alleluia chant, the melismatic *jubilus*. Along one millennium of development, the segments of the Christian mass have become increasingly independent from their original context within the liturgy, so that at the beginning of the 18th century, for example, Antonio Vivaldi was able to write a *Gloria* (catalogued as RV 589) for orchestra, solos and choir that did not need to be performed during the Mass already according to the composer's intentions. This path of "mundanization" seems to have moved away from the purpose to exceptionally set a bridge between our daily experience and the Heavenly Jerusalem of "music", to being restituted back to the "noise" of the semantic context, of the spoken parts that are nowadays prevalent in the contemporary liturgy: that is, in the end, the experience of sound as the vessel of the extrinsic. At the same time, it seems that, by mostly relegating the words to the mundane net of linguistic meanings, the increased divide between "music" and "noise" has prepared the path to the emancipation of instrumental sound, firstly by

inheriting the rhetorical structures of speech and afterwards by subjugating the text itself to the structures of music.[17]

In the end, the mundane practice of "music" emerges from putting "noise" into a background, so that it can never be neglected and it is always there as a reference, even during the music performance, because music is no longer functional to a relation with the extra-mundane, but to the mundane structure of the power it serves, such as the aristocracy, whose members possibly attend the performance themselves.

There is no wonder then, if in the romantic-era the recovered relationship with the extra-mundane, which is often emptied from the institutionalized religious content, there is a tendency to unify the "music" segments in a continuous stream, in an extended *durchkomponieren*, which is intended to drive the listener towards an "other side" with respect to the "noise" that surrounds him. In this way the process goes so far, with the idea of *Gesamtkunstwerk,* as to include in "music" those aspects of "noise" that, while being experiential, which is not only acoustical, used to be considered the background of "music", such as lights, shapes and even perfumes (Alexander Scriabin's "Prometheus: the Poem of Fire" and Stockhausen's "Düfte-Zeichen", "Scents-Signs", are two well known examples of the introduction of multi-sensoriality in music).

---

[17]See for example the discussion about the interaction between music and text in the last century in Zeller, 1964.

If "noise" has the role of leading "music" segments within a frame of meaning (the theatre, the concert), we should nevertheless ask ourselves what the segments included in the same work have in common.

Obviously, it is the context that establishes a connection in the first place. It is the coincidence of performance time and listening that sets the connection, even when the unity of space is somehow fragmented (see for example, the works realized by simultaneous performances in different locations, that while being far from each other, are sonically present in both or in a third place where the listeners are attending the event, as it happens with Bill Fontana's "Satellite Ear Bridge Cologne-San Francisco"). Firstly, electronic technologies and subsequently digital ones have been able to fragment the time unity that connected the performer to the listener as the medium (be it the vinyl, the CD or the album distributed on the Web) that is now in charge of providing a frame for the experience of segments. With the arrival of distribution services that are based on the extemporaneity of music fruition and ultimately on the unpredictability of the succession of the single segments/tracks, such as Spotify and YouTube, the cohesion between performer/producer and listener is finally disarticulated and the meaning of frame is led back to everyday "noise". Nowadays, music is often experienced as a part of "noise", it is somehow more similar to something you wear, a clothing, than something to be listened to carefully in order to distinguish its elements. As a result, the daily routine, which is related to "noise", is transfigured and modulated by "music", which is in turn assimilated to the listener's identity.

The occasion of listening is often enough to realize the common frame that encompasses the different "music" segments listened to in succession, because it is the "noise" from which listening emerges, that connects, in a social/cultural net, the different practices with each other that produced those segments. After all, the music we can listen to on Spotify, in spite of its global shared character, gathers around well-specified statistics, which expresses the intersection between what the connected people, which is a subset of the entire population, are able to find within Spotify's catalogue and what Spotify was able to find, and get, from all the sound expressions of the world. Nevertheless, we need still ask to ourselves what are the common sounding elements that are shared by the different segments of a collection, which are the sounding identities on which the sounding differences that articulate the collection are built.

For this purpose, I restrict myself to mentioning two planes. The first plane refers to the possible attributes that might be shared by the segments included in a collection, therefore constituting the identity of the collection. For example, timbre (such as works for a common instrumentation), qualities related to some emotional/sensorimotor character (such as a collection of waltzes), but also, on a more explicitly linguistic-semantic side, the attributes that are related to text (such as a collection of songs based on the poems by a same author) or use (such as a collection of "Etudes").

The second plane is related to the differences between the segments of a collection, which we recognize as being articulated in a succession that follows a well-learned path. This might be related to the attributes such as the ones included in the first plane, which instead of being more or less constant along the segments, vary in a predictable way—for example, the character of the different movements in a classical symphony or sonata, or of the dance movements in a baroque suite and so on. In other cases, the path of differences might be, so to speak, a parasite of an extramusical scheme, such as a story, for example in Eric Satie's "Sports and Divertissements" (1914). In the end, any articulation of the differences in some attribute, even if along the predicable and well-known path, is usually integrated by some other attributes that are constant throughout the segments. For example, it is uncommon for the movements of a Classical-era symphony, which is usually scored for full orchestra, to be scored for very different setups, for example with one movement for solo flute, another one for the full orchestra and a third for kettle drums and the clarinet section alone.

However, it is not difficult to imagine a collection of music segments, a macro-form, that shares the least possible elements besides the time frame of performing/listening: for example, a collection of works with very different instrumentation, duration, style and also performance location. I think of a sort of an online—that means shared by performers scattered all over the world—

contest of remixes of one song, which can be realized without constraints of duration and instrumentation.

Why then is the presence of constant elements so common within collections of musical works? We might argue that the homogeneities, for example in the instrumentation, must be related to the institutional needs in allocating resources, which are therefore to be related to the socio-(political)-economics of music.

But it might also be for the simple fact that we do not perceive the common elements as something changeable, but rather as necessary parts of the (cultural) objects that surround us in our world. And as soon it has been realized by somebody that something can be changed in what we were taking for granted, then, after a while, it becomes part of the difference that can be composed and is no more part of the identity.

5.4 How to end a sonata

In this path of distinctions from the "top" of separating music from to "noise" to the "down" of dealing with the identities and differences within a collection of music works, we need to proceed further to the distinctions we make within single music works.

For this purpose, let's imagine now that we are listening to a recording of a music piece, let us say for solo piano, that we have never heard before and without knowing whether we are listening to the whole piece or just a fragment of it.

How would we be able to determine by just listening whether the piece has been cut or presented in its integrity? Of course, our ability also depends on our knowledge of the piece's style. For example, in case of a *pointllistic* music style, such as in Pierre Boulez's "Structures I" for two pianos, we would hardly realize the possible cuts. On the other hand, in the case of a tonal work, let us say in a *galant music* style, it is definitely easier, at least for a trained listener, to realize whether the work has been cut or not.

In fact, we might need some clues, to which nevertheless we still need to be accustomed, in order to detect the proper beginning or ending of a music piece and these depend, more or less precisely, on the style and in general the music system we are dealing with. Clues can be found in different layers of abstraction and usually refer to the concept and rules of composition, such as the accent structure, the harmony, the organization of phrases and the arrangement.

Here are some elements that can be helpful in recognizing the final part of a recording as the proper ending of a typical Classical-era work:

- the recording ends with a chord on a time-moment that we have recognized as the most accented beat of the meter;

- alternatively the recording ends on a long, at least double, appoggiatura with a tonic note bass on the most accented beat resolving by single degrees on the proper final chord;

- the final chord is the tonic one in root-position;

- the final chord might be repeated several times, with different voicing, before

stopping on the accented beat:

- the final chord/appoggiatura is preceded in the immediately previous weak

beat, by the harmony of the dominant;

- last chords are presented with a slowed down pulse speed;

- if I have correctly determined the proper beginning of the work, the recording

must have a duration that follows the requirements related to the proper

duration for a work, such as being at least longer than ten seconds.

It is not important here to address the cultural specificity of the cues as the

product of a complex socio-cultural system of practices that involve

composition, but as discourses as well and therefore also theory. The

Classical-era sonata is here just an example, which is necessarily situated in a

historical context, of a sounding continuity in which we aim for cues to base

our judgments that might be, for example, directed to understand whether the

musical piece had come to an end.

Nevertheless, these clues are related to conventional endings of Classical-era

piano music and are expressed according to the appropriate musical grammar

that is used today to teach *common practice*[18] composition. Obviously, it not

possible to take for granted that a possible listener knows what a "tonic chord

in root position" is or that he or she is able to recognize it properly.

Nevertheless, a listener can definitely be trained to detect it, which is proven

---

[18]The idea of *common practice* labels the compositive practice between 1600 and the
beginning of 1900, under the explicative apparatus of *tonality*. The first proposal of the concept
is dated back to Walter Piston's *Harmony*, 1941.

when these clues are absent, as it happens in most works of contemporary music in which the general audience often has trouble detecting the ending of a music piece and ends up relying on the lone applause of some brave expert, who takes charge in leading the rest of the audience, as a clue that the piece has actually come to an end.

It would nevertheless be incorrect to claim that tonal music and non tonal music do not share any compositional schemes or rhetorical devices that can serve as cues for listeners who want to detect, for example, the end of a piece. I just want to mention here two typical ending schemes: the "pompous" ending and the "vanishing" one. The first one is obtained by giving the last moment of the music piece a surplus of energy, so to speak, so that the attention of listeners is driven to what is happening as a special moment, that is the ending of a piece, especially if it is the ending of the last movement or, as we called it earlier, the last "music" segment of a larger work.

At the very ending of Ludwig van Beethoven's 3rd Symphony, for example, the final chord is repeated up to about one minute in some performances, a very long time even for a traditional finale, so that not only the ordinary flux of harmony is interrupted, but the piece itself ends up developing a sort of an overgrown limb. And something very similar also happens in the ending of the 8th Symphony. Such a bombastic character for a piece's ending is so typical that it earned the mockery of Erik Satie, who made a parody of it with the

"never-ending" ending of "De podophtalma", the last piece of the composition "Embryons Desséchés" for piano.

Other examples of this "pompous" character include the ending of Johannes Brahms' 2nd Symphony and Dmitri Shostakovich's 5th Symphony, but also Stockhausen's "Punkte" or Iannis Xenakis' "Akea".

Regarding the second example of an ending pattern, the "vanishing" one, we encounter the progressive diminishing of the sounding energy, which is typically realized by a general diminuendo and by slowing down the tone movements as a sensory-motor metaphor of a "dying" or "fading into memory/ dream" character. See for example the ending of Gustav Mahler's 4th Symphony or Brahms' 3rd Symphony, but also Claude Debussy's "Fêtes" and even, a very unusual ending for a Baroque music work, the fading out shout of the "dying" soprano "I Come to You Lord Jesus" ("Ja, komm, Herr Jesu, komm!") in J.S.Bach's cantata BVW 106 "Actus Tragicus".

Examples of the "vanishing" ending in the 20th century include György Ligetis's "Lontano" and Brian Ferneyhough's "Carceri d'Invenzione" for orchestra, but the very long list of fading out endings in pop music songs should also be considered as valid examples of this ending pattern.

What is important here is that what we consider as the appropriate elements to signal the beginning or the ending of a music piece are part of a complex structure, in which they work at different levels and on different categories of

151

auditory discrimination. Single notes in their duration and pitch, groups of notes such as chords, groups of groups of notes such as the succession of ornament patterns forming the melodic line, harmony as the choice of the proper pitches and durations within a range of well-formed successions of *interpretants*, together with consistency rules that drive the organization of successions in time, such as metrics, forms and the constrained liberties of linear development, which are related to styles, genres and ultimately cultural sounding practices in general, their social (and economical) environment included: these are but some of the elements that are involved in the detection of the ending of a music work such as the first movement of a Classic-era sonata for piano. It does not matter whether words taken from common speech, musical grammar or even semiotics are mixed together without discrimination of the level of skill of the listener.

Whatever degree of objectivity of the labels we use possess, the simple fact that we use terms to point to distinctions is proof—according to the way we are able to describe our experience of listening with language—that we are caught in a complex net of possible distinctions, even if we are just trying to understand that a piano sonata movement is finished.

Of course, whether we are able to detect a proper and consistent system in the relationships among the elements listed above is uncertain to say the least. First of all, because the ideas that are behind terms, such as "note", "appoggiatura", "chord", "harmony" and so on, are related to a specific music

grammar that has neither been universally present throughout the world, nor even in the culture that generated it, it has been stable in time.

In this respect, the epistemological condition of terms taken from musical grammar is not so different from the terms taken from common speech, as is demonstrated by the commonality of using the words "tonality" of "harmony" for works written in an age that didn't know those words, such as 1600 or 1700, at least as we use them today (Hyer, 2008). The contingency of music practices, those related to the didactics of music composition included, forces us to dismantle the illusion of being able to use the notions taken from of our musical grammar as a passepartout for dealing with the world's sounding practices, even if it is the music of our past, besides projecting our contingency onto them. Not only because this is what actually happens, for example in all the treatises that are devoted to the *common practice*, but also because all the available documents can only express the surface of the musical practice that necessarily omits what is left out by the purpose of the document.

Secondly, even if we will someday be able to establish neural markers for these elements, as we have seen, we will still be dealing with the attempt to make a connection between the labels of our music grammars and the raw data of our objective measurements without being able to find a reference between the discontinuity we perceive and express with labels and the continuity of physical/physiological measurements without ladening them with our theories on the world.

Regardless, it is important to realize that we can detect different types of ending or beginning music pieces simply because we are able to distinguish them as sections in the first place, due to the fact that we have found some qualities or characteristic successions of qualities that, being consistent within those sections, allow us to separate them as (simple or composite) units from comes before or after them in the "music".

5.5 Changes

At this point we are dealing with an entire musical piece, a perceptive "parenthesis", which is separated, or better said it is framed, by the "noise" of the sonic stream we are constantly immersed in and is well identified as a whole, which can be even labelled and, so to speak, thrown into linguistic presence by a title.

The possibility for that musical piece to be identified as a whole relies on some constant quality, both on the sonic side and on the causal side. But it has to be noted that this "causal side" of music is the same side we are used to relating to as "noise". The sound of a clarinet is in fact bounded both by the experience of seeing or at least conceding that a clarinet, somewhere, somehow, is playing and by the sonic event that shows a constant quality that we have learned to recognize as the sound of the clarinet. This is why on one hand we consider the sonic event as proceeding from a cause and, on the other hand, we recognize the sonic event as the sound of a clarinet, which is consistent with the cause. We can easily imagine some situations in which one of the two

conditions are not held and therefore our identification might fail. In fact, we might be unable to identify the source of a sound because we don't see it or we are not expecting it (as in the case of hidden loudspeakers, for example) or because there is no constancy in the quality of the sound stream. Let's take for example the case of Stockhausen's "Klavierstuck XV" for electronic keyboard and electronic sounds. As the electronic keyboard part presents a lot of changes in the sound colours, which also happens in the electronic part, there is no easy way to distinguish only by listening what the synthesizer plays and what is played by the background track. Nevertheless, we are still able to put music in the foreground against the background "noise".

But what if we imagine a paradoxical music work in which the performers appear to us as people from the audience, sitting amongst the audience and making the customary noises that people make while waiting for a concert to begin? We need to imagine, of course, that these performers are following very precise instructions about the sounds they are asked to produce, which are written in a score and are performed by heart.

In this extreme case we would not be able to identify either the source of the "musical" sounds, because we would probably mistake the performers for people from the audience, or the "music" as a foreground against the background "noise", because we would not be able to find any constant sonic qualities that we can rely on in order to distinguish "music" from "noise". We would not even be able to realize when the "music" begins and when it ends. Of course, the topic of distinguishing "music" from "noise" has been widely

explored by artists and musicians especially in the artistic/musical paths following John Cage's ideas, such as the Fluxus movement.

In order to examine this situation more clearly, we will consider each element on its own. First, let us consider only changes within the sonic stream, hypothetically removing any experience of the causal source. In doing this, we can establish that if we are able to identify a change in some qualities of the sonic stream that let us separate the background "noise" from a foreground, we have, so to speak, something to deal with: a target for our attention which is somehow bounded both in time, by its beginning and its end, and in the sonic texture. Whether we identify this foreground as "music" and not as a layer of "noise", as it would happen if, when waiting for an electroacoustic concert to begin, all of a sudden an electric-like buzz would resound in the hall, it is up to our expectations. In this case, we would probably attribute the sudden electric-like buzz to an electric system malfunction, especially if we are not directed by some non-sonic change in the surrounding environment, such as the fading out of the lights in the hall, that let us identify any new foreground sound as "music". We would therefore assume the presence of a source, the system malfunction, even if we cannot see it or do not know anything about it. What is important here is that if we are appropriately oriented by our expectations to identify that foreground as "music", a simple distinction within the sonic experience between foreground and background is enough for us to separate "music" from "noise".

Concerning the "music" piece, as a "musical" segment of the sonic continuum, I have ultimately related our ability to identify portions of that continuum, on one hand, to a change in the sonic qualities we rely on to distinguish "music" from "noise" and, on the other hand, to a change in some sonic qualities that were constant within the "musical" continuum.

Even if I have not shown in detail precisely what those constant sonic qualities are that allow us to easily identify a beginning or an ending segment, it is the very possibility of distinguishing parts in "music" segments without having to rely on causal sources that grants us the possibility to look for some changes in aural experiences that we could call, in the most abstract and general way, *qualities*, be they the experience qualities that we can easily relate to terms of our musical grammar or not.

For example, the distinction of the humorous long lasting ending of Satie's "De podophtalma" from what precedes it, is provided by a sudden change in some complex elements of the piano part, such as the sudden interruption of the previous melodic textures, the juxtaposition of short segments that are related to a single chord, the tonic chord, together with the presentation of rhythmic and pitch schemes that are typical of classical music endings.

While we do not have a proper word for naming such changing qualities as a whole yet, a more or less skilled listener is able to detect those changes. It is not important whether an inexperienced listener might be able detect changes in some qualities or not, but the simple possibility for of other listeners that

have been trained to a certain level being able to do it provides us with the opportunity to relate distinctions within a "music" continuum to the change in some, perhaps complex, quality, as the requirement of these distinctions.

Yet, there must be some cases in which the lack of this requirement leads to a failure in segmenting. Let's take the case of the "infamous" La Monte Young "Composition 1960 #7". The score simply shows a stave with a fifth interval (B3 - F#4) "to be held for a long time", which, in some cases, might be extended to one hour and even more (Alburger, 2003), to be performed by an open instrumentation. If some distinctions can be made within the fifth interval continuum, it is not the score requires some related actions by the performer.

We can imagine a performance of this work for a sine wave generator, as it happens for some of Young's later sound installations, so that the performance extemporaneities are minimized. What will happen to a motionless and quiet listener during a two-hour performance of that work cannot be precisely predicted. We can only predict that, if the electronic gear does not malfunction, it will not be possible to relate any listening distinction to any change in a certain quality within the sound stream provided by the performance, because there is no way to detect any change in it.

In other words: if the "music" continuum does not change in any way, we cannot attach to it a distinction in time, a segmentation. But what do we mean, when we say that the "music" continuum does not change? Is that even possible? From acoustics we know that sound is the result of vibrations, so

there *must* be something changing in order to have sounds and ultimately a "music" continuum. Moreover, even if we accept the idea of a constant pressure wave, like an ideal sine wave, this wave is realized by constantly moving air (or any other elastic medium) molecules, which are even formed by particles the are a never stable: there is no possibility of the sonic continuum not changing at all, not even a microscopic one. But the sonic continuum is not the "music" continuum, as the first one is connected with our measurements while the second one is a mental construction that is realized by our perceptive system.

So, nothing changes in the "music" continuum when we cannot detect any change in it by listening: when we have not had any listening experience of change within it. In spite of the appearances, this is not a tautology, however.

In fact, as we cannot make any objective statement about experiences, if somebody claimed to have had the listening experience of a certain change in "music" that we haven't had, it is not possible for us to verify whether the person lied, had an hallucination or is simply more sensitive than us.

When we want to share our distinctions, we necessarily have to deal with setting the conditions for our discourse to be a public one.

Of course, not everybody has the same sensitivity for changing qualities within the "music" continuum. Indian traditional musicians, for example, are trained to detect intervals of 20 cents (Boep, 1999), instead of the western music's semitones' 100 cents. It is therefore possible that an Indian listener might

perceive as different what, in certain contexts, is analysed as two instances of the same interval by a western listener. Yet there is no need to bring up the diversities of western and non-western music traditions: the difference between *minimal techno* and *tech-house* genres in electronic dance music, for example, can be very hard to detect for a listener who has only been trained, even if highly trained, in the classical music repertoire.

When we want to create a public discourse about how to detect when qualities change, we are therefore inevitably forced to deal with the diversity of people's sensitivities in detecting changes. This means that we always need to refer to a context, to a community of people who agree on what sets the threshold for deciding whether a quality in a "music" continuum is stable or not.

5.6 - Extreme music

We can therefore find a community who agrees that a two hour version of Young's "Composition 1960 #7" for sine waves presents no changes that can be related to distinctions within the "music" continuum. This is definitely such an extreme case of a non-segmentable music work that we can hardly find other compositions that present such a changeless continuity. It is not that it would be a difficult task to create them: any held tone played by an electronic instrument with the most stable sound would have the same result. But of course extreme cases are for extremist composers, who are extremely infrequent.

Let's take another one: Phill Niblock. In his album *YPGPN* (1970), he presents a series of works based on the highest continuity of sound. "A Trombone Piece*"* for tape, for example is a work that is based on the superimposition of tape recordings of a trombone that mostly plays the same note in different octaves and with different micro-tuning variations. Even if all the tracks are mostly realized by a continuum that is based on variations of a static tone, it is very difficult not to hear changes here and there, such as the attacks of the instruments, due to their dependence on breathing, or some transitions highlighting different sound qualities, a tone-beat for example. Nevertheless, it is difficult to place these changes within a hierarchy that allows the listeners to distinguish the "music" segment in, let us say, a before and an after, as a first and a second part of the work. Another example, even if somehow at the opposite side of the spectrum, is Olivier Messiaen's "Épôde" for 18 strings, a section of "Chronochromie" which is technically written as a sort of a 18-part fugue in the so-called *style-oiseaux*. Even if in this highly contrapuntal composition everything moves and transforms all the time, it is apparently very hard to detect an overall change in quality during its four minutes of, so to speak, sonic chaos. It sounds like 18 birds singing all together in such a complex dialogue that it seems to overcharge human listening, which is no longer able to distinguish anything but the rapid articulation of bird songs. We are dealing here with a "music" continuum that, in spite of having the possibility of detecting several changes and events in it, we are not able, at least not without serious effort and a sort of a re-education of listening, to

161

divide it in parts, by simply detecting global changes. We cannot rely on the changing of qualities to segment the "music" continuum and yet we might not give up and turn to solely relying on causal sources: we can still engage a sort of attentive fight directed to "elevate", so to speak, the perceptive foreground to the "music" domain against its stowing to the background "noise".

Sheet music for piano that was found in Erik Satie's house after his death provides us with a final example of extreme music. The textual notes of this work, entitled "Vexations", state that: "in order to play the theme 840 times in succession, it would be advisable to prepare oneself beforehand, and in the deepest silence, by serious immobilities" (Satie, 1893). Even if it is not clear whether the composition was really intended to include 840 repetitions of the same sheet, which by the way already consists of a double repetition, this work has slowly become a classic of experimental and, so to say, adventurous music, starting from its first performance by 12 interchanging pianists, in 1963. Even if the double structure of the single repetitions already confuses the listener during the performance of this work, which lasts about 20 hours, by listening it is indeed easy to detect the single repetitions, so that we can actually distinguish a number of sections corresponding to the single repetitions in the whole performance. On the other hand, it is not possible to group the repetitions in larger sections by detecting more global changes in the perceived sound without putting into play the succession of the performing pianists and therefore relying on the single pianists interpretation as a changing quality.

5.7 - Qualities

But by connecting distinctions in time with repetitions we are already forced to question the idea of quality itself. In Section 5.1, I first described the ability of detecting the beginning of a specific event such as a concert by only listening in terms of the detection of changes in the sound quality of the sonic stream.

In the course of this text I have connected qualities to the instantiation of a sound property with a sonic character that is constant within a certain threshold: for example, a constantly loud dynamics, but also the constant sound colour of an instrument, such as the piano, or the constant presence of a causal source, such as the crowd attending a concert.

Also, a constant diminuendo or the constant rise of a note scale seems to be able to give rise to distinctions as soon as they change movements or direction.

I have therefore shown how qualities have a complex constitution and can result in qualities that seem to show a more abstract character. For example, the ability to consider the presence of a chord as a constant quality in a succession of notes with irregular attacks, different durations and changing note registers implies the cognitive possibility to gestaltically attribute that quality to elements like the possible pauses between notes, which do not directly express the chord.

The same is true for the possibility that even a music style might give rise to distinctions whenever it changes in an explicit and striking manner. In fact, there are works that use differences in style to articulate form, such as Bach's

163

"Mass in B Minor" or a long list of Frank Zappa's songs (like "Beauty Knows no Pain" from the album *You are What You Is*).

Finally, I have shown that even the recognition of a "music" segment, a repetition, is able to provoke a distinction within the "music" continuum, to the point that repetition is present in western and non-western music tradition in an overwhelming way,[19] which seems to point us to repetition as a fundamental aspect of music.

But what could possibly be the sound property, the aural quality, that we detect as a change in the moment we recognize a repetition?

The idea of quality seems to be too vague to be usable outside a very general approach to the distinctive aspects of perception and cognition. But I introduced the word "quality" as a marker for distinctions in the aural stream, so that it could play an ambiguous role of directing the process of distinguishing to the changing properties of the sound flow in causal terms, as a naturalistic approach to the aural world made of (sound) objects that exhibit properties. At the same time, the use of the word "quality" as an explicative term supported a naive ontology of sound expressed by talking about listening experiences with the vocabulary of music grammars, such as dynamics, or general music related terms, such as style.

Nevertheless, if we must take literally the claim according to which we distinguish in time as a consequence of detecting sudden changes in certain

---

[19]For general survey of the "ubiquity of repetition" see Ockelford, 2005, pp. 1-6.

qualities, we also need to account for the ability of experiencing all the different qualities that might be simultaneously present in a same sound flow. But as qualities might consist at least in indefinitely complex abstractions of those sound elements that are usually considered at the base of sound perception, such as tone height, intensity and timbre, they are in an infinite number.

There is therefore no way to understand how cognition might be able to identify each one of them at the same time in order to be open to all possible changes in the sound flow.

This problem is apparently a consequence of an idea of cognition that is not constituted in time from the materiality of the body, but is rather defined by the reversible and infinitely articulated space of linguistic and logical syntax.

But "quality" is just a linguistic invention, an apparatus that is instrumental within a narrative, with a role that is neither ontological nor foundational. It is there because if we distinguish in time, there must be something before and something after both the actual moment we make that distinction and the moment we address as the point in the sound flow that we relate to as that distinction. As these two moments cannot be coincident and the moment we realize a change always follows the moment of the change, we need to relate distinctions not to a syntactical complication of parametric values that we relate to qualities, but rather to the articulation of time in perception, which may ultimately give rise to the ideas of syntax and parametric values. "Qualities" are there to work as placeholders that are projected as properties of an object and are asked to assume a causal role in driving our perception by the

sensorimotor narrative of a sound flow that is projected around us and is populated by objects that, somehow, oppose us.

We rather need to start from distinction in itself as a starting point of a narrative that is possibly able to provide a backdrop against which emerges not only the idea of an external word, but the very syntagm "external word" as a segment in the sound .

6. INTERLUDE: INSTANT AND DURATION

The French philosopher Gaston Bachelard, in his short but dense book *L'intuition de l'instant* develops a metaphysics of the instant in opposition to Bergson's idea of *durée réelle*.

The exposition of the arguments are conducted by following *Siloë*, a book written by the historian Gaston Roupnel, in which the author presents an idea of time according to which: "time has but a reality: the one of the Instant" (translated from: Bachelard, 1932, p.12).

For Bergson, the true reality of time is its duration and the instant is just an abstraction that comes from projecting time into space. It is the connection between two durations that we perceive spatially as the point that separates a "before" from an "after".

On the other hand, for Roupnel and therefore for Bachelard

> the true reality of time is the Instant; duration is nothing but a construction that doesn't have an absolute reality. It is created from outside, by memory, the power of imagination par excellence, that just aims to dream and live again, but doesn't want to understand. (translated from: Bachelard, 1932, p.21)

Both Bergson and Bachelard claim that time is created by an act of the subject that involves attention, but while for Bergson they are actions that extend between the subject's decision and a goal, for Bachelard, instants are the

consequence of acts that are unique sudden decisions, which carry the burden of originality.

It is not a matter of proposing theories that are complementary with each other so that the instant is just the negative of duration. Relativity forces us to give up on the presumption of being able to address the simultaneity of events that happen in two different locations in space, without the need to adapt the notion of simultaneity to a measure of the relationship between the instants of time that are connected to the occurring events.

Not only are instants the psychological markers of moments in time, as they are unique, each of them singularly constitute time, which is therefore discontinuous: "life is the discontinuity of acts" (translated from: Bachelard, 1932, p.19). On the other hand, the experience of instants, with their dimensionless indifference, is exposed to the mercy of the world because "time is nothing if nothing happens" (translated from: Bachelard, 1932, p.19). Therefore consciousness awaits for an "attack from the world" as an "abrupt change" in which a creative act is in operation, in order to be capable of articulating itself in memories that can only be connected to instants, and in an attention that is the continuous renewal of expectations.

In the interplay of memory and attention as grounded in the Instant, we find Bachelard's idea of habit as a quality of the instant that is able to link within each instant, memory and tension towards the future as the time-forming act, whose symptom is identity:

being is the place of resonance for the rhythms of instants and, as such, we can say that it has a past in the same way, so to say, a voice has an echo. But this past is nothing but a present habit and this being present of past is just a metaphor. (translated from: Bachelard, 1932, p.43)

By being rooted in the present instant, habits are unique acts that are on one hand connected in memory with secondary habits in *routines*, but, on the other hand, are always renewed as expectations, so that a habit can be defined as a "repetitive integration of novelties" (translated from: Bachelard, 1932, p.43).

We will later see how habit can degenerate, so to say, into habituation, so that "novelties" are no more expected as such, but disappear from consciousness and become the qualities of the objects in the world we live in. Bachelard addresses this by saying that in order to maintain its engagement, a habit constantly needs to be disregarded to some extent, so that the sense of novelty is renewed. An effective habit is therefore the sign of progress. But we are the variable sum of conscious habits:

we recognize ourselves in our character because we imitate ourselves and because our personality is therefore the habit of our own name. (translated from: Bachelard, 1932, p.64)

Therefore, the idea of progress becomes a moral tension that asks us to ceaselessly transform the copies of ourselves that we instantiate in order to maintain the possibility to act in the world because "what endures is always what is able to regenerate itself" (translated from: Bachelard, 1932, p.65).

169

Along this path of ideas, we can find the essence of most of the ideas that will be presented in the following sections.

Bachelard has proposed a metaphysics of the Instant as distinguished, but not unrelated, from the psychological sensation connected to a sudden act of consciousness. We have seen in the very beginning of this text, that distinction is always a narrative that can be decomposed by other narratives in an articulation of instants: the instant in which the discontinuity happens, the instant we perceive, for example, a wince,[20] the instant we become conscious of it and relate that wince to the instant in the measured time in which we still were not experiencing that wince, and so on.

On the other hand we can develop sensorimotor metaphors in analysing the idea of instants as edges, so that we can distinguish Bergson's instants in terms of borders from Bachelard's instants in terms of boundaries. In fact, according to Edward Casey, borderlines, with their precision and arbitrariness resemble Bergson's "rigidly delimited nows" (Casey, 2008, p.6), which are invisible and yet are able to separate extended surfaces. Boundaries, on the other hand, are permeable and allow trespassing; they are not abstract stipulated lines on a map, but are undetachable from their organic substrate. Specifically, Casey discusses three aspects of Bachelard's instant that he also finds in his vision of boundaries. *Novelty*, the first one, is related to the uniqueness of each instant in the similar way an organic boundary-like edge is never static but always altering, always unique; *commencement* is related to

---

[20]The experience of a wince, here and for the rest of text, is used as an exemplary reference to the phyisiological (couscious or unconscious) relatus of the cognitive act of distinguishing, as it will be explicitly stated in Section 7.1.

the idea that an instant is always giving rise to an event as the crossing the boundary is always an invitation to something new, in opposition with the border that discourages trespassing. Finally, *verticality,* as the surprise of the suddenness of the instant, what I called the "wince", reminds Casey of the vertical position of boundaries with respect to the flow that crosses them.

The author admits that the "parallel is not perfect" (2008, p.11), but for us it is enough to recognize the potential of a sensorimotor metaphor, that will prove useful in intuitively visualizing structures of distinctions and, in a special way, in the didactics of the pattern composition.[21]

As we have seen, it is not possible to address listening without dealing with the way we talk about it. The same is true for dealing with time as well as for dealing with all the topics that are, willing or not, necessarily involved in discourses that are entangled with the structure of the language they use.

Maybe the distinction between duration and instant is also related to the different positions regarding the subjecting consequences of using a language of nouns. Maybe a discourse about listening should be made through a language that, in the attempt to avoid the *paradigm of the object* that we met in Section 3.6, is able to bring back naturalized labels to their process of constitution.

As far as concerns this text, I will necessarily have to use the language I have at my disposal, with all the ambiguities of noun-based clauses and the

---

[21]See Section 11.2.

unavoidable hypostatizations, in expressing, similarly to Bachelard, an idea of

the instant directly in terms of an act of distinction that therefore does not have

to carry the ontological burden of a metaphysical object, but is rather a

generative element of a narrative that involves, as a consequence, the

constitution of time.

7. OUR LOST DIMENSIONS

7.1 Distinction

When we talk about listening, as well as when we talk about human cognition, we put ourselves in the position of impersonating, so to speak, both the researcher and the object of research, in a way that the language we that we use to talk about ourselves is involved as an effect of our object of research. This is why, according to Humberto Maturana, we cannot avoid relying on an epistemology of the observer in our investigation, as we have seen in the first part of this text. Observing, according to Maturana (1988), is an operation that takes place in language and therefore already relates to a community. And here language is the fundamental requirement of knowledge, hence all the concepts we use and in the end all we can talk about is based on language, as a social extension of the body. The body itself, as well as cognition, is the result of language rather than its condition. In fact, we need not forget that "cognition" is a word in the first place, and it is inserted in a pragmatics, which includes all the operations, the experiments and data collections as a reference of our "praxis of living" when we talk or write about cognition. As a consequence, observation cannot but be founded on those emerging oppositions, as we have seen, when using language as a simulacrum of presence, so that the world ends up with being populated by objectified *différances.* Observation acts therefore as a process for propagating and articulating differences in the hall of

mirrors of naturalized semantics: "the basic operation that an observer performs in the praxis of living is the operation of distinction" (Maturana, 1988, §6.2).

"Distinction" in Maturana is the process of separating a foreground (a unity) from a background (a medium) including those operations carried out by the observer while implementing that separation, which are not therefore independent from the acts of the observer. While this is already a complex activity, especially if mediated by language, when we turn to listening we can relate it to an even more basic activity, without risking, when it is brought to more complex frames, losing its constructive potentiality. In the previous chapter, I characterized distinction as the moment when we realize that something happened, a moment which is often accompanied by an instantaneous "wince", the behavioural mark of the very moment of distinction.

We can relate this wince to physiological evidence that seems to be connected to it, such as the *orienting response*, the basic reflex of the head turning in the direction of the sound when something new happens, as shown by all vertebrates, including young infants.

Other direct physiological relata include the bradycardiac changes in the heartbeat, the reduction of heart rate, whenever a stimulus change is detected, and the so-called *mismatch negativity*, a peak in a particular electrical activity of brain cells, namely the N2 *evoked response potential* (Huron, 2007, pp.49-

52). When an explicit reaction to surprise is brought into play, other possible relata show up, such as *piloerection*, increased arousal or breath holding (Huron, 2007, p.362). There is nevertheless a problem in linking the wince of distinction to physiological data, because after some repetitions of the new stimulus the listener will habituate to it and it will not be possible again to find a change in the previously listed physiological relata. Consequently, it is possible to relate the physiological data to the perception of some degree of novelty. It seems we can therefore interpret the detected change as a violation of an expectation and that the sudden wince becomes the sign of surprise.

Let's go back to addressing what we are used to dealing with as a property of the world, in the shape of the sound flow, that is the external correlate of that wince. So far, I have used for it the most general word "quality" and I have said that the wince is related to the experience of change in the quality of experience. As I explained at the end of last chapter, we do not have a way to identify if the old quality was substituted by a new quality, because up to now we have not involved any ability to distinguish a single quality in the sound flow among the infinite numbers of qualities: the continuum of qualities, as we could call it, is related to any single moment of our experience of the sound flow. Yet, if what we really perceive, as far as we can say up to now, is just a change in a quality, without being able to detect a possible direction of that change, that change is experienced as a lack, affecting the experience of the sound flow after the change. In fact, even if this "contrasting phenomenon" seems to be

observed only in extremely simple conditions, it might be taken as the sign of what we experience changes in terms of differences, instead of absolute qualities:

> Listening to a stimulus with a particular spectral structure and then switching rapidly to a stimulus with a flat spectrum, such as white noise, may obtain a powerful demonstration of this effect. A white noise heard in isolation may be described as 'colourless'; it has no pitch and has a neutral sort of timbre. However, when a white noise follows immediately after a stimulus with spectral structure, the noise sounds 'coloured'. The coloration corresponds to the inverse of the spectrum of the preceding sound. (Moore, 1997, p.255)

In the previous quote, the terms of "coloured" and "colourless" express different qualities of "noise" and its "colour", which are characterized in terms of physical measurements, as sound spectra. A consequence of characterizing qualities with a dimensional metric of measurements, the sound spectrum, a quality ends up being identified within the continuum of possible measurements/qualities as a set of parameters, which is given the role of a hermeneutic probe within the sound flow. On the other hand, the description of the experience is made in evocative terms, which are related to the experience of qualities in other sensorial modes, such as with "colour", but it could also be made with an extended description or even with a label invented on purpose. So, the experience of a change can be associated with any kind of description that we are able to assemble, or better: if we can describe a change perceived in whatever way we do it, then that description has the right to be listed as a perceptive quality.

How many ways do we know in which we can describe a change within a music work, within a sound flow in the most general sense, which we can detect by simply listening to it with the full burden of our past experiences? Of course, even if it is not an infinite quantity, they are still in a very large number, because it is always possible to invent new ways to describe them. I mention here a few examples, starting from a very common one, at least for a skilled musician, and ending with one of the most abstract and blurred: pitch, harshness of timbre, position (in a stereo panning, for example), consonance, rhythmicity, tonality, pleasantness, deepness (a quality that is often used in some sub-genres of electronic dance music), presence of a voice, of an instrument, of a chord, the entrance of a theme, the genre. If we can say that a music has suddenly changed, let us say, in its genre, we can consider that it has a quality that, under certain conditions related to that quality, allows us to distinguish the moment of change within the sound flow. Nevertheless, qualities show different relationships with each other not only because their boundaries are always blurred and there might be some overlapping amongst them, but also because the way we talk about certain qualities occasionally connects them to other qualities, to which they relate by combination, inclusively or in other kinds of relations. For example, sound colour overlaps with pitch because the former influences the way we detect the latter and the lower the pitch, the higher the possibility for sound colour to be rich of partials. On the other hand, the case of harmony, as a way to deal with pitches, and

177

genre, as a possible combination of pitch, harmony, sound colour and so on, shows the high degree of complexity of possible relations among qualities.

7.2 - Dimensions

In her seminal article "Time, Our Lost Dimension", the cognitive psychologist Mari-Riess Jones proposes a way to deal with qualities from a parametrical point of view. Even if it is not the first or the last time that such a proposal has been made, especially in cognitive psychology and musical analysis, Jones' approach is particularly important for my research and therefore I will especially follow her article in the next sections. In fact, even if her text was intended to account for cognitive problems in the serial ordering of events and in particular to the psychological phenomenon of *streaming*, I have found her terminology so extremely useful in dealing with the distinctions of listening, that I have not only been able to apply them outside of the field of music psychology, in music composition and analysis, as I will show in Part Three, but it also seems to be promising in dealing with more general topics related to cognition in the context of society.

Jones' primary concern is to establish a shareable representation of the experiences of articulated changes in qualities, which she projects as properties of the world in terms of "world patterns" instantiating a physicalist "world structure":

this approach rests on the assumptions that world structure is built within the constraints of three dimensions of space and one of time and that humans lawfully reflect world structure. In particular our auditory environment is assumed to be subjectively represented by pattern relations along dimensions of pitch, loudness and time. (Jones, 1976, p.327)

Jones does not explicitly define the term "dimension", but it seems obvious from the context that she refers to its physicalist definition in terms of a coordinate within a mathematical space.[22] In this text the idea of "dimension" is simplified from its mathematical abstractions to simply intuitively refer to a parameter, as in the "height" or "duration" or even "pitch" and "mel" (or perceived ton "height") dimensions. A further development of the notion of dimension extends beyond the scope of the present text, but it is enough to say that a parameter is related to an "ordering structure" (Ito, 1993, p.1168) so that, for example, any set of values in the dimension of "height" can be ordered from the lowest to the highest one.

After having stated her assumptions, Jones immediately proceeds in defining a first set of assumptions that are the consequences of her physicalist approach. First of all, she defines "world patterns" as the "invariant relations along changing physical dimensions" (Jones, 1976, p.328), then she connects "world patterns" to the "relations of a finite number of subjective dimensions", so that it is possible to realize an isomorphism between the multidimensional physical space and a multidimensional subjective space. This is not too different from

---

[22]For a complete mathematical discussion on the concept of "dimension" see (Ito, 1993, p.448).

Oackelford's definition of *perspects*, the contraction of "perceived aspects", as "qualities of the musical fabric as we apprehend them" (Oackelford, 2005, p.10). After all, both Jones and Oackelford are talking about parameters that are in a, not necessarily one-to-one, correspondence with the basic acoustic parameters, such as frequency and amplitude, which in turn are theoretically laden by our sounding-aural social practices. The simplicity of these acoustic dimensions is only apparent though, as they have to be expressed as statistical abstractions ruled by thresholds in order to be applied to real world phenomena, and we perceive their subjective correlate only because our listening process is trained to detect them by simply ignoring the unfitting parts of experience. *Frequency*, for example, can be defined in absolute terms only if there is a perfect replication of a pressure oscillation, which is impossible by definition in real life (although it can be reached with digital instruments with a good but nevertheless approximation). In the same way, *phase* can be defined as an absolute value only in presence of a perfect and impossible replication of the oscillation period. It is true that *phase*, in spite of its basic character, is a sort of neglected parameter so that neither Jones nor Oackelford seem to be interested in it, because phase cannot easily be connected to any related subjective dimension (see Oackelford, 2005, pp.10-11). Yet the difference in phase has an important role, for example, in detecting the position in space of a sound source, which on the other hand is a subjective dimension. Sound spectrum as well, which was used as a acoustic evidence for detecting the "contrasting phenomenon", is just one of the many possible analytical

representations of the pressure-wave flow and, as being based on the unrealistic assumption that a pressure wave repeats itself in an precisely identical way, it needs some conceptual adjustments in order to be useful.[23]

Jones and Oackelford ultimately direct their analysis essentially to pitch, as the basic subjective dimension we have developed a strong sensitivity for, in western music practice, as opposed, for example to the strong sensitivity for timbre that bases our comprehension of natural languages. To be precise, in psychoacoustic terms, the subjective scale of tone height is usually expressed in *mel*, while the subjective experience of amplitude is usually expressed in the *phon* scale, nevertheless the dimension of pitch as investigated by the two researchers is quite surprisingly arranged according to the western musical grammar. In Jones' text, this is motivated by following Shephard's claim that his model (Shepard, 1982), which distinguishes between tone height from chroma, is more appropriate for analysing patterns structures. Twenty-nine years later, Oackelford relates the choice of using the pitch-scale instead of *mels* to Gilles Fauconnier's idea of "mental space" (Fauconniers, 1985), in the definition by George Lakoff as "a 'medium for conceptualization and thought', capable of representing 'any fixed or ongoing state of affairs'" (Lakoff, 1987; cited in Oackelford, 2005, p.12). The idea is that any descriptive structure of a conventional system of parameters/notations, such as the traditional western music grammar is suitable to work as a system of coordinates, on which it is

---

[23]See Huron, 2007, p.102 and Roads, 1996, pp.1073 and following.

possible to build an algebra, for the simple fact that it is accepted and used by a community.

I am now turning back to Jones, she has completed the passage from quality to quantity, from experience to the abstraction of numbers, while on the other, she has prepared the ground for detecting patterns, in the algebraic context of *symmetry*:

> relations that define subjective pattern structure can be represented by the mathematics of group of symmetry applied to the subjective dimensional representations. (Jones, 1976, p.328)

The path from the indescribable experience of qualities to the multidimensional set of parameters is a long one, from which there is no easy turning back: everything we listen to can be referred to the instantiations of a subjective dimension which is the result of some combination of more basic, still subjective, dimensions and is therefore the, again subjective, representation of their correlates in the multidimensional set of physical, objective, parameters. It is a long path because we have given up the constructive power of perception to the properties of the world as we find them as ready-made in our training process of listening. There is no easy turning back because in the meanwhile we have forgotten how we would listen before training, for example how we would listen to our mother tongue before we learned it, so that the effort to "listen differently and creatively" is harder and harder the older we get.

Nevertheless we can now characterize our distinctions as experiences of discontinuities along some perceptive scale, which can be articulated at will.

For example, Stephen McAdams's now classical proposal of a "timbre space", a multidimensional representation of the experience of timbre, includes the three dimensional components of "attack time", "spectral centroid" and "spectral flux". These components are already complex, so to speak, arbitrary hermeneutic dimensions and are related to the physical correlates in the complex multi-dimensional analysis of sound spectra (McAdams, 1999). "Timbre space" is therefore the example of a dimension that is analytically presented as a multi-dimensional space. But this is only one of the possibilities of articulating complex dimensions.

Mari-Riess Jones is interested in setting a proper numerical representation that is adequate to the complex relations among notes in the western music tradition, consequently her dimensions are *totally ordered*. This means that any couple of subsequent pitches can be ordered and consequently their interval can be described as ascending, descending or as a unison. In other words: if we consider two instances of that subjective dimension that we perceive as different from each other, we can always say if one instance is in a lower place than the other in the scale defining that dimension. However, if we take the list of subjective qualities I presented earlier, we can see that this is not always the case. For example, we can always say if the degree of rhythmicity, which is a quality that can indeed be defined and experienced in very different subjective ways, is different between two music excerpts, but it is not always easy to distinguish which one is more rhythmical than the other one. Nevertheless, the

locution "more rhythmical" is far from meaningless, because most of us can easily imagine two different music pieces or two different sounding events, the second of which is evidently more rhythmical than the first one.

The dimensions of style and genre, I am using these terms in their most general sense here, are even more puzzling as there are so many ways that a music piece can be perceived apart from the style, let us say, of XVIII Century *galant music* and therefore setting a degree of distance from *galant music* might seem pointless. On the other hand, if we are properly experienced, we are definitely capable of putting in a scale, according to the dimension of *galant music* style, a work belonging to the *galant music* style, a work which shares some similarities with that style let as say a piece of Beethoven's first period, and, for example, a contemporary music work such as Iannis Xenakis' "Achorripsis". It seems therefore that there might be conditions to be met in order to compare any two instances of some perceptive dimensions, which in this case are therefore only partially ordered ones.

I have to add that even if I did not define "style", nor I have presented a way to connect this dimension to the articulation of more basic subjective and physical dimensions, I can anyway say that style, however it is defined, must be at least phenomenally connected to sets of instances in dimensions such as pitch, harmony and timbre, and therefore seems to be a sort of dimension of dimensions. This is indeed the example of a maximally complex dimension that is different from the case of "timbre space", as it is not an explicit

multidimensional space, but is rather based on the presence of instances of some kind: chords successions, instruments, melodic or accompaniment types, that we could describe, in our terms, as domains of similarities collecting articulated instances in different dimensions. Let's keep this concept at a level of intuition for the moment and see what happens if we take away the reference to a particular style from the dimension of "*style*".

In this case, we can still discern if a style has drastically changed over time, but without a reference to a particular style we cannot set any scale. Nevertheless, we can still compare, under certain conditions, the "style distance" between two successions of music segments. In other words: in the example just above we can still say that the similarity in style between the *galant music* piece and Beethoven's piece, is greater than the similarity between Beethoven's piece and Xenakis' piece.

As the differences between the "style distances", the intervals of that dimension, are here macroscopic, we could quite easily think the "style" dimension to be an ordered one, but instead it is the differential dimension of "style distances" that can be organized in a scale ranging from couples of works that are most similar in style to couples of works that are completely different in style from each other. This dimension, which we could call the "style interval" dimension, is partially ordered and can therefore be used in building configurations or, in other words, successive discontinuities in that dimension, as it happens for example in several of Frank Zappa's songs, such as the already mentioned "Beauty Knows No Pain".

Let's take now another example, Luciano Berio's 3rd movement of "Sinfonia", which is based on a long succession of music quotations from very different moments of the history of the classical western tradition. In spite of the variety of styles and the speed change, it is often possible to detect the moments that the music changes along the "style intervals" dimension. But, we can more easily distinguish the repeated instances of the same work (for example the leading sections related to Gustav Mahler's 2nd Symphony), as far as our skills in the classical repertoire and our memory allow it, from the presence of music quotations that are related to different styles. Instead of the articulated dimension of "style interval" we might prefer to focus on particular works that we recognized, as they seem to vanish and appear again, lost in a music continuum. In dealing with the recurring appearance of an instance of a very complex dimension such as "style", we seem to turn to relying on a scale that is ultimately based only on two values: identical and different. This means that a reference style is posed every time we recognize an instance of it, as long as we are able to keep it in our memory. On the other hand, identity is dynamically defined according to the "style intervals" which, if they are wide enough, let us distinguish in the music flow the status of identical and different. Let us say, for example, that a music piece is made by quotes of Mozart, Beethoven, Xenakis, Haydn and the pointillistic Stockhausen. Most listeners will probably group Mozart, Beethoven and Haydn together as "identical" styles against the "different" Xenakis and Stockhausen, which will be nevertheless considered as expressing an identical style of their own.

Another example of this kind is the detection of a change, provided by the entrance of, let us say, an oboe. Let's imagine a music piece for electronic sounds and oboe. Let us say the electronic part is simply based on a constantly rumbling low frequency noise and after one minute of solo electronic sounds, the oboe starts playing a one minute part, of any kind, but in a way that it is clearly possible to distinguish the oboe from the electronic sounds by listening. The oboe part is then followed by another minute of solo electronic rumble and again this is followed by another entrance of the oboe that lasts one minute. Then the piece, which is four minutes long, ends. Of course, we are definitely able to detect changes in the music flow, which corresponds to the oboe's first entrance, the solo electronic part in the middle and the second entrance of the oboe. The particular organization of the work has also made clear to us that it has been possible to set a binary dimension expressed by the presence of the oboe, together with the binary dimension expressed by the presence of solo electronic music, which are perfectly complementary with each other so that we can also detect a repetition of the configuration realized by a section with solo electronics followed by the entrance of the oboe.

In the end, the idea that every distinction we make in listening to the sound flow, that we are also able to label somehow, has the right to enter the "realm" of perceptive dimensions has put us in contact with a variety of dimensional

structures, from the simple mono-dimensional to the multidimensional ones, from the complex to the binary ones.

Nevertheless, there might be a legitimate suspicion that this sort of explosion of dimensions of very different types, that can be dynamically changing and have the potentiality to be complicated at will, might be emerging from a description that is missing some more basic cognitive activity, that involves time.

Moreover, *difference* has assumed roles that vary according to the dimensions. In the totally ordered dimension of pitch, for example, differences express themselves in the symmetries of the repetitions of pitch configurations, that we recognize even if they are transposed and therefore presented with different values of the dimensional scale. Other dimensions, such as the one involved by the distinctions within a succession of rhythmic patterns, put differences in a simple binary opposition with some recognition. The *style* dimension seems therefore to be on one hand the result of complex components that are, more or less, constant in time, such as the presence of piano in a typical romantic *Lied*, but on the other hand, it also seems to be connected with abstract properties of configurations, as it happens with melodic and chordal schemes, for example, that need to be properly detected within the music flow because they are part of a style, such as an accompanying pattern, or a specific harmonic cadence. And this complex articulation of conditions, the range of possibilities that define a style, together with the virtually infinite possibilities of styles instances, is what impedes us to set a proper scale that can be put in

relation with numbers. In the end, we could consider a scale as a well-ordered collection of differences that have undergone the listener's training in the recognition of identities and in the ability to discard what could not be recognized by downgrading it to the role of simple variations (to be avoided or not) of the recognized identities.

Jones' physicalism binds her to space-time, so, as far as it concerns the description of the world, time is a physical dimension among the basic four that should not be denied. The article itself, already in its title "Time, our lost dimension", is mainly concerned with giving time its position back inside the parametric description of sounding phenomena and establishing its primary role in determining the perception of sound configurations. In order to integrate the physical dimension of time in a descriptive system that is able to account for the symmetries we detect in the sound flow, Jones introduces a time dimension that, in analogy with the pitch dimension, is based on time intervals arranged in an algebraic group. In this way, she is able to describe a configuration of pitches in time, in terms of their differential dimensions as a succession of pitch and time intervals. And consequently, the configuration does not lose its identity when pitches and/or durations are transposed or, in other terms, multiplied by a constant. While in 1976 the application of group theory to durations was a new experimental approach, Oackelford, in 2005, is already taking for granted that traditional music notation of rhythm is suitable to represent repetitions. This is not surprising considering that Jones is dealing

with the construction of controlled stimulus for her experiment in perception while Oackelford seems to be primarily interested in the analysis of music from the western written tradition, and therefore he approaches listening in the typically ambiguous way described by Fiske's *copy paradigm*, as we have seen in the first part.

Nevertheless, the status of time as dimension needs to be discussed. In fact, time onsets are instances of the physical time dimension, which relate to a reference value, the time-axis origin. On the other hand, the measurement of time intervals, or *inter-onset intervals*, is still derived from the physical time scale as an immediate consequence of its group structure. However, it is definitely possible to follow the idea of *mental spaces* that Oackelford invoked for pitches and to choose a metric structure over the time-scale in seconds. Moreover, it is usually possible, under certain threshold conditions, to detect symmetries in time, such as recognizing a rhythmic pattern when it is played at different speed, so putting a configuration in relation with instances of time intervals results in maintaining some symmetries. However, if a perceptive dimension is a scale that allows connecting to a perceived change in some quality as a point in the scale, the timeline in seconds is not a perceptive dimension, but rather that very quantified scale based on a spatial metaphor to which Bergson opposes a living experience of time as *duration*.

On the other hand, the differential dimension of "time intervals", especially if it is expressed in units of pulse as in the traditional western music grammar, can

be considered a perceptive dimension, as it is possible to distinguish within the sound flow by detecting a change in the time interval of some sounding event.

This might be confusing because a time interval is nothing per se, if nothing is happening to mark the beginning and the end of the interval: an instance of time interval needs to be associated with a change in some other qualities to be perceived as such. But the same is true for pitch, as a pitch that does not have dynamics or sound colour simply cannot exist. The same can be said for all the possible perceptive dimensions. In fact, we must remember that perceptive dimensions are hermeneutic probes, so to speak, that are launched in the continuum of qualities to select qualities we can talk about when describing our experience of listening.

So, if a change in a given quality cannot be separated from the time it happens, then time also cannot be separated from a change in certain qualities.

Let's then imagine a series of regular slow pulses of a sonorous click. Suddenly, the click stumbles. But it immediately resumes its regular pulse. This is an example of what could elicit a distinction in the sound flow, which is related to a change in the time interval dimension. But, as we have seen in the previous chapter, a duration is detected because something is always happening in the process of our enacting cognition, be it addressed as an event of the objectified "external world" or as a twitch of our cognition in operation.

7.3 Regions

We cannot disjoint time from a world event that happens in time: this is the most important idea fostered by Jones' article. In fact, she shows that configurations in some dimension, such as pitch, are experienced in different ways depending on their speed.

What the psychologist is interested in is the experience of a configuration in some perceptive dimension in terms of the recognition of its serial structure: the order followed by the instances of that perceptive dimension while realizing that configuration in sound. The eventuality that under certain conditions we are unable to preserve the serial structure of a configuration by listening, is exemplified by the *streaming* effect, a perceptive phenomenon according to which we experience, for example, a succession of sonic events, by grouping them not "according to their physical temporal order, but according to their attributes, e.g. their pitches" (Moore, 1997, p.260). We can easily find a typical expression of *streaming* in scores that use tremolo chords, the reiterated fast succession of two or more notes, which are not intended to express a fast repeating melodic figure, but rather the chord that results from the overlapping notes, together with a trembling sound.

As Jones' main purpose is to investigate the criteria underlying the recognition of a serial structure, she starts by pointing out that there are cases, as in *streaming*, in which the serial structure does not coincide with the listener's percept.

Her system of physical and perceptive dimensions provides her, as we have

seen, with the possibility of describing a succession of two or more events in

terms of pairs of values that reference instances of two differential dimensions:

the interval between the values of a perceptive dimension that are instantiated

by the two events and the interval of time they occur in. The pairs of intervals

can therefore be plotted on a two-dimensional space defined by a horizontal

axis referencing time intervals and a vertical axis for intervals of the perceptive

dimension. At this point, Jones proposes that the space of interval pairs is able

to account for the different ways that we deal with serial structures, and

therefore traces two lines on the graph plan that divide it as thresholds in three

regions, related to three subjective modes of distinction, as three maximal

perceptive categories (Jones 1976, p.343).
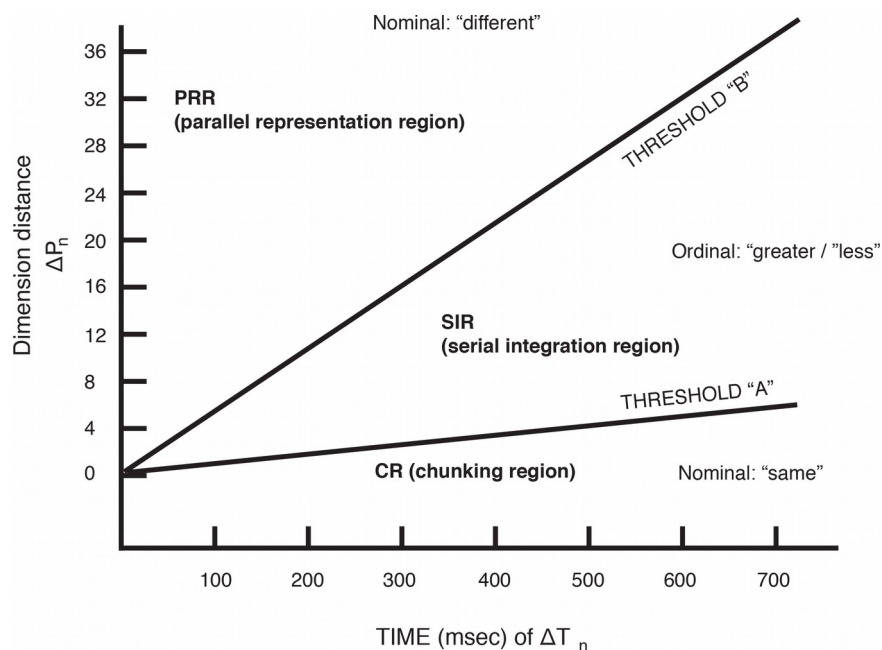
Figure 2: regional modes of distinction, reproduced from a graph in Jones, 1976.

The central region is the *Serial Integration Region* (SIR), the area including those intervals pairs that allow the listener to detect and compare changes in the perceptive dimension, in such a way that he or she is able to recognize their position within a succession. In other words, this region is connected to the possibility for the listener to detect the changes in terms of intervals—in terms of recognizable identities—so that he or she can memorize them in the order of their occurrences and possibly recognize that order in another configuration.

Above the SIR, there is the *Parallel Representation Region* (PRR). This region is bounded downwards by a threshold plotting a line that moves from a dimensional interval of zero and a very short interval of time towards wide intervals of both time and perceptive dimension. When the point defined by the pair that is related to two successive events falls in this region, the listener is no longer able to recognize the order in which the events occur and they will be therefore perceived as overlapping. This is the region in charge of *streaming*, among other perceptive phenomena, and is therefore connected to the perceptual category of *diversity* or, in a wide sense, polyphony.

Below the SIR and consequently below the lower threshold, which is again expressed by a raising line with a lower slope than the higher threshold, we have the third region.

This is the *Chunking Region* (CR), in which the different instances are "no more treated as events to be serially related, but as unified 'chunks'" (Jones 1976, p.343). When events fall into this area we cannot really detect changes, but

rather perceive them as the articulation of a single instance. This is why Jones refers this region to the subjective mode of distinction, which addresses different events to the "same" category.

So far, I have related distinction to the detection of a change in some quality. Jones' graph provides an analytical framework that defines "quality" in terms of a perceptive dimension. It relates changes to an interval of time between two consecutive instances of that dimension and it articulates distinction in three modes that depend on time and dimensional intervals. On the other hand, she aims to provide a framework to establish the parameters for a correct experimentation.

Even if Jones' theoretical proposal seems to be open to applications in different contexts, such as in vision perception (Skelly, Jones, Goodyear and Roe, 2003), her purpose primarily lies within the disciplinary boundaries of the cognitive psychology of music.

The present text, on the other hand aims to find a viable narrative that can address a variety of topics, as the third part will prove. Therefore, it is not important here to provide a precise parameterization of dimensions, regions and thresholds. What is essential here is rather the possibility of distinguishing thresholds within the continuum of objective data, a possibility that can be related to the constitution of human experience.

In this respect, it is possible to extend Jones' modes of distinction to three paradigmatic directions for constituting the perceived world.

The CR can therefore not only be considered a sort of blindness in terms of detecting changes, but is also a condition of distinguishing composite units as "chunks", in which the distinctions we cannot make become the textural properties of the perceived continuum.

On the other hand, the SIR accounts for the possibility of articulating chunks in figures: it expresses the possibility of recognizing and therefore is fundamental to the process of dealing with a dynamic world.

Our perception seeks to find changes that we can use in a meaningful way to act in the world, and these changes are made by events that fall in the SIR. But we cannot use the SIR events if we cannot distinguish the inherent differences among them: this is the role of the PRR.

The parallel representation is a spatial metaphor to express the perception of differences, the discontinuities that break chunks into pieces. Jones's PRR is at the core of the perception of simultaneity and acts by misplacing events that from an objective-scientific side are considered as consecutive in time, so that an objective articulated stream is broken into two, or more, simultaneous streams. And these, so to speak, hallucinated streams, which are the product of an illusion of listening, appear to "behave" as single streams that are independent from each other up to the point that they can be related to subsets of their objective measurements as a whole. They can therefore fall in the CR or in the SIR, which appear as nested within the PRR that stimulated their distinction as single streams.

In fact, we could even consider, for example, *virtual polyphony*, a composition technique which is directed to realize contrapuntal works for instruments that are not capable of substantial polyphony, such as the first Bourrée in J.S. Bach's cello Suite n. 3, as a system of streams, which are nested in a PRR and falling in the SIR. Of course, in Bach the single voices are too slow to undergo the *streaming* illusion and therefore cannot be perceived as simultaneous. Nevertheless, the use of terms such as *fugue*, which is typical of polyphonic textures, or the detection of chords, both in relation to analysis and composition, when dealing with works that are incapable of emitting more than one sound at a time, testify to the presence of some parallel representation in action.

Moreover, I am suggesting that the PRR can also be appealed to account for the experience of simultaneity not only between two objectively successive events, but also between two events that objective measurements consider as simultaneous. The PRR is therefore possibly gathering the conditions for any group of events to be perceived as simultaneous. The cognitive faculty of distinguishing simultaneous events is what in psychoacoustic terms is called *segregation*, although some authors call it "perceptual grouping", which in my opinion may give rise to a confusing overlap with the different phenomenon of grouping in time and segmenting, or "auditory scene analysis" (Bregman, 1990).

In order to explain how we can distinguish different simultaneous layers of sound only by listening, psychoacoustic research has detected several perceptive mechanisms, which operate by detecting conflicting fundamental frequencies, onset disparities, the spatial location of sources, contrasts with the previous perceived qualities and correlated changes in amplitude or frequency (Moore, 1997, pp.249-259). All the above criteria rely on the detection of a timing contrast between two streams, which happens at different time spans, such as the micro-time of waveforms in the phase locking of harmonics of fundamental frequency, the interaural time differences we use to determine positions in space and the conscious time span related to changes in sound qualities. Even if the above-mentioned mechanisms might also involve processes that are not explicitly related to the detection of time differences, as it happens for the "contrasting effect", this seems to me enough to experimentally consider segregation as the effect of some perceptive dimension that falls in the PRR.

I need to recall here that even if the inner ear works, in a huge simplification, as a sort of spectrum analyser, by translating the mechanical oscillations of the tympanic membrane into the spectral components, as the result of the organ of Corti in the cochlea, and therefore apparently losing in translation the original format of oscillations in the continuity of space, that information is still present in the neural decoding of phase-locked firing rates of the inner hair cells (Moore, 1997, pp.38-40). The role of phase locking is not yet completely understood and this research is not directly interested in looking at

198

physiological data for correlates of listening as a cognitive act, yet it seems

that, even according to the analysis of the auditory system, we can expect that

changes in very fast timespans, up to about 4-5 kHz, are used to distinguish

simultaneous streams in the sound flow.

It is also possible to interpret the relationship among the three perceptive

modes by following another important perceptive phenomenon: the dialectics

between figure and ground. In fact, as

> It seems that we are not generally capable of consciously attending to
> every aspect of the auditory input (…) the complex sound is analysed into
> streams and we attend primarily to one stream at a time. (Moore, 1997,
> p.268)

In this respect, *streaming* or segregation in general is the perceptive device

that separates a figure from a ground, so that in a given instant we attend to

one of the two or more streams, resulting from the parallel representation, as

the foreground against the background of the remaining streams. While the

selection of the stream that is in charge of bringing out the "figure" to attention

might also depend on the conscious aspects of the listening context, the PRR

is definitely a primary condition of being able to distinguish a figure from the

ground in the sound flow. We have already seen how the SIR and the CR can

be nested within a PRR, it is therefore quite an obvious consequence that

relates, in a given instant, the SIR to the figure, which is defined by the ordered

values of some perceptive dimension, and the CR to the ground, which is

identified as a whole, a "chunk", in opposition to the layers of the sound flow that fall in the SIR. In this respect, perception as a dynamic process of isolating a foreground from the background is the never-ceasing effort of breaking the sound flow by looking into the possible events falling in the PRR for the perceptive dimensions that fall in the SIR, so that all the remaining changes fall in the CR. What I have previously said about *virtual polyphony* is therefore not totally correct because in order to be compatible with the "figure-ground" principle, the voices need moment by moment to be arranged by separating a focus voice, which needs to be put in the SIR, from the rest of the voices, which need to be put in the CR. And this is why *virtual polyphony* works, as the monophonic instruments are allowed to leave out most of the CR parts, which are then simply alluded, and commit themselves to the SIR parts, so that the score essentially provides a figure, which is always related to different voices, while omitting the ground.

7.4 Expectations

Let's turn to the events that fall in the SIR and follow their definition to its consequences. As above mentioned, events in the SIR can be ordered according to the intervals in the perceptive dimension that define them, in such a way that their order can be recognized in other instances. Already starting from its title, "Toward a New Theory of Perception, Attention and Memory", Jones' article aims to frame perception within a wider context of cognition,

which includes, besides memory and attention, perceptual learning and issues related to age and illness (Jones, 1976, pp.347-352).

In fact, any theory of perception cannot avoid framing the analysed topic within the context of cognition, which in turn has to be related to the context of social and cultural practices, in order to be able to account for the wide range of related phenomena. As we have seen in the first part, it is not possible to avoid, for example, the topic of speech perception when dealing with listening, by simply stating different listening modes described in terms of their objects, instead of proposing how the process according to which the different modes are constituted in listening.

On the other hand, any experiment about listening cannot avoid dealing with human beings as living systems, with all the details of cognition as part of their "praxis of living". So, from the very beginning, Jones' proposal is explicitly contextualized within "the research with auditory patterns" (Jones, 1976, pp.323), but the theoretical background itself is also laden with the eminent role of patterns in our linguistic operations, scientific experimentation and mathematics. This is not a trivial claim, as in fact Jones' article seeks to find patterns in pattern finding, but, on the other hand how is it possible to deal with patterns without assuming their existence? Jones resolves this circularity, so to speak, by only addressing the research to auditory patterns and by characterizing them in terms of the possibility of recognizing the order of its components. In my research, which does not claim any scientific status and therefore does not need to rely on an epistemology of measurements,

circularity is resolved by setting a narrative, which even if it still needs to show

its *viability*, cannot avoid dealing with its constructive status, which results from

its birth in language and from its aim of addressing cognition.

I can therefore specify the idea of "configuration", which has been defined thus

far as a succession in time of intervals, or discontinuities, in some perceptive

dimension, in terms of the result of a complex cognitive activity involving

memory and attention. As far as it concerns memory, the ability to determine

the order of intervals within a configuration implies the ability to recognize their

position on a scale. If "the concept of order embodies our discovery of pattern

and regularity in the world" (Jenkins, 1968, p.429), then recognition already

happens within the horizon of cognition by constituting perceptive dimensions.

I will outline the consequences of this later in the text; however, for the

moment, it is enough to say that intervals, and eventually the recognized order

as well, need to be part of the past experience of the perceiver and

consequently need to be instantiated in some sort of time-lasting reservoir,

which we usually think of as memory. On the other hand, attention is also

connected to the idea of perceptive dimension as cognitive faculty that is in

charge of selecting that dimension within the continuum of sound flow

qualities. We have therefore met, if not a cause, at least an actor of selection: a

label that addresses the cognitive faculty of distinguishing qualities as

perceptive dimensions. At the same time by involving memory and attention,

listening has been given back its place in time as a cognitive practice. In fact,

the act of listening is here interpreted as the attempt of selecting within the sound continuum the perceptive dimension that allows the listener to recognize an order, which needs to be already present in memory. If we accept that memory and attention are faculties that work as functional blocks within the system of cognition, then memory is related to dealing with representations of the past, while attention is related to dealing with representations of the future, as the faculty in charge of selecting the dimensions in which the expected change will likely occur. We do not necessarily have to project a naturalization of our description of the mind so that we have functional units that represent in some way the objects of our external world. Still, we need to deal with our existence in time, at least for the sole reason of building a narrative that shows how time is at the base of the stories we tell.

The first attempt of describing music listening in terms of a complex process that involves a cognitive projection towards both the past and the future is usually ascribed to Leonard Meyer's *Emotion and Meaning in Music* (1956). But the roots of the idea according to which the present moment of the perceptual act is constituted in the conscience as a structure of moments that unfold in time and is not given to the subject's conscience have to be referred back to Husserl's phenomenological analysis of the constitution of time and the constitution of the objects of perception through passive synthesis (Husserl, 2001).

There is a fundamental difference in the level between the two analyses, as Husserl's aim is to show how a world of objects is formed in the conscience starting from sensory data, while Meyer takes for granted the system of music elements of western tradition from the outset, so his analysis is not concerned with showing, for example, whether his conclusions also work for other music traditions.

It is therefore curious that Schaeffer, who takes some relevant ideas from Husserl's research, never places listening within a frame that involves what Husserl calls retension, primal-impression and protension, or at least in an explicit relationship with expectation and memory, even though his "Treatise" is dated ten years after Meyer's book.

Meyer's research moves from the discussion on how emotion is related to the structure of music from the western classical tradition and how it is connected with meaning. It is therefore a typical implementation of Fiske's *copy paradigm*, as he assumes "that relationships between stimuli and events or objects they indicate 'are real connections existing objectively in culture'" (Fiske, 1996, p.110 quoting Meyer 1956, p.34) and can be directly found in objectified and objective form in the music scores.

Nevertheless, all the systems of expectations of memorized content and reactions, together with the idea that expectations are based on probability generated by the statistics of experienced occurrences, is already there.

According to Meyer, the relationship between expectation and its following reaction is able to account for meaning as enacted by the emotion of the listener. In fact

> affect or emotion-felt is aroused when an expectation - a tendency to respond - activated by the musical stimulus situation, is temporarily inhibited or permanently blocked. (Meyer, 1956, p.31)

In particular, there are three possible ways according to which the listener's expectations stimulated by an antecedent in the score might not be satisfied by its consequent and therefore might give rise to affect arousal, the degree of which depends on the strength of the expectations: a delay in the presentation of the expected consequent, an ambiguity of the antecedent that is therefore not implying an exact consequent, and a consequent that is completely unexpected.

Further developments of this primary analysis of listening have moved mostly in two directions. The first has been primarily concerned with articulating the content of expectation, by typically directing the research to the expectation of melodic structures and to its implementation in formal languages. Narmour (1990), for example, who was a scholar of Meyer, proposed in his *implication-realisation* model, an extension of Meyer's approach by distinguishing *implicative* and *non-implicative* music situations and by defining rules of melodic implication as implied by Gestalt principles, such as the principle of

proximity or of good-continuation. On the other hand, the same Gestalt principles have been addressed in the attempt to create an alphabet of symbols that could express a generative description of expectancies in terms of pattern induction processes. The original model (Simon and Sumner, 1968) has been developed by Diana Deutsch and John Feroe (1981) by completing the original repertory of operators and structures so that, in line with the cognitive paradigm of Human Information Processing, it was possible to propose an algorithmic model. This, on the other hand, is intended to produce an efficient computational representation of (western) music, but it is not concerned with putting music listening in the wider and complex context of cognition as the result of the "praxis of living".

On the contrary, a second direction of research has been more concerned with proposing a deeper analysis of the structure of expectation. In particular, the idea of *dynamic attending* has been used to extend Meyer's and Narmour's concept of expectation as not driven by metre and rhythm towards the idea of a complex process that is driven by cues extracted by an articulated perceptive context. While the first proposal of the re-introduction of time in the analysis of listening can be referred to the already widely quoted article of Mari-Riess Jones (1976), it is with the *ITPRA* model of David Huron (2006) that we meet the proposal of a structure of expectation that is most articulated. Instead of a simple dialectic of two stages of attending, that occupy the moments before and after the sonic event, as an expectation that relays on past experience and the reaction that forms new experience, we have now a

five stage model. This is formed by the succession of emotional response systems that are related to different biological functions and are evoked at different times during the cycle of expectation. The five stages of the expectation cycle, whose name's initials form the acronym ITPRA, are divided in pre-outcome and post-outcome responses in relation to the sonic event.

The pre-outcome responses are: the *imagination* response and the *tension* response, the former having the purpose of motivating an organism by letting the organism to feel "vicarious pleasure (or displeasure)—as though that outcome has already happened" (Huron, 2006, p.8), while the latter has the purpose of charging an organism in terms of motor preparation (arousal) and perceptual preparation (attention)".

The post-outcome responses are initiated by the *prediction* response, which causes the emotional response to be positively characterized in case the event-stimulus was expected and to be negatively characterized in case the event-stimulus was unexpected. Then, two types of response to the actual sonic event occur: a fast one and a slower one. The fast one is called the *reaction* response, and consists in a reflexive unconscious reaction that is defensive or protective in function. The slower *appraisal* response follows, providing a positive or negative reaction that "can involve conscious thought that often draws on complex social and contextual factors" (Huron, 2006, p.15).

On the other hand, interest has risen in the last years regarding a model of perception that emerged in the context of connectionist theories of rhythm

detection and is usually referred to as the *resonance* model (Large, 1996). The central idea on which the proposal of this model is based is that perception is enacted by the organism instead of simply being the (although articulated) response to a stimulus. In particular:

> the general idea of resonance theory is that an external auditory rhythm can be represented by the amplitude of internal oscillatory units. These oscillatory units are coupled to the external rhythm and are by definition periodic while the external rhythm does not have to be periodic. (Bååth, Lagerstedt and Gärdenfors, 2014)

While expectation is still in charge, in this context it is expressed in terms of the entrainment between world patterns and cognitive patterns. Cognition is therefore enacted because it is the result of the encounter between the motivated actions of the subject and an external world objectified by our measurements. But, as measurements are also the result of cognition, what we ultimately call "world" is a character in the narrative that moves from the sensations that are promoted by our subjectivity and is therefore an embodied narrative.

We can now come back to our distinctions, which are the markers of instants of time, to realize that the experiences of both distinction and time are far from being the *originary* (I'm borrowing here a Husserlian term) simple moment that can found a theory of listening. They are rather the start of a chain of descriptions caught in a circular narrative, within which our language, as an instrument for hypostatization, constrains us. It is therefore with great caution

that we connect successions of discontinuities with labels that on one hand may allow us to deal with them as a whole, and on the other hand will tempt us in considering them more than the occasional result of the intrusiveness of linguistic metaphors, as objects that inhabit our world.

8. STRUCTURAL LISTENING

8.1 In the name of the pattern

The hypostatization has indeed already begun when I used the terms "dimension", or "instance", or again "configuration", if not even before, when I was using generic labels to address the unique experiences of instants. But this is inevitable and therefore we have to accept it, or better, we need to embrace objectified experiences as illusions, so that we can highlight their inconsistency in our attempts to deal with experiences through language. After all, if the objects of perception are constituted as interferences between the cognitive action of a subject and a transcendence that we can never meet naked from our enacted cognition, we have to argue that we are entangled with some sort of cognitive hologram-like world that reflects our (en)actions.

We have therefore no other chance, if we want to build a tool to use in our world of ghostly objects, than to deal with successions of instants, as configurations of instances, as a fundamental constructive result, that seems more onto- and phylo-genetically connected to our existence than being related to a conceptual process, such as abstraction. The three regions of perceptive modes, which are at the root of the process of grouping, are therefore an elementary system of distinction, maybe the most elementary process that can be found in cognition. Nevertheless, they are not inscribed in our body, in the connection of neurons or in an emergent property of the brain,

but they are rather within our narrative as a scheme that depends on the way we describe our experiences of listening.

I will therefore use the word *pattern* to address the result of recognizing successive distinctions that fall in the SIR. I prefer this term over inventing a new one or adopting different locutions that have been used for similar contexts or purposes, such as Oackelford's "zygonic relationship" (Oackelford, 2005) or Fauconnier's "connector" (Fauconniers, 1985) for several reasons. The first one is that the word "pattern", as a general term of the English language, involves the ideas of repetition and resemblance. In fact, already in its Middle English original form, the term *patron*, which comes from the Old French *patron* from the Medieval Latin *patronus*, was used to designate a model of behaviour or something to be copied. When in the XVIII century the modern form *pattern*, created from the metathesis of the old term, was completely established, it was essentially used, especially in dressmaking, for addressing a model to be copied (Patridge, 1958, p.1062). Therefore, since the first uses of the word, a pattern is related to practices involving a reiterated connection to a reference that is based on resemblance: there is no pattern without a recognition of it and the possibility of a re-presentation.

On the other hand, "pattern" has been used as a technical word in a variety of disciplines.

In particular, I want to connect my research with Howard Margolis' adoption of "pattern" in addressing topics such as political and scientific judgments

(Margolis, 1987), hence a path to the cultural and social implications of using "pattern" as an hermeneutic key to cognition is somehow already traced. By choosing the term "pattern", I ultimately connect my research to all of the disciplines that use it as a technical term, by suggesting how all of them are variations of the general concept of pattern that I use in this text. Finally, although "pattern" is used in music composition more as a general term than as a technical one, it is still very present in different forms within the whole western tradition, up to the 20th century's minimal style and popular music, as well as in describing the surviving sonorous practices of other cultures. Therefore, the use of the word "pattern" in writing about listening already connects and interacts with its various applications in composition and musicology, possibly providing an interesting interrelation of meaning.

In "In the Name of the Pattern" (Viel, 2014), I have detected four different technical uses of the term "pattern", each showing a distinct application of the idea of the recognizing distinctions, as the process of organizing the units that form a composite unit.

A first use is related to *pattern recognition* and involves disciplines that can be related to formal languages, to statistical decision theory, fuzzy logic, digital hardware, pure math and computer programming. While the term "pattern" is often taken for granted in texts about "pattern recognition", it is often assumed that

a "pattern" is the result of finding some kind of regularities, that can be drawn from a set of models or "ideals", within some sort of ready-made data. This is a phenomenon that has to be there already in the first place, in the reduced form of raw data, ready to be picked up by the recognition process. (Viel, 2014)

In the end, *pattern recognition* can be defined as the "search for *structure* in *data*" (Bezdek, 1981, p.1) and consequently a *pattern* is a structure in terms of regularities.

A second use of the term "pattern" is found in the context of design, in the locution "design pattern". This is the description of

a particular recurring design problem that arises in specific design contexts, and presents a well-proven generic scheme for its solution (Bushman, 1996, p.8),

provided that the solution is in terms of an ontology of classes that can be used in a computer programming environment, and is therefore connected to the paradigm of Object Oriented Programming.

While *pattern recognition* is about linking raw data to an abstract model instantiated in the data, *design patterns* are the answer to a design problem, recognized as a singular case within a general one, for which there is already a solution.

A third use is related to art, craftsmanship and music. From this perspective, a *pattern* is a configuration in a given (visual, sensorimotor, sonorous) domain that is instantiated several times as contiguous (in space, in motion, in time). This means that an observer must be able to recognize the repetitions of pattern in the medium they are presented, otherwise the starting model is not recognized as a pattern.

As Oackelford pointed out, repetitions are paramount in music (2005). Yet, in spite of a wide use of the word "pattern" also in dealing with the music of the western tradition, it is difficult to find it as a technical term in texts concerning music. In my article I argue that precisely due to the pervasive presence of patterns in music,

> repetitions of different elements and with different features have been called with various names along the history of western music, following the evolution of styles, genres, composition practices and so forth. […] And the more names there, the more difficult it is to be aware of the common underlying principle of repetition, which seems to be hidden in plain sight. (Viel, 2014)

Finally, the fourth use of "pattern" as a technical term can be found again in music, and especially in jazz, but with a different definition than the previous one. What we can call *training patterns* are

> finger habits. [...] Something that is not necessarily pre-heard, but [...] is *understood to work* (by cognizance of the theoretical reasons and/or by previous experience), or [...] that feel comfortable to the fingers and hands. [Something] [...] to be practiced diligently by serious students. (Coker 1970, p.1)

Training patterns are ultimately never used in the actual practice of music performance, yet they are repeated during the training phase, so that when there is an opportunity to use them during a performance, even just once, the musician is ready to play them almost automatically.

We have encountered four technical applications of the word "pattern". They indeed share with each other, to a greater or lesser extent, the concepts of recognition, similarity, repetition, domain, previous knowledge, type of relation/ structure and habit. A *pattern* is a model that has to be instantiated repeatedly in the external word of objects or in the internal world of actions in such a way that the instances are perceived as contiguous in the dimension that realizes the development of its components. A *pattern* is therefore hardly perceived as a whole, but it is rather a collection of relationships that carry out an organization (*recognized pattern*), a function (*design pattern*), a shape (*visual pattern)* or a gesture (*training pattern*) that constitute the *pattern* as a composite unit, which is realized by certain components.

Finally, we can define a *pattern* as something we recognize in terms of (the order of) its components. Within the context of our narrative, a *pattern* is a configuration of instances in some perceptive dimensions that we can recognize by their order. As previously mentioned, the concept of "pattern" already holds within it the idea of repetition, which from the point of view of perception becomes recognition, and in turn recognition brings the idea of

resemblance. On the other hand, the claim that a pattern is something that is composed of something else connects with the idea of hierarchy. This emerges from the relation of inclusion between a "container" unit and the units it contains itself as a bigger structure that is suitable to include all the units. We can express inclusion by labelling all the units with a number that corresponds to the hierarchical level they belong to, so that if the container unit is said to belong to some level number, the contained units belong to the immediately inferior level. Once introduced, hierarchy cannot avoid replicating itself in a potentially pervasive tree-like structure of units and levels.

Starting from the idea of a structure of patterns (Section 8.3) and its interpretation within the scheme of perceptive regions (Section 8.4), the chapter moves towards the progressive construction of a world of objects. First of all, some sonic events that do not seem to be easily interpreted in terms of pattern are introduced in terms of statistical dimensions (Section 8.5), so that an interpretation of the cognitive role of attention and an idea of learning as a process driven by attendance that mirrors the statistics of occurrences can be presented (Section 8.6). The introduction of the "misattribution effect" (Section 8.7) opens the way to the naturalization of patterns in terms of objects (Section 8.8). The final section (Section 8.9) is dedicated to the unavoidable problem of meaning in music, which is first summarized in its most relevant issues and then is finally discussed within the framework of the present research.

But before introducing the idea of a pattern structure, we still have to discuss, in the next section, the thorny problem of resemblance, as it emerges naturally from the idea of recognition, and can be interpreted within the concepts of intention (Husserl) and *autopoiesis* (Varela)

8.2 A matter of resemblance

A *pattern* is the effect of recognition, but recognition does not constrain the simulacrum. On the contrary, it implies the discovery of identity within diversity. According to the notorious sentence ascribed to Heraclitus, "you cannot step twice into the same stream", where the impossibility of "sameness" is here just another aspect of the contention between the logocentric promise of eternal reference and the semantic solitude of the uniqueness of experiences. In between the two contenders, the notion of *resemblance* seems to bear the philosophical responsibility of bringing cognition towards its purpose of building a knowledge that is able to leap over the instants of time.

According to Hume, for example, all we can know or in the end, all we are, is made by just one kind of experience that is modulated in "degrees of force and liveliness" (2001, §1.1.1.1) from the highest degree of sensations and subjective states, to the lowest degree of higher order "reflexions". It is a population of impressions and ideas, a sort of weaker "internal" version of impressions, whether simple or complex, which belong to different orders, and come from the proliferation of primary sensations. From the world of

sensations, we have therefore moved to the relata of concepts: we are now in the world of abstract ideas. These rise from a sort of skilled activity of the imagination. In fact whenever we find a resemblance between several objects we label them with a same word, creating a habit that has the possibility of evoking all the ideas connected to a word whenever that word is heard and vice versa.

In the end, Hume's mind is a vast collection of just one kind of object with infinite variations of a sort of transparency under the care of a machine that has the ability to create, link and analyse objects with the magnifying glass of resemblance. This impersonal machine is stuck in an endless routine that generates objects with the rule that every time an object is created another one is created that is linked with the first one and is a little bit more transparent. It also looks for resemblance between objects so that when found, it creates a new object linked with all the similar found objects.

We could say that Hume depicted a Turing machine of some sort, with all of the related problems on the nature of reality and identity; problems that emerge from Hume's devotion to Newton's mechanicism and pay back, as underlined in the conclusion of the first book *A Treatise of Human Nature* with the "forlorn solitude" of an irreconcilable contradiction with everyday life experience.

In Hume's vision, as in his predecessors Locke and Berkeley, resemblance and its active counterpart, the habit, play an important role. If it seems, following Gamboa (2007), that Hume has avoided the perils of a one-dimensional

definition of resemblance as the possession of common attributes by showing an openness directed to a polymorphous conception, the Humean idea of resemblance is nonetheless not without problems.

One main issue in judging the resemblance between objects is that, as Goodman (1972) has shown, there are infinite ways in which two objects can resemble each other or, stated otherwise, there are infinite ways to detect subsets of a given set of objects based on resemblance. When we rely on a mechanism of resemblance detection, we need to face the problem of relevance; meaning that we need to establish some criteria to get the resemblance we need in a given moment.

While it seems that Hume was aware of the problem,[24] there is still no explicit answer to it. And this is not difficult to understand: apart from the activity of resemblance detection, whose analysis is not developed at all, according to Hume the path from sensation to knowledge moves only in one direction and there is no way for the knowing mind to turn to sensations and, in some way, to shape them. No wonder then, if it has been claimed, for example by Sellars, that for Hume, together with Locke and Berkeley, resemblance seems to be a "primordial, non-problematic feature of 'immediate experience'" (1997) instead of being the result of the knowing mind activity.

Another important issue in Hume is the ontological status of the objects that populate the world we know. It has been observed by Popper that Hume was a "a passionate realist that was led from his subjectivist theory of knowledge to

_____
[24]See for example, Hume 2001, §1.1.7.21.

metaphysical results which, while feeling obliged to accept on the basis of logic, was constitutionally unable to believe even for one hour" (2013). Husserl talks about a "fictionalist theory of knowledge" (1970) according to which "all categories of objectivity" are seeped in immanency as fictions of mind. And again, this is due to the knowledge arrow that, in the attempt to found an objective "physicalistic" knowledge, moves to the knowing mind from what we know of the transcendental objective world, which are the sensations. "No inference is thinkable [...] through which conclusions could be drawn from these sense-data about anything but other such data" (Husserl, 1970, §23).

Again, according to Husserl "Hume ends up, basically, in a solipsism". And already since Kant, but with roots in the Cartesian "primal self-evidence to be the self-evidence of 'inner perception'", it is the reverse of the knowledge arrow, that is proposed as a solution for the foundation of an objectivity of the world, as the legitimation of its transcendental nature.

With the introduction of the concept of "intention" the knowing mind is no more a neutral vessel that receives and organizes sensations, but it becomes an active instrument that moves towards the mere sensorial content and organizes it in what appears to our consciousness as "perception". In Husserl's vision, by applying a phenomenological method of judgement suspension of a sceptical nature, we have the chance to become aware of our "unconscious inferences", as Helmoltz would have called them, which unveil the hidden activity of the knowing mind.

This is a complex geometry of structures of intentions, which unifies a series of single sensations as adumbrations of transcendental objects in the same way that different visual and instantaneous sensations are perceived when a table is looked at whilst moving around it. But a table, as an intention, is never perceived in a final way, as it is primarily a cognitive horizon of possibilities open to infinite specifications in time and space. That is why the intentional process is taken in a sort of endless interplay of predelineation and fulfilment between apperception (the sensory data) and apprehension (the act of unifying through space and time different apperceptions in a transcendental object). Ultimately, the transcendental object itself remains "at an infinite distance", while what the intention intercepts just "pretends to be its essence, and it is it too, but [...] in an incomplete approximation, [...] that constantly grasps in an emptiness that cries out for fulfilling" (Husserl, 2001, p.59). Caught in this endless loop, the sceptical refutation of the possibility to meet the transcendental word is just a step away, so that Husserl needs to involve regional "ideal limits" of fulfilling that bring us back to our body and its physical thresholds.

In the relation to the body, with its limits, functional asymmetries and transformations, Husserl opens his system to the investigation of the ontogenesis of intentions as a so-called "genetic constitution". First of all, the body provides a "constitutive duet", that is composed first by "the system of [...] free possibilities of movement" that "is intentionally constituted as a practical kinaesthetic horizon", secondarily by the process according to which

every sensation is "ordered with respect to consciousness, to the current situation of the consciousness of the parts of the lived-body, creating an horizon of further possibilities" (Husserl, 2001, p.52) and consequent possible "appearances" linked to the possible movements of the body.

This loop seems nothing else but the predelineation-fulfillment loop that is applied to the matrix of possibilities of the body as a transcendental object among transcendental objects.

When we strip that loop from the contingencies of regional intentions, we meet the core and most abstract mechanism of the dynamic process of knowledge as an endless loop of protention and retention. These, when applied to predelineation and fulfilment, engage the loop of expectation and reaction.

In this infinite loop, habits are created and once again resemblance is what allows us to expect something that already happened in the past, or rather that resembles what I have already experienced. Once again, resemblance is at the centre of an unconscious mechanism that manages repetitions, the habits, as a basic pattern for constituting knowledge and, in a phenomenological vision, human reality itself.

What has changed since Hume? The notion of intention seems to have the chance to manage the relevance issue: the relevant characters of resemblance are selected by intention itself. Moreover, the issue of transcendence of the word is taken into account even we ultimately cannot meet it.

We also find the idea of cognition as a circular process, in Maturana and Varela's *autopoietic* system, according to which the organism is stuck in an endless loop of, let us call them, stimuli and reactions, where stimuli are modifications of internal biological states, due to the cognitive closure of *autopoiesis*, and reactions are the processes of maintaining itself as a unit, that in the end is "surviving" as an *autopoietic* system. This, exactly as the human being in Husserl, cannot reach the world appearing as a sort of "unrequested" modification of the system itself. The only context available of the *autopoietic* system is itself. As observers, we can see the living being in the context of an external environment in a structural coupling made of modifications of the environment and modification of the *autopoietic* system. As the structural coupling goes on, regularities emerge that we could interpret as habits or in the end intentions.

As Varela says: "signification arises in the emergence of a viewpoint proper to the autonomous constitution of the organism at all its levels, starting with its basic *autopoiesis*" while "the constitution of an autonomous unit provides the means for regularities to appear which are the bases of [syntactical] compositionality" (Varela, 1992, p.13).

Resemblance too emerges from regularities in the structural coupling, in the encounter of the living being with the mirror of the self, what we call external reality. Within the structural coupling in action, however, there is no need for such a concept. It emerges in the domain of descriptions that we make through language when we evoke a transcendental world of objects.

Therefore, it ultimately seems that the quasi-esoteric complications of Husserlian terminology arise from the contradiction between the vision of a world of objects, which is definitely made easier by our languages of nouns, and the genetic perspective of a world constituted by actions.

Resemblance lives in the descriptions that we make and therefore is open to the uncertainty of the infinite possibilities of selecting the qualities that realize it. The choice can only be made within the context of our descriptions, which in turn depend on the context that, within our "praxis of living", gives rise to those descriptions.

Goodman is indeed correct to be suspicious about the use of resemblance to validate knowledge beginning from our descriptions as observers. But what we call resemblance is rather the hypostatization of a symptom that is the consequence of the articulated process of constituting ourselves as subjects with our repetitive actions and their expected results.

Provided that these results do not perfectly match our expectations, but rather challenge, up to a certain threshold, our ability to expect, we could say that our possibility of action in the world, our refusal of habituation, is linguistically represented by that difference we call *resemblance*. We therefore do not have to deal with the infinite qualities in which to look for resemblances, but it is our repertoire of experiences that is inscribed in our world and drive our actions towards expectations that follow the world and can be described in terms of searching for resemblances.

8.3 Pattern structures

We can finally introduce one of the most important topics of the chapter: pattern structure. In fact, if a pattern is something we recognize in terms of its components, then an ordered configuration of patterns is a pattern as well, provided that it has been recognized in a representation. Jones was well aware of this, as it is a consequence of the group structure of dimensions. The associative property, which is a requirement included in the very definition of the "group" structure, provides the theoretical conditions for a succession of recognizable configurations to be understood together as a higher level recognizable configuration. Consequently (I am following here: Jones, 1981), two consecutive presentation-instances of the same pattern, in terms of intervals, form together a composite pattern unit, the simplest one, and can be connected to the interval between the first values of each one of the instances, as a *generator* that describes the interval of transposition between the two instances. This generator interval can therefore be taken as the compositional value describing the dimensional relationship between the two instances of the same pattern.

We therefore have a tree-like structure, in which the lowest, "surface" level comprises the units corresponding to the double two pattern instances, labelled by the intervals of the related perceptive dimension being generated through repetition and transposition, the single instances of the patterns. Immediately above the lowest level, we have a middle level that includes one composite pattern unit composed by the two pattern instances.

226

If the highest possible level, the "total" level, comprises only a single composite pattern unit, encompassing all the composite pattern units at the immediately lower level, that in turn encompass all the composite pattern units at the immediately lower level and so on down to the single instances of the pattern, then in this very simple tree-structure the total level coincides with the middle level.

TOTAL LEVEL

SURFACE LEVEL

P_n 1   4

ΔP_n 3
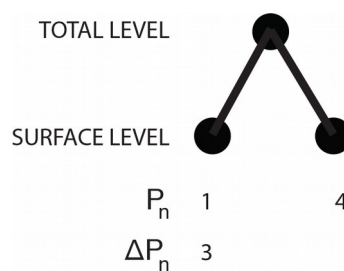
Figure 3: the simplest composite pattern unit with dimensional generators.

Of course, as we have seen, a perceptive dimension is not enough to characterize a configuration that, in order to be a pattern, needs to fall under the SIR. If we want to build a diagram of regions we also need to specify the time interval dimension for all the levels of the pattern structure, except for the highest "total" one. This can be done in a manner similar to the diagram used for the perceptive dimensions, so that the surface level units are connected with the time interval between their onset and the onset of the successive instance, while the units at the higher levels are connected with the time interval between the first surface unit of that higher level unit and the first surface unit of the successive higher level unit, which corresponds to the sum of the time intervals of all the surface units nested in that higher level unit.

227

So, if all the instances have, let us say, the same duration of one pulse, the surface level units will be labelled by that pulse, while the middle level units of component patterns will be labelled by the time interval between the onsets of the two presentations.
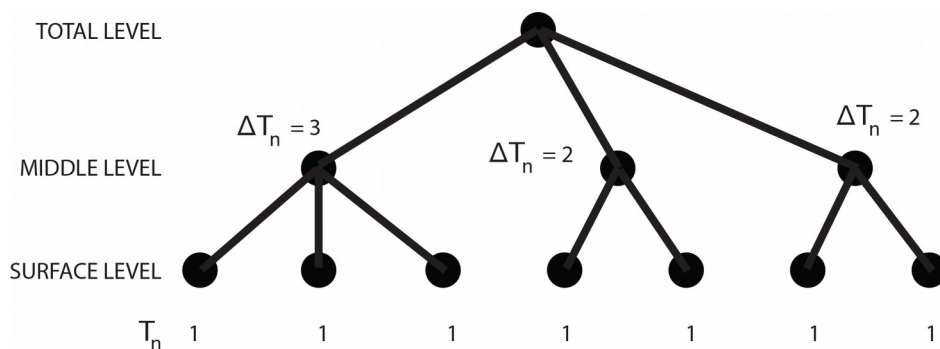
Figure 4: a simple pattern structure with time generators.

Of course, once we have determined how to describe a multi-level pattern, we can invent or discover extremely complicated patterns in the sound flow.
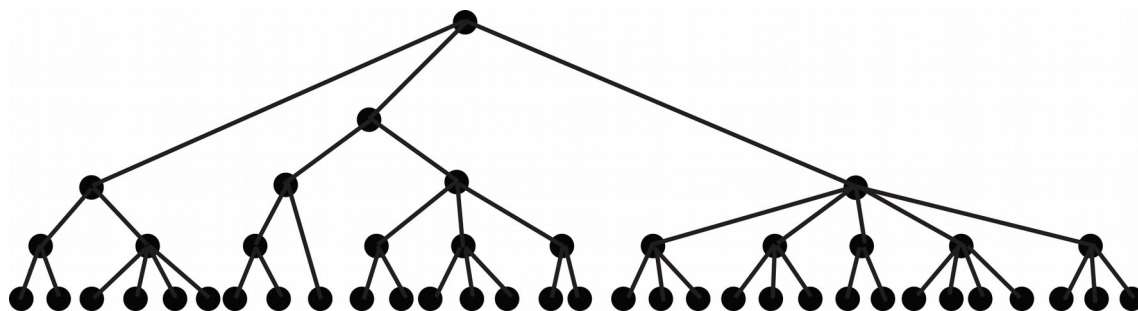
Figure 5: a more complicated pattern structure.

We can also apply statistics to determine the generators in cases where it might be useful for us: after all, as we have seen, every steady phenomenon is constituted as a chunk from its undetectable variations and in our sounding practices it is impossible to realize two time intervals that are exactly and

perfectly the same. Every measurement related to our listening experiences are de facto connected with error thresholds that make them statistical data. We can therefore, in some cases, describe each level of a complex composite pattern unit with a pair of values, the dimensional generator interval and the time generator interval, so that we can plot them in the space of perceptive modes. While, as I have said, this is not a research on psychoacoustics so it is not important here to determine exactly the regional ratios, that would anyway be different for each perceptive dimension, the possibility of relating the levels of a pattern structure to the regions of perceptive modes has very important consequences.
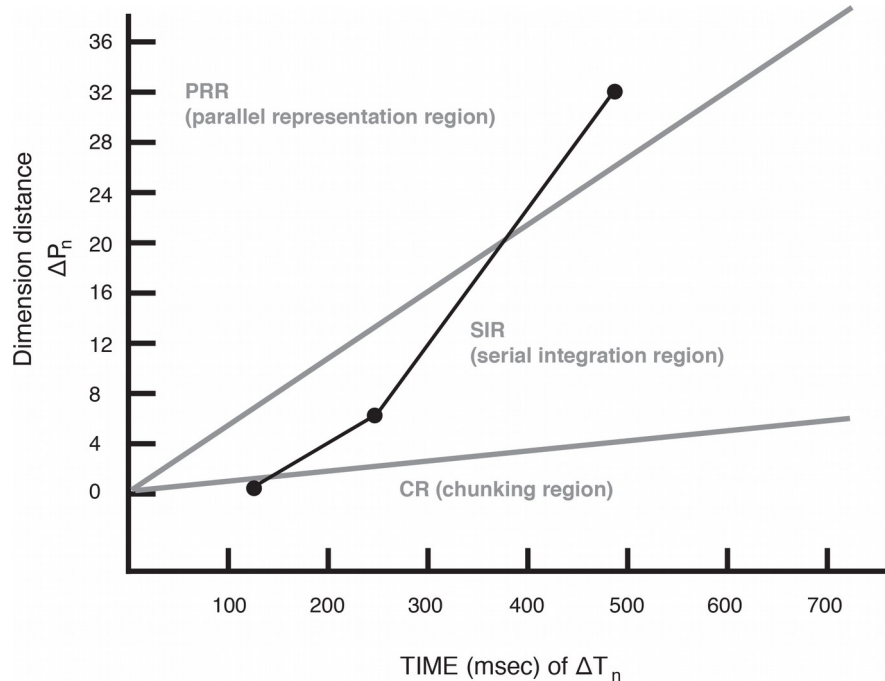
Figure 6: generators of a simple pattern structure as points in the space of the modes of distinctions. The surface level falls in the CR, the higher level falls in the SIR and the highest level, which is immediately under the total, level falls in the PRR.

8.4 Regions at higher levels of pattern structure

However, we can try to understand what regions mean when the idea is applied not to the single pairs of instances in the surface level, but to the generator pairs of the higher levels.

If we consider a dimensional generator equal to zero, we are dealing with a CR realized by the exact repetition of the component patterns at the level expressed by the generator. When we deviate from the value of zero, as long as we stay within the CR, by definition we should not be able to detect any relevant differences between the repetitions of the pattern units, composing a unit belonging to the immediately higher level, so that they realize it as a chunk. The composing units then become the elements of the texture that characterizes the chunk to which it belongs. We could therefore say that CR deals with *identity*.

$$\Delta P_n \; \blacksquare\!\blacksquare\!\blacksquare\!\blacksquare\!\blacksquare\!\blacksquare\!\blacksquare\!\blacksquare\!\blacksquare\!\blacksquare$$

Figure 7: succession of instances that fall in the CR in the dimension P.

Nevertheless it is still possible that the chunks in some level might work as units falling in the SIR in some higher level in the very same perceptive dimension. In fact while a composing unit in a certain given level stops its process, of constituting a higher level unit at the highest level so that it falls in the SIR, which does not mean that its lower levels could not fall in the CR or even in the PRR.

230
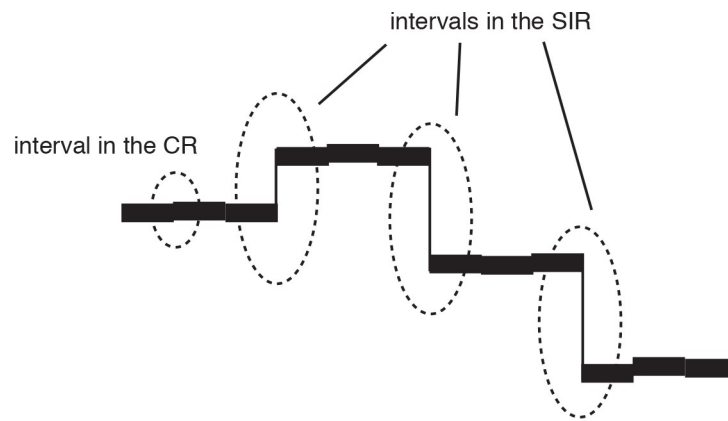
Figure 8: succession of instances that fall in the CR at the surface level, but fall in the SIR at a higher level.

The SIR is the region expressing the pattern as composed by ordered, recognizable components. Therefore, to be a composite unit a pattern structure needs at least to fall in the SIR at the highest level defining its components. The number of levels that fall in the SIR define the articulation of a composite unit so that the differences between the units in the lowest SIR level are the material, so to speak, from which the units in the higher SIR levels are built of by abstraction. This also means that the possible articulation of a composite unit must depend on the limits when dealing with our ability to order differences in memory. However, the complexity of dealing with world patterns, which are articulated in multiple dimensions with nested levels related to different perceptive mode regions based on time intervals, might explain why a determination of the perceptive thresholds involved is so difficult to reach, as shown by the history of the debate around this topic (see, for example: Miller, 1956; Honig, 1988; Gobet and Clarkson, 2004). In this context, it is enough to say that whenever the limits of distinction are reached, we fail to order the

configuration of instances on some structural level, which therefore falls back in the CR.

In sum, the relationship between CR and SIR in composite pattern units show us that the CR seems to be in charge of perceptively forming the units, which, by means of their ordered differences, realize patterns.

Moreover, CR levels that are higher than the maximum level related to the composite unit in dimensions that are different from the perceptive dimension where the pattern is constituted, are connected with the experience of what is constant in the sound flow during the enunciation of the pattern. For example, a pitch pattern played legato on a synthesizer with a sine wave timbre exhibits a CR for the whole duration of the pattern in the timbre dimension, which in other words means that the timbre dimension does not define the pattern as such, but its steady character helps us constitute it as a changing foreground against a stable background. Therefore, in opposition to CR's identitarian role, we could say that the main role of SIR, even in multiple level patterns is to address *difference.*

PRR intervenes when the generator interval of some level related to a certain perceptive dimension is too high to fall in the SIR. In a fast surface level, we have seen how PRR gives rise to the phenomenon of *streaming*, so that successive instances are perceived as simultaneous. In higher abstract levels, which are by definition slower than the surface level, this might not happen. Jones' example of a structure with a PRR level shows how, with surface

instances lasting 0,25 msec, even a simple pattern of eight pitches such as C-D-E-F-C3-D3-E3-F3 cannot be perceived as a repeating segment of eight notes, but it is broken into the two overlapping segments of four notes that fall in the SIR: C-D-E-F and C3-D3-E3-F3 (Jones, 1976, p.343).

The influence of PRR on perception does not stop when the generator time interval is slow enough to avoid *streaming,* although when two sections of the sound flow are evidently perceived as one after the other, it might seem unclear what a "parallel representation" could mean. But the perception of simultaneity before being the constitutive occasion for unity, expresses the fragmentation of the continuum of perception: it is a *segregation*. This continuity that is broken into pieces and cannot be brought back to unity without a wider context in which such pieces reach unity while maintaining their diversity is the general effect on perception that PRR accomplishes, even in higher levels of a pattern structure. If two successive composite units exhibit a generator dimensional interval between them too wide to fall into the SIR, then these units can hardly be merged into a single unit of a higher degree, but need to be perceived in a higher level as components of two different units that follow one after the other. Let's take the last example and increase the pitch intervals between the fourth and the fifth instances so that the pattern is now C-D-E-F-C7-D7-E7-F7, with a surface instances duration of 200 msec. We mostly will not be able to integrate, without effort, the whole pattern as a series of eight pitches, but we will rather easily fragment it into two alternating

patterns of four notes each. This is therefore an example of the *diversity* issue addressed by PRR in higher and slower levels of the pattern structure.

Let's see other examples. Bobby McFerrin's arrangement, of The Beatles' "Blackbird", for solo voice is a typical example of how the technique of *virtual polyphony* can be used to suggest, with just the aid of the monophonic voice, the simultaneous performance of the different sections of a band, for example the rhythmical section against the melody, by using extensively PRR among units in slower levels of the pattern structure. A different case results by "identifying" lower level units with CR units separated by PRR intervals. For example, let's take the "echo effect" which is typical of the polychoral style of *cori spezzati:* when it is reduced to the sole sounding flow resulting from its performance, we could say that the opposition between choirs is simply based on PRR intervals between identitary CR levels as their position in space. The same is true for other kinds of musical oppositions such as the PRR in timbre and dynamics that characterizes the distinction between *ripieno* and *concertino*, in the Concerto Grosso form, or between a soloist and the orchestra, or also between two manuals of, let us say, a work for cembalo. All the dimensions that fall in the PRR in the end are directed to suggest a diversity in the sources of what, in terms of the pitch dimension, is falling in the SIR and therefore constitute the pattern levels.

There is another aspect of PRR that can be connected to the *continuity phenomenon* in psychoacoustics. This happens when one signal (typically

called the "probe") is followed by another signal that is potentially able to mask the previous one. The phenomenon manifests itself in the difficulty of interpreting in different ways a *probe* interrupted by the following signal and a probe that is actually masked by the following signal. In both cases "the perceptual system 'assumes' that the signal continues" (Moore, 1997, p.254) because, according to the Gestaltic principle of "good continuation",

> as changes in frequency, intensity, locations or spectrum tend to be smooth and continuous rather than abrupt […] a sudden change indicates that a new source has been activated". (Moore, 1997, p.264)

This phenomenon is consistent with the distinction of diversities within a sound flow, deduced from the presence of necessarily abrupt changes within the PRR between sequences of units falling in the CR. And in the end, the dimensions that are related to the masking effect, such as dynamics, first of all, will be in charge of driving our perception, so that it can attribute the sequences with the roles of foreground or background.

CR, SIR and PRR are therefore the three poles of a dialectics of perception as an active process of moving in the world, as constitutive of the "praxis of living". I have chosen three words to interpret the functional roles of the three regions, *identity*, *difference* and *diversity*, which are not neutral at all. They are in fact deeply rooted in western thought: they have appeared as a fundamental topic since pre-Socratic philosophy and cover the whole path of western

history up to the contemporary era where they emerge in a variety of contexts such as the theory of mind and the cultural, post-colonialist and gender studies.

I do not dare here to deal with all of the connections that these terms might throw into the vast ocean of western thought and therefore I leave them as a constant background for possible references.

In this context and from the perspective of perception, *identity* refers to what cannot be distinguished, to what is invisible and yet conditions our experience; *difference* refers to the distinction of movement, of variation; *diversity* refers to the splitting of movements into *identities*. The process from CR to PRR is therefore almost circular as each PRR level implies a number of CR or SIR nested levels.

Consequently, we could also say that *identity* is the expression of unity; *difference* is the tension of *diversity* towards unity, while *diversity* is the tension towards breaking the unity.

8.5 Statistical dimensions

After having introduced and developed the idea of pattern structure, we need now to deal with the kinds of sonic events that do not seem to easily fit within this framework. In fact, I have introduced patterns as a fundamental character of my narrative of listening, but not every configuration is recognizable and consequently not all possible configurations are patterns. Let's imagine for

example a virtually endless series of unpredictable instances of a certain perceptive dimension, let's take as usual pitch, which falls in the SIR.

Figure 9: a randomic succession of instances that fall in the SIR.

We can still recognize the intervals so that we can detect their order, but as the configuration of orders never repeats, we will be able to await a possible re-proposition of instances only as long as they fall in a constantly moving memory window, the short term memory, which also depends on our cognitive skills. But as soon as we can recognize the re-proposition of even a small number of instances, we will experience that wince of distinction, allowing us to separate the configuration between a before and an after; and a pattern, so to speak, will be born.

Let's imagine this configuration is realized by ordinary tempered pitches changing every 250 msec within a global range of an octave. After a while, the pitch range moves to the upper octave without changing time interval and without losing the unpredictability of configuration. Yet we can distinctly detect the change in a certain given quality. Afterwards, first the pitch range returns to the original octave and then it moves back to the upper one, repeating the

237

sequence every four seconds. Again, we can easily detect the pattern of the range movements, even if we cannot recognize any pattern of pitch instances.
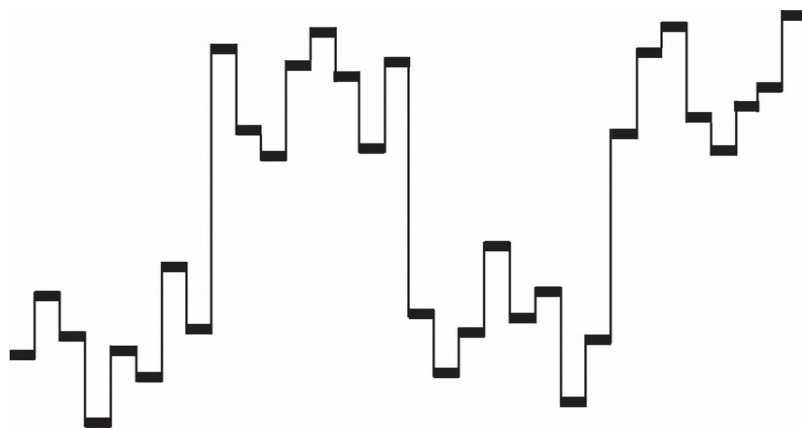
Figure 10: a random succession of instances with a duration of 250 msec, that fall in the SIR, changing their range of intervals every four seconds.

What is happening?

If we detect a change, it means that there must have been a constant instance of some perceptive dimension that changed into another constant instance of the same dimension and that the pair of dimensional and time intervals must have fallen in the SIR. What is the perceptive dimension that is involved in the SIR? I have mentioned pitch range, but because it does not change in width in our example, it must instead be a value that is related to the average distribution of pitches. It is not even the "simple" pitch dimension, nor its differential derivative, pitch interval, because we can imagine a transition between the two ranges that does not need to imply a wider interval in connecting them, and nonetheless it is still very perceptible.

238

It is indeed possible to define a dimension of pitch range in terms of rules regarding the successions of intervals, such as the impossibility within a CR unit to connect two or more intervals in the same direction whose total width is bigger than the defined width, for example one octave, and so on. But, as this is a rule for building a target sound sample, rather than a description of listening, it would presume that our perception focuses on the absence of something, the forbidden series of intervals, instead addressing the verifiable presence of something, on the positivity of recognition. Therefore, as the totality of what is possibly absent is virtually infinite, while the recognition of a presence can be shaped by memory, it seems more likely and consistent within the present framework to connect this quality change to the detection of an area over the dimension of pitches: a register.

If we can numerically describe what at this point seems to be a statistical dimension of an interval range around a central pitch, nevertheless it is just an artifice of mathematics that might not correspond to the target of perception. In fact, how can we define the generator interval between two instances of these dimensions: as the difference between the central pitches in the interval range or as the difference between the lower or upper pitches of that range, or again as a difference between the value resulting from a linear combination of them? Maybe the experimental research on the amount of change experienced between units with different central pitches and ranges will help to answer the question. In the meanwhile, we surely can build complex structures out of

variations of the dimensions of pitch register, both in terms of pitch range and central value that fall in the SIR.

I claimed earlier that this dimension is a statistical one. I do not want to say that statistics are constitutionally inherent to this dimension, as this would imply that statistics are constitutionally inherent to cognition. Statistics instead emerge as a discipline in the complex path of social practices that enact cognition. On the other hand, the dimension of register is perceptively constituted from the succession of unpredictable instances of pitch. It is a predicted succession of constant, within the CR, degrees of unpredictability. In this sense, it can be described statistically as the set of probabilities of a range of pitches. Yet, statistics lies in descriptions and it has to be related, within the dynamics of cognition, to the result of its functioning rather than to the intrusion of a (metaphysical) property of the sound flow.

In the end, statistical dimensions seem to emerge from a cognitive activity that aims to detect patterns with any means. When no pattern can be found in the SIR along a dimension, the simple possibility of turning into CR units what cannot be distinguished, gives the cognitive process a game in looking for patterns along other derivative dimensions.

SIR is therefore not only the threshold condition for recognizing patterns, but it is also the leading heuristic instrument for dealing with the dynamic world. On the other hand, the CR is not only the threshold condition of recognizing

identities, but it is also the heuristic instrument to build the very instances that form configurations and possible patterns.

Conversely, we can consider statistical dimensions as the way that CR works. In fact, falling in the CR means, for a collection of instances, that it is not possible to distinguish an order because the instances appear as equivalent with each other in expressing a quality. If the octave range in our example were reduced to a half-step and the pitches were consequently chosen microtonally within that range, the successions fell in the CR.

Therefore, the distinction between a CR and its statistical SIR version seems to lie only in the relative range, together with the impossibility of detecting repetitions.

In conclusion: the difference between CR and SIR seems to be in the openness of a window related to the perception of differences. We could therefore consider a couple of instances that fall in the SIR as a successful detection through a window that is closed enough to only let the two instances pass.

This window, with its selective role, seems to be a proper image for attention.

8.6 Learning

The cognitive process of transforming patterns in objects needs a mechanism that involves memory, attention and expectation, and is able to transform the formless sonic flow in a complex structure of recognizable elements, ie. patterns.

The psychological phenomenon of *habituation* seems to be key to this mechanism. This is usually defined as "a decrease in responsiveness resulting from the repeated presentation of an eliciting stimulus" (Huron, 2013, p.8), that Huron qualifies as "the most important mental mechanisms organisms have for ignoring stimuli". *Habituation* is therefore an effect of attention, which takes place when it addresses a constantly recurring stimulus, as it happens within a CR level of the pattern structure.

While this topic is typically related to learning, especially in infants, everybody has had the experience of erasing from consciousness a car alarm constantly playing in the background, so that we suddenly realize its presence only when it finally stops. *Habituation* is such an important instrument for constituting our world, that, for example, the AC frequency of the general distribution of electricity is able, by means of its continuous sounding yet habituated presence in our homes, to condition our emission of spontaneous pitches, as claimed by Murray Schafer (1993).

It is a complex phenomenon that is ruled by perceptive thresholds and latencies, which can be modulated by the processes of *dishabituation* and *potentiation*, depending on the details of pattern structures within the sound flow (see: Huron, 2013).

What is of interest here is that *habituation* is a phenomenon that arises in presence of a CR level that lasts long enough to stimulate it. From the perspective of expectation, on the other hand, it activates when the sonic event perfectly matches our expectation, so that there is no longer the

necessity to anticipate a change. *Habituation* is the sign of certainty or better: safety. So it is no wonder that when change happens we are hit by a strong physiological wince[25] that is often accompanied by a sense of alarm. We could describe our expectations in terms of probabilities represented as a function, which after some habituation time, reaches the zero for all the possible generator intervals in the focused perceptive dimensions, except for the one that falls in the CR, which has a probability of one.

Curiously, such a *habituation* function might well describe the process of building a statistical dimension, provided it is interpreted as the progressive opening up of the attentional window of the range of pitches. In fact, from the perspective of expectation, we can imagine that as long as the anticipated repetition of a pattern does not show up, the attentional window—that when dealing with SIR instances is open just enough to detect them in a possible repetition—will progressively open to intercept the instances of the configurations with the maximum success.

This vision of expectation implies that the cognitive system primarily aims to minimize the use of resources, namely energy in the first place, and therefore is committed to maximizing the results of detecting changes even at the cost of expanding the focus to any possible change. This idea is based on the assumption that a precise detection, as obtained by a barely open attentional window is more valuable that a loose one, in which the window is wide open.

---

[25]See note 17 in Section 6.

What happens in the initialization of a statistical dimension is that as soon as a configuration falling in the SIR has been detected a precise series of possible repetitions is immediately expected. On the contrary, as soon as it is not possible to detect any repetitions, an alternative system is engaged, to the energy-costly process of getting a precise but unexpected pitch. The range of possible pitches is therefore gradually increased until all the pitches in the actual configuration are detected, so that they start falling in the CR under the mark of new statistical dimensions and the process of habituation can start.

What is interesting here is that if we turn to describing the path of expectation in terms of probabilities, the related function would once again start from the probabilities set by the memorized pitches, but after a while all the possible generator intervals of the statistical dimension will reach the zero, except for the one that falls in the CR, which will have a probability of one.

As we have seen, in order to interrupt this process, some repetitions in the dimension of pitch intervals need to happen. Or we might have a change in the dimension of pitch register or indeed, any number of events may happen. But what I want to highlight here is that both *habituation* and the formation of "statistical" dimensions share a similar process of dealing with CR levels, the first one by dismissing them from consciousness, the second one by creating them.

In both cases a process starting from an uncertainty aims to minimize it. From this perspective, they both are related to learning.

As concerns the wince, it seems that it can be connected both to the detection of an unexpected change in some perceptive dimension and to the detection of the unexpected similarity that gives birth to a pattern. While from the physiological point of view it seems that the sudden allocation of resources due to unexpected events might well be responsible for some of the correlates of that wince, nevertheless both cases relate to a discontinuity in the function of the probability of expectations.

As already mentioned in Section 7.4, the idea, according to which listening is adapted to the statistics of its object, has already been introduced in Leonard Meyer's research. This idea, in the terms presented by David Huron (2006) states that learning is the preeminent source for auditory expectations and that the statistical proportions of sound expectations seem to match exactly and sometimes even against cultural common sense the statistical proportions of what has been listened to during an entire life. We can therefore call *statistical learning* a type of learning that is based on how frequently a particular event occurs, its non-contextual probability, or how tightly two or more events are correlated, which involves a conditioned probability typically expressed in terms of a Markovian stochastic process.

According to Huron it seems also plausible to link unspeakable subjective qualities to the statistical properties of the stimulus in particular contexts. The most important properties are:

(1) the non-contextual probability,

(2) the first-order tendencies between stimuli and

(3) the likelihood that no further stimuli will occur.

We are normally told that there are five senses: taste, touch, smell, sight and hearing. Over the past century, physiologists have established that in fact many more than five exist. Equilibrium, for example, proprioception and the four distinct systems that seem to be entailed by touch: heat, cold, pressure and pain. Huron suggests in addition to them, the existence of a sense of future anticipation. Thinking of a feeling of probability, as a distinctive quality, is an interesting hypothesis that has the potential to unify the perspectives on perceptive behaviours that apparently seem different from one another.

We have seen that a change is detected in one dimension when both the dimensional and time intervals between two instances, which are CR units, fall in the SIR. If we consider the dimension of the probability of expectation related to the single instances in a configuration and we assume that we can develop a subjective dimension for it, we can consider both a perceived change and a perceived repetition as a change in that dimension which falls outside the CR. If we are expecting that the "habituated" probability will last longer, both a perceived change or a repetition will violate that expectation, causing that thrill of surprise that comes with it. And the discontinuity of time, which is ultimately a part of the narrative device, can be connected with the discontinuity of winces expressing the unexpected.
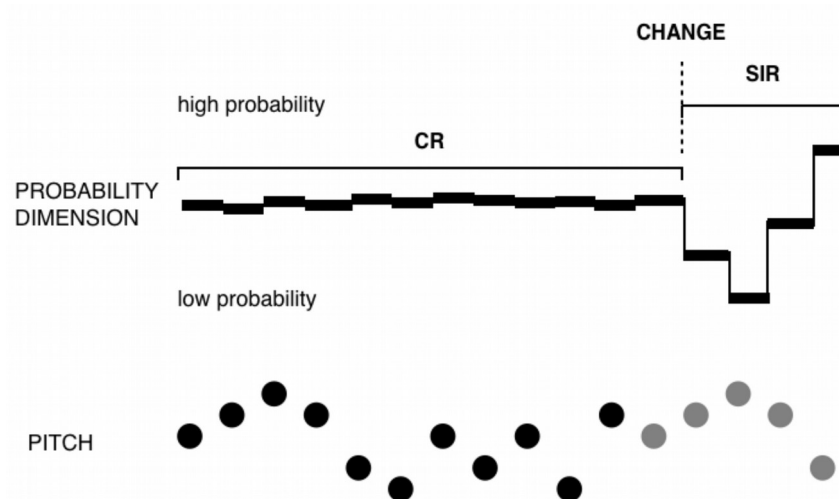
Figure 11: the dimension of probability.

8.7 Misattribution

In this section a key-concept is introduced, which seems to be able to connect *statistical learning* to the construction of a world of object. But first we have to acknowledge that the idea of a fundamental dimension of probability of expectation suggests that the two phenomena that depend on it, namely the detection of a pattern and the detection of a dimension, might actually share a common root in spite of being apparently distinct from each other. If we want to determine its possible conditions and consequences, we need to further develop the idea of dimension.

First of all, we already know that a dimension is expressed by an ordering relation within a set of possible instances, so that we can put two different instances related to that dimension in a succession from lower to higher in the dimensional scale. A first consequence is that we need to set an equivalence

relation between two instances in a dimension that we consider, more or less, equal to each other. In particular, if we have learned to use a dimension to detect qualities in the sound flow then we must have developed the ability to determine if two instances belong to the same equivalence class that is, otherwise said, the ability to recognize the two instances as expression of the same dimensional value.

However, as the sound flow is continuous and never-ending, it means that in order to compare two equal instances along the dimensional scale, we have to detect two changes in the sound flow, which are necessarily not successive, otherwise they would "melt" in a single CR instance. The key term here is "recognizing": it implies that an instance of a dimension is itself some sort of a pattern, with the consequence that, what has been said so far about patterns of instances, can possibly be said about the instances themselves.

Moreover, an instance is detected as the expression of a dimensional value, if it is recognized as equal to the possible sample of that dimensional value, where "possibility" is here the promise of an experience, the same kind of promise that links the locution "the sound of a piano" to the experience of listening to a performer playing something on a piano. A sample of a value in a dimension is therefore a fragment of the sound flow, that we can instantiate in sound, so that we can compare it to another fragment of the sound flow, in order to make sure that if both were occurring in immediate succession, the resulting configuration of the two fragments/instances would fall in the CR.

Up to now I have dealt with the recognition of units in a level as a segment within the same pattern structure. In the attempt to go beyond the context of a "local" structure, Oackelford distinguishes between *intra-opus* and *inter-operative zygonic* relationships. The former being related to the formation of patterns within a single music piece, such as a repeated motif or a theme, and the latter being involved among the recognizable patterns that can be found in different music works, such as the Alberti Bass but also the form of, let us say, a *minuet* (Oackelford, 1996, p.38). On the other hand, Candace Brower (2000) in her nearly unique attempt to develop Howard Margolis' *pattern theory* (Margolis, 1987) in musicology, lists *intra-opus patterns*, *musical schemas* and *image schemas* (Brower, 2000, p.324). While the first kind of patterns have to be identified with configurations connected to Oackelford's *intra-opus zygonic* relationships, Brower's interest in finding embodied meaning in music brings her to distinguish patterns that are common to different works, also referred to as *inter-opus* patterns, between schemas that refer to music conventions and schemas that, while being part of music conventions as well, are "abstracted from bodily experiences".

Image schemas in particular, by mapping bodily experiences of the physical world onto music, carry out the metaphorical concepts (Lakoff is here implied) of *musical space*, *musical time*, *musical force* and *musical motion*.

By developing and combining *image schemas* with each other, Brower pushes their range of implications up to include melody, harmony, phrase structure and form.

I will prefer in this context to use the term *intra-opus* patterns for addressing patterns that are introduced/discovered and repeated/recognized within single music works and the term *inter-opus* patterns to refer to patterns that are instantiated/recognized in different music works. These can be elements that are also repeated/recognized within the single pieces, such as the Alberti Bass, or can involve larger time spans and higher pattern articulation such as music form.

While the connection of the concepts of music theory with somatosensory experience is indeed full of interesting implications, it has to be noted that the verbal description of experience should not be confused with the institutional formalization of a practice, in the same way as the rules of music composition in some cultural context are different from their actual and individual realization by a composer, which also follows unwritten, and sometimes unwritable, rules. Metaphors are themselves linguistic patterns that can be used in addressing subjective experiences, as we have seen, but they are definitely different from the subjective experiences they label on one hand and with the practices that entail the experiences on the other. The concept of "harmonic tension" between a dominant and tonic chord indeed shows the use of a somatosensory metaphor when labelling the practice of connecting the two chords by justifying it in terms of the characterization of the subjective experience of listening to the result. Nevertheless, if we would try to follow the somatosensory metaphor of looking for the place and shape of that "tension",

we might be tempted to build a metaphysics of harmony, such as Hugo Riemann's opposition of dominant and subdominant as expression of respectively harmonics and subharmonics (Riemann, 1991). In fact, the only way to deal with "harmonic tension", without relating it to some "natural causes" with the consequence of considering the music expressions that contradict that "tension" as unnatural, is through the cultural and social practices that have involved it in history and in the consequent expectations that have been statistically learned from the attendance to these practices.

Huron's ITPRA model of expectation, as we have seen, involves some rewarding mechanism involved immediately after the confirmation of an expectation, in the *prediction* and *reaction* responses, in function of the degree of predictability of the expectation. As a consequence, we meet the *exposure effects*, stating that the probability of satisfied predictions are proportional to the intensity of the reward. This means that the formulation of a correct expectation is open to the influence of the *misattribution* effect, that is the "tendency to associate the emotional state with whatever salient stimuli exist in the environment" (Huron, 2007, p.136) whenever we experience a strong emotion. Huron's proposal is therefore that every time a prediction is satisfied, on one hand the predicting subject is awarded in such a way that he feels a possible violation of the prediction as unpleasant, and, on the other hand, he may tend to wrongly identify the pleasure of the satisfaction as a property of that particular prediction. The whole 9th chapter of Huron's *Sweet Anticipation* is devoted to the "venerable topic" of tonality and indirectly exposes the

foundational theories of western harmony as the result of an emotionally driven *misattribution*. But, we should not look for its effects only in music theorizations.

In fact, the listening analysis of both Oackelford and Brower, together with a very long list of musicologists and researchers, does not consider what happens outside of (mostly, western) music as relevant to understanding listening, and in this way they end up excluding from the analysis of listening its most relevant part.

8.8 Objects

Finally, we need to bring the constructive path begun with Chapter 7 to some advanced consequences, which includes the perspective of relating our world of objects to the complex structure of pattern that each one of us, yet entangled in intersubjective interactions, has built through *statistical learning*.

First of all, we have to put music back into the sound flow it belongs within: if we come out of the music context, little remains of the distinction between *intra-opus* and *inter-opus* patterns. In fact, when we consider that the music patterns we find as notational symbols in the score are intended to be instantiated in performance, we can connect the recognition of patterns that we find in the scores of different works with the recognition of the related patterns that are distinguished within the sound flow when we listen to those different works in our life. The difference between *intra-* and *inter-opus* patterns are therefore based on a shift in the levels of the pattern structure

between the level of the composite unit of the single "opus" and the higher level of composite units that instantiate the *opus* pattern within the sound-flow. But the onsets of the instances of the *opus* pattern point to distinctions that can be very distant in time from each other. If all the instances of *inter-opus* patterns belong to the same pattern structure, we could say that the onset of each instance is included among the onsets within the composite unit including the instance as the composing unit that falls in the CR along the binary dimension of "music presence".

It seems uselessly complicated to consider what happens between the attack of a music work and the next time we listen to a music work as the configuration of a pattern, because in the time between the two music instances everything can happen. Moreover, the same can be said about *intra-opus* patterns. In fact, between two instances of the same chordal pattern, such as the *authentic cadence,* everything can happen in a music work. But this is the consequence of avoiding dealing with objects, which need spatial bounds in order to be dealt with, in a sensorimotor way. Patterns are the hypostatization of instants that bind our expectations one after the other in a tremendously articulated way, but still without losing their instantaneous character.

In the wide context of the sound flow there is therefore no distinction between *intra-opus* and *inter-opus* patterns, but rather a possible difference of temporal range in which it is possible to detect instances of patterns. From the mere point of view of the most general pattern structure, or the "total personal sound

flow", the sound flow in which a given person has been immersed during his or her whole lifetime, there is little difference if patterns can be found within the highly composite unit that we call a music work or in a wider structure that includes the recognized instances of a musical genre, but also of a particular "sounding" verbal expression, in a given moment of life of the listening subject. Nevertheless, we are used to dealing with pattern instances as objects, not only because our language of nouns constrains us to this, but also because sounding patterns drive our somatosensory metaphors.

In fact, when we deal with a sound object, we usually express it as a "bounded pattern", a fragment of the sound flow that is distinguished as a "chunked" unit in the dimensions different from those making it a component of a unit in the higher level. In other words, we are dealing with qualities that are perceived as not changing for a while, in such a way that the fragment characterized by the non-changing qualities can be isolated from the sound flow by the discontinuities that bound it. A note, as a sample of the "note" pattern, is usually expressed by a tuned sound that is attacked from "silence" and is abruptly returning to "silence" when it is decided that it has been held enough. This exhibited discontinuity at the "time borders" somehow mimics the discontinuity of matter in space and therefore seems to be suitable to be considered as a proper sample.

But the "real life of patterns" is usually completely different. A thematic pattern is usually not bounded at both "extremities" because what is important in the

first place in the music flow is to recognize the beginning of a pattern so that it can give rise to a distinction. In fact, if we change the head of a motive we might not be able to recognize it any more, while the end is often varied within a music piece, without affecting recognition.

With the possibility to instantiate a pattern as a bounded one, we have granted the possibility to objectify patterns and to label them. As a consequence, everything we can name by listening is a pattern or better: it is a composite unit made of composing units that define it by changing some dimensions in the SIR.

Usually, in the grammar and the theory of the western music tradition, patterns are addressed differently according to the disciplines they are related to, such as harmony, form, analysis or the different instrumental practices. Sometimes discipline specific lexicons are so separated from each other, that it may happen for a same word to be connected to completely different patterns, as it happens with the very ambivalent term "cadence". "Pattern" itself, as we have seen, is sometimes used with different meanings in music, and therefore in music theory the word "schema" is often preferred to it.

Nevertheless, *schemas* are usually referred to units that are in a given level within the composite units that instantiate the "opus" pattern (Gjerdingen, 1988). As I am dealing with the sound flow as a whole, patterns might be also present at higher levels of the "total personal sound flow" as well as, like we

will see soon, at levels that are lower than those dealt analytically by traditional notation.

We ultimately have patterns of all kinds. First of all, the named patterns that correspond, in case of music, to the technical terms of music theory and grammar. We could start, for example, from the biggest forms, with the highest number of levels, such as cyclical forms, and proceed to explore all the terms that define the sounding elements, which all together give rise to the final result of the cyclic form. From "expositions" to "metres", to "themes", "motivic cells", "chords".

Secondly, we have patterns without a name, which are in most cases the instantiation of named patterns: themes have specific roles within specific musical forms, but a specific theme is the typically unnamed melodic pattern that on one hand instantiates the *theme* pattern and on the other hand needs to be instantiated several times within the specific segment of music, as a requirement for being a theme. But from the point of view of the sound flow, not all the specific themes are instances of the *theme* pattern, but all the specific themes that have been instantiated along the "total personal sound flow" are connected within the specific music works in which they are instantiated as composite units that instantiate the *theme* pattern.

The same is true for the infinite multitude of unnamed accompaniment patterns, such as the Alberti Bass, which populate the vast majority of the music works inside and outside the western tradition. However, there is no

name for them as a general phenomenon, besides the simple and most general word "pattern" and some specific denominations that are related to specific genres, such as the "Amen break" in electronic dance music, "cascara" in Salsa music and "Alberti Bass" in Classical period music.

I do not want to delve into the details here of an analysis of the different types of patterns in music and in the sound flow in general, as I am more interested in this context in proposing a general narrative that is open to composition and analysis, rather than in accomplishing the impossible and contradictory task of building a general theory of music.

Nevertheless, I still would like to discuss a few interesting cases of patterns in order to deal with pattern objectification.

The first type of pattern is related to terms expressing basic elements that we are accustomed to taking for granted as primitive concepts, such as "note" and "timbre". It is somehow unusual to express what we experience as a whole in terms of a composite unit, made by ordered components. But "if the listener is to correctly anticipate the progression of acoustical events, then she or he must somehow bracket or segregate two different sets of expectations" (Huron, 2007, p.204). So, for example, it is often difficult to recognize the sound of a piano, when represented by a note played backwards. In general, the recognition of musical instruments depends quite strongly on onset transients and on the temporal structure of the sound envelope. The research on sounds with temporally asymmetric envelopes has shown quite convincing results on

the existence of patterns that act under our threshold of consciousness. Listeners were presented in an experiment with a sine wave that was modulated in amplitude by following an exponential function, in such a way that the final sound was presenting an abrupt raise (a "ramped" sound) or, in the backward version, a slow decay (a "dumped" sound).

The experimenter

> reported that the ramped and damped sounds had different qualities. […] The damped sound was perceived as a single source rather like a drum roll played on a hollow, resonant surface […]. The ramped sound was perceived as two sounds: a drum roll on a non-resonant surface […] and a continuous tone corresponding [to the sine wave]. (Moore, 1997, p.248)

We should not be surprised at the existence of unconscious patterns because our narrative of cognition cannot but be based on our narrative of the self, which is not capable of reaching the meta-position of the observer in order to detect what we are missing in our descriptions.

It is therefore almost inevitable that unconscious patterns, as the indicators of a world that we cannot analyse in terms of articulated experiences, are constituted as objects/words, as a consequence of the *misattribution* effect, in such a way that it has become almost impenetrable by our analysis: it is not easy to think a note as a process instead of as a thing.

When a pattern is objectified, it becomes available to enter different domains of descriptions. A note can be listened, written, played, read and all the actions

seem to share the sound, the notation and possibly also the performing gesture as subsumed under the same word "note".

This ambiguity, which consists of the violation of the descriptive boundaries between domains of descriptions, is therefore a consequence of weakening the link between a pattern and the dimension that defines it, because it has entered the realm ruled by sensorimotor metaphors, the realm of objects.

But we have never exited such realm because, even if in the "isolation tank" mode we have pretended to act as an "ear without body", we are in fact immersed in the synesthetic continuum, so that we might extend the idea of regions of modes in distinguishing streams of senses, according to the correspondence of simultaneous rhythms in different qualities of experience.

If we accept the narrative that enacts the separation of senses and turn to the relationship between visual and sonorous rhythms, *identity* (CR) would therefore be entailed by distinctions in vision and sound matching with each other in time, so the two sensorial streams end up being connected as one sounding object.

This multi-sensorial pattern is at the base of Chion's idea of *synchresis* as "the spontaneous and irresistible weld produced between a particular auditory phenomenon and visual phenomenon when they occur at the same time" (Chion, 1994, p.63).

Diversity (PRR) on the other hand results from non-matching distinctions in the two sensorial streams, so that the instances presented have no connection with each other.

Finally, d*ifference* (SIR) connects the distinctions between the two sensorial streams as a composite unit made by the two ordered distinctions, in this way related with each other by an interpretation that depicts, for example, a causal correlation.

What interests me here is that objectification is directed by sensorial perception as a whole so that the sound pattern is directly linked to the sensorimotor phenomenic domain, without the need to go through a sonic bounding process. This is the case of "naturalized" sonic patterns that are labelled with reference to objects apprehended through other sensory modes such as vision. We have seen in the beginning of the text how "timbre" in acoustic instrumental music is often described in terms of the instrument that produced it. It is also the case of objectified patterns such as "notes". In the very moments they are dealt with, they are connected with visual-sensorimotor objects, such as the notational symbol or the instrument that plays them.

This is ultimately true of all the sounds and noises that are normally not considered as music, unless they end up in a music performance of *soundscape music,* in which the audience has been convinced to discard their appearance as objects in the world to become what they already were: interference patterns resulting from the process of distinguishing ourselves as a figure (a diversity) from a background (an identity).

In the process of being naturalized or objectified, sonic patterns have acquired an indexical property so that whenever I instantiate (in language, in vision or

other phenomenic domains) the objects connected to them, I can instantiate the related expected sound pattern as a specific anticipation, that is "the subjective experience accompanying a strong expectation that a particular event will occur" (Huron, 2007, p.409). This communion between senses has been realized in the first place because sound and vision are part of a wider synesthetic continuum and therefore it is already there, connected with each other as two notes of a chord that we have (statistically) learned to expect.

The indexical relationship between object and sound is not that different from the ostensive gesture that associates a word to a specific action in the world. But a word is a sound pattern in the first place, which connects with other units that together form composite units and finally realize an entire pattern structure with many possible levels that are composed with rules, that, while having more constraints and therefore being more complicated are not all that different from the ways that we compose music. In the context of pattern structures within the synesthetic continuum, a sentence can be described as a naturalized pattern in the sonic-sensorimotor phenomenic domain, which is connected with a pattern in another enacted sensorial domain in such a way that their discontinuity falls in the SIR, so that the connection between the two patterns is interpreted as an implication that we address as *meaning*. In the history of our exposition to language, we have reached such a strong habituation to the composite pattern made by the two sensorial patterns that we have objectified it and have started to think that the pattern in the sonic-sensorimotor domain is identical to the pattern in the other enacted sensorial

domain, therefore evoked in the same way as, when we see a steam train, we can easily imagine its noise.

Music is ultimately not a language, but it is rather language that is built on the same matter music is made of. If language is a kind of music, it is therefore an objectified and naturalized one that we have forgotten to listen to as such, so that we rather direct ourselves towards its evoked ghosts of anticipations.

8.9 Meaning and emotion in listening

We have started, in Chapter 5, from a sound flow that was split by a well-trained listener, although constrained as a "brain in a bowl", in music emerging from the "noises" of a full theatre. Then we began a top-down path in which the listener was shown as able to distinguish details that were increasingly miniscule and were meant to show the pervasiveness of our ability to distinguish when it is well trained by a practice that is culturally situated. In Chapter 7, we followed a bottom-up path that, moving from the most elementary experience of listening, distinction, engaged a constructive approach in order to build increasingly complex experiences. When we introduced the pattern, in Chapter 8, we reached the tools to bring that simplest experience of distinction to account for the complexity of distinguishing a world of sounds in the sonic flow and, in perspective, to distinguish a world of objects in the synesthetic continuum.

At this point, I would like to take a step back from and dedicate a few words on how the narrative I proposed deals with musical meaning. Before doing this, I

need to summarize some important steps in the history of the discussion about the semantics of music. In fact, this topic has always been one of the most debated problems in the history of discourses about music, almost as if the music's sonorous sister, that is natural language, was always trying to impose on music its logocentric version of naturalization by looking for some possible referent that could be linked to verbal descriptions of listening experiences, or to the terms of music theory and notation.

This is still true today, with the introduction of a more culturally oriented musicology, the so-called "new musicology", and the consequent fragmentation of musicological research in pursuing local contexts as rooted in social structures and leading to a possible premature death of a theory of music in place of a sociology of music.

On the other hand, the criticism has pointed out that the assumption according to which the

> social structure has some kind of objective existence, which is represented through homology with the patterns of music […], giving in this way the "impression that social meaning is inherent in music [should be supported by an investigation on] how particular pieces of music might support particular meanings, and indeed whether there are constraints on the meaning that any particular piece can support (Cook, 2001, p.173),

that is by a general theory of musical meaning.

In fact there are at least three reasons to discuss the status of musical meaning, the first one being that music, as early as the beginning of the

263

musicological tradition in the half of the 19th century, has been regarded as a

strictly formal phenomenon, as we shall see, in spite of the opposite popular

belief.

Secondarily,

> common sense often assumes the production of meaning in verbal
> language as an appropriate model, while it is rather obvious that, with
> respect to such a restrictive definition of meaning, music is not something
> that 'speaks'. Finally, even in case we can admit that a certain musical
> work has a 'meaning' or is characterized by intersections of different
> meanings, these appear to be too evanescent and fluid to deserve a place
> at the same level of direct constitutive elements of music. (translated
> from: Nattiez, 2002, p.206)

If we examine how musical meaning has been dealt with from the historical

point of view, we can easily notice that since the very beginning of the

discussion around music, musical aesthetics seemed to oscillate between

"conceptions centred on music in itself, that privileges its autonomy, and

conceptions that emphasize it as a carrier of extrinsic meaning" (translated

from: Nattiez, 2002, p.206).

On one hand, according to Pythagoras,

> music is the shadow of numbers and there is a certain correspondence
> between the harmony of universe and the harmony of music [so that,] as
> the soul is also harmony and numerical ratio, music can similarly influence
> and purify it. (translated from: Nattiez, 2002, p.208)

On the other hand, Plato, while maintaining the continuity with the pre-Socratic

Pythagorean tradition, limited himself

264

at discussing [the] ethic and aesthetic character [of music] and its role in the *paideia*, the Greek education; for Plato, music was considered more a (perhaps pleasurable) threat than a virtue for a healthy society. (Oliveira and others, 2005, p.46)

Plato's concerns are mainly addressed to connecting music with social behaviour, which thus becomes, so to speak, its reference, even if a theory of meaning was defined only half a century after his death. But if musical meaning is implicitly or explicitly involved in almost any work about music and philosophy, we have to wait until the 17th century to have the first explicit debate about music relata. The reference is now changed from *ethos* to text, and this is also due to a change of the social role of music, together with the exceptional importance that was granted to vocal music, as a consequence of its usage within the rites of Christianity.

This debate opposes the followers of the so-called *seconda prattica* to the followers of the *prima prattica*.

For old composers, who were all polyphonists and were faithful to the traditional laws of counterpoint, the text was meant to be just an occasion, or at most a cold structural platform, for polyphonic composition; for modern composers, both in the context of polyphony and monody, each compositive choice - even the most transgressive one - should have been absolutely determined by poetical content. (Translated from: La Via, 2007, p.31)

When the most important function of music was to provide a support for sacred texts, the main concern in dealing with words was clarity, in order to support faith. And this was true also, in spite of the cases that did not fulfil such request, as in the polytextual motet.

Once set free from dogma, the texts on which music was still mainly based started to demand to be accompanied by a music that was able to follow their content in order to enhance their expression.

La *Camerata de' Bardi,* just a few decades before the debate between the *pratiche*, develops an explanation for the connection between music and text, that relies on Plato and claims that the

> power of music to arouse the human emotion lays in the representations, by melody, of human speaking voice when expressing the various emotions. (Kivy, 2002, p.17)

Ironically, the problem of following the text when composing music was drawn to public attention, meaning that it was no more under the decree of the Church, just at the point that vocal music started to give way to instrumental music, where the problem lost direct relevance.

With the dismissal of text as responsible for a reference to music, the only candidate left in the equation, so to speak, was emotion. But first of all, a theory was needed that would justify the position of music within fine arts, under the criterion of representation, now that text had been side-lined.

The first attempt to found a new paradigm can be traced to Kant's "Critiques of Judgment", according to which music, with its representational content, is still included in the realm of fine arts, but it is "unable to excite a chain of aesthetic ideas by engaging the free play of the imagination and understanding" (Kivy,

2002, p.59) because it is addressed only to the body, in the form of its expressive properties, namely emotions.

From then on, music offered itself as the art of emotions. It was no longer in connection with cosmos and man through numerical ratios, no more *mundana*, but rather *humana*, to the point of being considered a secondary art that, according to Kant, was closer to ornamentation than to painting. But at the same time, and thanks to its distance from the aim of representing earthly objects, music is attuned to emotions and asked to represent as a "copy" of the will (according to Schopenhauer).

The privileged connection between music and emotions received its first aesthetic strike only in 1854, in the form of a little book by the Viennese musician and music critic Eduard Hanslick, *Vom Musikalisch-Schönen*, who earned the role of representing a whole legacy of musicologists and music theorists who wanted to recover the other pole of music thinking, which took the name of *formalism* after Hanslick.

In short, Hanslick's thesis was that the ability of music to evoke emotions is just a collateral feature because:

> a perfectly expressed musical idea is already an autonomous *beautiful*, it is its own purpose and not just a means or a material used to represent feelings and thoughts. The content of music is *forms put into motion through sounds.* (Hanslick, 1891, p.32)

In a certain respect, by refusing the idea that music can have a precise emotional content, Hanslick seems to turn his back on centuries of confidence in the semantic possibilities of music, including vocal music. On the other hand, his work has been interpreted as directing the reference of music to music itself (what Meyer calls "inclusive meaning" and Nattiez calls "intrinsic meaning") instead of directing it to elements of different nature (the "referential" or "extrinsic" meaning). Nevertheless, Hanslick's ideas seemed to anticipate the musical formalism that became an important feature of 20th century music, from Schoenberg (who claimed music speaks in its own language made of mere musical matters) to Varèse, Webern and Stravinskij (who held that expression was never immanent to music).

Leonard Meyer tried to mediate between the two tendencies, starting by articulating the positions of theorists with respect to the music reference as "absolutists" (those who insist that musical meaning lies exclusively within the context of the work itself) and "referentialists" (those who contend that, in addition to such abstract intellectual meanings, music also communicates meanings that in a certain way refer to the extra-musical world of concepts, actions, emotional states and character").

One the other hand we have formalism (musical meaning is primarily intellectual) and expressionism (musical meaning is primarily emotional). These four categories create a square of possibilities among which Meyer chooses for himself the position of an absolutistic expressionism.

Emotions are present, but they are generated by the formal structure of music. The question is: are we still talking about emotions? Meyer relies on a behaviouristic update of John Dewey's theory of emotion when he states, "emotion of affect is aroused when a tendency to respond is arrested or inhibited" (Meyer, 1956, p.14). Moreover, he claims the existence "of intangibles, non referential affective states experienced in response to music" (Meyer, 1956, p.20).

In the context of this text, I prefer to characterize the debate around musical meaning in terms of "objectivism", which looks for musical meaning in an abstraction of the measurable properties of a *neutral level* such as the score or the sound recording; and "experientialism" (Lakoff and Johnson, 1980), which relies on experience in terms of the description of the experience of listening.

A third tendency has emerged more recently that is usually traced back to Adorno. The claim that music

> presents social problems through its own material and according to its own formal laws lies at the heart of his approach; Music contains these problems within itself in the innermost cells of its technique. (Adorno, 2006, p.393)

According to Cook, as the tensions and contradictions of society are defined as technical problems, then social meaning can be decoded by the appropriate analysis of musical texts. While it must be pointed out that Adorno never ceased emphasizing the autonomy of art in the dialectical

opposition/integration of autonomous work and social context, the development of his ideas within the analysis of musical works, especially in the context of what was named "new musicology", have often been criticized for letting an ideological view surpass the musical text. And in fact, the paths of study around musical meaning can be mutually exclusive to the point that, as highlighted by Cook, there may be no reference to the research in one direction in research that follows another direction (Cook, 2001 p.176). Cook himself, in his article "Theorizing Musical Meaning" omitted entirely research following the "experientialist" path.

In the end, if *objectivism* relies on the score, while *experientialism* relies on subjective experiences of looking for meaning, we have a third direction, that considers both the neutral levels of music and the experience of listening as entangled in social practices, so that, in semiotic terms, music semantics has to be looked for in the pragmatics. But when we take the perspective of the ontogenetic constitution of the synesthetic continuum in the "total personal pattern structure" possibly objectified and sensorimotor-visually coupled (meaning that it has been naturalized) on some levels, there is no place for an ontology of meaning. In fact we need to take "meaning" out of a hyperuranic world with no space and time, and bring it back to the "praxis of living" in order to realize that "meaning" is a rhetorical device, an objectified pattern, that we involve when describing a specific action to which we are highly habituated. This action, as I said earlier, consists of filling an objectified composite unit that

is realized by a pair of units in different phenomenic domains, one of which is missing, with a specific expectation of the missing instance in the form of anticipation. This action is not necessary, so it is not realized whenever we meet such an incomplete unit: we do not necessarily have to imagine the noise of a train whenever we see one. But we can engage the anticipation, whenever we need it. The same can be said about language; and when we talk about meaning we are always talking about language. But meaning in terms of anticipation is not always present in language: on the contrary they are there only when that specific action is required by the context. On the other hand, syntax is almost always present as the rules of connecting words, as patterns, within a complex system of liberties and constraints.

Again, meaning in terms of an act of anticipation is eminently related to the sounding instances we use to describe our experiences of listening. But, if we can definitely connect some instances in the music flow to instances in some other phenomenic domain, such as the vision of the performer outside of the need to deal with listening experiences through description, we mostly do not rely on verbal connections with aural experiences in our social interactions, at least if we are not musicians.

Therefore, in the context of my narrative there is no difference between semantics and pragmatics: this can instead be found in the domains of description that are involved in dealing with the experience of listening in language. As a consequence, music does not need the idea of meaning, because, the experiences of listening do not need language in the first place.

As Huron has shown, emotional responses merely need a system of expectation to reveal themselves. Music does not necessarily need to rely on words to stimulate emotional responses because the system of expectations, with its related emotional connections, articulates and evolves along the whole "total personal sound flow". Moreover, music does not exist per se: it is always involved in a certain type of patterns, habits and rules that necessarily connect it to experiences in other sensorial domains, and to verbal descriptions that come from social interaction with their own repertory of emotional responses.

It is therefore difficult to detach the experience of listening from descriptions that involve metaphors, objectification and ultimately the naturalized constraints and liberties, the ontology that we have built during our entire life. Nevertheless, we can deal with them by engaging an explanatory description or a narrative, as I am trying to do with my research, which may help us to get rid of the objectification, on one hand, by means of proper distinction of the domain of discourses, and of naturalization on the other hand, by means of the reintroduction of the subject in our discourses, in order to possibly escape the processes of subjection and engage non linguistic practices of listening and sounding.

9. CODA: NOISE PATTERNS

In this final section of the second part, I would like to address a specific listening case that can be approached in terms of the proposed narrative.

Let's modulate in frequency a sine wave centred at 440 Hz, with a random signal that continuously interpolates random values that are chosen with the frequency of, let say, 500 Hz. When we slowly increase the range of random values so that the frequency of the sine wave is stable at 440 Hz at one extreme and changes randomly between 1 and 1000 Hz at the other extreme, what we hear is a crossfade between the sine wave and some coloured noise, instead of hearing a continuous transformation between a sine and a noise. The problem arises because we can distinguish two overlapping streams in the middle of the process instead of just one—the sine wave and the noise—in spite of the fact that the process of synthesis only addresses one sound stream, the modulated sine. The perceptive illusion, so to speak, is relevant because at slower frequencies of random generation, let us say at 200 Hz instead of 500, the transformation process is perceived as a continuous and unitary one. At no point is there the impression of hearing two overlapping sound identities.

Before going into the details of analysis, it is important to distinguish the different definitions of "noise", so that we can properly address the problem.

According to Nattiez (1990, pp.47-48), the word "noise" typically relates to any sound that shows an unpleasant, unacceptable character, whatever that means to us. This is indeed a definition that is in the first place based on a subjective uncontrollable sensation that is related to thresholds of tolerance and can lead to physical suffering. We can therefore say that it is a *vegetative*[26] use that is often informally expressed as an opposition between *noise* and *sound*. Nattiez also mentions the case "in which certain sounds accepted as 'musical' by the composer are classified as 'unpleasant' by the listener (1990, p.46)". We might say that the unpleasantness of noise is connected here to sentences like "this doesn't say anything to me, it's not music: it's noise". If a sound "doesn't say anything to me", I can only relate it to the cause the produced it. In this case, *noise* seems to therefore be connected to what we called in Section 5.2 the *extrinsic meaning* of a sonic event, which frames the sonic event as the effect of some cause, as opposed to the *intrinsic* one, which is rather related to the appreciation of sonic properties and relationships, as it typically happens with a music work. As being connected to judgments regarding the appropriateness of describing a sonic event or even a concert performance as music, we can characterize the use of the word "noise" as a *semantic* one, in *which* noise is opposed to *music*. While this second use is often conjoined with the first one, it is not necessarily a consequence of it.

We now encounter two technical usages: the first one originally comes from electrical engineering, but is also widely used in information theory and signal

---

[26]The term "vegetative", as in the "vegetative system", is here used to address the uncontrollable and instincual character of the unpleasantness of noise

processing, while the second one is mainly applied within the branch of physics that studies sound, namely acoustics. As the two definitions may share a lot, mostly due to the fact the acoustic measurements are realized by electronic devices, we will focus on the differences between the two usages. According to the first technical definition, *noise* is "any interfering signal adversely affecting the communication of the clean signal" (Dorf, 2000, p.406). This means that *noise* is opposed to *signal* and has to be considered connected to the intention of transferring a "clean signal" between an emitter and a receiver. We can thus consider this definition to be a *finalistic* one. On the other hand, acoustics, in the attempt to be consistent with the related field of architectural acoustics, explicitly defines *noise* in terms of a subjectively unpleasant noise (Nattiez, 1990, p.47), but it often takes for granted that *noise* can usually be described in terms of oscillations that occur in a statistically aleatoric way (or in the corresponding terms of a continuous spectrum). By following Nattiez (1990, p.47), we attain a *physical* definition that opposes *noise* to *sound* as aperiodic versus periodic variations in the pressure of an elastic medium such as the air.

Among the different usages of the word noise, from the psychological one, to the semantic one, from the finalistic one to the physical one, the only definition that is not depending on a context, be it social, cultural, psychological or intentional that is related to the listener is the last one. *Noise* or better the phenomenon that represents it at best, *white noise*, is the way we label

275

something that is physically measured as the most irregular variation of sound. So, *white noise* is the most variable sound event and yet we have an aural experience of it as quite stable. How is this possible?

There is a simple answer. In white noise there is something that is statistically stable, at least to a certain extent: the spectrum. This is decoded by passing through the non-linear peculiarities of cochlear organic processing and is transmitted to the brain by the organ of corti, with its hair cells. Nevertheless, we do not really have a direct experience of the spectrum, as this is essentially a mathematical abstraction. Moreover, as we have seen, the information of the sound pressure wave is maintained by the internal ear and is also transmitted to the brain, so, even if we accept the connection between the description of an experience of listening and the physiological measurements of the auditory system, there is not a direct and easy link from the variations of the air pressure to the experience of it.

According to the definition of *pattern*, the acoustic event of *white noise* is something that we can recognize by comparing the memory of a previously experienced *white noise* event with what we are listening to in the present. But what do we recognize when we listen to two samples of white noise? If we choose to rely entirely on the experience of the sound flow, we can answer that the high statistical dispersion of unconscious expectation is what characterizes the perceptive statistical dimension related to the *white noise* event, in contrast not only to sound, but also to other events that are related to different

276

probability distributions such as *coloured noise*. Yet, the duration of the single instances must be very short because if we reduce the frequency of the random generator from 500 Hz to 100 Hz the specific sensation of noise transforms into something completely different: in a sort of fast chirp. This means that in order to recognize a *coloured noise* the instances need to vary within the statistical range that defines the *colour*, at least every two thousandths of second.

Of course, it is very difficult to find in the vocabulary connected to listening experiences a dimension that is supposed to be distinguished at a speed that is definitely not only under the threshold for distinguishing two successive attacks, but also under the threshold for detecting pitch.

It seems therefore that the perception of (unconscious) patterns is engaged with very fast clocks, so that the sensation of *coloured noise* can be related to the distinction of a composite unit, whose components fall in the CR in a unnameable dimension that we could call the *click* or *impulse* dimension, a dimension in which we cannot distinguish anything but a gestaltic quality, related to variations of a continuous spectrum. But we can assume that this very fast succession of distinctions is likely always present, as the surface level on which all the dimensions, all the pattern structures, are built and learned.

As a consequence, the experience of a cross-fade between a pitched stream and a noise stream can be interpreted in terms of a PRR of a statistical dimension because we can split the sound continuum in shorter instances, so

that not only a very fast succession of instances is thinkable, but even a statistical dimension that, due to their statistical character, needs to involve the instantiation of several sub-units can foster distinctions.

The transition that we experience as a cross-fade is therefore describable as the PRR alternation of instances of a statistical dimension that capture large random deviations from the centre frequency and instances of the same dimension that are strict enough to capture small random deviations. As we can observe from Jones' map of perceptive regions, at very fast speeds close to that of the *impulse* dimension, PRR is engaged even with small discontinuities, so that the description of the experienced cross-fade between sine wave and noise in terms of a PRR seems viable. The perception of the centre frequency as a stream is ultimately the signature of a cognitive process that tries to maximize expectations events at very fast unconscious rates of distinction.

Finally, starting from the idea of a multi layered pattern structure of distinctions based on statistic dimensions and turning to the phenomenon of *white noise*, we can extend our observations to higher levels of pattern structure that are realized by low degrees of precise expectations.

In the end, the sonic identity of *white noise* seems to be characterized through the idea of the *indistinct*. At the lowest level of the pattern structure we do not even detect discontinuities, but on higher levels, the *indistinct* implies that we cannot distinguish the patterns that populate the lower levels of the structure

from each other. For example, if we listen to our heartbeat we can detect the typical repeating pattern of two close beats, but usually we are not able to find strong differences between each repetition of the pattern, unless we devote a pretty strong effort to this task. But as soon our attention fades, we quickly get used to it until we no longer hear it, unless we enter an anechoic room and the heartbeat again claims its presence in our sonic world. This happens for sonic events that have a constant presence, such as computer fans, the buzz of the alternate current or the noise of a stream of water we live nearby, what Schaffer calls in his analysis of soundscape a *keynote sound* (Schafer, 1977).

This happens for everything that falls in the CR for long enough in some level of the pattern structure. Therefore, the utmost unpredictability of white noise on higher levels becomes the utmost unpredictability of certain patterns that need to be defined by statistical dimensions. Let's consider traffic noise. It may be highly variable in the sonic content and it is unpredictable because its structure depends on the highly unpredictable structure of life: we cannot usually predict what vehicle will pass through the street at any given moment and which precise sonic event it will produce with its engine. And we don't care. But we can easily recognize the noise of a car, and we also know sound patterns like an accelerating car and the screeching of the brakes very well. Some people even develop the ability to recognize particular car or motorbike models by their sound and the engine noise can even become "music" for them.

However, the statistical character I addressed to deal with the experience of "noise" is not only typical of acoustically noisy events: the same happens for

pitched sounds as well, as in the case of people talking in another room or with the sound produced by a drove of oxen with their cowbells.

In general, these unpredictable layers of the sonic flow are put in the listening background and labelled as noise. And this term is more appropriate than it may seem at first, exactly because, as with *white noise*, the experience of the very low degrees of precise expectations of these sound textures is a perceptive dimension that falls in the CR of the pattern structure. It only happens in higher levels of the structure than the ones involved in distinguishing *white noise*. This is the domain of the *indistinct* with its character of "predictable unpredictability", which corresponds to what we called the *extrinsic meaning* of sound events, which is extrinsic merely because the only criterion of expectancy we possess, as a cultural practice, is demanded by real life contingencies. This case seems to express the *semantic* meaning of noise, that is something that does not say anything to us, apart from its functional extrinsic relationship with the object that produced it.

On the other hand, music is "organized sound" that concedes to us higher degrees of structured possible expectations; it is what deserves a proper attention and allows a satisfying relationship between expectations and results. This is what is said to exhibit *intrinsic meaning*.

When something that we are not used to conceding our "intrinsic" attention to claims its right to resound in front of our ears, we feel easily daunted: not only for the effort of putting attention where we are not used to, but also, sometimes, for the indignation of doing it. This is not only the case of a

psychological definition of noise, but also, for example, the motivation for the audience's sounding opposition during the contemporary music concerts of the last century. And this is due to the fact that noise is often considered as the opposite of the signal, we are mentioning here the *finalistic* definition: something that has to be eradicated in order to devote our attention entirely to something that somebody else, the emitter, decided worthy of consideration. Finally, no wonder that noise, both present in the lower levels of the pattern structure as noise sample, in the higher levels as soundscape or in the many other possible forms has been charged with a subversive character (Nechtaval, 2011). Noise always challenges us to cognitively break it down in patterns, in a process that subtracts noise from the naturalized world that surrounds us, in order to increase the consciousness of what we take for granted, in a never-ending path of statistical learning. But whatever we do to contain the lack of sense, noise will always be there, to remind us that there is always something that was left out of our patterns.

PART 3
MUSIC

10. ANALYSIS

10.1 Esthesic analysis

Much water has flowed under the bridge since the first attempts to investigate music listening. Since its very beginning, the analysis of listening has been linked to music and driven by the assumption that listening par excellence is conducted by tracing the composer's imagery in terms of traditional musical concepts. If starting from 1874, the year that marks the very first musicological research explicitly dedicated to listening, by Riemann (2010), the analysis of both music and listening was conducted by the traditional music system with its notation, at least since 1959, with Besseler's classic book (1993), we are accustomed to considering listening in terms of a reception that is not only culturally and historically dependent, but also individually modulated by circumstances and intentions.

However, it is only since 1966 that it has been possible to pursue an investigation around a possible grammar of listening, to be used with music, independently from the technical and notational practices related to its institutional aspects. With Pierre Schaeffer (1966), listening reached a sort of independence from what we could call the academic dictatorship of composition and performance and was finally able to follow John Cage's famous claim about the independence of composing, listening and performing from each other (Cage, 1961, p.15).

Schaeffer's research was indeed the inevitable consequence of the technological and technical development of music production, that, by no longer being based on traditional notation, could not but appeal to an analysis of listening in order to recover the objectivity, that was lost together with the materiality of the score.

In this respect, the most important achievement in technologically driving the process of listening is Pierre Schaeffer's concept of *reduced listening*, framed, as many of Schaeffer's concepts, as one pole of an opposition whose other side is "ordinary" listening. R*educed listening* is an attitude of "listening to the sound for its own sake [...] by removing its real or supposed source and the meaning it may convey" (Chion, 1983, p.30).

The act of removing all our habitual references in listening is a willed and artificial act (a sort of phenomenological reduction) that allows us to clarify many phenomena implicit in our perception. That's why *reduced listening* has to be learned as a process of "deconditioning" from "ordinary" listening, with exercises and the possible help of technology.

Listening repeatedly to fragments, sections or to a whole piece of music may indeed work as a sort of a magnifying glass: the interplay of memory and attention puts in the background what we can already predict, the higher CR levels, and focuses on changes in the details, the SIR levels.

Schaeffer's analysis by listening is directed, as is already clear from the title of his "Traité", to detect sound objects within the frame of their internal components, and their function as components of bigger structures, as the

result of the practice of the so-called *typomorphology*. This could be already a problem, because there are as many objects as the number of ways in which we can distinguish them. Therefore, a selection has to be made and this is already an aesthetic statement. No wonder that, as Nattiez observes (1990, p.95), Schaeffer's objects seem to be relevant mainly in strategies of composition, as "they do not emanate from the perspective of habituated ordinary hearing, which Schaeffer himself favours as very different from specialized hearing". And in the end, Schaeffer's "Traité" is therefore more similar to a handbook of composition than to a research on listening.

As the one who brought Schaeffer's work to the English-speaking world, it is worth mentioning Denis Smalley, who studied in France with the pioneer of concrete music, Francois Bayle.

In his famous article "Spectromorphology and the structuring process", Smalley (1986) introduces an important shift to Schaeffer's system under the name of *spectromorphology*, claiming it as an improvement. In the whole article there is no referral to the notion of intention or to *reduced listening*. Neither there is an intensive use of the word "object", often substituted by the word "unit", and when it is present, the concept of "object" is not defined and it is taken for granted.

There is not even a trace of some sort of *typomorphology*. Rather, a different taxonomy is proposed that uses acoustical analysis to expand the "terminology

of qualitative descriptions in order to deal more comprehensively with aspects of spectral space" (Smalley, 1997, p.119).

While appealing for the need to refer to the listening experience, Smalley always clings to the visual paradigm of the spectrum as it appears in sonograms, and to words and concepts that come directly from acoustics. In this respect, his work, while very important for the electroacoustic music analysis, is not entirely relevant to the issue of an analysis by listening.

Almost a decade after Schaeffer's classic, Jean Molino's semiotic model (Nattiez, 1990) succeeded in restoring a semantic independence between composing and listening, by characterizing the production of meaning in terms of strategies depending on social and cultural roles.

His *semiotic tripartition* changes the direction of the arrow, which in the classic model of communication proposed by Shannon and Weaver, used to directly connect the sender to the receiver in one direction. Molino breaks the communication arrow into two connectors that, starting from the two exemplar typologies of actors involved in the process of musical signification, the composer and the listener, point to the score as a *neutral level* of semiosis. The unity of signification, which in the classic model was related to the optimization of the coded transmission of messages through ether, is therefore lost in two strategies of signification, at least. A *poietic* strategy is directed to the construction of the material, and therefore *neutral*, element, the score or the support containing the sound data, according to meaningful practices. An

esthesic strategy is, on the other hand, directed to the articulation the process of listening, as inserted in meaningful social practices.

But what can be told about a music work, by an *esthesical analysis*, if each listener is different from all the others and the analysis does not want to privilege a particular one over all the possible ways of listening, the correctness of which is certified by an homogeneous set of listeners, such as the one composed by musicians and musicologists? Nattiez himself has expressed some doubts about the epistemological value of *esthesical analysis* and he often turned himself to the analysis of specific works in terms of their *neutral levels*, by following a *paradigmatic analysis* of score fragments.

As a possible solution, François Delalande proposed to open music listening to a taxonomy of "conducts" (*conduits musicales*)[27], addressing analysis to the adaptive reactions of listeners in responding to sound, as strategic behaviours that are capable of driving the listening practices and engaging, for example, emphatic, figurative or taxonomic listening (Delalande, 1993).

Specifically, *taxonomic listening*, defined as the segmentation "on the fly" of a sound flow during listening, seems, as noted by Stéphane Roy (Roy, 2003), to resemble a practice that was introduced by Schaeffer in his treatise, under the name of *gestaltic* or *morphological listening*.

---

[27]We have already met Delalande's "conducts" when we were dealing with the references of a theory of listening in Section 3.5, where it is possible to find definition by the author.

This consists in the search for perceptive units, which are hierarchically organized by the repeated *reduced* listening of some music fragments. According to Roy, *gestaltic listening* has the potential to base an epistemological value of an analysis by listening. Otherwise said: appealing to a conduct that is directed to the distinction of perceptive (gestaltically determined) units in the sound flow might be able to found an analytical discourse on music listening that is shareable. This is because it can connect the detected units to their trace in the *neutral level*. On the other hand, Delalande is sceptical of the possibility for a repeated and fragmented listening such as the *reduced* one, which is therefore extremely skilled and directed, to detect those elements that are pertinent to listening as a live social practice and are therefore detected, for example, by occasional listeners.

In the end, both Delalande and Nattiez, who focuses on the musicological consequences of Molino's semiotics, relate the analysis in terms of the *esthesic* process to the information gathered from listeners' answers to the experimenter's questions. In particular, the investigation of music conducts is drawn on a maximum of three listening sessions (Delalande, 1993, p.193). As a consequence of such a provisional relationship with the music work, the adherence with the text is likely to be lost. Consequently, the object of research seems to be more the analysis of listening, than the analysis by listening. For example, an interesting problem is related to how much a listening strategy changes depending on the number of times it is applied to

the same work. In fact, as soon as a conduct goal has reached a satisfying threshold, any subsequent application of the same conduct might result in deviating the listening to secondary goals, such as the memorization of the work (a goal that Delalande only assigns to the *taxonomical* conduct). Moreover, *music conducts* seem to rely too much on a cultural concept, the one of music, in such a way that when dealing for example with soundscape or background music genres, the number of detected conducts might increase in an unpredictable way and their targets might fade from music to all the possible targets that are addressed by listening in our general "praxis of living" (in Maturana's terms).

The idea of sound pattern may suggest an alternative approach in a listening strategy directed to music analysis than Schaeffer's and Smalley's search of a sound vocabulary of objects/words. If we focus *reduced listening* on discontinuities, instead of on objectified "bounded" patterns, we may avoid the constructive appeal of *poietics* on one hand, and the arbitrariness of individual conducts on the other hand. At the same time, the structure of discontinuities in levels, whether nested or not, emerges spontaneously without privileging a level for objects. That is because a discontinuity in one parameter (as a set of perceptive qualities that can be related to a, not necessarily strict, order structure) immediately qualifies what is there before and after it, as continuities in that parameter, which in turn can be further distinguished by discontinuities in another parameter or minor discontinuities in the same parameter.

291

In the terms of Mari-Riess Jones' labels for discriminative regions, an *esthesic analysis* of a work might therefore consist in the detection of: the elements that are relevant to break down the work in a multi-layered structure (i.e. the dimensions that change in the SIR); the elements that characterize the work as a chunk that is distinguished from what is there is before and after it in the sound flow (i.e. the dimensions that change in the CR); and, if pertinent, the elements that are useful to distinguish the work in streams (with changes in the PRR).

First of all, it is evident that detecting discontinuities by listening might be arbitrarily driven. For this reason, the adherence to the *neutral level*, is always necessary: a change must always be linked to something in the score or in some acoustical analysis of the sound flow, that may lead to the instantiation of that change, in any parameter at any degree of complexity, in the sound flow. Therefore, if a perceived change in the recording of a music work cannot be traced back to any element of the score (according to its semiographic/semiotic status), it must be referred to something else, as background activity that does not need to be accounted in a strict analysis of a music work, such as the unwanted creaking of a musician's chair. The same guiding role can be given to other *traces*, such as the spectrogram, the sonogram or the sound wave data, which represent the objective material side of a music work.

In this respect it is easy to understand why Nattiez considered *reductive listening* as part of the *neutral level*, as this kind of listening seems to avoid the variety of approaches of conducts in selecting the relevant parts of experience they react to, so that it settles in a matter-like objectivity open to semiotic interpretation.

Indeed, a score or a spectrogram must be perceived as such (for example visually) before it can be used as a score, in the same way that a sound vibration must be perceived and transformed in aural sensations before we can talk about music. Nonetheless, we are definitely in the domain of listening in precisely the way described by Delalande's paradigm of conducts as selections of pertinencies—only that pertinencies are here aural distinctions in the most general way.

We have thus a listening strategy among the others, but a special one, which is strictly linked to visual *traces* (scores and so on) and therefore can serve as a platform to build other listening strategies as functions of this *reduced* strategy, a meta-conduct.

On the other hand, an analysis is necessarily a translation of the pertinent distinctions in a phenomenic domain that is able to surpass the limits imposed by the irreversibility of time. It would be interesting indeed to communicate the results of an analysis, as a sounding instance, matching in time the relevant discontinuities of a music work that can be listened to as a simplified version of the work. Nevertheless, it's the graphic transposition of time, as Bergsonianly projected in the perceptive dimensions related to space, that is often invoked

in communicating analyses of music. If the communication of analyses has now earned the possibility to submit the arrow of time to a repertory of transformations that include reversibility and enlargements, the qualitative experiences of listening are now objectified in pictograms, ideograms and also logograms. There is therefore a *graphemology* of inscriptions (Levy, 2004), that needs to be properly instantiated as a materiality resulting from the *poietic* strategies of the analysing subject, in order to orient the reader within the interference between the specific peculiarities of the music flow and the specific peculiarities of the repertory of listening habits, that constitute the analysing subject as a listener.

In this section, I proposed how focusing on discontinuities, i.e. changes, within a layer of the sound flow, i.e. the recording of the performance of a music work, instead of appealing to the gestaltic units of morphological listening (Schaeffer) or gestaltic conduct (Delalande) has the potential to build a discourse on music listening related to a specific music work (i.e. an esthesic analysis), that is shareable, open to discussion and validation, and therefore has a potentially high epistemological status. This is the result of being able to link the experience of listening, in terms of the detection of discontinuities, to specific points on the neutral level, i.e. the timeline related to the audio track, without having to rely on a sensorimotor paradigm of sound objects.

In the following sections, I will present two examples of esthesic analysis that will show different ways to set up a possible visual representation of discontinuities in music listening.

10.2 Ligeti's "Lontano"

The first example is related to Giörgy Ligeti's "Lontano", a work for orchestra that was composed in 1967. This piece is well suited to exemplify this kind of analysis as it was conceived to study the mediation between continuity and discontinuity by using a sort of sound blurring effect, so that the task of detecting discontinuities is challenged from the very beginning. The observations about listening are related to the famous recording by Baden-Baden's Sinfonie-Orchester des Südwestfunks under the direction of Ernest Bour, which was published by Wergo.

First of all, this is a work for orchestra, that, as usual, separates itself effectively from a silent background thanks to the presence of orchestral instruments. This might seem obvious, but the deceptive use of fading-in or out respectively from or to the listening environment, as we have seen in the second part of this text, has been used several times along the history of western music as a quasi-theatrical effect. Even if "Lontano" is not difficult to distinguish from a silent listening background, it nevertheless shows a rhetorical use of such an emergence of sound from the background in the form of a slow ending fade-

out *al niente* and a fade-in-like beginning, which emerges from a solo flute and two cellos' harmonics in unison (with four *p*).

In between, a collective orchestral texture grows and transforms in a rather continuous way, giving us only two major discontinuities (bar 56 and bar 120), that, by falling in the SIR, divide the work in three parts with compatible durations: 3'28", 4' and 3'.

What is at the base of the two discontinuities is a big change in the dimension of presence, in which something ceases and suddenly something new appears. They are really evident discontinuities, which only the concert hall's reverberation can try to hide. And they appear to happen within the general continuity of the work, together with the major discontinuities of the very beginning and the very end of the work, in such a way that their chunking process bound the music flow into durations of a same degree. We can therefore perceive the whole work as realized by an ordered succession of three parts. In other words, we can *integrate serially* the discontinuities, which therefore appear to be in their SIR region.

I have said that the change happens in the dimension of presence, but in this complex context it is easy to analyse the general dimension of presence as the composite result of simultaneous changes in different dimensions. For example, we can easily detect the first discontinuity at least in its timbrical qualities (brass and contrabassoon vs. strings), in register (mainly low vs. medium) and in the unison vs. chord opposition, whose dimension we should
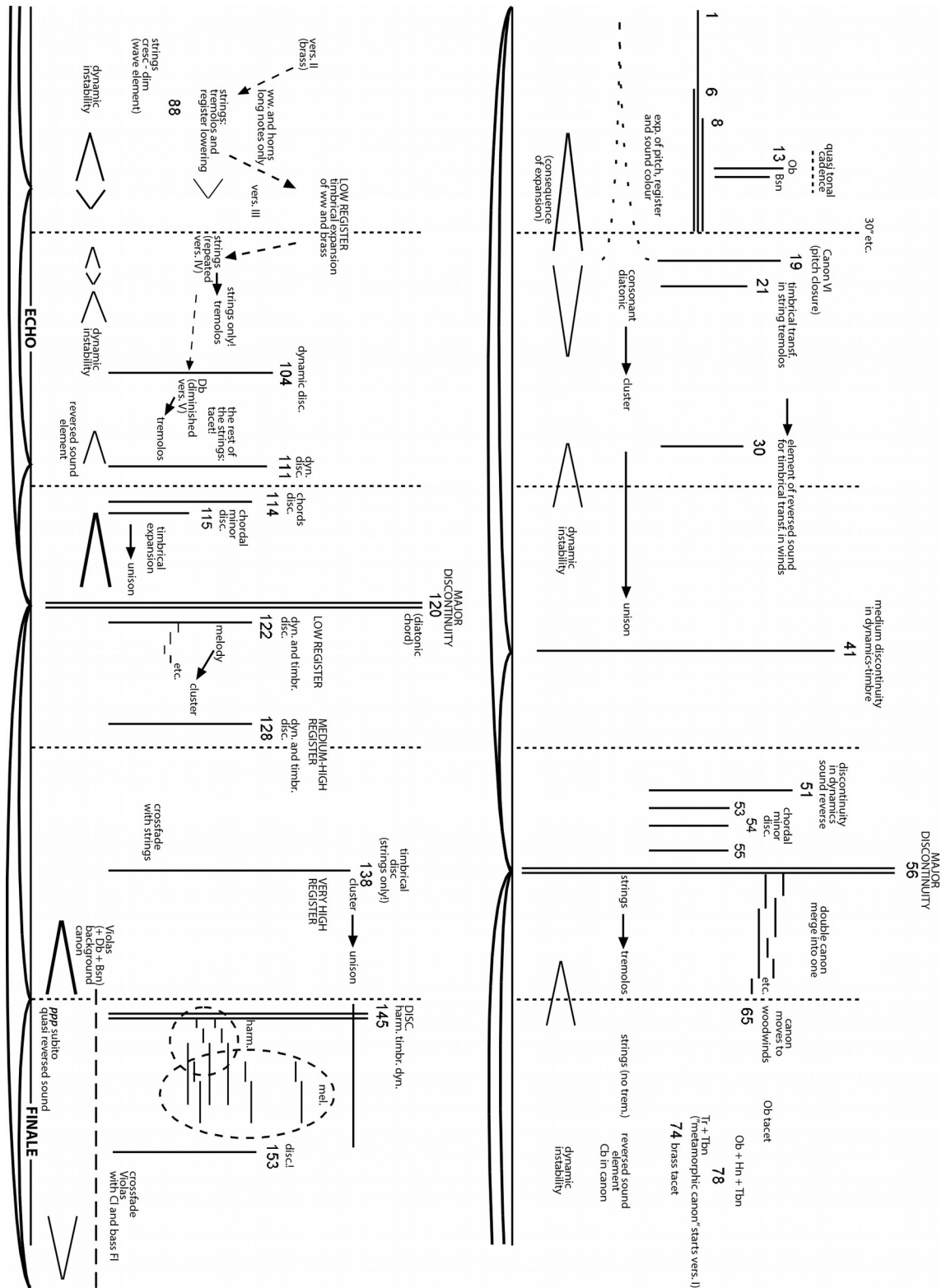
Figure 12: the analysis of macro-formal discontinuities in György Ligeti's "Lontano".

not call "degree of polyphony" or "degree of consonance" or even "harmonic function" without understanding how this "nameless" dimension works before and after the discontinuity.

If we take the first segment, from bar 1 to 56, we can see that on one side this forms a whole in opposition to the other parts, while at the same time we can find discontinuities inside it, which might let us distinguish parts in it, belonging to a lower level of the general structure.

We meet the first considerable discontinuity at bar 41 on the dimension of dynamics and extension of register. What is important here is however not the dimension in itself, but rather the character of discontinuity, that, as being much stronger than what happened before, if we exclude the attack at the very beginning, we can use to mark the distinction between what is there before it and what is there after it.

This discontinuity separates the first part in two sections lasting 2'30" and 1', that is in two durations, which express a ratio of 1/2,5.

Let's take the first section. Here the discontinuities are softer and less evident. Nevertheless we can often hear discontinuities down to the emission of the single instruments. We could therefore put the discontinuities heard in a hierarchical structure, so that we could possibly distinguish sections and sections within the sections, down to what in our analysis we consider the lowest meaningful level of discontinuities, the level which is the nearest to the surface and fastest level.

In the analysis of "Lontano", we could well express the detection of single note discontinuities, or at least a number of them.

Nevertheless this would be relevant only within an analytical narrative that includes those discontinuities in that dimension, as important ones for communicating our understanding of the work. So, it is obviously not necessary to express all the discontinuities we would possibly detect, because in that case the "map" would coincide with the "territory" and the reader would be lost in data.

So, in this first section we can consider as relevant discontinuities, the sudden pitch overture at bar 13 (which might remind us of a perfect cadence), the voicing closure (preceded by a static moment) at bar 19 and the appearance of the note that will result in the general unison at the end of the section. About this last discontinuity it must be said that it is nearly impossible to position it correctly at the first listening session. In fact, it gains relevance only within the perspective of the final phase of the process towards the unison and of course it is not possible to predict during the first listening where a process will bring the listeners. But at a second or a third listening session the discontinuity is already there, exactly in the same way the discontinuities caused by the recognition of a theme would not be there without already knowing that theme. In the end, it appears that one of the main segmenting dimensions in this first part is the pitch range of a collective sound texture that starts with a single

note, gradually opens up to the first noticeable discontinuity, the one in bar 13, and from bar 30 gradually turns back to unison.

Another important dimension is the one of loudness. This is useful in detecting the discontinuity in bar 19, resulting from a clear global inversion of trend, a diminuendo after the initial crescendo that will bring, after a static region in loudness again to a crescendo starting from bar 30. So these four sections (related to bar segments: 0-13, 13-19, 19-30 and 30-41) are mainly marked by two different dimensions, such that gestaltically it is quite difficult to understand whether the three discontinuities are able to group together the adjacent sections in different ordered degrees of a same dimension. It is likely to hear a continuity from the beginning to bar 19, with a minor discontinuity at bar 13, and then another chunk from bar 19 to bar 41 with a minor discontinuity at bar 30, but the problem is that the discontinuity at bar 19 does not seem to be big enough to satisfactorily separate what is before from what is after. In other words, the discontinuity in bar 19 falls at the border between the CR and the SIR.

Of course, there are also other dimensions in play, such as the diatonic vs. chromatic harmonic fields and the smooth vs. rough (with tremolo) textures. These are helpful at various degrees in distinguishing the sections at different levels of the sonic structure, but they do not seem to be treated in such a way to make them essential to distinguish the structure, that we have simply traced in this brief presentation. But if we can hear them, they are there somewhere

and help us characterize the structure at some level; therefore, they could be relevant in distinguishing chunks within the work, but also in linking the work to other works in terms of quotes, forms, styles, repertories, genres and so on.

Finally, I have mentioned the difference in duration between the two sections of the first part of the work, that is the one from the beginning to bar 41 and the other one from bar 41 to bar 56. As the discontinuity at bar 56 is stronger than the discontinuity in bar 41 we can quite naturally consider them as units of a same whole, namely the first part. But the difference in their durations might not let us feel like we can consider them as having the same importance. We might rather consider the shorter unit as a sort of articulation of the bigger one: something that is very clear at faster speeds, such as in an *acciaccatura* or a *mordente* at the head of a longer note or a decay at the end of a note. In this slower case we might choose to consider bars from 41 to 56 as a sort of a coda section, but we could also leave the shorter section its autonomy with respect to the longer one. As Kofi Agawu points out (2004), analysis is after all a kind of art in itself and at the same time a way to direct listening. Therefore, perhaps the decision to link the short section to the longer one in such an ambiguous temporal range has to be done after a broader analysis of the whole work in order to see if the detection of the coda in the first part of the work is a relevant listenable pattern in "Lontano", regardless of whether it has been made on purpose or not.

10.3 Romitelli's "Nell'alto dei giorni immobili"

After this generic approach to *esthesic* analysis, which has served the purpose to introduce some general issues, I would like to show a more detailed analysis. This is taken from my analysis of Fausto Romitelli's "Nell'alto dei giorni immobili", a work for flute, also piccolo and alto, B flat clarinet, also bass clarinet, piano, violin, viola and cello that the composer wrote in 1990. The source of the analysis is the recording of the work by the American Talea Ensemble, that is included in their 2012 album *Anamorphosis*.

Let's start from an excerpt, that in the score corresponds to the bars from 21 to 33, starting from about 55"5 from the beginning of the recording. These bars are delimited by two strong discontinuities in dynamics, the first one being very uncommon in the work as the sound emerges abruptly from silence with a slap tongue played by the flute. The second one is a pretty loud chord and a sort of reverberation-like chord, while being much more common within the work than the former, it is also very strong. The duration of the performance of this section is about 16 seconds.

The figure 13.1, represents the discontinuities detectable by listening, with a reference to their possible source in the score. I decided not to show those discontinuities that are too soft to be considered as relevant or with an interonset interval shorter than 100 ms, such as in tremolos. These are in fact accounted for in the graph, more as a sound quality than as a rhythmical figure.

Some distinctions are marked in white as they could not be connected to distinctions in the score and they are most probably artefacts resulting from the complex instrumental writing and from the consequent unpredictable resulting sound, as it happens for harmonic tremolos. While these, so to say neglected discontinuities are still important in shaping the listening experience, it is paramount for the analysis to connect listening detections to the *neutral level* of the score, in order to reinforce the epistemological status of the analysis on one side and, as we will see, to define the work's dialectics between score notation and performance.

Therefore, for now, these white discontinuities generally will not be considered in the analysis, even if, within the turbulent section marked with the dashed oval, they can be pretty strong.

The distinctions marked in the figure 13.1 are mostly based on discontinuities in three sound characters: loudness, as a continuum between silence and the loudest sound, movement, as a continuum between being still and being the most turbulent as possible, and pitch range, as the continuum between just one tone and a white noise cluster. A discontinuity in one quality, or better "dimension" may come from a sudden change in quality or, in case of a continuous change, in the direction of change.

The figure 13.2 shows how distinctions are related to the three dimensions.

These dimensions are far from being obtained by objective measurements as they are the result of evaluations, by listening to this section over and over.
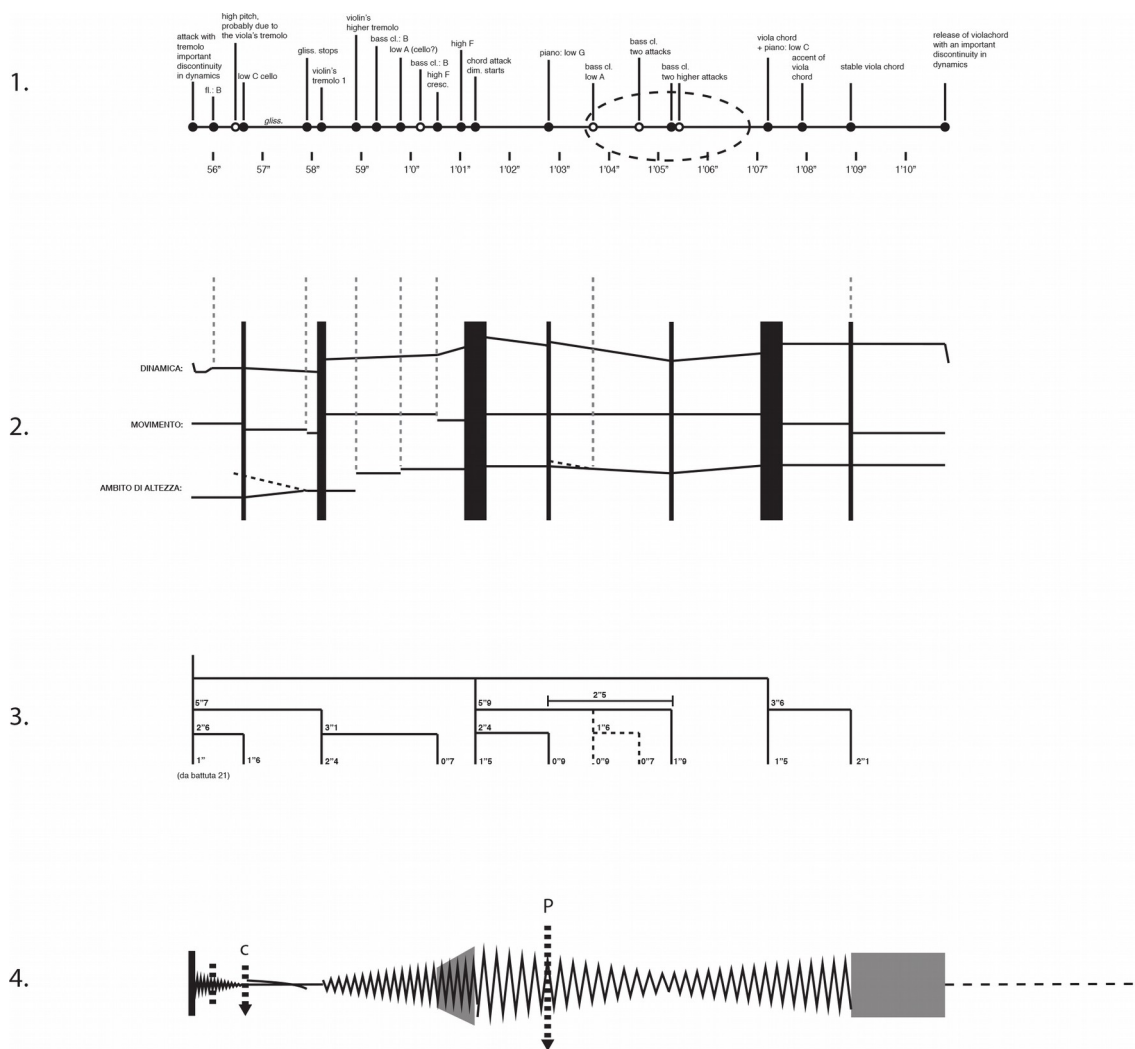
*Figure 13: four analytical schemes of bars 21-33 of
Fausto Romitelli's "Nell'alto dei giorni immobili".*

They rather serve the purpose to express the intuition of the listener on his experience and to stimulate the intuition, and of course the criticism, of the researcher. Moreover, they are certainly not the only dimensions involved in the experience of listening to this section. Timbre, for example, is also very important in defining the pattern structure. But they are what I have considered as most relevant in defining a common ground for changes in listening qualia, so that in the end they can account for a perceptive structure of the section.

In fact, we can try to organize the discontinuities in a hierarchy, according to the number of dimensions involved and to the degree of changes. This is not an easy task, as some distinctions are more comparable than others. Moreover, in this particular section something apparently strange happens besides the already confusing part included in the dashed oval. From 1'7" onwards, some distinctions appear that should not be there according to the score. For example, the viola chord is already there while the turbulent movement is still going on. In the score, the chord appears as a static presence only when all the other instruments are also still, while in the seconds before, the chord should be played with fast tremolo. I do not know if this can be ascribed to an imprecise performance, but this can be confusing when comparing distinctions. For this example, I decided to accept this unexpected sound result as part of the work.

In the figure 13.2, I decided to show the main hierarchical distinctions, similarly with the analysis of "Lontano", with vertical lines crossing the graph. The thicker the line is, the more important the distinction.

Now we can organize the hierarchies in a graph that corresponds to the figure 13.3, showing the pattern structure in terms of units and levels.
When comparing this section to what follows, we can immediately recognize the repetition of some elements that become actual components of a pattern.

To show this, I used a very different kind of graph. The figure 13.4 still shows the same section, but with a different graphic interpretation. The vertical line represents the strong attack in the beginning, but in general the thickness of lines and shapes shows intensity and pitch range. Zigzag lines are turbulent moments while shapes represent static chords even when they are overlapped with zigzag lines. Dashed vertical lines are actual signals, single notes that emerge from the sound flow and may lie within its range, as it happens with the first one and not marked by any letter, or as they isolate themselves in the lowest register, as it happens with the cello (marked by the letter C) and with the piano (letter P). They are the most recurrent elements of the work, something like a one-note figure that grounds the development of the sound continuum to its foundation—or better, to its fundamental tone. The curved line represents a simple glissando element. The thicker line at the end represents the static chord performed by strings, this time in the right position that is shown in the score.

What we have presented so far seems to be a pattern that is repeated three times along the bars 21 to 64 of the score (figure 14).

What we previously interpreted as a sort of reverberation of the last chord seems in fact to be the beginning of the repetition of this sort of pattern. This is a kind of rule for the whole work, as patterns/sections connect with each other in a continuous manner that is not clearly distinguishable. However, we have a very soft change of timbre that can help us to distinguish the end of a section,

306

represented by a dashed line, from the beginning of the next one, which is represented by a continuous line.

Therefore, in the second section (figure 14.2) we see the disappearance of some elements, such as the attack and the beginning tremolo, the glissato moves upwards and the central part is realized by static chords instead of a turbulent texture. Considering that the notes and the chord are in general different, we can still find a strong resemblance between the two sections, especially in the general shape.

The third section (figure 14.3) is still quite similar to the first one, but at the same time it starts to go in other directions, so to speak. Moreover, the introduction of a loud quasi-solo part of the bass clarinet, which can be
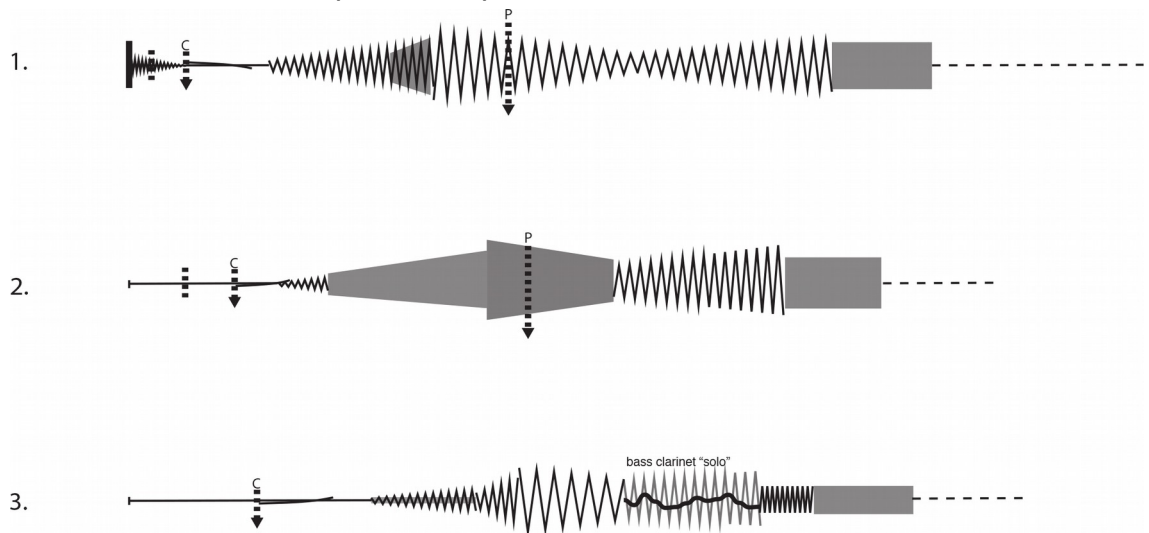

Figure 14: bars 21-64 analysed in terms of the repetition of a pattern.

analysed in terms of a repeating pattern itself, adds a new level to the sound structure.

We have seen thus far a set of distinctions that on one hand seems to be constraints for a transformation of sound qualities in what in electronic music is called an *envelope*, while on the other hand it appears as recognizable elements, signals along the timeline. The envelope's shape with its six segments is a repeated pattern whose starting point can be considered both as the ending point of a section and the beginning point of the next section, while signals reinforce the identity of the pattern and the recognition of its repetition.

In the fragment we have presented, the pattern is instantiated three times. But along the whole work we can recognize a number of patterns like this. In the following figure, with the analysed section in relief, the entire lowest level grouping is done according to a common pattern. Nevertheless, in the units from bar 166 to 185 it is not easy to recognize an explicitly repeated pattern, thus they are grouped following a different criterion.
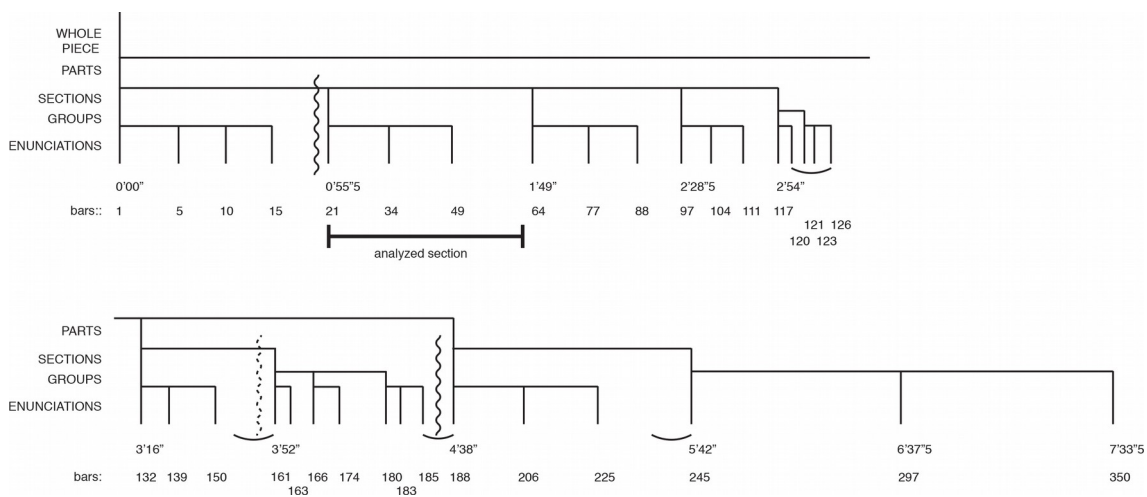

Figure 15: the pattern structure of the whole piece.

308

Perhaps the most striking aspect of this work is the dialectic it exhibits between flux and structure, the former emerging from the continuum of sound transformations in a blurring gestaltic texture and the latter fragmenting the sound surface in crossing references by the recognition of resemblances. We can see this dialectic on all levels, such as in the opposition between envelopes and signals, or better between metamorphic turbulent textures and well recognisable events such as single notes with a sharp attack.

But at a higher level too, we find the continuity with which the climax chords rise in the first part of the work up to the master climax chord in bar 132, followed by an attempt of a descent in the first following sections from bar 161. From the structural side there seems to be a general tendency in discovering that both envelopes and signals are patterns, only they belong to different time levels.

And this is the premise to the emergence of all kind of temporary structures between the fastest signal level and the slower section pattern level, such as the quasi bass-clarinet solo in bar 59 to 61. This process will bring us to the final section, in which the pattern first presented in bar 245 and then repeated and varied in bars 297 and finally in bar 350, is already structured in pattern levels so that, for example, its first component from bar 245 to bar 254 is already made by three presentations of a single pattern originally lasting four and a half measures.

Therefore, the dialectics between flux and structure is also dramatically realized as a transformation from flux, passing through a short blurred zone

between bars 166 and 185 to a structure, defined by an increment of the pattern levels.

But it would be completely incorrect to think of "Nell'alto dei giorni immobili" as a work built on a narration: every passage of this dialectics is hidden. All we have are clues, often quite clear, but mostly in disguise, often contradicted, as happens with the major discontinuities expressed by waved lines. For example the two seconds of silence that separate two sections in bar 188 with a strong discontinuity is compensated by a quasi-chromatic descending connection between the climax chords of the two sections in a way that suggests continuity.

In the end, despite the tension between spectralist thinking on one hand, as it is explicitly stated in the foreword of the score, with its idea of sound continuum, and a more traditional way of building form from pattern structures such as repeated cells and phrase organization on the other hand, this work seems to find a third way in a chaotic and yet static world, the constant low G takes care of that, where form constantly struggles against the flowing matter in order to emerge and to build temporary shapes that are destined to be absorbed again in the ever changing flux. And in the meanwhile, the dialectics between continuity and discontinuity offered me the opportunity to illustrate how graphic elements may engage different visual relationships with each other in order to express an analysis whose elements are specific to the single works.

11. COMPOSITION

11.1 Sonifying Perseus

Over the years, I have applied the framework of pattern structures on a variety of occasions related to the composition of works for acoustic and electronic instruments, to the production of installations and to the didactics of composition, during my years of teaching at the Conservatory "C.Monteverdi" of Bolzano, Italy, at the Conservatory "G.Verdi" of Milano, Italy and in workshops and courses both for professional and amateur musicians.

I will start by presenting a very general topic related to the pattern structure, namely dimensions. While the definition of the composing dimensions, in terms of the "out-time" parameters (Xenakis, 2001) are primary in setting up a pattern structure, there are situations in which the "in-time" structure is derived from the parametric analysis of non-musical systems or depends upon the organization of the "out-time" parameters.

I have dealt many times throughout the years with the elements of the celestial vault as a source of data to be somehow translated in sounds. In particular, the installations "Ordo Coelestis" (Celestial order) and "Andromeda", and the works "Costellazioni" (Constellations) in the versions for alto flute or piano and electronic sounds and "Carillon" (Music Box) for Piano four hands were directed to translate the positions of the stars into musical elements that could be perceptively related to them. In these cases, we could speak of

"sonification", which is the "the technique of rendering sound in response to data and interactions" (Hermann, Hunt and Neuhoff, 2001, p.1) and is at the core of Auditory Display, a field that aims "to enable a better understanding, or an appreciation, of changes and structures in the data that underlie the display".

However, these projects are not intended to help the detection by listening to patterns that were not easy to detect in other phenomenic dimensions, as it would have been if I were following the directives of Auditory Display. Nevertheless, the data set I used when dealing with stars, challenged me to find a structure of the sonic dimensions that could possibly be compatible with the peculiarities of the visual dimensions related to star data. For example, *magnitudo*, which is related to the brightness of the stars, is perceived differently from the dimension of *loudness*, which we likely would automatically involve in the translation of *magnitudo* due to the effects of the objectification of sound as the sensorimotor paradigm that connects brightness, size, proximity and loudness with each other. In fact, our sensitivity to the differences of loudness cannot easily be related to our sensitivity to different degrees of brightness without involving sound masking. As a consequence, a possible translation has to follow different criteria.

I will only present the case of "Ordo Coelestis" here. This audiovisual installation for eight screens and eight loudspeakers was realized in 2014 at the ExpoGate in Milano to celebrate the passage of the Perseid Meteor Showers

(Viel, 2017a). The sound part of the work is the sonification of the 20 most visible stars in the Perseus constellation, for which the stars' position and brightness become the sounding elements of a planetarium for listening. As in the visual part, occasional electromagnetic swarms in the form of radio interferences interrupt the rational order of the cosmos.

The sound was realized by mixing the stars' sonification, realized by a Max software patch, with some interference material deriving from signals and noises that were recorded from an AM radio receiver.

The data set to be sonified includes the position of the stars in the sky, in terms of their spherical equatorial coordinates (*declination* and *right ascension*) and their *magnitudo*. The in-time structure is based on the effects of translating the equatorial coordinates, which are independent from the observer and are based on the celestial equator, into the horizontal coordinates that, on the contrary, depend on the position of the observer on the Earth. This means that once the latitude and the longitude of the listener have been set in Milano for the whole duration of the exposition, the apparent position of the stars changes in the sky depending on the time. In order to let the listener perceive a significant movement of stars in about half an hour of listening, I decided to accelerate the 24 hour cycle of star movements in four layers that were related to the four different couples of loudspeakers, so that it was possible to hear at the same time the complete revolutions in 41, 89, 149 and 181 minutes.

This sonification is based on two main ideas. According to the first one, the horizontal position of the stars is translated by their position within the stereo panorama, so that each star, during its cycle oscillates along a segment whose width depends on its position along the vertical coordinate.

On the other hand, *declination* involves the dimensions both of frequency and timbre. In fact, the sound of a single star is obtained through an amplitude modulation between a sine carrier and a modulating signal that is obtained through frequency modulation. The frequency of the carrier oscillator was constant during the exposition, as it is related to the latitude of listeners. The modulating signal is realized by a sine wave whose frequency depends on the star declination. This is also rescaled in order to engage a cycle of frequency modulation that can represent the star's *azimuth*, which is the vertical coordinate that depends on the position of the observer/listener and the equatorial *declination* of the star.

The second idea is that the vertical position of the star, which changes along the line that connects the centre of the earth to the celestial pole, the *axis mundi*, is expressed as a harmonic partial of a fundamental that changes with a one-year cycle and therefore depends on the day of presentation. The idea of connecting the *axis mundi* to harmonic partials is taken form Marius Schneider's analysis of the sounding practices in shamanic cultures (1970). In this case, the position of the stars is translated with a complex perceptive dimension that lets the data related to the listener (latitude and time) interact

with each other, so that a range of spectra are offered to the stars, to be selected by their vertical position.

Magnitudo is finally expressed in two ways: the first one is by trivially connecting it with a small change in loudness. The second and most important one, is realized by relating *magnitudo* to the speed of the repetitions of an amplitude envelope. While the proportion between *magnitudo* and speed is a direct one, the proportion with loudness is inverse, so that the fainter stars cannot be easily masked by the faster repetition of brighter stars.

As a result, the starry sky is translated by a multitude of sorts of bells that, with their repeated short rings represent a variety of brightness.

This installation is just one example of how dealing with perceptive dimensions can be important not only in setting up a pattern structure, but also in translating the most various data sets into sonorous instances.

Of course, I did not limit myself in just having to do with the simple data/dimensions related to position and brightness: in "Andromeda" for example, an installation for electronic music boxes in which each box is related to a star of the Andromeda constellation, I used, besides position and *magnitude*, the distance from the earth, the spectral class and a sample of the Cosmic Background Radiation.

11.2 Pattern Composition

If we move from dimension to patterns, an immediate development of pattern structures cannot but go in the direction of writing and producing music with patterns, in the most common definition.

In my courses, the topic of pattern is framed within the wider topic of what I call the *linear function*, a perhaps unfortunate term that should not be confused with the correspondent mathematical concept. My idea of *linear function* emerges in the context of Boris Porena's concept of *poli-linearity* (1983) as a fundamental character of music. This concept refers to the idea that music is primarily comprised by overlapping linearities, ordered series of musical elements that are not yet Jones' SIR levels, but are rather conceived in terms of *streams* that can be represented by the staves on a score.

Starting from this, I decided to organize my teaching path in composition by following the possible roles that these linearities can play in the full *poli-linearity* of music. I am merely listing the full set of *linear functions* I deal with in teaching "music composition for pedagogy" or other basic courses/workshops here, without presuming to present any sort of ontology: they are simply a first tool to allow the student to engage in the practice of composition. Nonetheless, this approach has proven some degree of effectiveness over the years.

We start with the *melodic* function, as enacted by the main melody in a poli-linear texture. We can have then a secondary melodic function that can act as a synergic melody, expressing the *heterophonic* function, or as an antagonist

melody, that aims, so to speak, to overstep the melody in order to become the main one. This is the (pedagogic) way in which I teach *counterpoint* as a general function enacted by several antagonist melodies. Then we have the *harmonic* function, which has the sole purpose of expressing and supporting harmony without any melodic or rhythmic intent. Finally, we have *pattern*, as a function that is eminently expressed by consecutive repetitions of the same figure.

Each one of these functions has its own rules and peculiarities. When we come to *patterns*, we explore these rules by writing or producing music solely with the use of patterns. We start from writing music for the simplest instruments such as hands beating on a table. Then we move to voice and body percussion. And finally we may involve the entire range of pedagogic instruments first and then the traditional instrumentation of the entire orchestra. Depending on the occasion, it has also been possible to use electronic instruments like drum machines (I have built a sequencer application just for the purpose of teaching *pattern composition*) and even circuit bended radios.

A variation of the TUBS (Time Unit Box System) notation can be used to a certain extent, by people not able to write or read traditional western notation. On the other hand, TUBS prepares to bring the concept of *pattern composition* to drum machines. My sequencer application has been created in order to link the TUBS-like notation of drum machines with the letter notation of the traditional analysis of musical form, so that once a repertory of patterns has

been created and each pattern is identified with a letter, the user can write a sequence of letters in the proper text field that will represent the pattern structure of the composition and, at the same time, will serve as the reference for instancing patterns that will be played by the sequencer.

The rules I provide for learning how to deal with patterns are ultimately aimed towards developing a technique of composition that mimes the esthesic strategies of the listener, provided that there must be a model of the listener. In this respect, I try to force the students, at least in the beginning, to follow some rules that define, with some strictness, the idealized listening strategies.

For example, I propose a certain set of requirements for building patterns "correctly", such as: a pattern cannot be made by the exact repetition of subpatterns, so that it can exhibit a figural significance as a unit, or that a pattern, at least the first pattern that is instantiated in a composition, cannot begin with a pause.

Secondly, as recognition does not necessarily mean exact instantiation, I require that, at least for the first experimentations, an instance must be exactly identical to the pattern, at least in the way it begins, in order to avoid the risk of being improperly recognized.

Finally, the different layers that realize patterns of the same length have to be considered as a multi-layered pattern, so that the students possibly avoid the temptation of exchanging layers between patterns and therefore of addressing the topic of resemblance among patterns prematurely.
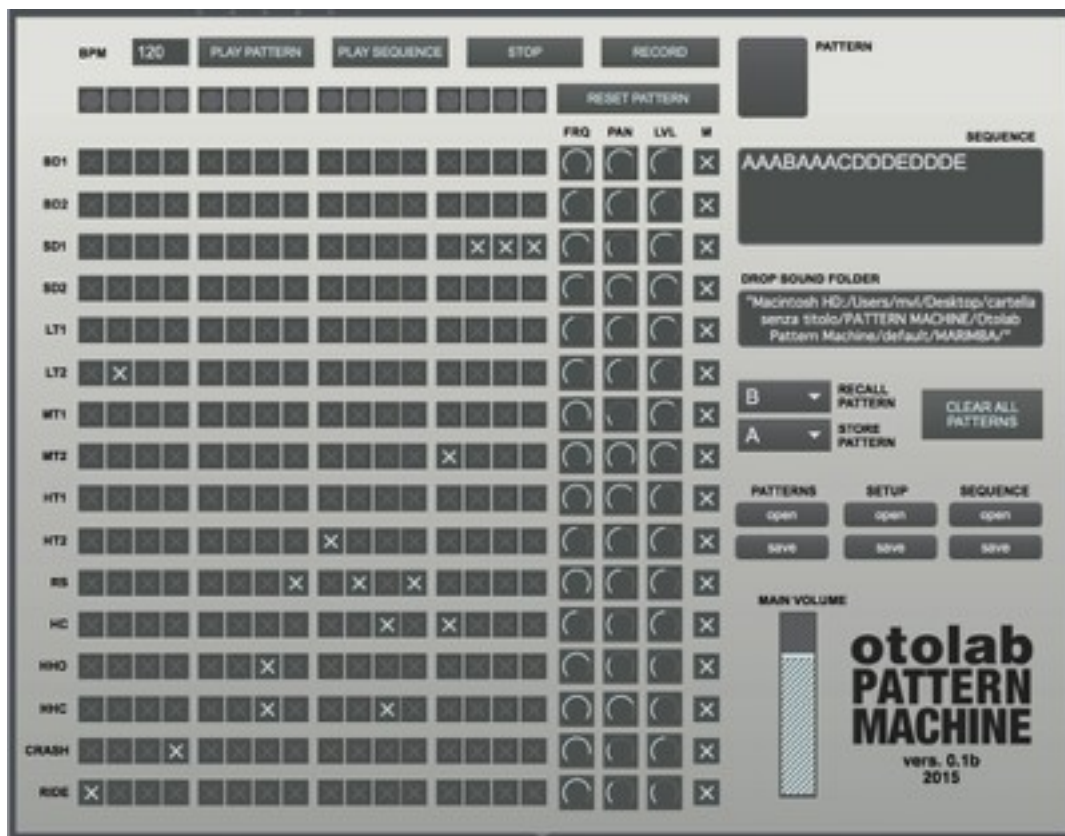
Figure 16: the pattern machine I have realized for my courses.

As soon as possible, I introduce the first technique that is used for building pattern structures, namely the *contrasting technique.*

This is realized when a series of instances of a pattern are followed by instances of another pattern. The technique aims to create in the listener a distinction between a before and an after that is represented by the instant of the beginning of the first instance of the second pattern. Clearly, what is presented here is a narrative for engaging a composing practice. As we have seen, the discontinuity that is shown by distinctions always occurs before the actual moment of distinction and is therefore the result of an articulated process rather than being the simple expression of an instant. We always have to keep this in mind and we always have to repeat it to the students: we are

following a narrative for pedagogical purposes that needs to be abandoned as soon as we have assimilated it in order to maintain the process of pedagogical construction-deconstruction.

We develop the *contrasting technique* when building structures, especially the binary structure, in which each unit is made by two sub-units from the total layer down to the surface layer, by first involving the idea of patterns made of instances of patterns. We therefore spend some time in verifying that the structures built by students use the instantiation of patterns in such a way that a possible listener cannot distinguish the pattern structure as a different one.

I have learned not to be too strict in giving rules on how to combine patterns and letting the students simply find their own rules by deriving them from discussions in the classroom.

However, I have developed a generative system that is able to automatize the construction of a pattern structure by following a grammar of substitutions, which proceeds in a top-down direction. Nevertheless, for teaching purposes it might be too strict to foster a creative approach to composition.

In developing the *contrasting technique*, we gradually move from rhythmic patterns that are made by non-tuned instruments to figural patterns that are based on rhythms and pitches.

As concerns these aspects, the discussion about the difference between *scales* and *modes* is introduced together with the criteria that can be used in creating one's own *scales*.

Then it is time for the second technique: the *marking technique*. This is typically realized when a series of instances of a pattern is followed by a single instance of another pattern, afterwards followed by a series of instances of the first pattern or of another one.

The idea of *marker* comes from the practice of breaks and fills in popular drum techniques, so that its role is to mark the end of a unit at a certain level. In brief, a *marker* might also be instantiated as a *marker* in other units, possibly on the same level. But it also can be presented just once, as it happens for the *marker* that is dedicated to the end of a work, which I call, more as a tribute to formal language theory than to science-fiction movies, the *terminator* mark.

Of course, on one hand an on-time-only *marker* is not a pattern, but, in the composition narrative, it can be (objectively) treated as one, and for that reason it has earned the right to be named after a categorical label such as "fill". On the other hand, *markers* are the occasion to talk about *inter-opus* patterns, and to show how styles and genres are typically involved in using repertories of patterns that characterize them.

The next step is related to the discovery of different levels' *markers*, which are not only used specifically for units on some level, but that are made by patterns themselves. In this way, a connection is made between the two techniques of *pattern composition*.

The final part of the course is dedicated to one of the most important problems that presumably emerges along the whole path and regards the relationship between what we have called in the last part the *difference* and the *identity* of patterns.

A pattern is defined in terms of the ordered elements that build it, the units that fall in the SIR along the dimension (or dimensions) that define their configuration as the pattern. This is what we can call the *difference* that is invoked by SIR. On the other hand, there are certain qualities that do not change along two contiguous instances of the same pattern. First of all, they constitute the *identity* of a pattern, a set of qualities that, as we have said, are more relevant within a series of repetitions and are involved in the process of *habituation*. These are, for example, the instrument playing, the general loudness, the position in space and so on, with the infinite procession of Goodmanian resemblances.

We could think of a number of patterns that might be different only in the instrumentation, or that might be composed for just one instrument that occupies a place in space that is different for each pattern, or again they could be written for only a whistle and differ from each other only in their dynamics. And the resulting pattern structure would equally engage the proper distinctions, allowing the listener to detect the same structure, although differently realized.

What we have here is the contrast between the dimensions that fall in the SIR among the pattern components and the dimensions of the same pattern

components that fall in the CR and therefore constitute the *identity* of patterns as simple units, as chunks.

As a professor, I can only leave the students with a possible rule, according to which, in order to preserve the distinction of a pattern structure, the discontinuities in identitary "chunking" dimensions should be proportionally related to the highest level in which they appear, that is the highest level inclusive of the unit that starts with a discontinuity as a composing unit. And the lower that level is, the closer it is to the "surface" level, the smaller the discontinuity should be.

Of course, the path of pattern composition ends with the reassurance that in the end there are no rules, but only codes of conduct that are sometimes useful in reaching some goals and sometimes need to be broken in order to invent new goals.

11.3 Presences

If the *pattern composition* is essentially conducted by interactively connecting the composition of single patterns and *markers* with an abstract structure of patterns both in writing scores and in letting a computer automate the procedure, there are still different possibilities of dealing with the elements of our narrative.

"Presenze" (Presences) is a concept related to the construction of an electronic instrument for live performance and to a possible narrative to be implemented in performances and fixed media. In "Presenze", the pattern structure becomes a system of triggers that launches patterns in the form of samples. The live performance of "Presence" is done by operating an instrument, which is an application realized with Max, and by choosing on the fly both the structure and the set of patterns that realizes it.

A pattern structure in this context is a tree-like structure in which each unit is composed by two or three sub-units. The extension of the structure is up to seven levels, from zero to six, starting from the surface level units, which have no sub-units, up to the sixth level, which is the highest one. We therefore have 64 possible structures, according to the choice for each level to be expressed by units that have two or three sub-units.

Moreover, it is possible to set the degree of certainty of having exactly the chosen number of sub-units for the units of all the levels from one to six, so that a minimum degree of certainty will result in having the number of the relative sub-units, randomly changing between two or three, for each unit of a given level.

A sub-level number of three on a given level means that when the sample related to a unit of that level is triggered, the samples of all the lower levels are triggered as well, but it is necessary to wait for two more triggers of the units

belonging to the immediately lower level, in order to for it to be possible to trigger the sample related to the next unit, that, again, will be simultaneous in all the triggers on the lower levels.

In "Presenze", we can use up to two pattern structures at the same time, which can also be synchronized with each other in beat.
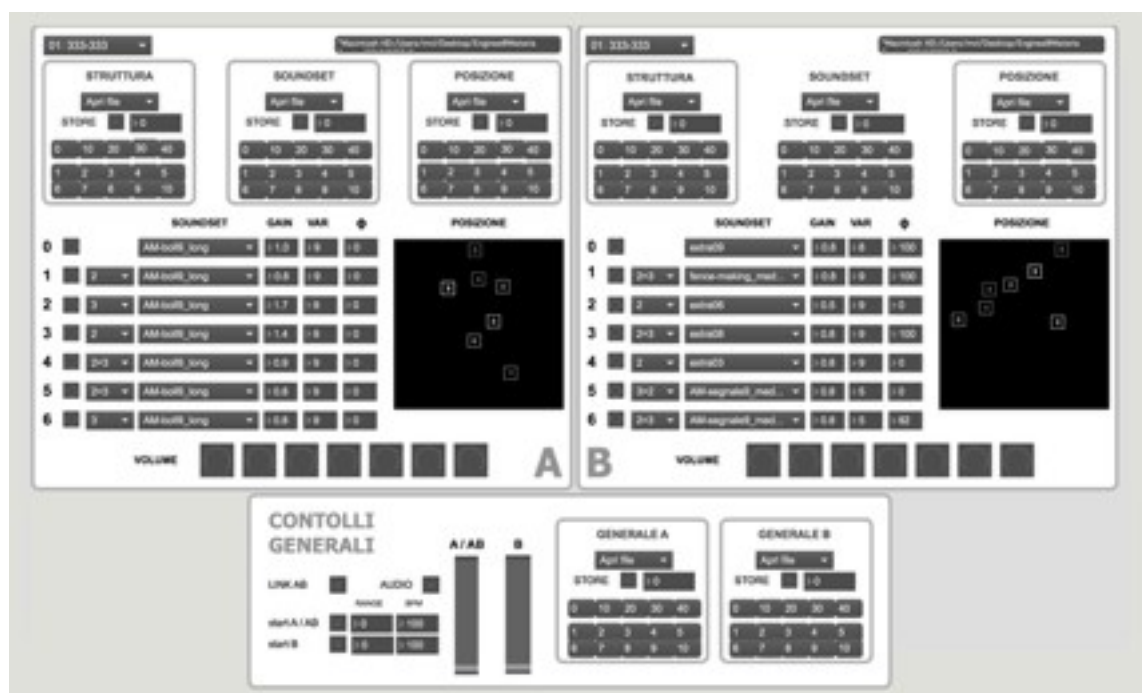

Figure 17: the user interface for the live realization of "Presenze".

For each set of seven levels (a pattern structure) we have:

- a structure number (from 1 to 64) that will set the number of sub-units in the units of each level;

- the bpm value that will be related to the units of the surface level, the level zero;

- the range of possible variations in time of the bpm value so that the clock driving the structure can be expressed on a range from the most precise one to an almost random one, within a certain degree that maintains an average speed;

- the general volume of the structure.

For each level, we can set:

- the certainty degree, according to which the number of sub-units will be exactly in the number related to the structure;

- the set of samples that will be triggered by the structure units of that level;

- the number of samples (from one to nine) that will be randomly chosen and triggered by the structure units of that level;

- the volume of that level;

- the position in a quadraphonic panorama of the sound layer of that level;

- the degree of precision according to which the sample trigger will match the clock, so that, for example, the triggers that are supposed to by synchronous with each other can be varied in that degree of synchronization.

In this way, it is possible to mediate between a most clear structure of triggers and a totally chaotic structure that still maintains the organization of levels as independent streams. Moreover, by setting the volume independently for each level it is possible to instantiate the structure only partially so that some levels are unexpressed. With the possible erasure of a level that is, for example,

between two instantiated levels, we can obtain a level including units that have between four to nine sub-units. In fact, let us say that the sub-units number for the three levels, from the highest to the lowest is 2-2-2. This means that each unit of the highest level will have two sub-units, each one of which will have two sub-units. The highest-level unit will be connected to four sub-units of the lowest level. If the middle level is not instantiated then each sample of the highest level will be triggered, with every four triggers corresponding to the samples of the lowest level. In case of a 3-3-2 structure with a non-instantiated middle level, on the other hand, for each sample of the highest level we will have nine samples of the lowest level.

With the "trick" of non-instantiating one or more levels, we can reach an extreme pattern structure of seven levels. While having a sub-unit structure of 3-3-3-3-3-3, by instantiating only the highest and the surface levels we will obtain that for each trigger/unit of the highest level we will have 2187 triggers/units of the surface level.

Finally, as each sample has its own structure that can be more or less articulated, the possibilities of realizing pattern structures seems to be limitless. In the example I provide (Viel, 2007b), which is the beginning of a performance I realized in July 2016 in Brescia, Italy for the "Materia Sonica Festival", the structure number four, 322-232, is instantiated in such a way that the fifth and the fourth levels show a minimum degree of certainty of the sub-units number. Each level uses an entire set of nine samples, but in different ways: the highest

level, that is used alone at the very beginning of the excerpt, has a very electronic sound, mostly made by wide spectrum energy noise that is filtered in different ways, and a internal structure that is independent from the general pattern structure. The other levels that are instantiated one after the other in this excerpt, are the lowest four, each one with a sample set that is relative to an acoustic percussion sound. The random and ever changing choice of the sample to be played within a set for a specific unit has the purpose here of rendering an acoustic-like performance, in which every beat of the percussion is slightly different in timbre and volume.

What are the dimensions at play in this work? They might be volume, and pitch or even timbre: it really depends on the character of the sample sets. The most general dimension is therefore the "presence of a pattern". In fact, in the excerpt, especially when percussions enter, it is the presence of a particular instrument/sound/pattern, as possibly falling in the PRR and therefore distinguished as streams within the whole music flow, that allows the perceptive articulation of the structure in such a way that listening entrainment, that is, as we have seen, expectation in terms of enaction, is possible.

But, of course, if the samples are chosen so that each sample in a set is completely different from the others of the same set, we would need to use other elements to preserve entrainment, such as a length of the sample that is shorter than the time interval between two successive units in the same level, or other "tricks".

Finally, "Presenze" has an extremely wide potential of realization, that depends on the formal setup of structures, for example in terms of the timeline, and on the sample set that can be used, from electric to acoustic sounds, to found recordings, music samples and so on. It cannot therefore only be a "work in progress", which is instantiated in fragments. I hope that in the near future I will have the chance to build a sounding "wheel of structures" as a range of possibilities that is able to entangle pattern structures to sample sets with each other, in order to present the work as fixed connections that can be instantiated and performed in sections or as a whole, depending on the contexts of performance and music distribution.

11.4 Cluster

Over the years, I have developed the narrative of pattern structure in every sort of compositional context from audiovisual performances to the realization of scores for acoustic instruments and the production of electronic music. To show the extent of using the concepts that I developed in the second part of this text, within the context of composition, I have chosen to analyse the poietic strategies of "Cluster (for Demetrio Stratos)", a 20 minute work for electronic sounds with optional video, that I realized between 2011 and 2014 (Viel, 2017c).

This is a work that is entirely based on recordings of Demetrio Stratos' voice, which I received permission to use. The starting idea is connected with the image of a star cluster, a sort of spherical mini-galaxy that in about 10 to 30 light years gathers from ten thousand to several million stars. The individual character of stars and their massive presence suggest to me the idea of a high polyphony, which on one hand is perceived as a whole and on the other hand is composed by layers that maintain a recognizable individuality. "Cluster" is therefore a study around the *permeability* of sound textures in the Ligetian sense (Ligeti, 1964). From the point of view of imagery, a cluster becomes a metaphor of the chaotic polyphony of the contemporary hyper-connected world.

At the same time, this work is a tribute to the voice of Demetrio Stratos, who was a singer and lead figure of the Italian experimental scene of the 70s and an expert in the extension of vocal techniques. The recordings of his voice are therefore best suited to represent an expressive range that is able to connect the human voice to electronic sounds.

For that reason, I decided to multiply Demetrio Stratos in a polyphony of 12 parts, organized in three quartets according to vocal types. These types metaphorically develop the Medieval categories of *musica humana* (the human voice, symbolized by the letter V), *musica instrumentalis* (the human voice that is still recognized as such but is somehow transformed in an instrument of in a non-human voice by sampling techniques, symbolized by S) and *musica*

mundana (the human voice that is no more recognized as such and is therefore transfigured besides the organic, symbolized by E).

Regarding the *out-time* organization of the dimensions involved, each quartet has a functional relationship with the other two quartets; moreover within each quartet the four layers/voices engage functional relationships with each other.

The relationship among quartets is defined in terms of foreground versus background that is connected inside each quartet with the distinction between active and inactive parts.

Within each quartet, one active voice performs the main signal, while the other possible active voices perform a range of functional relations with the signal that can be: homophony, parallel heterophony, temporal heterophony (homorhythm) or an independent and contrasting signal, which are respectively represented by the letters O, P, T and C.

As regards the *in-time* organization, "Cluster" comprises 20 sections, whose structure is differently realized depending on the dimensions involved. This means that the different pattern structures, all entailing 20 surface units, are instantiated in such a way that we have a sort of "structural polyphony", whose dimensional streams are chunked together only at the "total level".

The duration of each section is set in a rather flexible way so that it is chosen among the possibilities of 40", 50", 60", 70", 80" and 90", that correspond to the range of four to nine that we have already met as the possible sub-units of

a higher level unit, whose immediately lower level has not been instantiated. The only requirement is that the succession of two sections with a same duration should be avoided.

There are other elements that are left to the extemporaneousness of production, as dynamics and the transition between contiguous sections. Nevertheless, each passage of the "Cluster" realization is carefully conducted within the techniques of a post-serial, parametric thinking in order to realize the details of pattern structures in nearly the same way it has been done for "Presenze". I will show here only some exemplary details of the realization of "Cluster" that can possibly show the parametric instantiation of the pattern structure.

First of all, the roles among the quartets in each section are decided according to an arbitrary structure on four levels (total and surface levels included) that is expressed by the variation of a dimension that expresses the number of quartets in the foreground, and therefore leaves the remaining quartets for the background.

The structure is defined in the same way as in "Presenze" but follows the graphic convention of indicating the units of the first level, which is just above the surface level, with their sub-units number. The units of the second level are expressed by single parentheses. A further parenthesis includes the complete series of units in the total level.

The dimensional realization of the structure follows a different graphic convention, for clarity, in which the space separates the units of level one and the dot separates the units of level two.

The pattern structure of foreground quartets is therefore ((23)(233)(223)) and is realized by the dimension with the values 1, 2 and 3, that refer to the number of quartets in the foreground, as: 12 221.21 223 123.12 21 321. In total there are 20 numerical instances corresponding to the 20 sections.

How is the structure realized by the proposed instantiation? Each unit realizes a pattern with its sub-units that can be rising, steady or falling. If we take the units of level one we can see four rising units: 12, 223, 123, 12; and four falling units: 221, 21, 21, 321. If we go back to the instantiated structure, 12 221.21 223 123.12 21 321, and take the first value out of each unit of level one we can see the generator intervals of units of level two: 12-221-123. If we take out the first values of these last generators, we'll see that the resulting generator pattern of the unit of level three, 1-2-1, that indeed forms an arch-like pattern. This may prevent the possible distinction of the structure as composed by three units of level two. But I need to say that I realized these possible difficulties, that are met only in particular cases, only after having composed the piece, and moreover while they might be a problem for the implementation of the theory in the composition, they surely do not impede a fruition of the piece as a work of art.

The resulting scheme of foreground quartets is as follows:

```
Tot.   2  1  2  2  1  2  1  2  2  3  1  2  3  1  2  2  1  3  2  1
Sect. 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20
V         X     X     X        X  X  X  X        X        X  X        X        X
S      X        X  X  X              X  X        X  X  X        X        X  X
E      X        X              X  X  X        X        X  X        X        X  X  X
```

A similar process, which is realized independently for each quartet, determines the number of active voices within the quartets in a dimension from one to four, which directly corresponds to the number of active voices.

Here is the resulting scheme:

```
Sect. 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20
V 1    X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X
  2    X  X  X  X  X  X     X  X  X  X  X  X     X  X  X  X  X  X
  3       X  X  X  X           X     X  X  X     X     X     X  X
  4       X     X                 X     X                    X
S 1    X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X
  2    X  X  X  X  X  X  X  X     X  X  X  X  X  X  X  X     X
  3    X              X     X        X  X  X  X  X  X        X
  4                   X              X     X
E 1    X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X  X
  2    X  X  X  X        X  X     X  X  X  X  X  X  X  X  X  X  X
  3    X  X     X        X     X     X  X  X  X  X     X     X
  4       X                          X  X
```

In the next phase, the previous scheme is freely adapted to give each part the same average activity along the whole work.

```
Sect. 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20
V 1    X  X  X  X     X  X        X  X  X  X     X  X  X        X
  2    X  X  X  X  X        X  X  X  X     X     X     X     X  X
  3       X  X  X  X        X  X     X  X  X     X     X  X  X  X
  4       X     X  X  X        X     X  X  X  X     X        X  X  X
S 1    X     X  X     X     X  X  X  X  X  X  X     X  X
  2    X  X     X  X  X  X  X     X  X     X  X  X     X     X
  3    X        X  X  X  X  X     X  X  X  X  X  X     X  X
  4       X  X     X     X           X  X  X  X  X  X     X
E 1    X  X  X     X     X           X  X     X  X  X     X  X  X  X
  2    X  X  X  X     X  X     X        X  X  X  X     X  X        X
  3    X  X     X        X  X  X     X  X  X  X  X     X     X  X
  4       X     X     X        X  X     X  X  X  X  X     X  X
```

In general, the process is conducted in the same way for all the involved dimensions.

For example, the previous scheme is modified with the instantiation of the degree of polyphony in terms of different and contrasting signals that are present in the active voice of the single quartets.

We have therefore, once again, different structures and independent instantiations. However, in this case as the number of active parts and the number of signalling active parts are calculated independently from each other, it may happen that there are inconsistencies between the dimensions that are resolved by filtering: two signals in a section with only one active part means that there is just one signal.

Here are the structures and schemes relative to signalling voices.

```
V
Structure: ( ( 3 3 3 ) ( 2 2 3 ) ( 2 2 ) ): 112 112 123 12 34 221 21 11
S
Structure: ( ( 2 2 3 ) ( 2 2 ) ( 3 3 3 ) ): 21 21 321 21 32 244 221 211
E
Structure: ( ( 2 2 ) ( 3 3 3 ) ( 2 2 3 ) ): 21 12 112 123 234 31 21 321
```

```
Sect.  01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20
V 1     S  X  S  X        S  S        S  X  S  S  -  S  X  S        S
  2     X  S  X  S  X        S  S  X  X     S     X     S     X  X
  3        X  S  X  S        S  S     S  S  S     S     X  S  X  X
  4        X     X  X  S     S        S  S  S  S     S     X  S  X
S 1     S     S  X     S     X  S  S  S  X  S  S     S  X           S
  2     S  X     S  S  X  S  X     S  X     S  S  X     X     X
  3     S        X  S  X  X  S     S  X  S  S  S  S     S  S
  4        S  S     -  S     S        S  S  S  S  S  S           S
E 1     S  X  S     S     S        S  S     S  S  X     S  S  S  X
  2     S  X  X  S     S  X     S  -  X  S  S  S     S  X  -     S
  3     X  S     S        S  S  X     S  S  S  S  X     X     S  X
  4        X     X     X        S  S     S  S  X  S  S     S  X
```

In the end, the scheme is completed, after the involvement of a calculation that I do not show here, with the role of each one of the 12 parts within the 20 sections, which are expressed by the following symbols:

- ∑ is the main signal;

- ∑O, ∑P, ∑T is the heterophonic or homophonic version of the main signal that is expressed by some other part;

- S is a contrasting signal; o are long notes expressing inactive parts;

- >S means that a signal in the previous section is still going on in the present section;

- the hyphen "-" is related to parts that were filtered away in assigning signal and therefore might be expressing non-active though "restless" parts;

- X characterizes parts that are dedicated to O, T or P textures that are determined in another way;

- finally, empty parts are just inactive parts that are silent or present a continuous very soft tone.

Here is the completed scheme:

```
Sect. 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20
V 1    o  X  o  X     ΣO o        ΣT X  o  S  -  Σ  X  o        Σ
  2    o  Σ  o  S  o        Σ >S  X  X     S     X        o     o  X
  3       X  o  ΣT o        S >S     Σ  o  S     S        o  S  o  X
  4       X     X  o  S        ΣP     S  o  Σ  o     S     X  o     X
S 1   ΣP     S  X     o     o  Σ >S  o  X  S  Σ     S  o              o
  2    S  o     Σ >S  o  o  o     ΣP  o     S >S  o     o     X
  3    X        X  Σ  o  o  o     S  o  S >S >S  o     o  Σ
  4       o  Σ     -  o     o        o  Σ >S >S  o  Σ        ΣO
E 1    Σ  o  ΣT    o     S        S  o     S  o  X     Σ >S  Σ  o
  2    S  o  X  o     Σ  X     o  -  o  ΣT S  o     o  X  -        o
  3    X  o     o        Σ  ΣT o     o  S >S  o  X     X        S  o
  4       o     o  X        o  Σ     S >S  o  ΣT o     S  X
```

In the end, in the whole work we have 46 signals that are presented in each section in the following numbers:

```
Sect. 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20
       3  1  2  2  1  2  2  2  1  3  2  4  8  1  2  3  1  3  2  1
```

I should now introduce the criteria that were adopted in constricting signals. These are Chomskian structures, meaning that they are constructed from a set of nonterminal and terminal symbols with rules for connecting symbols in words and words in sentences. It is not important here to go into the details of signals composition. Nevertheless, I would just like to mention that signals are composed by a main note, that is different for each signal, and, when the signal is not simply made by the main note, a set of one to four secondary notes.

A main note, which only changes when the functional relationship also changes, is given to each part of all the three quartets, in each section and is chosen among one of three scales. The basic pitch structure of the work is therefore similar to a chordal piece made of 20 chords. The three scales are

obtained from the harmonic series on a E note in such a way that, given a pitch range that is more or less correspondent to the human voice's one for the V parts and is a wider one for the S and E parts, the difference from scales one to two and three corresponds to a proportional increase of dissonance, which comes from relating the notes expressed by the scales to a fundamental whose octave register is lower and lower.

Figure 18: the three scales of "Cluster".

Figure 19: the general scheme of pitches connected to main and secondary roles.

Of course, there is much more to say. The reason I mention this is because the choice of the scale for each part in each section follows an arch-like path through the piece that is still modulated by a pattern structure, both as a dimension and as the interference of the parts' functional role in the quartet.

Finally, the process of composition, in this context, even when it starts from a moment of sonorous imagination, is involved in a process of analysing the evoked sonorous image in terms of the qualities, be they simple or complex, that it exhibits so that they can be resolved as (parametric) dimensions. Some of these dimensions might be involved in characterizing the composition as a whole, and therefore are instantiated by discontinuities that fall in the CR. On the other hand, other dimensions, the ones I want to change in the course of the work, will necessarily fall in the SIR and, by doing this, will be presumably involved in a pattern structure. This pattern structure, however articulated, is in charge of managing the discontinuities so that, for example, the articulation of the work in sections must deal with the instances of those dimensions that, by falling in the SIR, allow the listener to distinguish those sections. In order to obtain this, a top-down path of composition could be engaged by starting from articulating the general form in two or three sections, then articulating each section in two or three composing sections and so on up to the shortest units that are supposed to enter the planned pattern structure. But a bottom-up path could be engaged as well, so that units are connected by instances of dimension at higher levels and so on. Moreover, a pattern structure could be

complicated at will up to involving polyphonic or even counterpoint-like structure textures. In the end, what is important in this technical frame of pattern structures, is that the poietic strategies of the composer/producer/artist are directed to mime the esthesic strategies of the listener. This means that the compositional practice is driven by an "esthesic reference" in terms of a structure of the detection of changes, which is based on dimensions and patterns that can be set, although always experimentally, according to a specific perceptive target. The idea of an "esthesic reference" seems to be important in teaching composition because it helps to capture all the possible creative outcomes of the students within a net of meaning—the changes in the score/production that can possibly be detected by a listener—that can help in selecting a direction for creativity. But this is also important for artistic practice. On one hand, it fosters a proper analysis of the target listener in terms of the repertory of learned patterns and dimensions, which can help the composer to deal with the ivory tower of the repertoire of patterns and dimensions that she has developed in a lifetime of practice without forcing her to feel that he is hostage to a mass audience. On the other hand, it can help to positively determine the criteria for analysing the connection between poietic and esthesic strategies, and in setting the basis of a meta-composition practice.

But whatever aesthetic or commercial strategy is involved by the composer/producer, a technique of composition based on pattern structures will inevitably engage an interaction with the listener, with his or her past experiences and his efforts in distinguishing, in terms of the interference

between the pattern structures of music and of the listener; and in the end it will compel the artist to take a position on his relationship with the listener in a media culture, by constricting with him a relation of difference, instead of the happy subjection of identity and the solipsism of diversity.

12 WHAT WE TAKE FOR GRANTED IN THE SOUNDS AROUND US

To engage a general discourse about the soundscape, we have to start by remembering John Cage's description of his experience inside the Harvard University's anechoic chamber in 1951 with the discovery of the endless sonic flow we are immersed in.[28] As Cage stated, an immediate consequence is that any experience of silence is functional and emerges through a cognitive process of distinction between a sonic plane, that is different each time, and the plane of its absence. Silence is the result of a way of listening.

When we turn to the soundscape, John Cage's experience brings some important consequences. As it has been pointed out (Kelman, 2010), in spite of the success in a wide area of studies of using and abusing the term "soundscape" since it was introduced by Murray Schafer (Schafer, 1969), not only a general consensus about a possible standard definition has not yet been reached, but even the question has been raised whether it is time to abandon the term for good (for example: Ingold, 2007 and Helmreich, 2010).

Defining a keyword in human sciences is often more the final result of research than its starting point. No wonder then, if a number of texts includes the word "soundscape" without referring to a precise definition and leaves the outlining of its semantic field to deduction. Such use of a word may resemble the implementation of an apparatus (I refer to its definition by Giorgio Agamben, which was included in section 1.2.3) in spite of its involvement in discussions

---

[28]For a more detailed description of the experience, see Section 2.2.

about music and environment as a technical term. The word "soundscape" acts in these texts as a device or a label, if not a brand, that on one hand has the capability of capturing meanings that determine the blurred boundaries of research, while at the same time identifying a corpus of works, scholars and institutions that strategically define a disciplinary domain that is inevitably inscribed in a play of power.

On the other hand, the attempts to define the soundscape answer to the needs of institutionalization of the term, and for that purpose aim to legitimate it, by connecting it to supportive disciplines, that have possibly a strong epistemological statute.

Murray Schafer's well-known definition of soundscape as "any acoustic field of study" (Schafer, 1993, p.7) already points inclusively to the academic world as a topic that apparently encompasses all the studies of acoustic related subjects. If the utility of a definition consists in its ability to exclude the possible misuses of a term by delimitating its semantic domain, we could say that this function is here accomplished by referring to a "field of study" in the first place and secondarily by the word "acoustic". The possibility of a clear distinction between the appropriate and inappropriate uses of the word is here guaranteed by the discipline related to the term "acoustic", namely physics.

Another example comes from the composer Pauline Oliveros, who describes the soundscape as "all of the waveforms faithfully transmitted to our audio

cortex by the ear and its mechanisms" (Oliveros, 2005, p.18). This definition updates, so to speak, Schafer's board of guarantors by adding physiology and leaves the role of managing the connection between the semantic domains of medical and physical sciences to the word "faithfully". However, it is not clear how the one-dimensional parameter of faithfulness can match the complex transduction operated by the listening system and in what form it is able to account for sound cognition. But the main problem is that not only does this definition encompass everything from a Beethoven symphony, to the noise of a passing car passing, to the cochlear echo caused by them, but it also sacrifices on the altar of the "objective" aura of scientific data the way we use the word soundscape independently from the topics of sound waves or music.

The Careggi Declaration[29] seeks a more usable and shareable definition of the soundscape as "the acoustic property of every landscape according to a species' specific perception".

In this case, the relativism of a "species' specific perception", which is open to post-human directions, is balanced by the word "acoustic" that on the other hand seems to allude to the physical properties of sound. Moreover the difficulties of defining the soundscape are delegated to the concept of "landscape", which has its own history of debate. Nevertheless, what is important here is that the relationship between soundscape and perception is

_____

[29]The "Careggi Landscape Declaration on Soundscapes - June 2012" is a document that was made available on the occasion of a Third Careggi Seminar on Soundscapes organized by UNISCAPE, the Network of Universities dedicated to the implementation of the European Landscape Convention.

finally stated in an explicit way. This moves the whole question of defining what is a soundscape to a matter of distinctions as the perceptive domain of two species may in fact overlap to a certain degree or be disjointed, but above all it poses the issue of the presence of an observer, who measures "acoustic properties" and relates to the perceptive experience of a species.

Bernie Krause tries to solve the dilemma by shifting the definition of soundscape towards the definition of sources and splitting what we usually deal with as a whole in *geophony*, *biophony* and *anthrophony*, respectively the sounding signatures, or better the traces, of non-biological natural sources, non-human and non-domestic biological sources, and human sources (Krause, 2012). On one hand, this distinction bounds itself to serve as a criterion of categorization that, by implying a distinction between human and non-human beings, refers explicitly to an unavoidable human perspective, as observers of the sonic world, without necessarily drawing into play essentialist or religious justifications. But, by relating to sources, this classification points to the practices that generate sounds and noises as a criterion to distinguish them. In this respect, it is not clear how relying on the dynamics of an engine to understand the accidental, so to speak, *antrophonic* noise of a passing car, is equally as significant as relying on the dynamics of a bow stroked on a violin's string to understand the not at all accidental *antrophonic* sounding of J.S.Bach's "Ciaccona".

A definition of *soundscape* cannot avoid stating the difference between what we call "music" and what we call "soundscape", without risking a poor use of the word "soundscape" to simply refer to all possible acoustic events.

Let's finally take a last classic definition. Barry Truax defines *soundscape* as

> an environment of sound (or sonic environment) with emphasis on the way it is perceived and understood by the individual, or by a society. It thus depends on the relationship between the individual and any such environment. (Truax, 1978)

We can notice here several improvements on the previous definitions.

First of all, the observer is submitted to quantification: from a single one, an "individual" (the researcher?), to "a society", which is a collectivity that exhibits not only an internal organization ruling the interactions between individuals, but also a relationship with the space of these interactions, which constitutes an environment. We are far, here, from relying on an objectivist theory that states the existence of an external word that is independent from observers: the word acquires a structure, in the form of a relationship, as the result of an act of observation. Secondly, this relationship extends its scope from perception to understanding—that is, from a relationship in terms of the bodily organization of sensations to one in terms of the organization of meanings. This last part of the definition points to the active role of observation that not only creates the soundscape as such, but also inscribes it in a narrative and ultimately within an academic path.

We can now get back to John Cage and his lesson about listening. If acoustical silence can never be experienced, not even in an anechoic chamber, then our ordinary experience of silence, together with any other aural experience, needs to be constructed out of a never ceasing aural sensation. *Geophonies*, *biophonies* and *anthrophonies* are only possible if we are able to refer to the sources of the aural sensation because we have investigated them in the first place or because we have learned to detect them during our life experiences or listening exercises. John Cage ultimately teaches us that if we want to define what a *soundscape* is, we need to understand in the first place the way we distinguish it within the *sound-flow*, which also includes the sounds produced by the listening system itself.

Of course, Gestalt laws and the principles of the so-called, *auditory scene analysis* state the strategies we activate in segmenting and segregating the endless sound-flow in order to distinguish in it layers (events that overlap in time) and sections (events that occur in time one after the other), that only later we might be able to list in categories such as *biophonies* and so on.

But how do we choose the correct strategy to apply case by case? There is a gap to be filled between the physiological acquisition of a continuum of auditory stimuli and the psychological application of strategies of perception.

The idea of *Umwelt* shows us why addressing ontologies of sources and objects that are a complex final result of the cognitive process is not of any

help in understanding the constitution of the soundscape from within the *sound-flow*. The *Umwelt* (a German word often translated as "environment") is here intended as the technical term first proposed by the biologist Jakob von Uexküll that responds to the observation that "'no animal can enter into relation with an object as such' but only with its own carriers of significance" (Agamben 2002, p.42). We could here substitute *objects* with *sound objects* or even with, so to speak, *sources related sounds*. But what is the *Umwelt*? According to a definition of Giorgio Agamben, it is

> the environment-world that is constituted by a more or less broad series of elements that are […] 'carriers of significance' (*Bedeutungsträger*) or 'marks' (*Merkmalträger*), that are the only things that interest the animal. (Agamben 2002, p.40)

We have moved from a *sound-flow* to a *sonic-Umwelt*, the system of the units we can distinguish by segmenting and segregating the sound-flow, as far as they are relevant to the operations we enact with our body and therefore define the world we live in.

As a consequence, there must be a different *Umwelt* for each different living being: a particular man's *Umwelt* is different from a particular cat's *Umwelt* and both differ from a particular amoeba's *Umwelt*. At the same time we can group the *Umwelten* that emerge from similar bodily operations in such a way that the individual human being's *Umwelten* can be considered as singular expression

of the general human *Umwelt* that is determined by levelling all the individual differences.

> All animal subjects, from the simplest to the most complex, are inserted into their environments to the same degree of perfection. The simple animal has a simple environment; the multiform animal has an environment just as richly articulated as it is. (von Uexküll, 1934, p.50)

We are now back to the position of an observer who evaluates the degree of complexity of animals' *Umwelt* from its meta-*Umwelt*. This includes what the observer identifies as the surroundings of single animals and "simpler" species, what Uexküll calls *Umgebung*. As a consequence, our *Umwelt* seemingly shows an organization of *Umgebungen* that are partially overlapping and whose extension is related to the complexity of animal cognitive behaviours. These *Umgebungen* are what we can access of the animals' *Umwelt*, the image of their *Umwelt* form our point of view. If we turn to the experience of sound, we can therefore define a sonic-*Umgebung* of a living being as our aural experience of its surroundings. But the organization of *Umwelten* like a series of Chinese boxes, with us representing the top-level box, is just an illusion that comes from the fact that we cannot escape our *Umwelt*. We can therefore suppose that the real organization of *Umwelten* must be more similar to a scattered rhizomatic system over different planes of *Umgebungen* amongst which there might possibly be an unknowable meta-*Umgebung*.

This means that the position of an observer is never neutral, not even when he appeals to disciplines with a strong epistemological statute, because

"everything said is said by an observer to another observer that could be him or herself" (Maturana, 1988, p.27).

But the *Umwelt* is not really something we are in: it is rather the space recursively defined by the interactions between our bodily actions against, we could call it, a *substrate*, that is unreachable outside our possibilities of action.

In the words of Humberto Maturana and Francisco Varela, the *Umwelt* is our cognitive membrane, whose integrity, the defence against the possible disruptive results of our actions, is the very concern of our life. This corresponds to the analysis of living beings that Maturana and Varela proposed in 1974 as the *autopoietic* model (Maturana, Uribe and Varela, 1974).

Moving from the *Umwelt* to the *membrane* is here an important passage because it allows us to characterize the constitution of the *Umwelt* as an endless process of surviving in the substrate by acting recursively to repair it from the damages caused by the unknown.

Maintaining the membrane is therefore a continuous operational loop of detections and reactions that we, as observers, can describe on one hand as the structural coupling between the organism and its *Umgebung*, and on the other hand, in considering it as the process of constituting the organism's *Umwelt,* as a constellation of loop patterns or cognitive habits that give rise to signification. As stated by Francosco Varela:

> signification arises in the emergence of a viewpoint proper to the autonomous constitution of the organism at all its levels, starting with its basic autopoiesis. (Varela, 1992)

351

This is the level at which automatic mechanisms of cognition, such as Gestalt laws, emerge as complex epiphenomena of *autopoiesis* in order to satisfy the need of survival so that distinguishing, as the tool for meaning construction, is finally possible.

But *distinguishing* is just a word for an extremely intricate process that in the cognitive domain of description involves high order processes such as attention and memory. Jones' regions of distinction seem in the end to be attentional spans driven by the dynamic and recursive process of attending in which no cognitive effort is made unless a significant change is detected that eludes the cognitive prediction, exactly in the same way as the autopoietic organism entails repetitive routines of maintaining the membrane until an emergency occurs. The cognitive condition related to the routines of maintaining the membrane is known as *habituation* and is related to a reduction of attention at the point that a threshold of consciousness may be crossed, in which case the object of perception may disappear in front of our eyes, like the so-called "lazy eye" or *amblyopia*, or, what is a well-known aural experience, when we might stop hearing, as it happens with a constant noise or a car alarm playing for a long time. On the other hand, the cognitive condition related to emergency situations is known as *arousal* and is accompanied by an increment of attention and a strong consumption of

energy. The two conditions engage a complex, mobile and, so to speak, competitive articulation of foreground and background that develops during our entire lifetime and constitutes (always already) the *Umwelt.*

As we have seen: we do not perceive, to a certain extent, what we are used to, because the constants of our aural experiences, on any level of our structure of distinctions, fall in the CR and are submitted to *habituation*. But *habituation* conceals itself as an action that unfolds in time: as the action that defines the limits of our *Umwelt* and offers itself to us as a presence, is the process of *misattributing* (Huron, 2006) our distinctions to the world's objects as their properties.

Our recursive experiences in listening are stretched between the disappearance of conscious sensations and their constitution as presence in our sonic landscape. We are therefore oriented to the new by forgetting what we are used to in terms of events, in order to constitute them as objects ("chunks"). This means that, by losing consciousness about what we are used (or habituated) to, we embody it and project it as a presence in our *Umwelt*, in our surroundings. In other words, we are (read as "we build our identity on") what we cannot (consciously) listen to any more (CR) and consequently what we started to perceive as a quality of what we consider our environment. On the other hand, what we (consciously) listen to (SIR) points us to the different, to the other, to what in fact opposes our identity and eventually changes it as

soon as it becomes a recursive experience and therefore reaches the CR at some level.

The most important critics of the "abusers" of the term "soundscape" accuse them of engaging the evanescent experience of sound in a process of objectification mostly driven by an incorrect association of the "scape" part of the word to a visual explanatory paradigm[30]. In general, an opposition is drawn between *soundscape* and *sound* so that as a solution it has been proposed to re-define the research on *soundscape* as a research on "listening against", a sort of a new listening mode (Helmreich, 2010), or on "background noise" (Kelman, 2010).

As we have seen, these critiques have a strong foundation in the objectifications operated by the principle of *misattribution* that is enacted by our description of the sonic-*Umwelt* when it naively incorporates the illusion of a word of objects.

Besides, the proposal of turning to the "background noise" also has a legitimate reason: we usually consider music as opposed to the noise that surrounds us, as the SIR opposes to the CR or better the constitution of a foreground opposes to the constitution of a background. This is the opposition that is often removed, as we have seen, from the definitions of *soundscape* and yet it is vital to understanding the subject.

---

[30]While being associated with the visual term "scope", the actual etymology of "scape" is the Dutch *schap*, "cognate with the English suffix '-ship', referring to a fellowship or community of persons with a commonality of land, law and custom" (Ingold, 2007).

We listen to music because we seek the thrill of surprise that is provided by the *arousal* condition, the process of detecting within the SIR, even if it is nevertheless oriented by the possible routine of going to concerts, listening to the mp3 player and so on, that express a CR at some pretty high level. On the other hand, what we usually consider as the background noise disappears as the result of the process of *habituating* to a CR in a relatively low level of the structure of distinctions: we usually do not pay attention to the ordinary noises that surround us, unless they deviate from the range of our predictions.

This does not mean that we are totally deaf to them: our capability to turn our attention to CR level depends on the time span of those levels so that a constant buzz, that expresses a very fast CR level, will be more likely to vanish, or better to be embodied in our identity, than a regular event that occurs, let us say, every year. This happens because the larger the span, the more SIRs will be included in the faster levels of the structure. On the contrary, the process of *habituating* takes time, so that for higher levels, for longer time spans, a lot of time will be needed to get used to recognizing the CR as a constant in our lives.

Nevertheless it is not yet the time to get rid of the word "soundscape". In fact, the discovery, so to speak, of the soundscape corresponds to the discovery of noises and sounds in the background that surround us, the aural experiences we tend to ignore because we are habituated to them and they have become part of who we are.

The soundscape is therefore the process of getting in touch with our cognitive membrane in the domain of language, the exploration of *sonic-Umwelten* in terms of the *sonic-Umgebungen* that surround individuals, cultures and species. It is the process of translating the CR levels of the structure of distinctions into SIR levels in some possible explicative dimension (verbal, graphical and also in scientific data). Not only this: the practices related to the research on the soundscape have brought us to listen to recordings of our and other's soundscapes. Of course, we are not habituated to exotic soundscapes, so no wonder if these practices have potential in bringing the CRs of other cultures into our SIRs and eventually to the creation of a music genre, *soundscape music*. But *soundscape* is opposed to *music* and the research around the *soundscape* is different from *musicology* because it considers music only in relation to its CR levels, such as in background music or in general acoustic content of music.

On the other hand, the topic of *soundscape* is aurally related to the effort of bringing CR levels into SIR levels, in our own structure of distinctions, in our own *sonic-Umwelt*. This is the most difficult process of all and is the most revolutionary aspect of *soundscape studies* because they aim to recover what we cannot hear any more, what we have taken for granted and has become, willingly or not, part of our identity; they aim towards change and for that reason they have a fundamental visionary attitude towards the development of self consciousness.

We have interpreted the constitution of the subject, similarly to Bachelard's metaphysic of the Instant and to Howard Margolis' analysis of the subject in terms of strategies of choices (Margolis, 1987), as the result of a narrative that is grounded on a bundle of habits. Our word is ultimately built from the patterns that have been habituated and have "solidified" during our whole lifetime, so that they have become the background, which is always renewed as such, of our action in the world. Consequently, this world/background is different for each subject. But by being at the same time entailed by the recursive interactions of a social context of subjects, linguistic interactions included, and by being entangled by them, we can deal with our world in the language we share as well as with the subjective experiences of our "praxis of living", so that in the end we end up being involved in a Wittgensteinian subjective understanding, according to which we attach to our discourses subjective experiences that we erroneously suppose we are sharing. And in the end, we can all pretend to live in the same world of objects that we can address with our public language.

However, the narrative of pattern puts the constitution of the subject, of our identity, in the context of the interplay between memory and expectations, as statistically learned. This also means—I am following here the conclusions of my article "What we recognize" (Viel, 2016)— that our identity is somehow related to the frequency of the occurrences of our experiences.

We are therefore defenceless in front of the forced exposition to practices that constrain us to recognize, to embody and finally to slot in our consciousness, the patterns that have become objects and consequently are guarantors of objectivity, of an irrefutable Truth (translated from: Viel, 2016)

In the end, who has the control of the statistics of our experiences, also has the control of what we learn from them and therefore is able drive us in the constitution of our identity. In such a condition within a media culture, Bachelard's call for *progress* is more important than ever.

In order to maintain the possibility of acting in the world as creative subjects, it is always necessary to engage new patterns, to drive our attention towards new recognitions, so that on one hand we can keep looking for SIR configurations where a narrative of CR is proposed to us, and on the other side we can neutralize, at least in part, the objectifying power of coercive processes of identitary formation and subjection.

It is therefore necessary to realize the active role that each one of us has in building our own repertory of patterns, by driving the statistics of our experiences. We need to recover the responsibility of our distinguishing acts, by driving them towards targets that are alternative to what is offered by media. We need to engage strategies of perceptive subversion that redesign, in an unpredictable and experimental way, the pattern structures that we consider objectified and naturalized. We need to throw ourselves into the most insignificant statistical occurrences. And, surely, no matter what obsessions

and memes media will tempt us with, nobody will ever be willing and able to take them away from us.

Bibliography

Adorno, T., 1976. Introduction to the sociology of music. New York: Seabury Press.

Adorno, T., 2002. *Essays on music*. Berkeley, Calif.: University of California Press.

Agamben, G., 2004. *The Open. Man and Animal*. Stanford, CA: Stanford University Press.

Agamben, G., 2009. *What is an Apparatus.* In Agamben, G., 2009. *What is an Apparatus and other Essays*. Stanford, CA: Stanford University Press.

Agawu, K., 2004. How we got out of analysis, and how to get back in again. *Music Analysis*, 23, pp.267-286.

Aikhenvald, A. and Storch, A., 2013, Linguistic expression of perception and cognition: a typological glimpse. In Aikhenvald, A. and Storch, A., eds, 2013, *Perception and Cognition in Language and Culture*. Leiden: Brill.

Alburger, M., 2003. La Monte Young to 1960. *Century Music*,10(3), pp. 3–9.

Atkin, A., 2013. Peirce's Theory of Signs. In Zalta, E.N., *The Stanford Encyclopedia of Philosophy*, [online] Available at: https://plato.stanford.edu/archives/sum2013/entries/peirce-semiotics/ [accessed 4 Apr. 2018].

Augustine, 1955. *Confessions and Enchiridion*. Translated by A. C. Outler. In: *The Library of Christian Classics, VII* Philadelphia, Westminster Press, [online] Available at: <https://www.ccel.org/ccel/augustine/confessions.txt> [accessed 23 Aug. 2016].

Bååth, R., Lagerstedt, E. and Gärdenfors, P., 2014. A prototype-based resonance model of rhythm categorization. Iperception, 5(6), pp.548-558.

Bachelard, G., 1932. *L'intuition de l'instant*. Paris: Éditions Gonthier

Beard, D., 2009. A Broader Understanding of the Ethics of Listening: Philosophy, Cultural Studies, Media Studies and the Ethical Listening Subject. *International Journal of Listening*, 23(1), pp.7-20.

Besseler, H., 1993. *L'ascolto musicale nell'Età Moderna.* Translated by M. Giani. Bologna: Il Mulino (Originally published in 1959).

Bezdek, J., Keller, J., Krishnapuram, R. and Pal, N., 1999, *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*, Norwell, MA: Kluwer Academic Publishers.

Bickel, B., 2007. Typology in the 21st century: Major current developments. *Linguistic Typology*, 11(1).

Bierce, A., 1999. *The Devil's Dictionary*. Oxford, MA: University Press.

Boep, J., ed., 1999. *The Raga Giude. A Survey of 74 Hindustan Ragas*. Wyastone, UK: Nimbus Records.

Bohm, D., 1981. *Wholeness and the implicate order*. London: Routledge & Kegan Paul.

Bregman, A., 1990. *Auditory scene analysis*. Cambridge, Mass.: MIT Press.

Broschart, J., 1997. Why Tongan does it differently: Categorial distinctions in a language without nouns and verbs. *Linguistic Typology*, 1(2), pp.123-165.

Brower, C., 2000. A Cognitive Theory of Musical Meaning. *Journal of Music Theory*, 44(2), p.323.

Cage, J., 1961. *Silence : lectures and writings*. Middletown, Conn.: Wesleyan, University Press.

McLuhan, M., 1960. Acoustic space. In Carpenter, E. and McLuhan, M., eds, 1960. *Exploration in communication*. Boston: Beacon Pres.

Casey, E., 2008. Taking Bachelard from the instant to the edge. *Philosophy Today, 52*, pp.31-37.

Candlish, S. and Wrisley, G., 2014. 'Private Language' in The Stanford Encyclopedia of Philosophy (Fall 2014 Edition), Edward N. Zalta, ed., [online] Available at: <http://plato.stanford.edu/archives/fall2014/entries/private-language/> [accessed 22 June 2016].

Chion, M., 1994. *Audio-vision*. New York: Columbia University Press.

Chalmers, D., 2010. *The Character of Consciousness.* Oxford, MA: Oxford University Press.

Chion, M., 2009. *Guide To Sound Objects. Pierre Schaeffer and Musical Research*. Translated by J.Dack and C.North. [online] Available at: <http://ears.pierrecouprie.fr/IMG/pdf/Chion-guide/GuideSectionI.pdf> [accessed 18 Jan. 2017]

Chion, M., Gorbman, C. and Murch, W., 1994. *Audio-vision*. New York: Columbia University Press.

Clarke, E.F., 2005. *Ways Of Listening*. Oxford, MA: Oxford University Press.

Coker. J., Casale. J., Campbell. G. and Greene., J., 1970. *Patterns for Jazz*, Lebanon, Ind.: Studio P/R.

Cook, N., 2001. Theorizing Musical Meaning. *Theory Spectrum*, 23(2), pp.170-195.

Corbin, S., 1987. *La musica cristiana dalle origini al gregoriano*. Translated by A. Crespi Bortolini. Milano: Jaca Book.

Cruse, D., 1986. *Lexical semantics*. Cambridge: Cambridge University Press.

Datteri, E., 2012. *Filosofia delle scienze cognitive*. Roma, Italy: Carocci.

Delalande, F., 1993. *Le condotte musicali*. Bologna: Clueb.

Delalande, F., 2013. *Analyser la Musique, pourquoi, comment?i*. Paris, France: INA Editions.

Deleuze, G. and Guattari, F., 1987. *A thousand plateaus*. Minneapolis: University of Minnesota Press.

De Natale, M., 2015. *Per una teoria dell'ascolto musicale. Tra soglie vegetative e immaginario eccentrico.* Milano: Mimesis.

Dennett, D., 1991. *Consciousness explained*. Boston: Little Brown and Co.

Derrida, J., 2010. *Voice and Phenomenon: Introduction to the Problem of the Sign in Husserl's Phenomenology.* Translated by L. Lawlor. Evanston, Ill.: Northwestern University Press.

Deutsch, D. And Feroe, J., 1981. The internal representation of pitch sequences in tonal music. *Psychological Review*, 88(6), pp.503-522.

Di Stefano, N., 2016. *Consonanza e dissonanza*. Roma: Carocci.

Dorf, R., ed., 2000. *The Electrical Engineering Handbook*, Boca Raton: CRC Press LLC.

Eimert, H., 1954. *Manuale di tecnica dodecafonica.* Translated by L. Rognoni. Milano: Carisch.

Evans, N. and Wilkins, D., 1998. *The knowing ear: An Australian test of universal claims about the semantic structure of sensory verbs and their extension into the domain of cognition.* [Arbeitspapier 32]. Cologne, Germany: Institut für Sprachwissenschaft.

Fauconnier, G., 1985. *Mental Spaces: Aspects of Meaning Construction in Natural Language.* Cambridge, MA: MIT Press.

Fauconnier, G., 2003. Cognitive Linguistics. In: Nadel, L., 2003. *Encyclopedia of Cognitive Science.* London: Macmillan.

Fiske, H.,1996. *Selected theories of music perception*. Lewiston: Edwin Mellen Press.

Fiske, H., 2008. *Understanding musical understanding*. Lewiston: Edwin Mellen Press.

Foerster, H. von, 2003. On Self-Organizing Systems and Their Environments. In: Foerster, H. von, 2003. *Understanding Understanding: Essays on Cybernetics and Cognition*. New York: Springer-Verlag.

Ford, A. and Peat, F.D., 1988. The Role of Language in Science. *Foundations in Physics*, 18(12), pp.1233-1242.

Forte, A., 1964. A Theory of Set-Complexes for Music. *Journal of Music Theory*, 8(2), pp.136–183.

Foucault, M., 1970. *The order of things*. London: Tavistock Publications Ltd.

Foucault, M., 1972. *The archaeology of knowledge*. New York: Pantheon Books.

Foucault, M., 1995. *Discipline and punish*. New York: Vintage Books.

Foucault, M., 2005. *The hermeneutics of the subject*. New York: Palgrave-Macmillan.

Frova, A., 2006. *Armonia celeste e dodecafonia*. Milano: BUR.

Gamboa, S., 2007. Hume on Resemblance, Relevance, and Representation. *Hume Studies*, 33(1), pp. 21-40.

Gardner, H.E., 1985. *Frames of mind: The theory of multiple intelligences*. New York: Basic books.

Geeraerts, D., 2010. *Theories of lexical semantics*. Oxford, MA: Oxford University Press.

Gjerdingen, R., 1988. *A classic turn of phrase*. Philadelphia: University of Pennsylvania Press.

Glasersfeld, E. Von, 1981. The Concepts of Adaptation and Viability in a Radical Constructivist Theory of Knowledge. In Sigel, I.E., Brodzinsky, D.M. and Golinkoff, R.M., eds, 1981. *Piagetian theory and research.* Hillsdale, N.J.: Erlbaum.

Gobet, F. and Clarkson, G., 2004. Chunks in expert memory: Evidence for the magical number four … or is it two?. *Memory*, 12(6), pp.732-747.

Goldman, J., 2010. Structuralists contra Serialists? Claude Lévi-Strauss and Pierre Boulez on Avant-Garde Music. *Intersections: Canadian Journal of Music / Intersections : revue canadienne de musique*, 30(1), pp.77-94.

Goodman, N., 1972. Seven strictures on similarity. In Goodman, N., 1972. *Problems and projects*, New York: Bobbs-Merrill.

Graziano, A.B. and Johnson, J.K., 2006. Richard Wallasche's Ninetheenth-Century Contributions to the Psychology of Music. *Music Perception: An Interdisciplinary Journal*, 23(4), pp.293-304.

Hanslick, E., 1891. *The Beautiful in Music*. London: Novello.

Helmreich, S., 2010. Listening against Soundscapes. *Antropology News*, 51(9), p.10.

Hewitt, M., 2008. *Music theory for computer musicians*. Boston, MA: Course Technology, CENGAGE Learning

Heidelberger, M., 2003. Theory-Ladenness and Scientific Instruments in Experimentation. In: Radder, H., ed., *The philosophy of scientific experimentation*. Pittsburgh, Pa.: University of Pittsburgh Press, pp. 138-151.

Hermann, T., Hunt, A. and Neuhoff, J.G., eds, 2001. *The Sonification Handbook*. Berlin, Germany: COST Office and Logos Verlag.

Hirst, G., 2009. *Ontology and the Lexicon*. In Staab, S. and Studer, R., eds, 2009. *Handbook on ontologies*. Berlin: Springer.

Hofstadter, D. and Sander, E., 2010. *Surfaces and essences*. New York: Perseus Books Group.

Honig, W., 1988. The magical number four, plus or minus one: Working memory for numbers of items in animals. *Behavioral and Brain Sciences*, 11(04), p.587.

Hume, D., 2001. *Trattato sulla natura umana*. Milano, Italy: Bompiani.

Huron, D., 2002a. Listening styles and listening strategies. Society for Music Theory 2002 Conference. Columbus, Ohio.

Huron, D., 2002b. A six-component theory of auditory-evoked emotion. In: proceeding of ICMPC7, pp.673-676.

Huron, D., 2006. *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge: MIT Press Books.

Huron, D., 2013. A Psychological Approach to Musical Form: The Habituation–Fluency Theory of Repetition. *Current Musicology*, 96(2), [online] Available at: http://currentmusicology.columbia.edu/article/a-psychological-approach-to-musical-form-the-habituation-fluency-theory-of-repetition/ [Accessed 7 Jan. 2017].

Husserl, E., 1970. *The Crisis of European Sciences and Transcendental Phenomenology.* Evanston: Northwestern University Press.

Husserl, E., 1983. *Ideas Pertaining To A Pure Phenomenology And To A Phenomenological Philosophy*. The Hague: Martinus Nijhoff Publishers.

Husserl, E., 2001. *Analyses Concerning Passive And Active Synthesis*. Dordrecht: Kluwer Academic Publishers.

Hyer, B., 2008. Tonality. In Christensen, T., 2008. *The Cambridge History of Western Music Theory. Cambridge, UK: Cambridge University Press.*

Ibarretxe-Antuñano, I., 2002. Mind-as-body as a cross-linguistic conceptual metaphor. *Miscelánea. A Journal of English and American Studies*, 25, pp.93-119.

Imberty, M., 1986. *Suoni, Emozioni, Significati. Per una semantica psicologica della musica.* Bologna, Italy: Clueb.

Ingold, T., 2007. Against Soundscape. In Carlyle A., ed., 2007. *Autumn Leaves*, Paris: Double Entendre.

Ito, K. (ed.), 1987. *Encyclopedic Dictionary of Mathematics: The Mathematical Society of Japan*. Cambridge, Mass.: MIT Press.

Kelman, A.J., 2010. Rethinking the Soundscape. *The Senses and Society*, 5(2), pp. 212-234.

Keuler, J., 1999. The Paradoxes of Octave Identities. *Studia Musicologica Academiae Sciantiarum Hungaricae*, 40(1), pp. 211-224.

Kolinski, M., 1962. Consonance and Dissonance. *Ethnomusicology*, 6(2), p.66.

Kopytko, R., 1990. Verbs of sensory cognition. A contrastive analysis of a lexical field in the lexicon of Polish and English. *Papers and Studies in Contrastive Linguistics*, 25, pp.59-70.

Kostelanetz, R., 2003. *Conversing with John Cage*. New York: Routledge.

Kirk, R., 2005. *Zombies and Consciousness*. Oxford, MA: Oxford University Press.

Kivy, P., 2002. *Introduction to a Philosophy of Music.* Oxford, MA: Oxford University Press.

Kramer, J.D., 1988. *The Time of Music.* New York: Schirmer.

Krause, B., 2012. *The Great Animal Orchestra: Finding the Origins of Music in the World's Wild Places*, London: Little, Brown.

Kuhn, T., 1970. *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Jenkins, I., 1968. The Modern Distemper: The Failure of Purposivness. In: Kuntz, P., ed., 1968. *The Concept of order*. Seattle: Published for Grinnell College by the University of Washington Press.

Jones, M., 1976. Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83(5), pp.323-355.

Jones, M., 1981. Only Time Can Tell: On the Topology of Mental Space and Time. *Critical Inquiry*, 7(3), pp.557-576.

Large, E.W., 1996. Modeling beat perception with a nonlinear oscillator. In *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society,* pp. 420-425.

Lakoff, G. and Johnson, M., 1980. *Metaphors we live by*. Chicago: University of Chicago Press.

Lakoff, G., 1987. *Women, fire, and dangerous things*. Chicago: University of Chicago Press.

La Via, S., 2007. L'espressione dei contrasti tra madrigale e opera In: Borio, G. And Gentili, C., 2007. *Storia dei concetti musicali. Espressione, forma, opera*. Roma: Carocci.
Lévi-Strauss, C., 1969. *The raw and the cooked*. New York: Harper & Row.

Laudanna, A. and Voghera, M., 2002. Nouns and verbs as grammatical classes in the lexicon. *Rivista di linguistica, 14, pp.9-26.*

Lévi-Strauss, C., 1950. Introduction à l'oeuvre de Marcel Mauss. In Mauss, M., *Sociologie et anthropologie*. Paris: Les Presses universitaires de France.

Lévy, F., 2004. *Complexité grammatologique et complexité aperceptive en musique. Etude esthétique et scientifique du décalage entre la pensée de l'écriture et la perception cognitive des processus musicaux sous l'angle des théories de l'information et de la complexité.*[online] Available at: <http://www.ircam.fr/equipes/repmus/Rapports/TheseFabLevy.pdf> [Accessed 6 Mar. 2016].

Ligeti, G., 1964. Metamorphoses of Musical Form. *Die Reihe* 7, pp. 5-19.

Lilienfeld, S.O., Schwartz, S.J., Meca, A., Sauvigné, K. And Satel, S., 2015. Neurocentrism: Implication for psychotherapy practice and research. *Behavior Therapist*, 38(7), pp.173-181.

Lilly, J. 1988. *The Scientist: A Metaphysical Autobiography*. Berkeley, CA: Ronin Pub.

Margolis, H., 1987. *Patterns, thinking, and cognition*. Chicago: University of Chicago Press.

Maturana, H., 1988. *Ontology of observing: The biological foundations of self consciousness and the physical domain of existence.* Conference workbook: Texts in cybernetics. Felton, CA: American Society for Cybernetics Conference. [online] Available at: <https://www.merriam-webster.com/dictionary/pattern> [Accessed 3 Jan. 2017].

Maturana, H. and Varela, F., 1980. *Autopoiesis and cognition*. Dordrecht, Holland: D. Reidel Pub. Co.

Maturana, H. and Varela, F., 1992. *The tree of knowledge*. Boston: Shambhala.

Maturana, H., Uribe, R. and Varela, F.,1974. Autopoiesis: The organization of living systems, its characterization and a model. *BioSystems*, 5, pp.187-196.

Mazzeo, M., 2000. *L'origine tattile della geometria. Merleau-Ponty e il triangolo*. In: Antomarini, B., ed., 2000. *Le tattiche dei sensi*. Roma: Manifesto Libri.

McAdams, S., 1999. Perspectives on the Contribution of Timbre to Musical Structure. *Computer Music Journal*, 23(3), pp.85-102.

Meyer, L., 1956. *Emotion and meaning in music*. Chicago: University of Chicago Press.

Miller, G., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63(2), pp.81-97.

Möller, J., 1991. The domain of description, a systemic, action-oriented approach to modelling. [online] Available at: <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=EEF57151ACC5A13115C1FB2A23450D1B?doi=10.1.1.31.3652&rep=rep1&type=pdf> [Accessed 28 Jan. 2017].

Moles, A., 1966. *Information theory and esthetic perception*. Urbana: University of Illinois Press.

Moore, B., 1997. *An introduction to the psychology of hearing*. London: Academic Press.

Nagel, T., 1974. What Is It Like to Be a Bat?. *The Philosophical Review*, 83(4), pp. 435-450.

Nancy, J., 2007. *Listening*. New York: Fordham University Press.

Narmour, E., 1990. *The Analysis and Cognition of Basic Melodic Structures: The Implication-realisation Model*. Chicago: University of Chicago Press.

Nattiez, J.J., 1990. *Music and discourse*. Princeton, N.J.: Princeton University Press.

Nattiez, J.J., 2002. Musica e significato. in Nattiez, J., ed., 2002. *Enciclopedia della Musica* (Vol. 2, pp. 206-238). Torino, Italy: Einaudi.

Nechtaval, J., 2011. *Immersion into noise*, Ann Arbor: Open Humanities Press.

Negretto, E., 2012. Expectation and Anticipation As Key Elements for the Constitution of Meaning in Music. *Teorema*, 21(3), pp.149-163.

Nietzsche, F., 1996. *On the genealogy of morals*. Oxford, MA: Oxford University Press.

Nietzsche, F., 2006. *Thus spoke Zarathustra*. Cambridge: Cambridge University Press.

Nöe, A., 2008. Précis of *Action In Perception*. *Philosophy and Phenomenological Research*, 77(3), pp.660-665.

Noë, A., 2009. *Out of our heads*. New York: Hill and Wang.

Ockelford, A., 2005. *Repetition in music*. Aldershot, Hants, UK: Ashgate.

Oliveira, L.F. and others, 2010. Musical Listening and Abductive Reasoning. Contributions of C. S. Peirce's Philosophy to the Understanding of Musical Meaning. *Journal of interdisciplinary music studies*, 4(1), pp.45-70

Oliveros, P.F., Haselager, W.F.G., Manzolli, J. and Gonzalez, M.E.Q., 2005. *Deep Listening: A Composer's Sound Practice*, New York: iUniverse, Inc.

O'Regan, K., 2011. *Why Red Doesn't Sound Like a Bell: Explaining the Feel of Consciousness*. Oxford, MA: Oxford University Press.

Patridge, E., 1958. *Origins. A Short Etymological Dictionary of Modern English*. London: Routledge

Patel, A., 2008. *Music, language, and the brain*. Oxford, MA: Oxford University Press.

Popper, K., 2013. *Realism and the aim of science: From the postscript to the logic of scientific discovery.* New York: in doing this.

Porena, B., 1983. Nuova didattica della musica : n. 2 per la composizione: questioni grammaticali e sintattiche. Milano, Italy: Ricordi.

Riemann, H., 1991. Riforma della teoria dell'armonia. Translated by L. Azzaroni, In Azzaroni, L., ed., 1991. *La teoria funzionale dell'armonia.* Bologna, Italy: QUEB.

Riemann, H., 2010. *Über das musikalische Hören*. Herausgeber: Nabu Press

Rimsky-Korsakov, N., Shteĭnberg, M. and Agate, E., 1964. *Principles of orchestration*. New York: Dover Publications.

Roads, C., 1996. *The computer music tutorial*. Cambridge, Mass.: MIT Press.

Rocchesso, D. and Fontana, F., 2003). *The Sounding Object*. Firenze: Edizioni di Mondo Estremo. [online] Available at: <http://www.e-booksdirectory.com/details.php?ebook=28> [Accessed 15 Jan. 2012].

Roden, S., 2005. *Active listening*. Bronx Soundwalk05. [online] Available at: http://www.soundthreshold.org/download/roden_active.pdf [Accessed 5th Aug. 2016].

Roy, S., 2003. *L'analyse des musiques électroacoustiques: Modéles et propositions*. Paris: L'Harmattan.

Satie, E., 1893. *Vexations*. [online] Available at: <http://www.mayfieldfestival.co.uk/wp-content/uploads/2016/01/Vexations-Score.pdf> [Accessed 8 Feb. 2017].

San Roque and others, 2015. Vision verbs dominate in conversation across cultures, but the ranking of non-visual verbs varies. *Cognitive Linguistics,* 26(1), pp.31–60.

Schaeffer, P., 1966. *Traité des objets musicaux*. Paris: Le Seuil.

Schafer, R.M.,1969. *The New Soundscape; A Handbook for the Modern Music Teacher*, Don Mills, Ont: BMI Canada.

Schafer, R.M.,1993. *The Soundscape: Our Sonic Environment and the Tuning of the World*, Rochester VT: Destiny Books.

Scheffler, I., 1982. *Science and Subjectivity*. Indianapolis: Hackett Publishing Company.

Schenker, H. and Salzer, F., 1969. *Five graphic music analyses*. New York: Dover Publications.

Schneider, M., 1970. *Il Significato della Musica*. Milano: Rusconi.

Searle, J., 2008. *Philosophy in a New Century*. Cambridge, UK: Cambridge University Press.

Sellars, W., 1997. *Empiricism and the Philosophy of Mind*. Cambridge: Harvard University Press.

Serafine, M.L., 1988. *Music as cognition*. New York: Columbia University Press.

Shepard, R., 1982. Geometrical approximations to the structure of musical pitch. *Psychological Review*, 89(4), pp.305-333.

Simon, H.A. And Sumner, R.K., 1992. Pattern in music. In Schwanauer, S.M. and Levitt, D.A., eds, 1968. *Machine Models of Music*. Cambridge, Mass.: MIT Press.

Skelly, J.J., Jones, M.R., Goodyear, C.D. and Roe, M.M., 2003. *Attentional pacing and temporal capture in slow visual sequences*. *Air Force Research Laboratory, AFRL-HE-WP-2003-0078*, Wright Patterson Air Force Base, OH.

Smalley, D., 1986. Spectromorphology and Structuring Process. In: Emmerson, S., ed., 1986. *The Language of Electronic Music.* London: MacMillan.

Smalley, D., 1997. Spectromorphology: explaining sound-shapes. *Organised sound*, 2, pp.107-126.

Sterne, J., ed., 2012. *The sound studies reader*. New York: Routledge.

Subotnick, M., 1988. Towards a deconstruction of structural listening: a critique of Schoenberg, Adorno, and Stravinsky. In Narmour, E. And Solie, R., eds, 1988. *Exploration sin Music, the Arts, and Ideas: Essays in Honor of Leonard B. Meyer*. Stuyvesant, NY: Pendragon Press, pp.87-122.

Sweetser, E., 1990. *From Etymology to Pragmatics. Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: Cambridge University Press.

Szendy, P., 2017. *All Ears: The Aesthetics of Espionage.* New York: Fordham University Press.

Truax, B., 1978/1999. *Handbook for Acoustic Ecology* [online] Available at: <http://www.sfu.ca/sonic-studio/handbook/Soundscape.html> [accessed 01 Jan.2016].

Tuuri, K., Mustonen, M. and Pirhonen, A., 2007. Same sound – Different meanings: A Novel Scheme for Modes of listening. In *Proceedings of Audio Mostly 2007*. Ilmenau, Germany: Fraunhofer Institute for Digital Media Technology IDMT, pp.13-18.

Tuuri, K. and Eerola, T., 2012. Formulating a Revised Taxonomy for Modes of Listening. *Journal of New Music Research,* 41(2), pp.137-152. [online] Available at:

<http://www.auditorysigns.com/kai/papers/Tuuri_Eerola_2012_Formulating_a_Revised_Taxonomy_for_Modes_of_Listening.pdf> [accessed 20 Sep. 2016].

Uexküll J. von, 2010. *A Foray Into the Worlds of Animals and Humans: With a Theory of Meaning.* Minneapolis: University of Minnesota Press.
VA, 2012. Careggi Landscape Declaration on Soundscapes - June 2012, [online] Available at: <http://www.uniscape.eu/allegati/Ref%20UNISCAPE%20CD2-14-06-12%20-Second%20Careggi%20Declaration%20on%20Soundscapes%20draft%20new.pdf> [accessed 20 Dec. 2015].

Ullrich, J., 2014. Animal Music. David Rothenberg, Dario Martinelli, and Martin Ullrich Exchange Their Views on the Topic. *Relations – 2.2 - Nov. 2014 - Minding Animals: Part II* [online] Available at: <http://www.ledonline.it/index.php/Relations/article/view/670> [accessed 2 Feb. 2017].

Varela, F.J., 1992. Autopoiesis and a biology of intentionality. In McMullin, B.,ed., P*roceedings of the workshop "Autopoiesis and Perception"*. Dublin: Dublin City University.

Viberg, A., 2001. The verbs of perception. In Haspelmath, M., König, E., Oesterreicher, W. and Raible, W., eds, *Language Typology and Language Universals. An International Handbook*. Berlin: De Gruyter.

Viel, M., 2012. Sounds and Objects. *Technoetic Arts*, 9(2/3), pp.233-238.

Viel, M., 2014. In the Name of the Pattern. *Noema - Essay, Ideas 22 June 2014* [online] Available at: <http://noemalab.eu/ideas/in-the-name-of-the-pattern/> [accessed 20 Mar. 2016].

Viel, M., 2015. Nell'Alto dei Giorni Immobili: une Analyse Esthésique. In Arbo, A., ed., 2015. *Anamorphoses. Études sur l'Œuvre de Fausto Romitelli*. Paris: Hermann.

Viel, M., 2016. Quello che riconosciamo. In Bianchi, A. And Leghissa, G., eds, 2016. *Mondi Altri. Processi di soggettivazione nell'era postumana a partire dal pensiero di Antonio Caronia.* Milano: Mimesis.

Viel, M., 2017a. *Ordo Coelestis (excerpt).* [online] Available at: <http://www.massimilianoviel.net/listening_patterns/ordo_coelestis.wav> [accessed 1 Feb. 2017].

Viel, M., 2017b. *Presenze (excerpt).* [online] Available at: <http://www.massimilianoviel.net/listening_patterns/presenze.wav> [accessed 1 Feb. 2017].

Viel, M., 2017c. *Cluster (per Demetrio Stratos).* [online] Available at: <http://www.massimilianoviel.net/listening_patterns/cluster.wav> [accessed 1 Feb. 2017].

Wilson, C., 2007. *Let's talk about love. Why Other People has such Bad Taste.* New York: Continuum.

Wittgenstein, L., 1953. *Philosophical investigations*. Oxford, UK: Blackwell.

Wittgenstein, L., 1961. *Tractatus logico-philosophicus*. New York: Humanities Press.

Wittgenstein, L., 1969. *On Certainty*. Oxford, UK: Blackwell.

Wright, C., 2000. *Listening to music*. Australia: Wadsworth/Thomson Learning.

Xenakis, I., 2001. *Formalized music*. Hillsdale, NY: Pendragon Press.

Zeller, H. and others, 1964. *Speech and music*. Translated by M. Shenfield and R. Koenig. Bryn Mawr, Pa.: T. Presser Co.

Websites

Changing Minds. Lexical database [online] Available at: *<http://changingminds.org/techniques/listening/types_listening.htm>* [accessed 20 Dec. 2015].

DANTE | A Lexical Database For English. Lexical database [online] Available at: *<Dante.skxmlbox.idm.fr>* [accessed 20 Dec. 2017].

Internet Archive [online] Available at: *<https://archive.org>* [accessed 8 Aug. 2016].

The World Atlas of Language Structures Online, 2014. [online] Available at: *<http://wals.info/chapter/81>* [accessed 25 Jan. 2017].

WordNet Search - 3.1. Lexical database [online] Available at: *<Wordnetweb.princeton.edu>* [accessed 20 Dec. 2015].