Musical Analysis of Sound Feature Emergences Using Acoustic and Psychoacoustic Descriptors

Análise Musical de Características Sonoras e Emergências de Forma Usando Descritores Acústicos e Psicoacústicos

Danilo Rossetti
Federal University of Mato Grosso, Cuiabá, Mato Grosso, Brasil
d.a.a.rossetti@gmail.com

Micael Antunes
State University of Campinas, Campinas, São Paulo, Brasil
micaelant@gmail.com

Jônatas Manzolli
State University of Campinas, Campinas, São Paulo, Brasil
jotamanzo@gmail.com

Abstract: In the 20th and 21st centuries, after the advent of electroacoustic music and its influences in instrumental music, the expansion of the sound palette employed in musical compositions is considerable. Moreover, with recording techniques and technological advances available, a new culture centered on sound became relevant. Along with the expansion of compositional possibilities, new problems in musical analysis arise, such as how to incorporate emergences of sound and form in the analysis work. We present a methodology of musical analysis that aims to represent these emergent phenomena that arise through perception by computer-aided tools based on acoustic and psychoacoustic descriptors. These tools developed are discussed, and the information achieved is interpreted from two biases: 1) the emergence of sound features related to the interaction of spectral components, and 2) the emergence of macroform of the musical works by perception. As a conclusion, this research aims to give interpretations of those emergence features in music by means of the analysis of the singularities (saliences and pregnances) that arise in the works analyzed.

Keywords: Emergence. Musical form. Sound perception. Computer-aided musical analysis. Audio descriptors. Computational musicology.

Resumo: Nos séculos XX e XXI, após o advento da música eletroacústica e suas influências na música instrumental, a expansão da paleta sonora empregadas nas composições musicais é considerável. Ademais, com as técnicas de gravação e avanços tecnológicos disponíveis, uma nova cultura centrada no som se tornou relevante. Por outro lado, surgiram novos problemas na análise musical, tais como incorporar emergências sonoras e formais no trabalho analítico. Apresentamos uma metodologia de análise musical que procura representar estes fenômenos emergentes que surgem pela percepção por meio de ferramentas computacionais baseadas em descritores de áudio acústicos e psicoacústicos. Estas ferramentas desenvolvidas são discutidas, e a informação obtida por essas análises é discutida de duas maneiras: 1) a emergência sonológica que acontece pela interação dos componentes espectrais, e 2) a emergência da forma musical pela percepção. Como conclusão, esta pesquisa busca fornecer interpretações destes tipos de emergência relacionadas à música por meio da análise das singularidades (saliências e pregnâncias) que surgem nas obras analisadas.

Palavras-chave: Emergência. Forma musical. Percepção Sonora. Análise musical com suporte computacional. Descritores de áudio. Musicologia computacional.

Submetido em: 30 de junho de 2022
Aceito em: 4 de setembro de 2022

Introduction

Near the end of the 20th Century, computational tools began to be developed and used in musical analysis, bringing new possibilities to this field. Among them are the use of computers in symbolic analysis (analysis of the score) and the analysis of digital audio recordings focusing on revealing acoustic and psychoacoustic features, in addition to timbre characteristics that can complement the information extracted from the score. Those tools were capable of describing the characteristics of the internal content of sounds, such as their temporal, spectral, harmonic, and psychoacoustic features (MÜLLER, 2015; COUPRIE, 2020).

Along with the development of new graphical ways of representing music, listening became central in music activity. Moreover, music perception and cognition were managed to lead towards analytical and creative approaches in different contexts such as in (among others) the phenomenological approach by Schaeffer's morphology of sound object and reduced listening (1966), in the systemic approach by Vaggione's operatory listening (2010), going through details of different temporal scales and searching for a morphological richness, and also in practices such as deep listening by Oliveros (2005), giving attention to sound features perceived both acoustically and psychologically in the whole space and time continuum of sound.

In musicological and epistemic terms, the new realm dialogues with the notion of the emergence of sound in music. As Makis Solomos has argued, thanks to recording techniques and technological advances, it has been possible to listen to music in any place, permanently. From this achievement, a new culture is born, centered on sound. On the other hand, in many cases, the music itself contributed to this focus on sound, such as in contemporary instrumental music or live electronic music. In these cases, this mutation is expressed by the work on timbre and the discoveries related to extended instrumental techniques, instrumental effects, and orchestration (SOLOMOS, 2013, p. 9-12).

Richard Parncutt pointed out that since the mid-20th Century the subdisciplines of musicology have become increasingly specialized, creating their networks and methods. On the other hand, as a unitary discipline, its success depends on the degree of subdisciplines' successful interaction (PARNCUTT, 2005). Parncutt also pointed out the possible bifurcation of systematic musicology in cultural and scientific musicology. The latter, which is where this analysis proposal is placed, is primarily empirical and data-oriented, involving empirical psychology and sociology, acoustics, physiology, neurosciences, cognitive sciences, and computing and technology (PARNCUTT, 2007).

David Huron affirmed that in the past two decades a rise in scientifically music research happened. A new empirical enthusiasm was seen in the areas of psychology of music and systematic musicology (HURON, 1999, p. 1). The reason is that music researchers have more access to computational and database resources, in addition to new reference tools, modeling techniques, and other innovations that are helpful in collecting, analyzing, and interpreting musical information (HURON, 1999, p. 25). Thus, this rise of empiricism is characterized by the motivation of learning as much as possible from this huge quantity of information available (HURON, 1999, p. 29).

The problem we are facing in this article and broadly in our research is related to the different possibilities to represent graphically emergent phenomena in music, depending on the specific characteristics of each musical work that is being analyzed. This question will be detailed and deepened in this article in different ways. In a previous article (ROSSETTI & MANZOLLI, 2017) we stated that musical analyzes based exclusively on the score cannot represent the sonic complexity attainted in the performance of those instrumental works that employ traditional or extended instrumental techniques. By using different recordings of the same piece, we can perform a comparative analysis aiming to reveal interpretative nuances chosen by each interpreter. In another article (ROSSETTI *et al.*, 2018) we defined the phenomenon of emergent timbre which is related to the interaction of sound structures in the microtime level especially in electroacoustic music, generating sonorities that emerge during the performance. The emergent timbre is formed by the amalgam of different sound textures that are perceived as a *Gestalt* unity.

In music, sound phenomena which are "not expected" (or present in the musical notation of the score) occurs in musical performance and are perceived by the listeners. Our hypothesis is that those phenomena can be understood under the concept of emergence (ASHBY, 1956; DI SCIPIO & PRIGNANO, 1996; DI SCIPIO, 2002). As a methodological choice to approach this problem, we find that those emergent sound features can be detected by computer-aided tools such as acoustic and psychoacoustic audio descriptors and explored in computational analysis (PEETERS, 2004; PEETERS *et al.*, 2011). With the aid of those tools, we can generate quantitative data related to spectral, temporal, harmonic, psychoacoustic, and other features of sound that are considered as the basis of qualitative musicological discussions and conclusions in the analysis field. We sustain that a complementary analysis of the audio recordings and the score (when it is available) is very useful since the more information we can recover from different sources, the more embracing the conclusions of the analysis will be.

Starting upon this recent musicological scenery, the contribution of this article is to present computer-aided tools and graphical representations to support analytical strategies of such repertoire, and to define the epistemological basis of such methodology we worked in the last few years (ROSSETTI, 2017; ROSSETTI & MANZOLLI, 2017; ROSSETTI *et al.*, 2018; ROSSETTI & MANZOLLI, 2019; ANTUNES *et al.*, 2019; ROSSETTI *et al.*, 2020; ANTUNES *et al.*, 2021). The scope of the methodology presented here is not restricted to electroacoustic music (acousmatic and live-electronic) but it can be applied to study other repertoires such as the contemporary instrumental music that makes use of extended instrumental techniques (ROSSETTI *et al.*, 2018), textural and sound mass music of the second half of the 20th Century (ANTUNES *et al.*, 2021) and the analysis of the performance based on recordings of different periods and aesthetics (ROSSETTI & MANZOLLI, 2017).

In short, our contribution is fitted around topics of musical analysis incorporating concepts such as process and emergence mediated by perception aiming to define temporal, formal, and structural characteristics of the analyzed works. Of course, none of those concepts are new and have already been previously discussed in this research area with different methodologies, as detailed in the next paragraphs. The novelty we are introducing is the manner to connect and inter- relate these concepts, which is connected to the authors' viewpoint expressed in this article. The possibilities of this methodology are amplified when the computer tools are associated with musicological information about the creative processes employed.

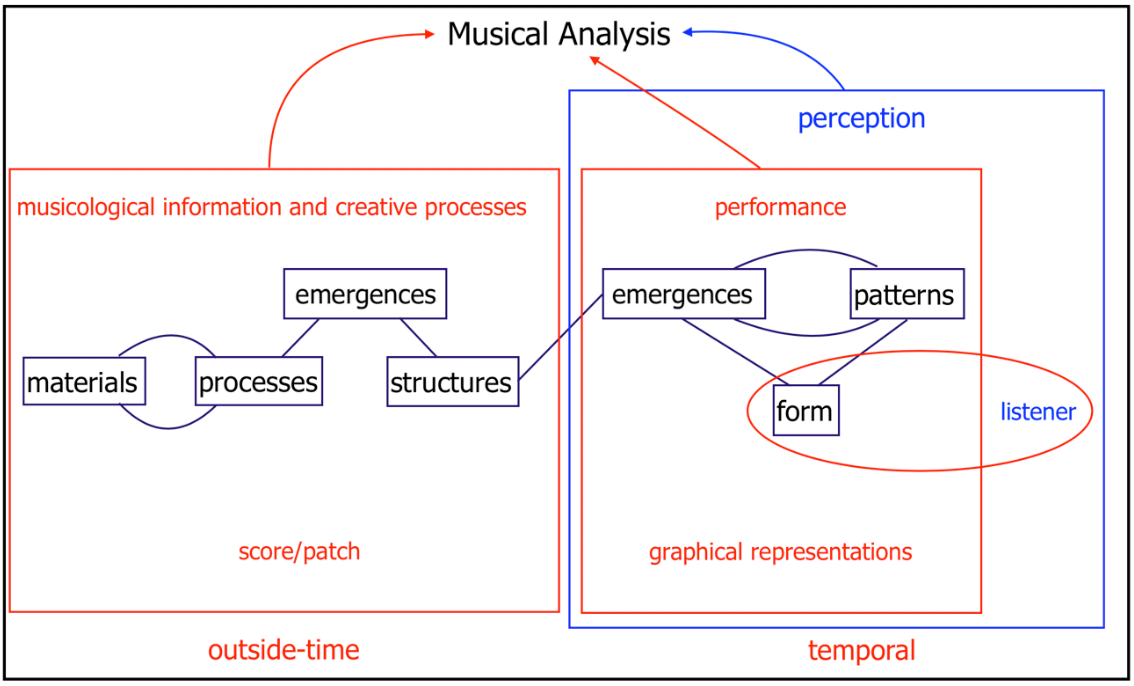
The contribution to scientific and empirical musicology we are aiming to discuss is the mediation between an analysis of the musical text and the analysis of the data recovered from the analysis of the emergent features that arise in compositional processes and performance, the latter carried out by audio descriptors, as well as the graphical representations constructed from this information. We do not seek a definitive view of the works analyzed, but mainly an interpretation of the perceptive emergence related to spectral features and musical form.

Concepts and notions related to the analysis method proposed

To represent emergence in music, our aim mainly lies in constructing and proposing alternative graphical representations of acoustic and psychoacoustic sound features we intend to describe and interpret, by applying audio descriptors for data extraction of digital recordings, as we will discuss afterward. Figure

1 is an essay to summarize how we conceive the methodology of analysis we will present.

Scheme 1: Schema of the methodology of analysis discussed in the article.



Source: The authors.
Image description: Sketch of the analytical proposal.

In the Scheme 1, two bigger boxes are found, described as outside-time (the creative processes of composition related to the score and patch) and temporal (the action of musical performance). Those categories are thought considering Iannis Xenakis's conception of compositional structures in relationship with time: outside-time (*hors-temps*), in-time (*en-temps*), and temporal (*temporel*) (XENAKIS, 1965, in XENAKIS, 1994, p. 68; XENAKIS, 1992, p. 183; ROSSETTI, 2012, p. 12-13). The outside-time category refers to categories of musical formalization and structuration and is connected with the idea of representation. In this sense, the score is the representation in two dimensions with instructions of how a musical work should sound, but it is not the music actually happening in the flow of time. The patch, as well as, is a digital representation of sound that is analogous to traditional musical notation. It is presented as a graphical interface that is at the same time an abstraction and a device for sound generation and transformation (VAGGIONE, 2010, in SOULEZ & VAGGIONE, 2010, p. 57-62). The patch can be an element that aids the musical compositional process and is also used in musical performance for sound generation or transformation, but it is conceived in both cases during the compositional process, as a structural element of the piece. Finally, the temporal category of Xenakis refers to the realization in-time of outside-of-time constructions in its real occurrence, during the musical performance itself (XENAKIS, 1965, in XENAKIS, 1994, p. 68).

Musical performance, which is temporal, engages the perception of form by a listener. The musical form incorporates the sound emergences that arise in the musical act, just like the presence of patterns that gives a sense to the musical flow (ROSSETTI *et al.*, 2020). Those emergences, patterns, and form are new elements that are not found in the outside-time representations such as the score and patch (the concepts and the differentiation between structure and form will be discussed in the next section). So, they need new graphical representations that will help their description and analysis. This analysis schema aims to identify musical information which is outside-time and temporal processes and to propose a methodological approach to deal with and relate both cases in a complementary view.

The main concepts and notions relevant to be considered in this methodology are 1. materials; 2. processes; 3. emergences; 4. structures; 5. patterns; and 6. form. In the discussion about the idea of process in musical composition and performance, it is interesting to examine the relationship between different processes, in addition to the relation between processes and the emergence of patterns leading to the perceptive form. The perceptive results and their emergences can be examined in terms of Gestalt psychology (SIMONDON, 2013 [1964]) and semiophysics' attributes of saliences and pregnances (THOM, 1981; 1988). All the investigations and interactions regarding musicological information and creative processes, performance, and their related applications are the general scopes of our contribution.

In our analysis, we consider "sound" relevant to the construction of music. Our starting point is that the score is very important but, in several cases, it is not sufficient to represent the whole musical phenomenon perceived by a listener. The outside-time structures found in the score lead to the emergence of patterns in time during the performance, constituting the audible musical form. Our analysis is focused on such auditory patterns perceived in the process of listening or the definition of perception indexes since each individual has his/her unique perception. To define these indexes, aiming to raise musicological attention and describe the sound phenomena, we apply acoustic and psychoacoustic audio descriptors to generate graphics and other kinds of representations. Acoustic descriptors extract spectral information of the audio signal, such as energy, frequency, partials, and its durations, and other acoustic measures, and psychoacoustic descriptors are computed using models of human earring processes (PEETERS, 2004, p. 2).

In that context, we propose a differentiation between structure, pattern, and form concepts. The structure of a musical work is given by the compositional processes employed and can be identified when we observe it as an out-of-time object. Structures can be identified in the score or in the electronic patch employed in live-electronic music. Patterns are temporal features that emerge and are identified in listening. They can be observed in graphical representations such as in phase-space graphics that give us spectral liveness and recurrent patterns representations (discussed later). In addition, audible form is also an attribute of perception and needs time to be identified while the piece is being performed, be it a live performance or a recording. The form perceived or audible form is dependent on the emergent phenomena and the organization of the perceived patterns and indexes (ROSSETTI *et al.* 2020, p. 164-166). Those features are perceived in the temporal musical flux, including elements that are similar or different, moments of continuity and abrupt cuts, silences, repetitions, and many other features that can be discussed. Our conclusion which will be detailed in the final part of the text indicates a possible emergence of musical signification from the mediation of both parts of the schema shown in Scheme 1.

The emergence of sound features and form mediated by perception

Here, we enumerate the key concepts in which our methodology of analysis is anchored. The analytical methodology we propose intends to meditate the musical writing (compositional operations and processes) and music perception from the employment of analysis tools such as audio descriptors and alternative representations of sound (which will be discussed next). We will define the concept of *emergence* and present relationships between this concept and perception, as well as how the recurrence of musical and spectral patterns leads to the emergence of form.

Emergence and music as a complex dynamic system

We consider emergence as the result of a process that is greater than the sum of its parts and has properties not shared by them (ASHBY, 1956, p. 109-113). As background, in music, we dialogue with the idea of morphological emergence as a result of microtemporal compositional processes (DI SCIPIO, PRIGNANO, 1996, p. 32) and emergent properties of dynamic systems in environmental sound textures and ecological systems (DI SCIPIO, 2002. MERIC & SOLOMOS, 2014). According to Di Scipio, emergent properties mean global features that appear in a system as a result of the dynamical interrelationship of system components (DI SCIPIO, 2002, p. 113).

Di Scipio argues that, in music, the sound is the result of a lower-level process designed by a composer. The interactions and interferences among components involved in the compositional process are transformed into sound (by acoustic or electronic instruments) as a process of sonological emergence (ANDERSON & DI SCIPIO, 2005, p. 18). In that sense, Meric and Solomos explain that emergent sonorities describe the hypothesis of 2nd order sonorities of the electronic music of the 1950s (MERIC & SOLOMOS, 2014, p. 5). On the same line, emphasizing common aspects of music and dynamic systems, Manzolli (1993) employed non-linear dynamics and fractals as a model for sound synthesis and real-time composition.

In this article, we will deal with questions about spectral emergence (emergence of sound features) and form emergence (emergence of a macroform related to perception). In that sense, we find it relevant to define synchronic and diachronic emergence. In synchronic emergence the higher-level, emergence phenomena are simultaneously present with the lower-level phenomena from which they emerge. It can be viewed as a "vertical" emergence. Diachronic emergence, on the other hand, is a "horizontal" emergence evolved through time, in the sense that "the structure from which the novel property emerges exists prior to the emergent" (VINTIADIS, 2013).

In the context of complex dynamic systems, time is considered as an independent parameter, above which other parameters are attached or vary. This idea appeared in the thermodynamics of the 19th Century, coming together with the idea of an "arrow of time", or time in "one-way-direction". Henri Poincaré, in *Les méthodes nouvelles de la mécanique céleste* (1892), pointed out that complex dynamical systems evolve in time in one direction, which means that they are unstable (PRIGOGINE, 1995).

Perception

The concepts presented above can be understood in musical practice from a mediation by the auditory perception. The perception of musical form most of the time deals with the presence of patterns and recurrences (ROSSETTI *et al.*, 2020). For Simondon, the object perceived (sound) is dynamic and is transformed without losing its unity. It is an autonomous object with relative energetic independence. There is a genesis of forms so as there is a genesis of life: in this sense, perception is the invention of a form and an act of individuation (SIMONDON, 2005, p. 233-235).

Perception can be understood as the invention of a form and an act of individuation[1] (SIMONDON, 2005, p. 235). The object perceived can be seen as a beam of differential relations to which an individual gives a coherence. Through the different gradient levels detected by our senses, the information gains an intensive sense (SIMONDON, 2005, p. 242). To perceive is to organize or to invent an organization: it is the act that links the forms contained in the individual to the signals received (SIMONDON, 2005, p. 243).

Gestalt theory presents useful keys to approach perception. Simondon, in his *Cours sur la perception* (2013, p. 88-95), states that *Gestalt* stands between a metaphysical approach (such as Bergsonism) and a psychophysical analysis. It can be understood as a mediation that integrates form and matter in one unique operation. According to this theory, primitive perception would already have a form and structure, and is considered as a totality, not just bringing elementary sensations. For that reason, *Gestalt* implies the existence of a knowledge of the object in the sense that all forms have a pregnance.

In the same sense, conceptualizing a theory of forms, transformations, and alternative representations, we can recall some ideas of René Thom. Thom defines semiophysics as the

1 For Gilbert Simondon (2005), individuation is an operation from which the material takes its form in a certain system with internal resonance. It is an operation that performs energy exchanges between material and form until the ensemble reaches its equilibrium. For a material to take its form, two processes should occur: transduction and modulation, the former occurring in the micro and the latter at the macro level. Transduction is a physical, biological, mental, and social operation in which an activity propagates step by step inside a domain, structuring it from place to place (Simondon, 2005, p. 32). Modulation is an operation that gives a perceptive form to an object, modeling it in a continuously and variably way (Simondon, 2005, p. 47). Simondon also formulated the allagmatics (allagmatique), a theory of operations that he considers the operatory side of scientific theory. The allagmatics defines the relationship between an operation and another or the relation between an operation and a structure (Simondon, 2005, p. 559-561). In addition, from that theory, Simondon proposed the definition of the analogical act as the connection between two operations, or the transference of operations (the identity of operatory relations). This method supposes that we can obtain the knowledge about things by defining the structures by the operations that drive them, instead of defining the operations by the structures between which they act, such as a mediation between two terms. The analogical thought reveals the identity of relations involved in processes and not relations of identity (p. 562-563).

research of significant forms, aiming to constitute a general theory of intelligibility (THOM, 1988, p. 11). To understand perception, he employed the concepts of salience (saillance) and pregnance (prégnance). The salience is the experience of discontinuity from the continuous background, belonging to the notion of individuated form. The pregnance supposes the condition of continuity. The propagation of pregnances would be the principle of the intelligibility of phenomena (THOM, 1981, p. 54).

To understand perception as the invention of a form or as an act of individuation, as Simondon argues, means to understand perception as an operation conceived individually. This act is formed by processes and operations that occur synchronically and diachronically (such as transduction and modulation – see footnote #2) leading to the emergence of form in different time scales, be it in a mesoscale (in terms of sound objects – sound feature emergence) or in a macro scale (the form emergence of a musical work). Moreover, Thom defines semiophysics as a method of intelligibility of the world founded on the notions of salience and pregnance. From those notions, we can search to build a knowledge of the studied objects from the processes linked to the emergence of their forms.

Form, representation, and pattern recurrences

The form and segmentation of musical works from a perceptive standpoint can be revealed by alternative graphical representations elaborated with information extracted from the recordings by audio descriptors. Those representations can aid the analyst to understand specific musical features of the work he/she is dealing with. We highlight, according to Mikhail Malt (2015, p. 243), that there is no single answer about what is a good representation in musical analysis. In fact, the element that should be considered is the objectivity of the representation. Referring to the production of knowledge, different types of representations of the same phenomenon are one of the keys to the answer since different representations can produce strong effects on cognition.

Segmentation of form can be achieved from a graphical representation when discontinuities and(or) oppositions are found in the musical discourse. From Gestalt or Thom's semiophysical point of view, we can speak of pregnances (or continuities of form) or saliencies (discontinuities or breakpoints of the form). In this sense, we assume that the huge majority of musical works are formed by a combination and interpolation of moments of continuity (pregnances) and discontinuity (saliences). Normally, breakpoints in form or discontinuity moments give space to the installation of new pregnances that stand for a longer duration. At some point, those new pregnances come to an end, when invaded by new saliences, which provoke new ruptures in form (THOM, 1988, p. 21).

Pattern recurrences regarding the sound material, on the other hand, can be associated with continuity of form and of the perception of time. It means that recurrence (or repetition) in sound material – that can also be interpreted as the return of a musical idea in a varied way – is related to the psychological responses of our sensation of time and with pregnances of form. In that way, the emergence of a continuous time perception is related to the presence of similar events and transformations of an original material. Thus, we assume that the form is given by the identification of order, recurrences, and periodicities. On the other hand, a non-identification of patterns and recurrences in a musical work leads us to the emergence of a discontinuous form perception where fewer or no periodicities are found (ROSSETTI *et al.*, 2020).

Computer-aided tools and graphical representation developed

Here, we present different tools we employed and developed for our analysis' method proposal that were employed in previous analyses. The graphical representations aim to highlight emergent musical phenomena that were not predicted in the score in instrumental and live electronic music. In fixed-media pieces of electroacoustic music, those representations can highlight specific characteristics of microtemporal events and how they can be related to the macroform, among other possibilities.

Audio descriptors, graphical representations, and their applications in musical analysis

Audio descriptors provide data and graphical representations that support analytical tasks. Since we have a variety of aesthetical and compositional procedures in the musical repertory, the use of multiple models and representations is necessary. The choice between one or another descriptor depends on the context of the musical material and the perceptual phenomena intended to be analyzed. For that, we present a framework with descriptors we employ in musical analysis and the musical context in which it occurs. The goal of this framework is to provide measures and representations that are linked with our perception, using mainly psychoacoustics as mediation.

In this section, we introduce the descriptors divided into 2 categories: 1 – descriptors that provide one-dimensional representations (loudness, spectral centroid, spectral spread, spectral flux, and spectral flatness) and 2 – descriptors that provide multidimensional representations (volume, spectral liveness, recurrent patterns, auditory filters and mean energy of auditory filters). In the next, section we discuss separately sound feature emergence descriptors and descriptors related to the emergence and perception of form.

To extract some features of the audio recordings, we briefly introduce descriptors[2] that are usually used in MIR, musical analysis, and music composition. Their definitions and models

2 All data of the analyzes presented in this article (in text files) is available at: https://drive.google.com/drive/folders/1cd6a7mBbfynIe286-zhh7f-rE3IrBArt?usp=sharing.

of calculation have already been introduced in previous works such as Rossetti *et al.* (2020), Antunes *et al.* (2019), and Rossetti & Manzolli (2019).

The subjective perception of the sound intensity is estimated by the loudness model. This model is based on curves of equal loudness in function of the frequency, as proposed by Fletcher & Munson (1933) and reviewed by the International Organization for Standardization (ISO, 2003). The spectral centroid is defined as the barycentre of the spectral content concerning each window of analysis (PEETERS, 2011, p. 2907). It is calculated as the frequency-weighted mean of the magnitudes of the Discrete Fourier Transform of the analysis window. The higher is the frequency values of spectral components, the higher the value of the spectral centroid. The spectral spread represents the spread of the sound components around its mean value (PEETERS *et al.*, 2011, p. 2907). Malt and Jourdan (2009) associate the spectral spread with the idea of the thickness of a sound (KOECHLIN, 1954).

Spectral flux is a measure of the level of change of power of the spectrum (PEETERS *et al.*, 2011, p. 2909). It means that the lower the changes in the power spectrum, the lower the level of the spectral flux. The spectral flux was already applied to investigate the perception of the emergent timbre in a live electronic music context (ROSSETTI *et al.*, 2018). Spectral flatness determines if the spectra have more noise-like or tone-like characteristics (PEETERS *et al.*, 2011, p. 2908) by calculating the homogeneity of the distribution of the spectral content.

Based on experimental tests of dissimilarity-rating tasks with sounds of musical instruments, McAdams *et al.* (1995) affirm that the spectral centroid and spectral flux (together with the spectral irregularity) are important dimensions of the timbre differentiation. Finally, the spectral centroid is currently associated in literature as a good measure for the perception of brightness (MALT, JOURDAN, 2009. MCADAMS *et al.*, 1995).

Multidimensional representations and their applications

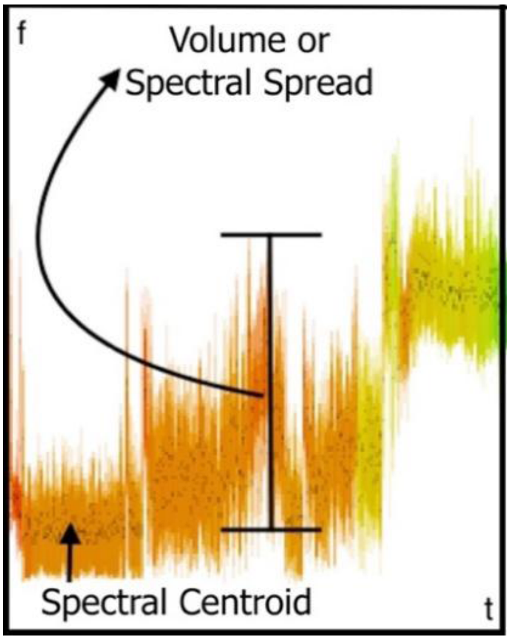
Taking into account the multidimensionality of the sound phenomena and their behaviour in time, it is useful to use multidimensional representations based on audio descriptors data. They provide a more meaningful approximation between the graphical representation and how we perceive sounds. So, we introduce two kinds of multidimensional representation: 1) descriptors that compile data from one-dimensional representations to represent: a) a subjective perception of the timbre magnitude, denominated volume (ROSSETTI & MANZOLLI, 2019), b) the sound quality of tonal/noise characteristics based on the amount of spectral activity in time, the spectral liveness (ROSSETTI & MANZOLLI, 2019), c) the spectral patterns' repetition of the sound spectra in time, denominated recurrent patterns (ROSSETTI *et al.*, 2020); and 2) auditory filters, that aim to represent the spectra linked with the properties of the inner ear, such as the mel scale descriptor (ROSSETTI *et al.*, 2020), and the perceptual accumulation in frequency-bands (ANTUNES *et al.*, 2020).

To construct those representations, we employed the Ircamdescriptors library to calculate the audio descriptors data (PEETERS, 2004) connected with the MuBu library (SCHNELL *et al.*, 2009) to build the graphical representations of kind 1, and the Libxtract library (BULLOCK, 2007) to calculate the data for the auditory filters of kind 2.

Volume

An important tool for the analysis of the spectral behavior of a musical work and its perception in a musicological approach is the volume descriptor. The volume descriptor actually connects the graphical representation of the BStD descriptor (MALT, JOURDAN, 2009), which combines three audio descriptors presented above (loudness, spectral centroid, and spectral spread) with the concept of volume by Truax (1992). Truax defines the volume as the "perceived magnitude of sound or the psychological space it occupies" (TRUAX, 1992). We implemented the graphical representation of the volume descriptor (based on the BStD descriptor) in Max/MSP using the Ircamdescriptors and MuBu libraries, as presented in Figure 1.

Figure 1: Graphical representation of the volume descriptor.

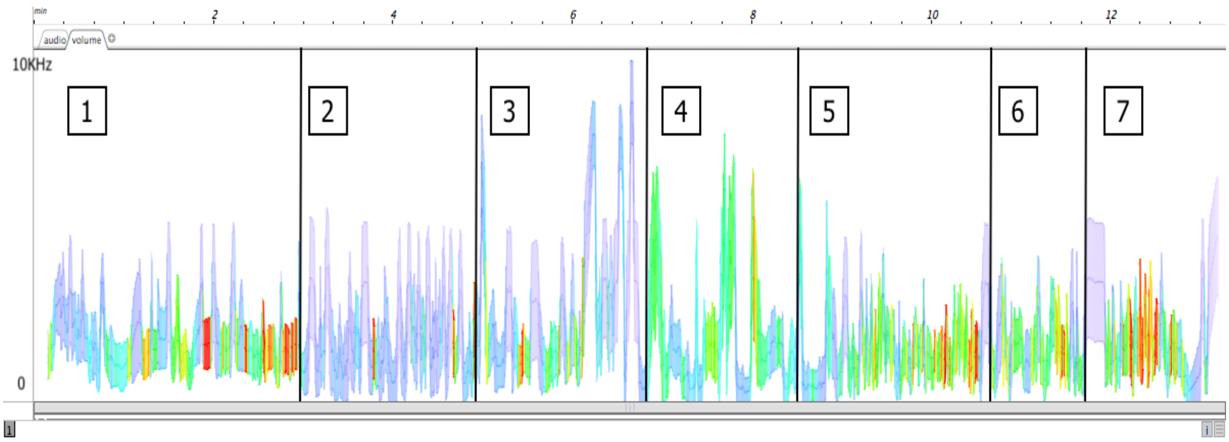


Source: The authors.

Image description: The central line is the spectral centroid, the spectral spread is given by the area around the spectral centroid, and the loudness is represented by the colors (ROSSETTI & MANZOLLI, 2019, p. 209).

We introduce an example of volume representation of the work *Gesang der Jünglinge* (1955-56), by Karlheinz Stockhausen, including its segmentation, as presented in Rossetti *et al.* (2020, p. 162). In Figure 2, we emphasize the idea of the variation of the spectral parameters in time. We have a low level of variation of the parameters in segment 1. On the other hand, segment 3 exhibits a high variation of the spectral features. By analyzing the profile of the descriptor, we can understand the development of the piece as a whole, proposing a segmentation by observing the similarities and saliences (breaking points) of the graph. Based on the volume representation, we propose a segmentation of the work into 7 parts. More details are found in Rossetti *et al.* (2020).

Figure 2: Volume descriptor representation of *Gesang der Jünglinge* (ROSSETTI *et al.*, 2020, p. 162)



Source: The authors.

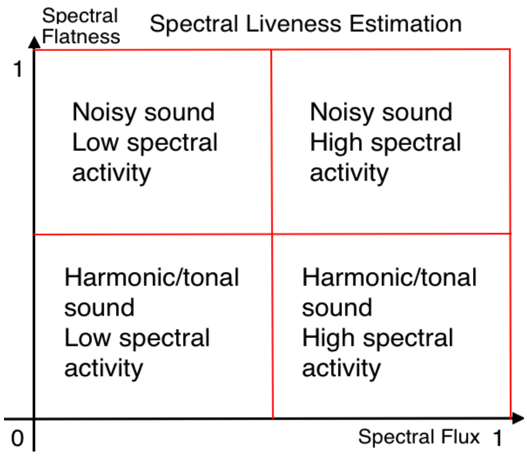
Image description: Audio format: Wave file (16 bit - 44.100 KHz); audio descriptors data extraction of spectral centroid, spectral spread, and loudness: Window size: 65.536 samples; hop size: 4.096 samples.

Spectral liveness (SL)

Spectral liveness (SL) consists of a two-dimensional phase-space graph, with two audio descriptors previously defined: spectral flux (on X-axis) and spectral flatness (on Y-axis) (Figure 3).

As proposed by Rossetti and Manzolli (2019) the interpretation of the analysis parameters of SL divide the graph into four quadrants, as shown in Figure 4: 1) the bottom-left quadrant represents sounds with higher harmonicity degree and low spectral activity, 2) the bottom-right quadrant represents sounds with higher harmonicity degree and high spectral activity, 3) the top-left quadrant represents noisy sounds with low spectral activity, and 4) the top-right quadrant represents noisy sounds with high spectral activity.

Figure 3: Interpretation of the spectral liveness descriptor (ROSSETTI, MANZOLLI, 2019, p. 209).

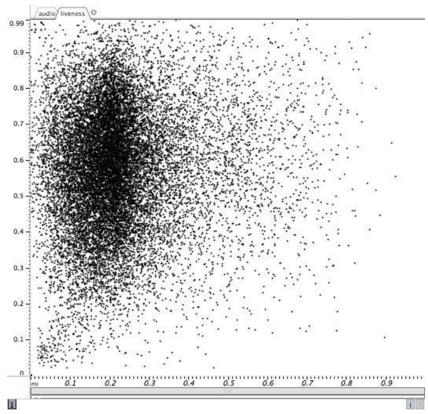


Source: The authors.

Image description: Information of harmonic/noise and low/high spectral features.

We introduce as an example, in Figure 4, the SL representation of the work *Schall* (1995), by Horacio Vaggione (ROSSETTI, MANZOLLI, 2019: 219). In the graph, we can find a high level of point dispersion in the SL representation. However, it is possible to visualize convergent areas characterized by an accumulation of points in the top left quadrant. This means that the perceived timbre of the piece has considerable inner movement and the sound grains have a mainly noisy spectral configuration. For more details about this analysis, see Rossetti and Manzolli (2019).

Figure 4: Spectral liveness representation of *Schall* (ROSSETTI, MANZOLLI, 2019, p. 213).



Source: The authors.

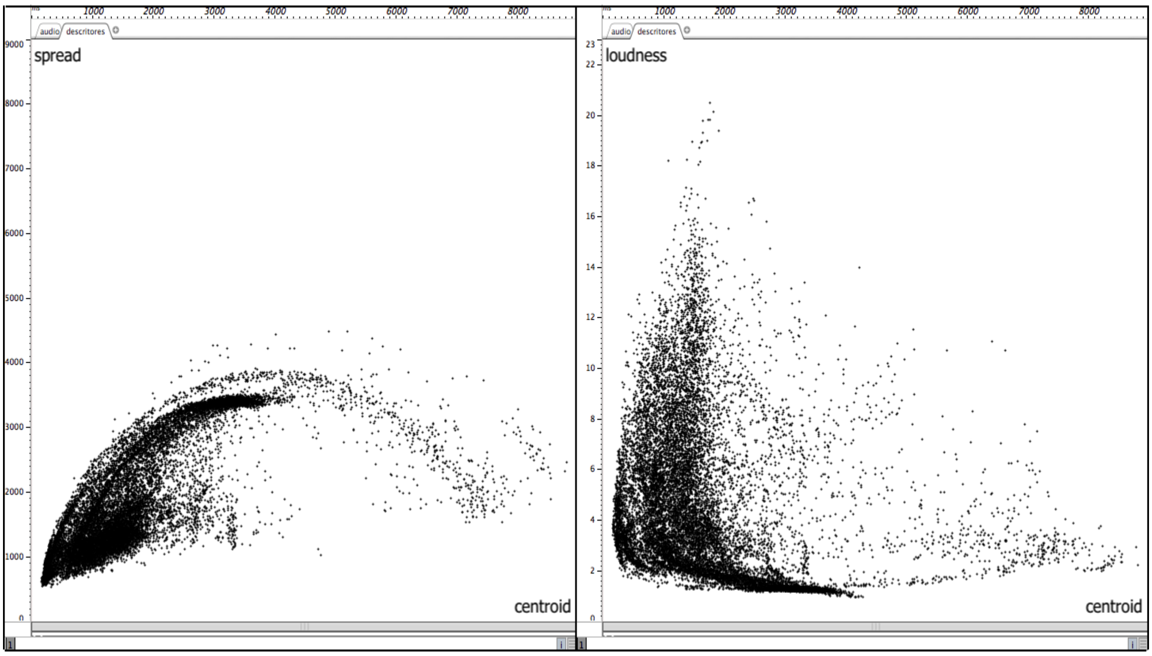
Image description: Audio format: Wave file (16 bit - 44.100 KHz); audio descriptors data extraction of spectral flux and spectral flatness: Window size: 2.048 samples; hop size: 1.024 samples.

Recurrent patterns (RP)

The recurrent patterns descriptor (RP) aims to estimate the potential emergence of perceptual form features by plotting the extracted data into a phase-space graph. The central idea is to represent the recurrence of spectral features indicating sound patterns through areas of accumulation of points in the graph.

In this case, we chose to present in Figure 5 two plots of RP of the work *Gesang der Jünglinge*, by Stockhausen, as presented in Rossetti *et al.* (2020, p. 165). The first graph has the spectral centroid on the X-axis and the spectral spread on the Y-axis, and the second graph has the spectral centroid on the X-axis and the loudness on the Y-axis. The RP representation suggests that the higher the accumulation of points in a determined area of the graph, the higher the recurrence of patterns of a musical composition. This approach is useful to evaluate and understand how the perception of time and musical form evolve in analyzed musical works, as explored in Rossetti *et al.* (2020).

Figure 5: Recurrent patterns of *Gesang der Jünglinge* (ROSSETTI *et al.*, 2020, p. 165).



Source: The authors.

Image description: Audio format: Wave file (16 bit - 44.100 KHz. Audio descriptors data extraction: in the Y-axis of the left square, we plotted the spectral spread data, in the right square we plotted the loudness data, and in the X-axis of both squares, we plotted the spectral centroid data. Window size: 2.048 samples; hop size: 1.024 samples.

Auditory filters

To simulate the behavior of the auditory system, some models based on the inner ear properties were employed. The basis of these models is the basilar membrane, which is part of the cochlea (VON BÉKÉSY, WEVER, 1960). As the sound stimulus with different frequencies simultaneously stimulates different regions of the basilar membrane (VON BÉKÉSY, WEVER, 1960), it is convenient to implement these models as a set of auditory filters. Here, we introduce two of them: the critical bandwidth of Zwicker and the mel scale.

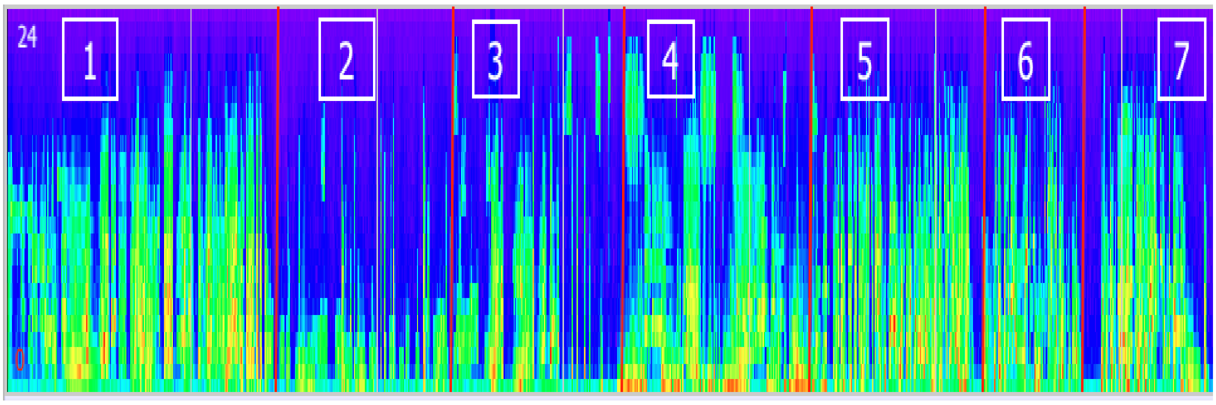
The critical bandwidths model is a consequence of a series of psychoacoustic experiments with individuals, testing thresholds on the perception of pitch and loudness, and is mainly based on the observation of the masking behavior of tones by narrow-band noises (FASTL, ZWICKER, 2007). This model divides the range of frequency between 0 and 15.500 Hz into 24 nonlinear bands, called barks. The Bark 1 has a size of 100 Hz, while Bark 24 has a size of 3.500 Hz. The critical bandwidths model has proved useful to study some musical phenomena. For example, starting with Plomp and Levelt (1965), many studies investigate the relation of the perception of dissonance and other harmonic effects with the critical bandwidth model (SETHARES, 2005. TERHARDT, 1984. VASSILAKIS, 2001).

Another model of auditory filters is the mel scale. It is based on the mel model, according to experiments that investigate pitch perception along with the auditory range. The mel scale is linear below 1.000 Hz and logarithmic above this threshold. According to Peeters (2004), the bandwidths of mel are calculated by dividing the mel curve into 24 equal parts.

From a methodological point of view, the advantage of using these two representations is to base the analysis on data of the spectrum considering the nonlinearity properties of the inner ear. These models support the investigation of emergent phenomena that can be a consequence of the elements' interaction of the musical texture.

An example of mel scale is introduced in Figure 6 with the representation of *Gesang der Jünglinge,* found in Rossetti *et al.* (2020, p. 162). The figure provides an overview of the spectral development in the frequency domain. We note segments with more densities, such as segment one, and others with lower densities, such as segment 2. We could also detect moments of segregation of the spectra, such as in segments 3 and 4. For more details of this analysis, see Rossetti *et al.* (2020).

Figure 6: Mel Scale representation of *Gesang der Jünglinge* (ROSSETTI *et al.*, 2020, p. 162)



Source: The authors.

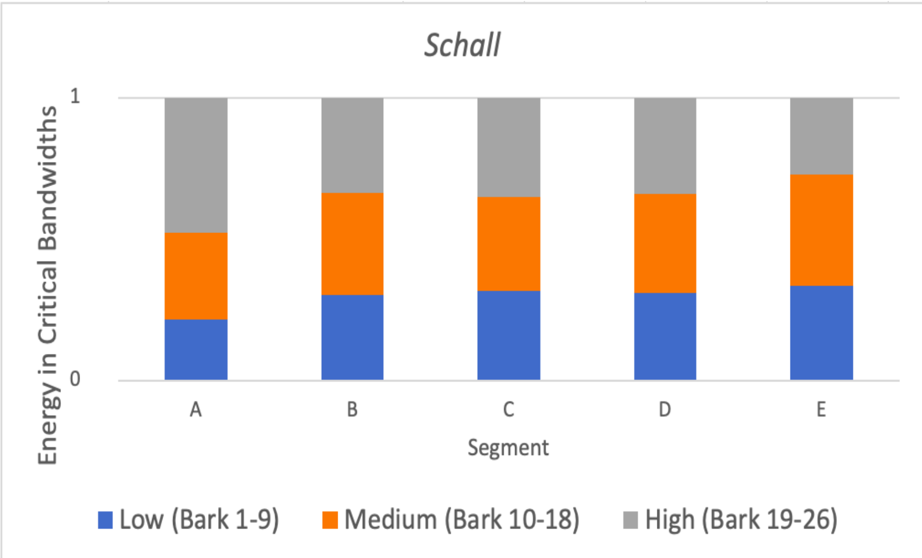
Image description: Audio format: Wave file (16 bit - 44.100 KHz). Window size: 2.048 samples; hop size: 512 samples.

Perceptual accumulation in frequency bands (PAF)

The Perceptual Accumulation in Frequency-bands (PAF) procedure consists of determining the normalized mean value of a region of any auditory filter descriptor as a function of a predetermined segmentation. This procedure was introduced by Antunes *et al.* (2019) and implemented by dividing the bark coefficients into three bands: Low (barks 1-9); Medium (barks 10-18); and High (barks 19-26) (BULLOCK, 2007). Configurations with different divisions were also tested in musical analyzes. The main goal of this representation is to understand how the musical texture privileges different regions of the auditory perception range. Moreover, because the mean value of the coefficients is normalized, it is useful to perform a comparison between two musical pieces or excerpts.

As an example, in Figure 7, the PAF descriptor is introduced through the analysis of *Schall*, by Vaggione, plotted with its segmentation (ANTUNES *et al.*, 2019). We note relative equality between the three bands, with a small increase of the medium band in time. Also, we observe a small contraction of the high band within the development of the piece. For more details of this analysis, see (ANTUNES *et al.*, 2019).

Figure 7: The PAF descriptor in *Shall*, using Bark Coefficients data (ANTUNES *et al.*, 2019).



Source: The authors.

Image description: Audio format: Wave file (16 bit - 44.100 KHz). Window size: 2.048 samples; hop size: 1.024 samples.

Discussion

After presenting the developed tools and graphical representations that can help us in our musical analyzes, we would like to discuss the results achieved regarding the representation of emergence and perception of form. This discussion aims to reflect on the emergent features expressed by the analytical tools developed.

The emergence of sound features' representation

The graphical representations we have been constructing are a possibility to help musical analysts to grasp temporal, spectral, harmonic, acoustic, and psychoacoustic features that dynamically emerge during a musical performance. It is important to highlight that there are yet many questions to be approached in that domain and many more graphical representations to be constructed and thought.

The kind of emergence that can be represented by those graphical representations is related to the interaction of spectral components in the microtime (textures of partials and grains) that are perceived as a Gestalt unity at the macro level. Here we are referring to the emergence of a form in the scale of sound objects (sound morphology). We believe that the range of these questions covers a considerable research field to be explored, and computer tools are useful in these tasks.

The possibilities are amplified when these tools are associated with musicological information about the creative processes employed. In those cases, might understand these descriptors under the idea of musical qualities' representation since their purpose is linked with musical and compositional concepts. It means that the acoustic and psychoacoustic features are combined to support musical ideas.

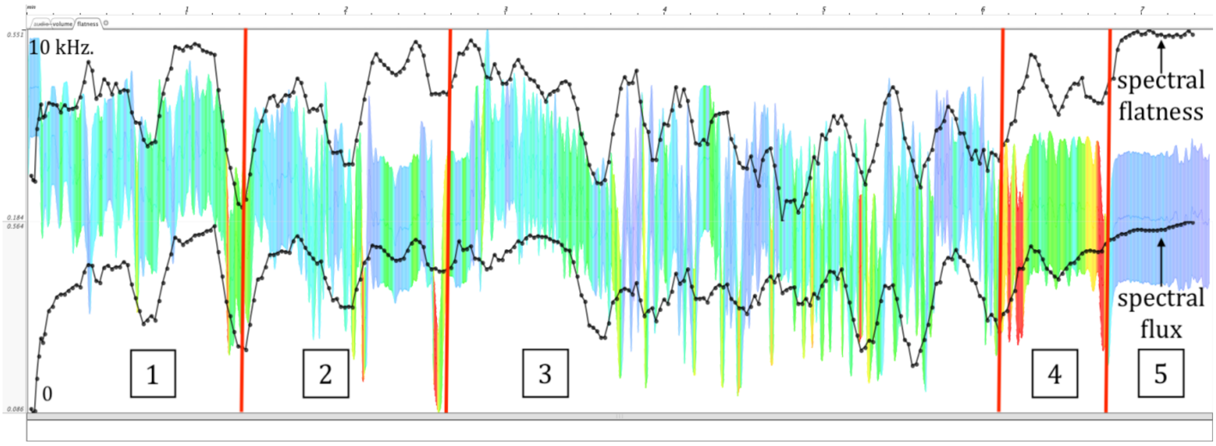
In this context, with the volume descriptor, it is possible to visualize the interaction between spectral information (given by the spectral centroid and spectral spread descriptors) and psychoacoustic perception of intensity (given by the loudness descriptor) of the sound texture in time. In other words, it is a possible representation of how we perceive the sound texture of a musical work in time concerning these features. Observing Figure 2 where we have the volume representation of *Gesang der Jünglinge*, time is represented in the X-Axis, and frequency is represented in the Y-Axis (0 to 10.000 Hz). The spectral centroid line is in the center of the colored mass and its thickness is given by the spread around the spectral centroid. The colors represent the level of loudness, from lilac (less perception of intensity) to red (high-intensity perception), passing through tonalities of blue, green, yellow, and orange). In this musical work, the volume descriptor indicates a higher spectral movement and variation in segments

2, 3, and 4, suggesting a differentiation of the sound material in the development of the spectra in time, compared with the other segments (1, 5, 6 and 7).

The spectral liveness descriptor intends to represent the emergent sonority of the musical works in terms of movement of the spectrum (spectral flux information) and tonal/noisy characteristics of sound (spectral liveness information), plotting the data in a phase-space graphic. Depending on the accumulation of the points in the graphic we can deduce the features of the resultant sonority of the works. In the case of the spectral liveness representation of *Schall*, by Horacio Vaggione (Figure 4), the accumulation of points occurs in spectral flux between values of 0 and 0.4, and from 0.3 to 0.9 in spectral flatness. This accumulation region indicates that the timbre has considerable inner movements, and the spectral components mostly present a noisy configuration.

Yet, we can combine the volume descriptor with curves of spectral flux and spectral flatness. With this configuration, it is possible to define segmentations of the musical works from considerable variations of the descriptors' values. In Figure 8 we present the volume representation of *Schall* with spectral flatness curves (above) and spectral flux curves (below). By the interaction of the information of all descriptors involved (spectral centroid, spectral spread, loudness, spectral flatness, and spectral flux) we defined points of segmentation of this musical work, totaling 5 sections. The points of segmentation were defined by taking into account considerable volume variations allied to important changes in the behavior of spectral flux and spectral flatness curves. The points of segmentation coincide with changes in perception of the sound texture.

Figure 8: Combination of the volume descriptor with spectral flatness (above) and spectral flux (below) curves.



Source: The authors.

Image description: Audio format: Wave file (16 bit - 44.100 KHz). For the audio descriptors employed in the representation of volume (spectral centroid, spectral spread, and loudness) we employed the value of 65.536 samples in the window size, and the value of 4.096 samples in the hop size. For the audio descriptors of spectral flatness and spectral flux we employed the value of 131.072 samples for the window size, and the value of 65.536 samples for the hop size.

Perception and emergence of macroform

If perception is an individual process, each individual has its singular perception that organizes and gives coherence to the information received. In terms of music, the listener organizes the form by the patterns' articulation perceived in the musical flow. This organization can lead to the emergence of a particular signification: a signification strictly related to the materials and processes employed in the composition and performance perceived in time. At this point, we recall the writings of Ligeti affirming that a musical passage can only have a signification concerning other musical passages (LIGETI, 2010, p. 143).

As pointed out, Thom (1988) considers perception a form of knowledge. Simondon defined the analogical act (2005, p. 562-563) as a method to obtain knowledge by defining the structures by the operations involved in the process (we can think of a creative process in music). As he states, a signification emerges in the individual and it is a space-time construction: the space is related to structures and time is related to processes, tendencies, and developments (SIMONDON, 2005, p. 263). In terms of our methodology of analysis, the emergence of a signification or a kind of knowledge about the musical work we are analyzing occurs through the relation between musicological information and creative processes with the identification of emergent features related to the performance (Figure 1).

In this article, we discussed the psychoacoustic audio descriptors and their representations that are related to sound perception. With the mel scale representation, we can visualize in which frequency region the spectrum is concentrated, highlighting that this scale is linear below 1.000 Hz and, above this threshold, a logarithmic scale. The mel descriptor in *Gesang der Jünglinge* (Figure 6) shows a denser presence of the sound spectrum in sections 1, 5, 6, and 7, while in the other segments (2, 3, and 4) the spectrum is more segregated along with the bands.

The perceptual accumulation in frequency-bands descriptor divides the bark coefficients into three bands (low, medium, and high). The bark model divides logarithmically the range of frequency between 0 and 15.500 Hz into 24 nonlinear bands. The PAF descriptor applied to *Schall* shows the predominance of medium and high bands in the 5 sections of the piece (Figure 7), with a tendency of increasing in low and medium bands, while the high band decreases in time.

The last descriptor developed, spectral recurrence, presents the recurrence of acoustic and psychoacoustic features in a musical work by plotting in two phase-space graphics data of spectral centroid (X-axis) and spectral spread (Y-axis) descriptors, and spectral centroid (X-axis) loudness (Y-axis) descriptors, the same descriptors employed in the generation of the volume representation. By observing Figure 5, *Gesang der Jünglinge* presents a significant concentration of the spectral features analyzed, especially in spectral spread and loudness. Then, we can infer that this musical work exhibits a kind of spectral redundancy in terms of the sound material employed, suggesting that similar patterns are explored in this piece in different moments, with a possible idea of recapitulation, besides an idea of cyclic and spiral time. These characteristics indicate a stability and continuity in the musical form.

Conclusions

In this research in musical analysis, the main achievements are related to the representation of emergent phenomena in musical performance, with acoustic and psychoacoustic features, especially in temporal and frequency domains. Those phenomena can be understood under the notion of emergent timbre. Also, we argue that the perceptive macroform of a musical work is related to the identification of patterns during the musical flow and may differ from a possible structure defined in the compositional processes or present in the score.

In the next steps of this research, we intend to increase the tools, libraries, and graphical representations we have been developing and expand the analysis repertoire of instrumental works. Furthermore, we intend to expand this methodology of analysis to the representation of musical timbre in the auditory cortex.

Referências

ANDERSON, Christine; DI SCIPIO, Agostino. Dynamic Networks of Sonic Interactions: An Interview with Agostino Di Scipio. **Computer Music Journal**, v. 29, p. 11-28, 2005.

ANTUNES, Micael, ROSSETTI, Danilo, MANZOLLI, Jônatas. A Computer-Based Framework to Analyze Continuous and Discontinuous Textural Works Using Psychoacoustic Audio Descriptors. Proceedings of the

BRAZILIAN SYMPOSIUM ON COMPUTER MUSIC, 17., São João Del-Rei. *Proceedings...* São João Del-Rei: UFJF, p. 4-11, 2019.

ANTUNES, Micael, ROSSETTI, Danilo, MANZOLLI, Jônatas. Emerging structures within microtime of Ligeti's Continuum. Proceedings of the **INTERNATIONAL COMPUTER MUSIC CONFERENCE,** 46., Santiago. *Proceedings...* San Francisco: ICMA, p. 271-274.

ASHBY, William Ross. **An Introduction to Cybernetics**. New York: John Wiley, 1956, 295p.

BULLOCK, James. (2007). Libxtract: A Lightweight Library for Audio Feature Extraction. *In:* Proceedings of the **INTERNATIONAL COMPUTER MUSIC CONFERENCE**, 34., Copenhagen. *Proceedings...* San Francisco: ICMA, 2007.

COUPRIE, Pierre. Quelques propos sur les outils et les méthodes audionumériques en musicology. L'interdisciplinarité comme rupture épistémologique. **Revue Musicale OICRM**, v. 6, n. 2, p. 25-44, 2020.

DI SCIPIO, Agostino. The Synthesis of Environmental Sound Textures by Iterated Nonlinear Functions, and Its Ecological Relevance to Perceptual Modeling. **Journal of New Music Research**, v. 31, n. 2, p. 109-117, 2002.

DI SCIPIO, Agostino; PRIGNANO, Ignazio. Synthesis by Functional Iterations. A Revitalization of Nonstandard Synthesis. **Journal of New Music Research**, v. 25, p. 31-46, 1996.

FASTL, Hugo; ZWICKER, Eberhard. **Psychoacoustics***:* facts and models. 3rd Ed. New York: Springer, 2006, 462p.

HEISENBERG, Werner. **Physics and Philosophy**. 3rd Ed. London: George Allen & Unwin, 1971, 176p.

HURON, David. (1999). The New Empiricism: Systematic Musicology in a Postmodern Age. *In:* **Lecture 3 from the 1999 Ernest Bloch Lectures**. Retrieved from https://citeseerx.ist.psu.edu/viewdoc/

download?doi=10.1.1.464.5455&rep=rep1&type=pdf. Accessed on 11/01/2021.

KOECHLIN, Charles. **Traité d'orchestration vol. 1**. Paris: M. Eschig, 1954, 336p.

LIGETI, György. **Neuf essais sur la musique**. Genève: Contrechamps, 2010.

MALT, Mikhail. (2015). **La représentation dans le cadre de la composition et de la musicologie assistées par ordinateur. De la raison graphique à la contrainte cognitive**. Habilitation à Diriges des Recherches. Université de Strasbourg, 2015. Strasbourg: Université de Strasbourg 2015.

MALT, Mikhail; JOURDAN, Emmanuel. Le "BSTD"–Une représentation graphique de la brillance et de l'écart type spectral, comme possible représentation de l'évolution du timbre sonore. *In:* **L'analyse musicale aujourd'hui, crise ou (r)évolution?** *Proceedings...* Strasbourg: SFAM, 2009.

MANZOLLI, Jônatas. **Nonlinear Dynamics and Fractals as a Model for Sound Synthesis and Real time Composition**. Ph.D. Dissertation. University of Nottingham, 1993. Nottingham: University of Nottingham, 1993.

MCADAMS, Stephen; WINSBERG, Suzanne; DONNADIEU, Sophie, DE SOETE, Geert; KRIMPHOFF, Jochen. Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. **Psychological Research**, v. 58, n. 3, p. 177-192, 1995.

MERIC, Renaud, & SOLOMOS, Makis. Analysing Audible Ecosystems and Emergent Sound Structures in Di Scipio's Music. **Contemporary Music Review**, v. 33, n. 1, p. 4-17, 2014.

MÜLLER, Meinard. **Fundamentals of Music Processing***:* Audio, Analysis, Algorithms, Applications. Cham: Springer, 2015, 133p.

OLIVEROS, Pauline. **Deep Listening**: A Composer's Sound Practice. New York: iUniverse, 2005, 78p.

PARNCUTT, Richard. **Introduction:** "Interdisciplinary Musicology". **Musicæ Scientiæ**, Special issue, p. 7-11, 2005-2006.

PARNCUTT, Richard. Systematic Musicology and the History and Future of Western Musical Scholarship. **Journal of Interdisciplinary Music Studies**, v. 1, n. 1, p. 1-32, 2007.

PEETERS, Geoffroy. A Large Set of Audio Features for Sound Description (Similarity and Classification) in the CUIDADO Project (Project Report). Paris: Institut de Recherche et de Coordination Acoustique-Musique, 2004. Retrieved from: http://recherche.ircam.fr/anasyn/peeters/ARTICLES/Peeters_2003_cuidadoaudiofeatures.pdf, accessed on 30/06/2022.

PEETERS, Geoffroy; GIORDANO, Bruno; SUSINI, Patrick; MISDARIIS, Nicolas; MCADAMS, Stephen. The Timbre Toolbox: Extracting Audio Descriptors from Musical Signals. **Journal of the Acoustical Society of America**, v. 130, n. 5, p. 2902-2916, 2011.

PLOMP, R; LEVELT, W. Tonal Consonance and Critical Bandwidth. **Journal of the Acoustical Society of America**, v. 38, n. 4, p. 548-560, 1965.

PRIGOGINE, Ilya. **The End of Certainty:** Time, Chaos and the Laws of Nature. New York: Free Press, 1995, 228p.

ROSSETTI, Danilo. A percepção do timbre em *Farben* Op. 16 n. 3 de Schoenberg: uma abordagem estética e psicoacústica. **Revista Música** v. 17, n. 1, p. 292-324, 2017.

ROSSETTI, Danilo. **O tempo e sua reflexão a partir da obra de Iannis Xenakis**. Master Thesis. State University of São Paulo, 2012. São Paulo: UNESP, 2012. 246p.

ROSSETTI, Danilo, ANTUNES, Micael, MANZOLLI, Jônatas. Compositional Procedures in Electronic Music and the Emergence of Time Continuum. **Organised Sound**, vol. 25, n. 2, p. 156-167, 2020.

ROSSETTI, Danilo, MANZOLLI, Jônatas. Analysis of Granular Acousmatic Music: Representation of Sound Flux and Emergence. **Organised Sound** v. 24, n. 2, p. 205-216, 2019.

ROSSETTI, Danilo, MANZOLLI, Jônatas. De Montserrat às ressonâncias do piano: uma análise com descritores de áudio. **Opus**, v. 23, n. 3, p. 193-221, 2017.

ROSSETTI, Danilo, TEIXEIRA, William, MANZOLLI, Jônatas. Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of *Desdobramentos do Contínuo* Performed by Audio Descriptors. **Musica Hodie**, vol. 18, n. 1, p. 16-30, 2018.

SCHAEFFER, Pierre. **Traité des objets musicaux:** essai interdisciplines. Paris: Seuil, 1966, 701p.

SCHNELL, Norbert, RÖBEL, Alex, SCHWARZ, Diemo, PEETERS, Geoffroy, BORGHESI, Riccardo. Mubu & Friends – Assembling Tools for Content-Based Real-Time Interactive Audio Processing in Max/MSP. *In: Proceedings of The* **International Computer Music Conference**, 36., 2009, Montreal. *Proceedings...* San Francisco: ICMA, 2009.

SETHARES, William. **Tuning, Timbre, Spectrum, Scale**. 2nd Ed. New York: Springer, 2005, 426p.

SIMONDON, Gilbert. **Cours sur la perception (1964-1965)**. Paris: Presses Universitaires, 2013, 416p.

SIMONDON, Gilbert. **L'individuation à la lumière des notions de forme et d'information** [1958]. Grenoble: Millon, 2005, 571p.

SOLOMOS, Makis. **De La Musique Au Son**: L'émergence du son dans la musique des XXe-XXIe siècles. Rennes: Presses Universitaires de Rennes, 2013, 545p.

TERHARDT. Ernst. The Concept of Musical Consonance: A Link between Music and Psychoacoustics. **Music Perception**, v. 1, n. 3, p. 276-295, 1984.

THOM, René. **Esquisse d'une sémiophysique.** Paris: InterEditions, 1988, 285p.

THOM, René. Morphologie du sémiotique. *In:* THOM, René. **Apologie du logos.** Paris: Hachette, 1990, p. 53-65.

TRUAX, Barry. Musical Creativity and Complexity at the Threshold of the 21st Century. **Interface**, v. 21, p. 29-42, 1992.

VAGGIONE, Horacio. Représentations musicales numériques: temporalités, objets, contextes. *In:* SOULEZ, Antonia; VAGGIONE, Horacio (Org.). **Manières de faire des sons**. Paris: L'Harmattan, 2010, p. 45-82.

VASSILAKIS, Pantelis. **Perceptual and physical properties of amplitude fluctuation and their musical significance**. Ph.D. Dissertation. University of California, 2001. Los Angeles: UCLA, 2001. 307p.

VINTIADIS, Elly. Emergence. *In:* **Internet Encyclopedia of Philosophy**, 2013. Retrieved from https://iep.utm.edu/emergenc/#SSH2aii, accessed on 30/06/2022.

VON BÉKÉSY, Georg, WEVER, Ernest. **Experiments in Hearing**. New York: McGraw-Hill, 1960, 745p.

XENAKIS, Iannis. **Formalized Music**: Thought and mathematics in composition. Stuyvesant: Pendragon Press, 1992. 387p.

XENAKIS, Iannis. La voie de la recherche et de la question (1965). *In:* XENAKIS, Iannis. **Kéleütha:** Écrits*.* Paris: L'Arche. 1994, p. 67-74.

Financing

This work was supported by the São Paulo Research Foundation (FAPESP) [grant numbers 2016/23433-8 and 2019/09734-3]; and the National Council of Scientific and Technological Development (CNPq) [grant numbers 304431/2018-4 and 429620/2018-7].

Publisher

Federal University of Goiás. School of Music and Performing Arts. Graduate Program in Music. Publication in the Portal of Periodicals UFG.