Revista Vórtex | Vortex Music Journal | ISSN 2317–9937 | http://vortex.unespar.edu.br/
D.O.I.: https://doi.org/10.33871/23179937.2022.10.1.3

Map, Trigger, Score, Procedure: machine-listening paradigms in live-electronics

Vinícius Cesar de Oliveira, José Henrique Padovani

Universidade Estadual de Campinas, Universidade Federal de Minas Gerais | Brazil

Abstract: Since the advent of real-time computer music environments, composers have increasingly incorporated DSP analysis, synthesis, and processing algorithms in their creative practices. Those processes became part of interactive systems that use real-time computational tools in musical compositions that explore diverse techniques to generate, spatialize, and process instrumental/vocal sounds. Parallel to the development of these tools and the expansion of DSP methods, new techniques focused on sound/musical information extraction became part of the tools available for music composition. In this context, this article discusses the creative use of Machine Listening and Musical Information Retrieval techniques applied in the composition of live-electronics works. By pointing out some practical applications and creative approaches, we aim to circumscribe, in a general way, the strategies for employing Machine Listening and Music Information Retrieval techniques observed in a set of live-electronics pieces, categorizing four compositional approaches, namely: mapping, triggering, scoring, and procedural paradigms of application of machine listening techniques in the context of live-electronics music compositions.

Keywords: live-electronics, machine listening, music composition, compositional paradigms.

Received on: September 29th, 2021. Approved on: October 27th, 2021. Online available on: April 30th, 2022. Presented at the 18th Brazilian Symposium on Computer Music (2021). Guest editors: Dr. Ivan Eiji Simurra (UFAC), Dr. Flávio Schiavoni (UFSJ), Dr. Cristiano Figueiró (UFBA), Dra. Isabel Nogueira (UFRGS) e Dr. Tiago Tavares (Unicamp).

In the last two decades, the search for new approaches related to techniques aimed at Machine Listening (ML) and Music Information Retrieval (MIR) in the live-electronics repertoire has expanded considerably. The growing interest in exploring the possibilities that such tools offer to music creation resulted in artistic productions that encompass the most varied aesthetics currents, technical approaches, and compositional strategies.

In this article, we seek to outline remarkable characteristics of the compositional use of ML/MIR tools in live-electronics/interactive music combining instrumental/vocal sounds and DSP processes. By observing some compositional strategies that have been explored for composers in the context of ML/MIR and DSP techniques for the extraction of musical data, we aim to circumscribe a series of approaches by organizing them into compositional paradigms.

The article comprises two main parts. In the first one (sections 1, 2 and 3), we seek to situate ML/MIR as interdisciplinary research fields, placing in context these areas and tools both in music researches/application as well as in a broader scope. In the second part (section 4), we seek to identify approaches to the use of ML/MIR processes in creative processes involving interactive systems, outlining four distinctive compositional paradigms in the field of interactive music: (1) mapping; (2) triggering; (3) scoring, (4) procedural.

1. ML and MIR as Interdisciplinary Research Fields

As interdisciplinary research fields, ML and MIR involve several areas, comprehending computer science, electrical engineering, acoustics, psychoacoustics, music theory, among others. In general, researches, and applications in these fields involve the investigation and development of algorithms and DSP models that allow for the estimation of distinctive sonic/musical attributes from audio signals/streams and the extraction of correlated quantitative data through audio descriptors (including, possibly, machine learning and statistical methods).

By employing different DSP methods and psychoacoustic models, ML/MIR techniques retrieve quantitative data that, through a correlation with some semantic characteristics of sound, may give relevant clues about qualitative and perceptual elements of sound and music. These techniques allow for data retrieving from audio signals that can be useful to outline both more traditional music concepts (pitch, tempo, key, and intensity, for instance) as well as psychoacoustic-based parameters (like harmonicity/inharmonicity, brightness, roughness, loudness, among others). To the extraction of these features, many higher-level ML/MIR algorithms rely on cognitive and psychoacoustic models, seeking to simulate organs, mechanisms, and specificities of human auditory perception (COLLINS, 2011, 2012).

According to COLLINS (2011, p. 439), most of these algorithms are based on hypothetical modeling of human auditory perception and seeks to mimic psychoacoustic phenomena. There are several research areas aimed at the study of the operation and behavior of the human auditory system. For instance, the *Auditory Scene Analysis* (ASA) studies how our cognition and auditory perception are capable to decompose and separate overlapping sounds and, consequently, identify them as their sound sources (BREGMAN, 1990). In this sense, recent studies in the ASA field are focused on computational modeling of these methods, giving rise to *Computational Auditory Scene Analysis* (CASA), a new research field with several interests, motivations, and goals in common with ML and MIR fields (GUTSCHALK; DYKSTRA, 2014; WINKLER *et al.*, 2012).

The field of *Music Information Retrieval* (also referred to as *Music Information Research*) has developed mainly in the 2000s. With the growing number of collections of music in digital format, MIR researchers were broadly interested in practical tools and strategies to automatically process these large sets of audio to manage them according to musical or perceptual criteria (FUTRELLE; DOWNIE, 2003, p. 121). More generally, what seems to unify the diversity of investigations that arose in this area is the interest in methods, tools, and techniques that make it possible to retrieve diverse secondary information from audio signals and other digital representations of music and sound. The retrieved data, which may be related to various perceptual and music theory concepts, is generally not directly represented or encoded digitally in audio files or streams: it is retrieved by using different DSP processes (and, more recently, machine learning strategies).

2. ML and MIR techniques and music composition

In the context of interactive music systems and computer music research, the expression *machine listening* was firstly used to refer to processing symbolic music parameters associated with the extensive use of MIDI protocol and its employment in the design of interactive music systems (ROWE, 1993)[1]. Relying on interactive music systems that depended mainly on the real-time processing of MIDI data (by the time real-time DSP of audio was not yet accessible for wide usage with personal computers), the idea was to create a more elaborate relationship between the human-machine interaction using MIDI-based digital instruments:

> My own program *Cypher* analyzes MIDI data on two levels: first, it classifies incoming musical events across a number of features such as loudness, speed, and register; and second, the behavior of these features within the current phrase is characterized as being regular or irregular. The user (the composer) constructs rules using the output of this analysis. Such rules determine which compositional algorithms will be invoked in response to which analyses and how the behavior of the algorithms will change over time. As a result of this process the program generates new MIDI output to be sent to synthesizers, thereby providing an additional voice to an ongoing musical texture. (ROWE, 1999, p. 84–85)

The development of audio-based *machine listening for* creative practices with interactive music systems, live-electronics, and real-time computer music started to be technically achievable for a more general public of artists in the late 1990s and the 2000s. By this time, personal computers began to have enough processing power to allow real-time DSP processes (PUCKETTE, 2002). At the same time, researchers/engineers released the first audio counterparts of computer-music languages and environments such as *Max/MSP* (today simply named *Max*), *Pure Data*, and *SuperCollider*. With the parallel development of research in MIR, it would be only a question of time for *machine listening* to expand its original MIDI-related connotation to the realm of DSP-based processes.

3. General characteristics of live-electronics in the context of ML and MIR technique

The integration of ML and MIR processes in live-electronics is, therefore, considerably recent. Different processes and techniques have been implemented in the form of *patches*, *objects*, *plugins*, *externals*, and *libraries* to enable the real-time retrieval of data from audio signals. This data is related to different sound attributes. The algorithms enable processes aimed at tasks such as the estimation of the fundamental frequency or the degree of harmonicity/inharmonicity of a signal, for example.

---

[1] https://wp.nyu.edu/robert_rowe/text/interactive-music-systems-1993/chapter5/

It is also possible to perform tasks such as: detection of attack/release of sounds; estimation of harmonic and pitch content (*chroma*)*;* partial tracking; temporal and spectral centroid analysis – which generally allows to infer the general behavior of the dynamics and the spectral balance (*brightness*) of a given audio segment. In addition, the increasingly in-depth knowledge about the characteristics of human psychophysiology has given rise to the construction of mathematical and computational models that allow the identification and correlation of sounds using high-level psychoacoustic descriptors based on techniques such as *loudness estimation*, *roughness* estimation, MFCC (*Mel-Frequency Cepstrum Coefficients*) and BFCC (*Bark-Frequency Cepstrum Coefficients*).

In this context, there is a direct relationship between specific sound and musical aspects that are at stake in performance — such as variations in tempo, timbre, rhythmic structure, articulations, among others —, and the structures generated by the processes of transformation and synthesis. The composition of live-electronic music employing ML/MIR resources deals with emerging sound and musical structures, which are result from an interrelation between gestures and ideas noted in the score; nuances in the interpretation and execution of each performer; intrinsic characteristics of instruments/voices; improvisational, open, and non-written musical ideas; and the transformations and generation of sounds by the computational systems coupled to the musicians. Therefore, we can explore these relationships at different levels of interaction from two basic flows of operation: (1) the performer controls the electronic processes, and (2) the electronic processes affect the instrumentalist's performance (GARNETT, 2001, p. 24).

The palette of possibilities of agency of the processes of interaction, transformation, generation, and sound spatialization that the use of ML/MIR tools bring to the practice of live-electronic music are the most varied, differing in terms of technical, poetic, and musical approach. Such resources allow structuring different strategies of creative use of the multiple information retrieved from the audio signal.

One can use the data to control/modulate parameters of sound synthesis and processing algorithms. The intensities can be used, for example, for different tasks such as modifying harmonic and spectral aspects of the live sounds or segmenting sections recorded in audio buffers. Likewise, the rate of variation of a given attribute, such as the harmonic content or the number of attacks detected, can control heterogeneous processes, such as audio spatialization and different synthesis algorithms.

Among the strategies used by different composers, one will find concatenative synthesis and morphing; detection of complex musical events — such as noisy instrumental sounds (MALT; JOURDAN, 2009) —, mapping, classification and ordering of timbres (BRENT, 2010), and mapping of information related to vibratos produced by the voice as source of data to control different synthesis/processing algorithms (EINARSSON; FRIBERG, 2015).

4. Compositional Paradigms

To understand the technical realizations, compositional approaches, and poetic conceptions that composers employ in the creative exploration of the set of tools of our interest, we propose an arbitrary organization of the different ways of thinking about its use into four groups, namely: (1) mapping; (2) trigger; (3) scoring; (4) procedural. We identified these paradigms regarding strategies of the conception of the algorithms, the organization and treatment of the extracted data, and the approximations on the creative use of these techniques to create the most diverse computational processes within the scope of live-electronic music.

We emphasize that the pieces discussed in this text are not meant to be representative of any aprioristic paradigm, which would precede them as a model, category, or aesthetic group. The identification of these paradigms is consciously arbitrary, built upon analytical interpretations. The objective, however, is not the categorization of these paradigms but to understand different poetic approaches regarding the use of specific tools in music creation. The idea is to identify poetical stances with regard to the creative application of the tools, aiming less at the segregation of particular aesthetics or compositional schools and more to comprehend how the insertion of these new tools into heterogeneous poetic practices takes place.

4.1 Mapping

We classify as mapping paradigm the creative applications of MIR/ML tools that rely on correlations and transformations between input parameters and the algorithms of sound transformation/generation. It is strongly related to the processes of "liaison strategies between the outputs of the gestural controller and the input controls of the sound generation" (MIRANDA; WANDERLEY, 2006, p. 3) units in the context of the design of New Digital Instruments, allowing for building more simplistic or more complex relationships between these values and the synthesis algorithms they control. This correlation happens through a mediation layer, which is in charge of defining the paths that the input information stream will take, and the function that they will acquire within the architecture of the synthesis processes (WANG; WANDERLEY; SCAVONE, 2020, p. 107). In short, mapping operates on transferring information and converting it into parameters in order to control computational actions and processes.

Mapping is a characteristic process of digital luthiery, i.e., the design of digital/expanded instruments and digital music interfaces (MIRANDA; WANDERLEY, 2006; MONTEIRO, 2012; RIMOLDI, 2018). While in acoustic instruments, the interface is linked directly to the sound production mechanism — the violin strings are an integral part of both the control and the sound generator mechanisms —, in digital instruments, in general, the interface is separated from the sound production device: the relation between them needs to be defined, what is done through mapping processes (HUNT *et al.*, 2002, p. 1).

In general, mapping processes are conceived through pre-defined strategies, which correlate in straightway a set of data to a series of parameters — e.g., the roughness rate extracted from a signal can control the grain size of granular synthesis. Therefore, these correlations can be divided into four categories: (1) one input data controls only one parameter (one-to-one); (2) one input data control more than one parameter simultaneously (one-to-many); (3) more than one input data controls only one parameter (many-to-one); (4) a combination of all three categories, in which more than one input data controls more than one parameters (many-to-many) (IAZZETTA, 2000, p. 20).

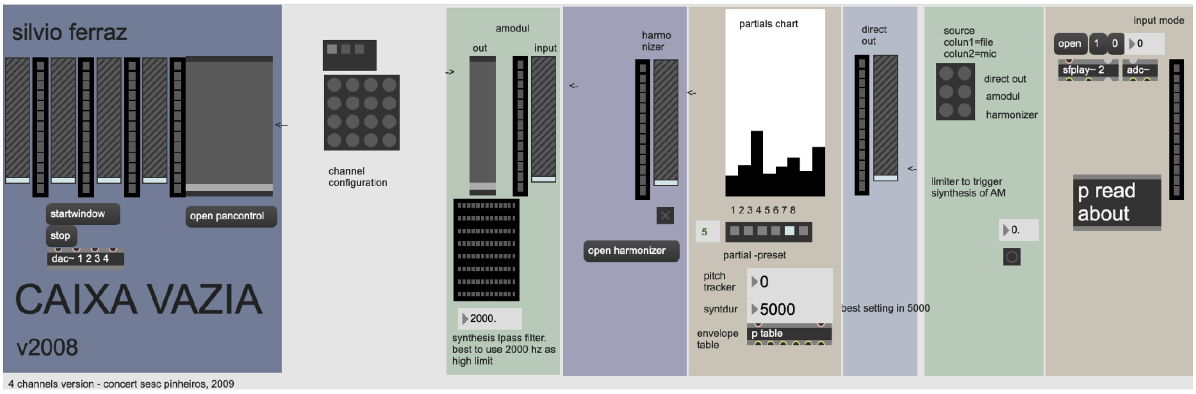
a) *Cortazar, ou quarto com caixa vazia* (1999), by Silvio Ferraz

An example of a simple direct mapping strategy can be seen in *Cortazar, ou quarto com caixa vazia* (1999), for piano and live-electronics, by Silvio Ferraz. The title of the piece, in addition to paying homage to the Argentine writer Julio Cortázar, reminds us of the image of a pianist who, locked in a room with an empty box, breaks the walls and frees the sound of the piano from western musical standards. According to PESTOVA (2009, p. 26), the "empty box" [*caixa vazia*] is also a metaphor for the Max/MSP real-time computer music environment, used to structure and perform the live-electronics processes.

In this piece Ferraz uses three different sound synthesis and transformation procedures: a pitch shifter with a random ramp control to create microtonal oscillations in the transpositions; additive synthesis from a set of eight partials that can be varied according to eight presets, and a ring modulation process.

Ferraz uses the piano's signals as a sort of dynamic interface to control the ring modulation algorithm that processes additive synthesis generated douns. It is worth mentioning that, although the composer has used the raw signal the piano – as it is captured by the microphone without any specific treatment to extract information –, attributes that compose and describe the physical behavior of the signal in the time domain (amplitude envelope and waveform) are used in the mapping of the process in question.

FIGURE 1 – Max/MSP patch of *Cortazar, ou quarto com caixa vazia*



Source: patch of *Cortazar, ou quarto com caixa vazia*, by Silvio Ferraz.

According to Ferraz, the ring modulation process was conceived based on the creation of an additive synthesis module, which in turn obeys a system of chord expansion and contraction and a certain degree of inharmonicity generated from the multiplication of the fundamental frequency extracted in real-time from the piano, by eight pre-established partials — this harmonic principle also guides the writing of the instrumental part (FERRAZ, 1999, p. 5). The result of this synthesis is then multiplied by the amplitude envelope of the piano's signal, resulting in a sound with a certain degree of inharmonicity and a spectrum distortion.

Ring modulation is a signal processing technique inherited from the telecommunications area in which the amplitude of a given signal or frequency (carrier) varies according to the oscillation of another frequency (modulator). As a result, the frequencies of the two inputs add and subtract. For example, if we consider two pure sine waves, one carrier and the other modulating at 500 Hz and 100 Hz, the result will be the superposition of two waves, 600 Hz (addition) and 400 Hz (subtraction).

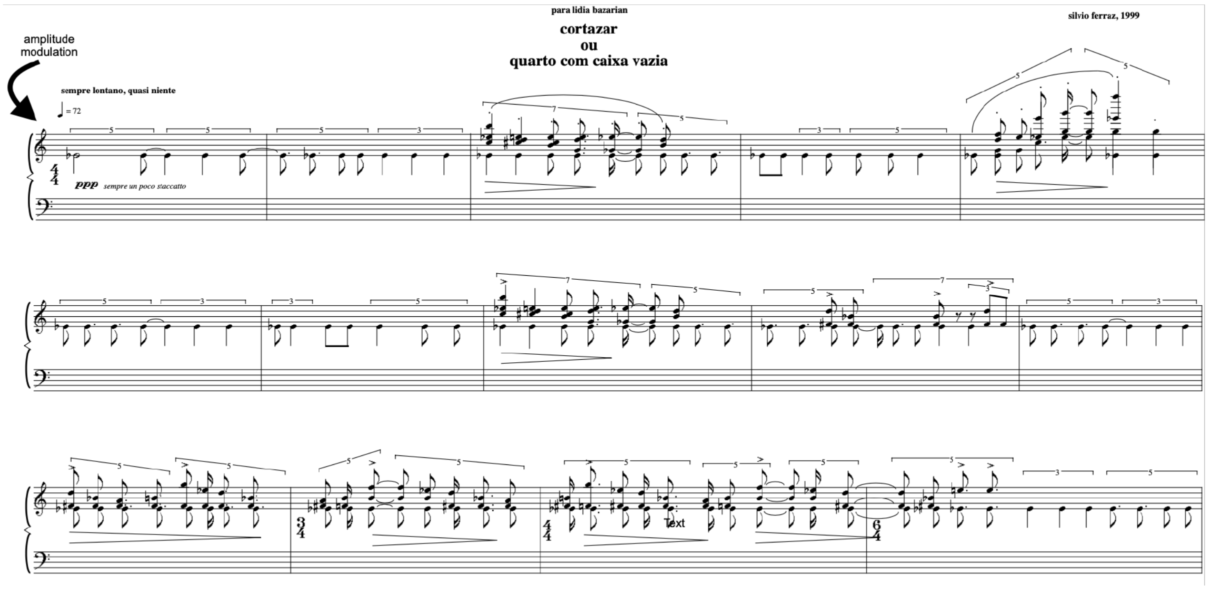
Therefore, the more complex the signal of the two inputs becomes, the more overlap of frequencies will result since the process of addition and subtraction starts to be applied to each component of the signal. For example, a piano note has a spectrum with numerous frequency components that present behaviors, to some extent, unpredictable and independent. Each spectral component will pass through this process of addition and subtraction (EMMERSON, 1977, p. 14), generating a large amount of overlapped frequencies. This process is widely used in electronic music, especially in live-electronics, since it allows the interval contraction and expansion of pitches as well the elaboration of different timbristic colors.

In *Cortazar, ou quarto com caixa vazia*, the piano's digital signal is mapped directly and used as a control parameter to modulate the amplitude of the additive synthesis output signal. An attack detector is responsible for triggering the modulation process. With this, the composer establishes a direct correlation between the piano signal (input data) and the modulating frequency (control parameter) of the ring modulation process in a one-to-one relationship.

With this direct mapping process, Ferraz creates an automatic and dynamic mechanism of control, in which the sound becomes its control interface, manipulating processes that will engender new sound structures. As the amplitude of the piano's attack increases, the greater the spectral curvature applied to the additive synthesis, thus, resulting in a more inharmonic sound.

Figure 2 shows the first bars of the piece, where this procedure can be observed. At the first moment, only the direct sound of the piano can be heard. With the increase of the presence of attacks and accented chords in the instrumental part, the modulation process gradually becomes more noticeable.

FIGURE 2 – First bars of *Cortazar, ou quarto com caixa vazia*



Source: score of *Cortazar, ou quarto com caixa vazia,* by Silvio Ferraz.

b) *PS. I will be home soon!* (2012), by Anna Einarsson

A practical example of mapping application can be observed in the chamber opera, *PS. I will be home soon!* (2012), for six instruments, four voices, and live-electronics, by Anna Einarsson. The libretto was written by Maria Sundqvist and is based on the life of the Swedish sailor, Calle Pettersson covered in the book *Kung Kalle av Kurrekurreduttön* (REGIUS; LANGER, 2002).

Considering the technical and musical aspects, Anna Einarsson developed tools and procedures aiming at musical/spectral information extraction from the singing voice, mapping these parameters to control the interactive processes of the piece. According to the composer, the idea since the beginning of the compositional process was to develop mechanisms of dynamic control of the electronics, without requiring physical interfaces or sensors to be coupled to the performers (EINARSSON, 2017, p. 58).

In collaboration with the researcher Anders Friberg, Einarsson has implemented in *Ps. I will be home soon!* an extraction method for retrieving attributes related to the vibrato range produced by the singing voice. Such a method relies on an algorithm called CUEX (Cue-extractor), developed by Friberg in *MatLab* and, posteriorly, ported for *Pure Data* (FRIBERG; SCHOONDERWALDT; JUSLIN, 2007). The extracted information from the voice is mapped and used to control four polyphonic synthesizers, each of them containing resonant filters and oscillator banks, which are excited by the noise signals (EINARSSON; FRIBERG, 2015, p. 19).

The extraction process can be divided into five steps: (1) fundamental frequency detection and signal amplitude measurement through the *Yin* algorithm (DE CHEVEIGNÉ; KAWAHARA, 2002); (2) creation of a gate using a specific pitch range and an amplitude threshold (only the pitches which fall within this limit will pass to the analysis); (3) the signal is divided into two parts, the first part is filtered by a low-pass filter and the second by a medium filter with a window size equals to a vibrato cycle — considering a vibrato rate equals to 6hz; (4) the two parts are subtracted in a way that the resulting signal will be zero when there is no vibrato; (5) the average of the absolute signal value is calculated through a medium filter with the double size of a vibrato cycle – thus the signal resulting will be proportional to the extent of a vibrato given in cents (EINARSSON; FRIBERG, 2015, p. 413; FRIBERG; SCHOONDERWALDT; JUSLIN, 2007, p. 19).

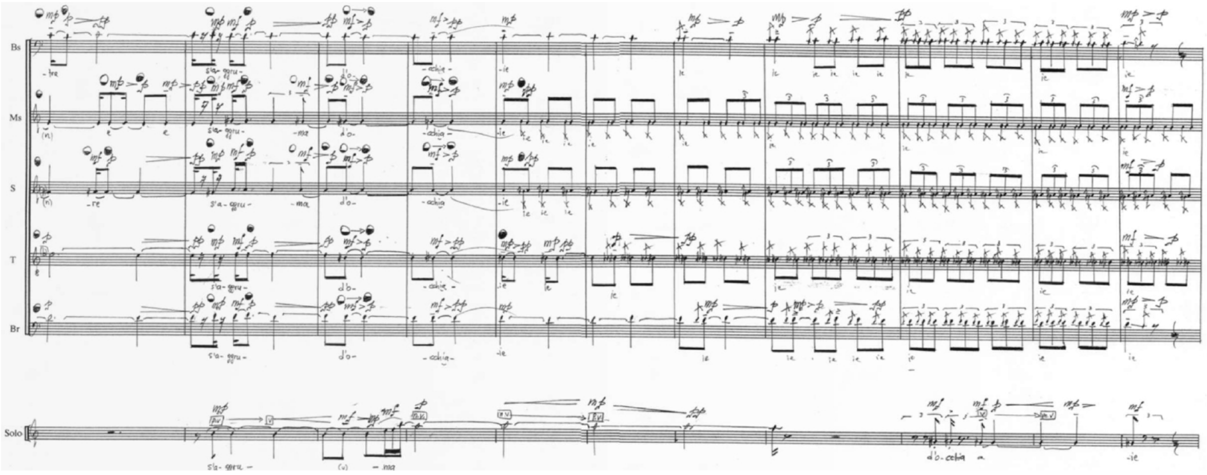
In *Ps. I will be home soon!* the mapping process is related to the notion of metaphoric mappings, such as described by Johnson (2013) . What is transferred, in this process, are not the attributes themselves, but the relationships between the source domain (origin) and the target domain (destination). According to Einarsson and Friberg, in this piece, there is a relationship of forces employed to the mapping method. The density expresses the force in the target domain, which represents a dynamic structure of grouping of pitches, and the vibrato is the expression of the force in the source domain, i.e., the singing voice. Thus, the greater the presence of vibrato, the denser the grouping of pitches resulting from the synthesis process.

c) *Da una crepa* (2011-2012), by Marta Gentilucci

Another piece that illustrates very well the employment of what we call here mapping paradigm, is *Da una crepa* (2011-2012), for choir, soprano, ensemble, and live-electronics, by Marta Gentilucci (Figure 3). Some of the central compositional aspects of *Da una crepa* arise from Gentilucci's desire to explore rhythmic structures, both at the macro-level — controlling and developing complex rhythmic structures throughout the composition — and at the micro-level —

relationships between the internal temporal structures of the sound, and the possibility of using these structures as a control interface to generate and/or manipulate other sounds.

FIGURE 3 – Excerpt with vowel tremolos referring to bars 122 to 130 of *Da una crepa*, by Marta Gentilucci



Source: score of *Da una crepa,* by Marta Gentiluci.

In this context, there arises the idea of correlating the rhythmic structures generated by vowel tremolos extracted from the singing voice — i.e., the repeated alternation of vowels over the same note, for example, “a-e-a-e” — to parameters that could manipulate pitches in electronic processes. In other words, the alternation of vowels would be applied to a given sound to control the amount of variation in the transposition of that same sound. In short, the composer intended to design a process that could modulate the pitch of a tenuto (sustained) sound through the rhythmic structure generated by the articulation of the voice.

Thus, the process of extracting and mapping the attributes used in *Da una crepa* is carried out in three steps. Throughout the piece, certain excerpts are recorded and saved in buffers. Then, these segments are analyzed with the help of audio descriptors (a combination of spectral deviation, odd to even second tristimulus, and total energy to extract the vowel tremolo (MALT; GENTILUCCI, 2015), and the resulting data of this analysis, being stored in a [coll][2] object. Finally, the stored data maps the transposition of two modules of pitch shifter and a harmonizer.

---

[2] Max/MSP object that allows the storage, organization, editing, and retrieval of messages.

4.2 Triggering

Triggering paradigm is related to the ML/MIR employment in the conception of systems and algorithms aimed at sound/musical event detection. This strategy deals with procedures that rely on logical processes. In other words, algorithms analyze the digital signal looking for specific characteristics, which can describe certain events, such as *pizzicato*; dynamic variation; trills; vibrato, among others. Upon the detection of such events, the algorithm may trigger some electronic process.

The triggering paradigm is widely used within the scope of live-electronics, as well as in the design of interactive musical systems and autonomous agents (COLLINS, 2006) , as a direct interaction strategy between the machine and the performer. By "listening" to the instrumentalist and detecting the presence of pre-defined events, the computer triggers processes, thus establishing a direct relationship between the instrumental gestural actions and the response of the machine to these impulses.

a) *Deconstruction Dowland* (2009), by Natascha Barret

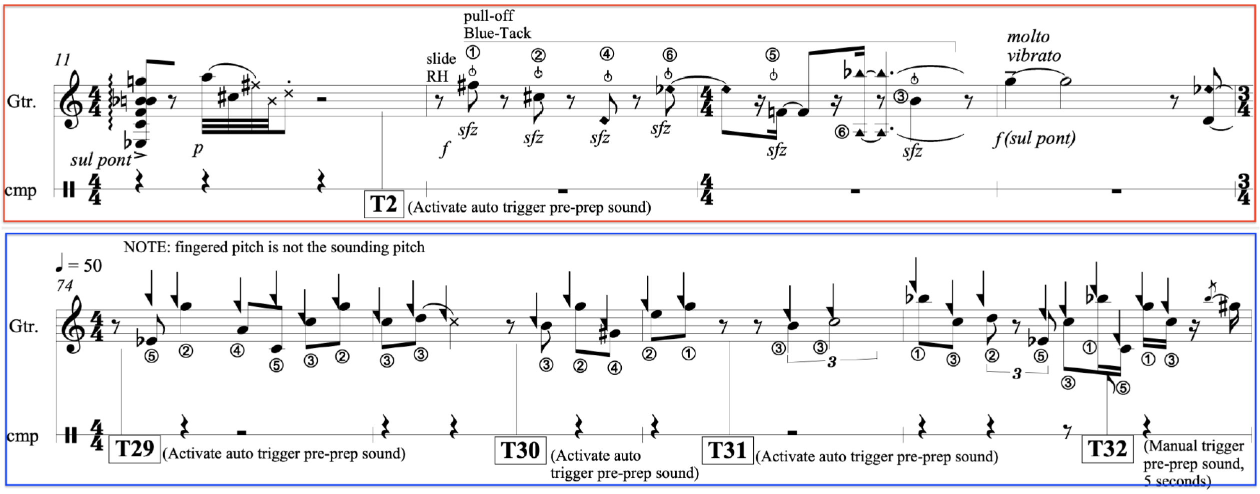
Decostructing Dowland (2009), for acoustic guitar and live-electronics, by Natasha Barret, illustrates in a simple and didactic way the practical application of the triggering paradigm (Figure 4). With the desire to unite ancient music with current musical practice, Barret takes, as a starting point, the three galliards for lute solo published in the first book of songs by John Dowland.

In this piece, the triggering paradigm occurs through the application of an attack detector to automatically and accurately trigger small sound samples previously conceived in the studio and saved in buffers. To realize this, Barret uses used the [bonk~] object, implemented by Miller Puckette, for Pure Data and Max/MSP (PUCKETTE; APEL; ZICARELLI, 1998).

[bonk~] works in the audio signal analysis to detect attacks, through the search for sudden changes in the energy envelope, which presents specific characteristics. The detection process is applied in bar 12 and in the excerpt that goes from bar 74 to bar 76 of Deconstructing Dowland. In these two excerpts, the [bonk~] responds to Bartok's pizzicatos (measure 12), and the attacks produced by the artificial harmonics (measures 74 to 76), triggering the pre-recorded sound samples.

Although such excerpts have relatively simple rhythmic structures, the automatic attack detection strategy results in more precise synchrony between the action of the instrumentalist and the sample triggering. This provides an efficient amalgamation between acoustic instrument sound and the pre-recorded samples so that the triggered sounds work as a spectral extension of the acoustic guitar resonance.

FIGURE 4 – Bar 12 (red) and bars 74 to 76 (blue) of *Decostructing Dowland*



Source: score of *Decostructing Dowland*, byNatasha Barret

b) *Tensio* (2010), by Philippe Manoury.

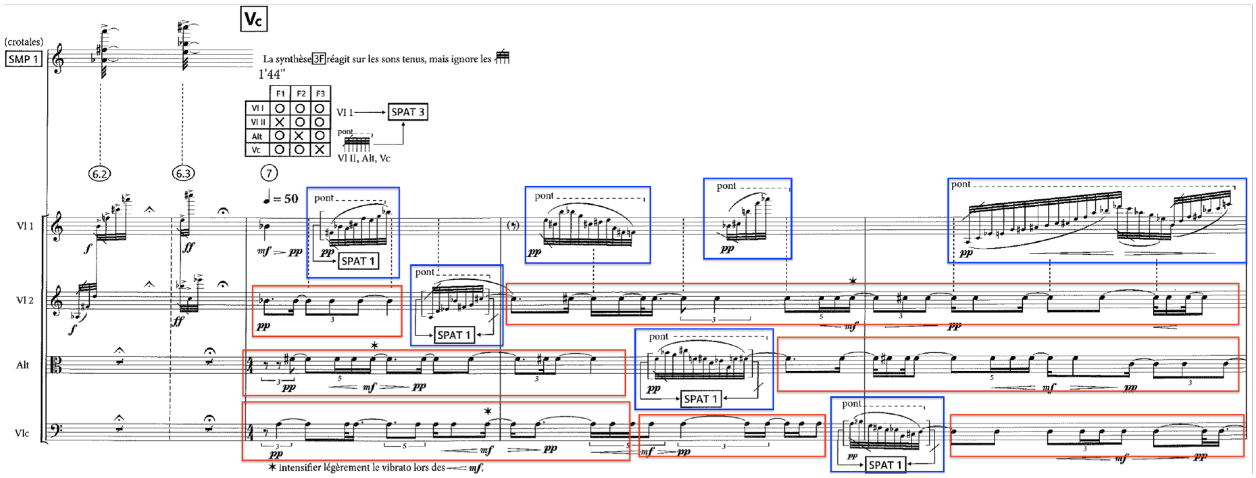
Another application of the trigger paradigm can be seen in *Tensio* (2010), for string quartet and live-electronics, by Philippe Manoury. The title of the piece refers to the word “tension”. According to the composer, this tension is the physical tension of the strings of the quartet that proliferate through the electronic part. “It seemed beneficial to me to resort to that primordial image of a string tensioned between two points, and to make it be played in extreme ways, which only technology allows us to glimpse” (IRCAM, 2010; MANOURY, 2013).

In *Tensio*, the composer uses a combination of ML/MIR techniques and strategies to construct mechanisms that work in parallel to detect different aspects of specific musical events. This procedure happens in section **Vc** of the piece, within an instrumental texture built from gestures composed of quick sounds in sul ponticello, and medium-length sounds played with the bow in a natural position.

To detect both events separately and trigger an additive synthesis module and a spatialization process, Manoury uses the object [sigmund~] associated with the extraction of the third tristimulus and the calculation of the signal amplitude performed by the total energy descriptor.

In this way, [sigmund~] is configured to report the fundamental frequency of the analyzed signal above 56 rms in an interval of 200 milliseconds. When events with such characteristics are detected, a calculation process is triggered, generating a set of frequencies and amplitudes that will feed a bank of oscillators, thus producing a synthetic spectrum (we will describe in detail the process of calculating these spectra in section 3.4).

FIGURE 5 –Begining of the sectin **Vc** of *Tensio*



Source: score of *Tensio*, by Philippe Manoury

On the other hand, to detect rapid sounds in sul ponticello is used a combination of between [sigmund~], third tristimulus, and the calculation of the total signal energy. While the [sigmund~] is responsible for detecting fast sounds — interval of 7 milliseconds between each note — the tristimulus and the power calculation are responsible for detecting the timbre generated by the excitation of the string with the bow in sul ponticello. Since the power of the signal is 1 — the signal power is scaled from 0 to 1 – the value of the third tristimulus is calculated. Events in sul ponticello are detected when the third tristimulus returns values above 0.5. Hence, when an event with these characteristics is detected, a spatialization process is triggered. Figure 5 shows the instrumental texture composed of notes with medium duration (blue) and fast notes in sul ponticello in the Vc section of *Tensio*.

Therefore, considering that the passage in question is written with a 4/4 time signature and with a tempo that corresponds to the quarter note equal to 50 bpm, the detector configured to report the fundamental frequency every 200 milliseconds will only react to rhythmic figures with a value above a thirty-second note, approximately. Reacting to the events highlighted in red in figure 5 results in a spectrum generated by the additive synthesis that presents a rhythmic structure close to that one of the string quartet. While the other detector, configured to report the fundamental every 7 milliseconds, will react to rhythmic figures close to the appoggiaturas played in sul ponticello, thus allowing only gestures with these characteristics to be subjected to the spatialization process.

4.3 Scoring

The scoring paradigm is a way of thinking and conceiving the creative use of ML/MIR tools that broadly encompasses procedures, technical solutions, and compositional approaches aimed at the implementation of score following systems — i.e., methods aimed at tracking the musical performance (audio stream) and its synchronization to a score. Score-following processes rely on symbolic representation of the musical events (score) and on musical data extracted by ML/MIR algorithms. The score and the ML data are compared, with different strategies, allowing for the estimation of the probability that the instrumentalist is performing a specific section or musical event.

The first implementations of score following methods depended on the retrieval of symbolic information, strongly dependent on the extensive use of MIDI-based instruments in interactive music systems. Since the emergence of real-time DSP-based ML/MIR algorithms, score-following systems have started to get to recognize musical events of musical scores by analyzing audio streams. These new algorithms explore psychoacoustic, stochastic, and machine learning models and methods, relying on probabilistic heuristic methods and on diverse descriptor-based ML/MIR processes. These tools allow for the prediction of the performer's tempo, yielding a variable metric variable related to the instrumentalist's agogic – such as happens with *Antescofo*[3] (CONT, 2008

---

[3] Antescofo is a modular score following system developed by Arshia Cont, which combines machine listening and machine learning methods.

By enabling the possibility of determining and prescribing with some degree of precision the actions that should be performed by a computer, synchronized to a live musician, the conception of morphologies resulting from computational processes begins to operate in the same domain as the instrumental writing process. Since musical notation has a symbolic and abstract character, it allows a musical gesture to be graphically represented and recorded on media — in this case, the score. Similarly, score-following-based processes bring close together live-electronics composition approaches and practices traditionally explored in compositions for acoustic instruments and fixed electronic media. By using score-based tools in live-electronics, composers may conceptualize sound morphologies that arise from the interaction between instrumental sounds and the computational processes (MANOURY, 1990)[4]

a) *Jupiter* (1987), by Philippe Manoury

One of the first pieces to creatively explore this set of strategies that we classify here as the scoring paradigm is *Jupiter* (1987), for flute and live-electronics, by Philippe Manoury. This is the first piece in a cycle called *Sonus ex Machina* composed in collaboration with Miller Puckette between 1987 and 1991, and stands out in the live-electronics repertoire, as it established new paradigms regarding the use of ML and interactive music systems — particularly, the synchronization and interaction strategies between electronic processes and instrumental performance (MAY, 2005, p. 147).

The score following mechanism implemented in Jupiter uses the [scofo~] object, capable of reading the MIDI pitches prescribed from a text file (accompaniment score) and comparing them with those extracted from the flute signal. For each note contained in the accompaniment score [scofo~] assigns a "theory". The *theory* is responsible for describing the strength of the evidence that the most recent pitch played by the flute matches the most recent pitch read from the accompaniment score. If they match, the [scofo~] returns the number of the event linked to it.

---

[4] http://www.philippemanoury.com/?p=330

If the pitch extracted from the live flute matches the pitch prescribed in the accompaniment score — which may or may not is linked to an event —, a control message contained in [qlist][5]. If there is no event linked to that note, the system will only receive the pitch and continue following the flow of data extracted from the flute. One of the main consequences of the use of score following in *Jupiter* reflects on the choice of the composer to represent the structures and sound/musical morphologies emerging from the flute-computer interaction through musical notation (representation of pitches, durations, and dynamics). Since the parameters of computational processes are precisely prescribed in [qlist] it is possible to predict the approximate behavior they will present when triggered, enabling the representation of these parameters using traditional notation.

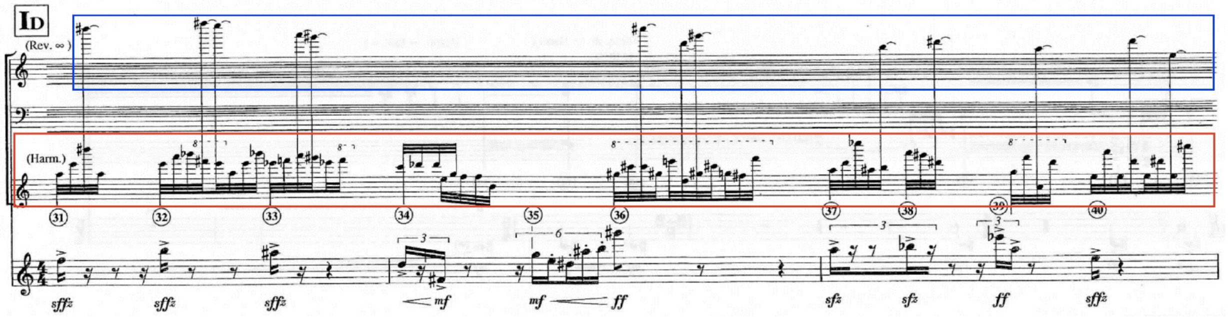
This process of notation of electroacoustic events in the musical score converges with the notion of Solfège of models. This concept consists of knowledge, intellectual and cognitive skills acquired by the composer when designing certain processes, allowing to control and master the musical structures resulting from generative models and to associate graphic/textual representations of certain software and musical systems and the final musical result (MALT, 2004).

The use of *score following* linked to the idea of *Solfège of models* allowed the composer, from previous knowledge of the nature and parameters of the processes of sound transformation/generation involved in the piece, to glimpse their behavior when integrated with the flute gestures. This leads us to believe that an important step in the compositional process of *Jupiter* — and that seems to be one of the most present characteristics in musical writing of Manoury — concerns, precisely, the conception of the electronic part in an approximate way of the instrumental composition — i.e., imagining musical structures and morphologies resulting from the amalgamation of computational processes and instrumental gestures, and reducing their sonic complexity to a symbolic abstraction that could be represented by traditional notation. The figure 6 shows the beginning of the section **ID**, of Jupiter, which shows the representation of the morphology resulting from the interaction between the flute (lower stave) and four modules of transposition of heights associated with a reverb (upper staves).

---

[5] [qlist] is a *Pure Data* object that reads text files containing time-stamped messages, which can be sequenced automatically or manually is sent to the transformation and synthesis modules, and a process is triggered.

FIGURE 6 – The beginning of section ID of Jupiter



Source: score of *Jupiter*, by Philippe Manoury

b) *Raggi di strighe* (2011), by Lara Morciano

The desire to bring the concept of live-electronics computational processes closer to the notational dimension of instrumental writing seems to us to be a central issue that permeates the compositional processes that explore the scoring paradigm, as shown by some aspects of the compositional process of *Jupiter*. This becomes even more accentuated when we observe some compositional thoughts used in pieces that explore the use of the *Antescofo* as a tool that acts in the integration between these two environments.

Raggi di Stringhe (2011), for violin and live-electronics by Lara Morciano, in addition to illustrating this approach, exemplifies some possibilities and compositional implications of the use of the *Antescofo* and consequently serves as a reflexive field to approach the use of the score paradigm. This piece was composed at IRCAM in collaboration with the musical informatics director José Miguel Fernandez. The basic idea, according to the composer, was to create a piece in which the electronic processes could be developed in a well-articulated way and were conceived as a complement to the virtuosity of instrumental writing, establishing a certain relationship with the violin through a special focus on reactions, and in the sound and gestural transformations of the instrument (IRCAM, 2011).

> One of the crucial issues at the base of my artistic reflection resided in the desire to produce a work conceived in a unitary and coherent way concerning the writing and materials used, as well as the conception and realization of the electronic part, with a view to expressiveness and interpretive freedom; the different problems linked to the temporal development of

> the musical acts at the moment of performance, and the synchronization between performer and machine (MORCIANO, 2018, p. 36)

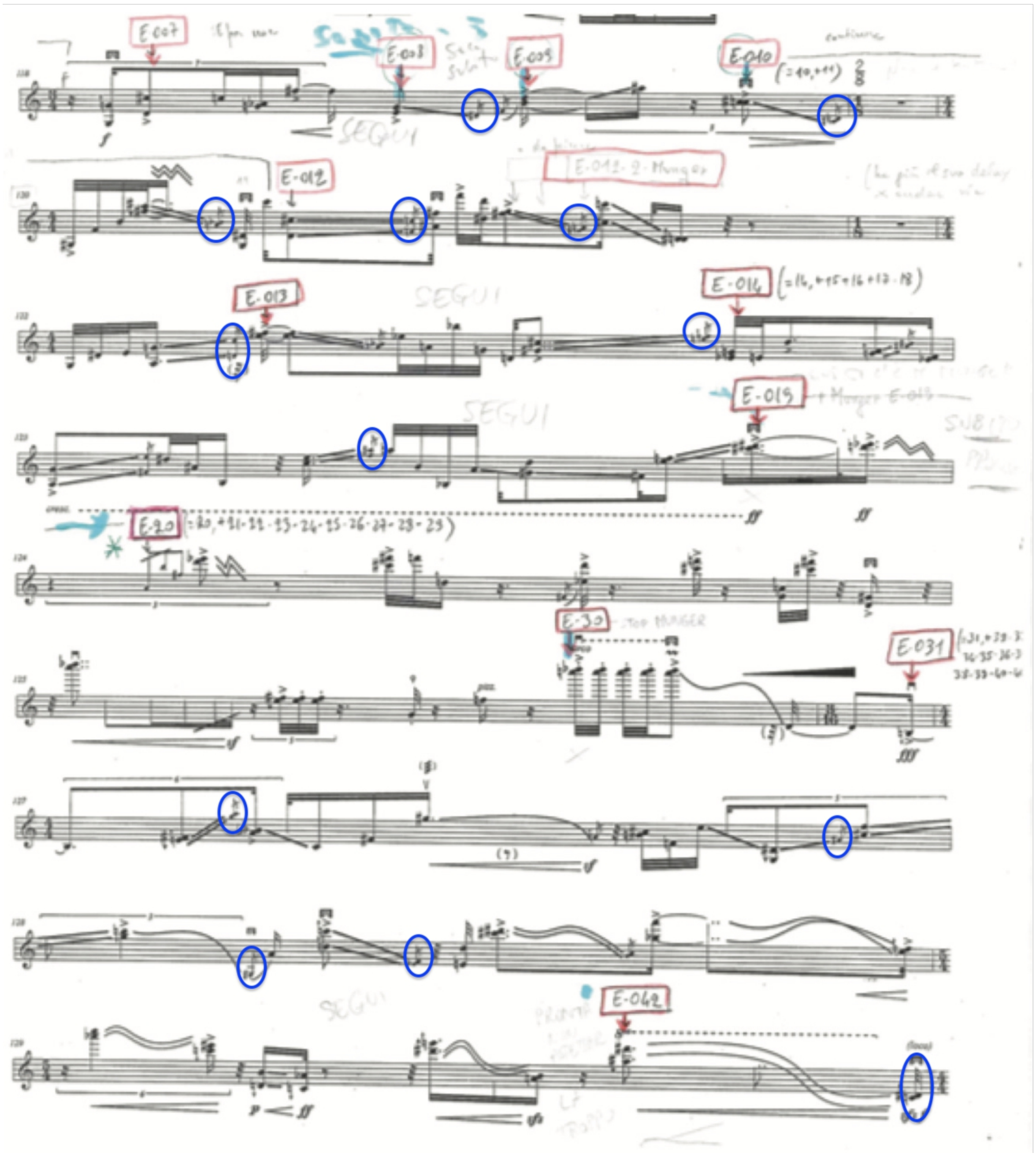
We emphasize here that this requirement to define and meticulously detail the processes of synthesis and sound transformation in this piece — in an attempt to emulate the nature of instrumental writing —, seems to have been motivated by a desire to derive and correlate the sound morphologies resulting from electronic processes to instrumental writing and to make it an integral part of musical discourse. Therefore, the use and exploration of the possibilities offered by the *Antescofo* is a technical choice that directly contributes to the realization of *Raggi di Stringhe's* poetic and musical premises.

To find a direct correspondence between specific ways of playing, rhythmic structures and instrumental gestures, electronic sounds are worked through various types of rhythmic and timbristic transformations, showing the influence that instrumental writing exerts on the conception of transformation processes and sound synthesis in the piece as a whole. On the other hand, it is evident that when using the *Antescofo* there is the opposite movement, in which the use of such a tool ends up transforming the instrumental writing itself in an attempt to adapt it to the use of the scoring paradigm and to integrate both sound environments.

According to Morciano, it was necessary to question certain elements of his rhythmic writing, such as *apoggiature*, which is often used to anticipate rhythmic figures, and also increasing the duration of the cells to which they refer or presenting the function of accelerating certain rhythmic gestures within regular pulsations established (MORCIANO, 2018, p. 83).

This process happens in the final section of *Raggi di Stringhe* (figure 7). In this excerpt, the composer uses appoggiaturas (in blue) at the end of the glissandos as a strategy to establish a rhythm and a pitch reference allowing the *Antescofo* to detect the beginning and end of each event and follow the instrumental gestures with greater precision — since the glissando does not provide a recurring temporal regularity that can guide the *Antescofo*. This writing strategy allowed maintaining an irregular and unpredictable rhythmic fluidity while allowing a detailed electronic writing score to be performed precisely in response to the player's actions.

FIGURE 7 – Final section of *Raggi di strighe*



Source: score of *Raggi di strighe* score, by Laura Morciano

c) *Dispersion de trajectoires* (2014), by Jospe Miguel Fernández

Another example of a compositional strategy that employs the *Antescofo* can be seen in *Dispersion de trajectoires* (2014), for baritone saxophone and live-electronics, by José Miguel Fernández. In this piece, *Antescofo* plays a central role in the process chain, as it is responsible for connecting, monitoring, and coordinating processes in different devices and computing environments. Like in *Jupiter*, there is in *Dispersion de trajeroires* a compositional thought that to some extent implies the use of some a media or technological apparatus aimed at the conception and registration of sound events in parallel with the instrumental score – again in an attempt to centralize the conception of computational/interactive processes and bring them into the realm of instrumental writing. This is evidenced by the architecture of the technical set created by the composer, which gives the *Antescofo* a leading role in coordinating processes throughout the piece.

As a technical strategy to work with instrumental writing and the conception of interactive processes in a centralized way, Fernández developed a library called *Antescollider*. This library integrates the *Antescofo* with the *SuperCollider*, allowing the conception of a centralized electronic score. The audio processing information is distributed to the *SuperCollider* servers using the *Open Sound Control* (OSC) protocol under the supervision and coordination of a score/script conceived in the *Antescofo* language. With *Antescollider* the audio processing can be distributed across multiple *SuperCollider* servers, thus maximizing CPU efficiency (FERNANDEZ; GIAVITTO; DONAT-BOUILLUD, 2019).

According to Fernández, the centralized score is relevant to the execution of the performance, since this model of employment of the *Antescofo* and the *Antescollider* creates support that makes it possible to explicitly write instrument-computer interactions. This setup also favors a dynamic approach to managing audio processing and synthesis – audio chains in *SuperCollider* can be easily created and destroyed dynamically, in direct response to the actions of the *Antescofo*.

4.4 Procedural

We classify as procedural paradigm every process that makes use of ML/MIR techniques to design algorithmic and compositional procedures implemented to operate in real-time. These procedures can be defined as a finite sequence of rules, reasoning, or operations that, once applied step-by-step to a data set, are capable of engendering structures as a final product.

Observing through the prism of the practice of Western instrumental and vocal writing, it can be seen that the notion of algorithm is already present in compositional procedures beyond the insertion of the computer in creative practice. The rules and formulae that guide contrapuntal writing or even the composition of dodecaphonic series, for example, are very close to concepts behind the definition of algorithm.

This set of compositional rules presupposes a particular formalization, in other words, a process of identifying a problem and choosing a solution. This formalization does not exempt the composer from searching for individualized solutions aimed at achieving specific aesthetic results. Although the need for non-rational thought coexists, such procedures aim at solving musical impasses through a set of logical steps, which demand the implementation of a congruent and well-formalized conception.

Therefore, the approximation between algorithms and compositional procedures is valid when they predict a symbolic and quantitative representation of music, and this is submitted to logical operations. That is, the procedures are defined by a formalization that demands the elaboration and definition of mathematical representations and automated operations on these representations (PADOVANI, 2009, p. 22).

Before the integration of these techniques into musical thinking, compositional procedures — such as structural manipulations; rhythmic; melodic, and harmonic, as well as those aimed at organizing and transforming the sound material from the perspective of timbre and spectral behavior — took place outside the time of the performance, requiring prior elaboration and calculation to be able to materialize. With the advent of live-electronics, and above all, with the instantaneity and interactivity brought by ML/MIR tools, such procedures became part of the musical performance itself (CHADABE, 1984), being able to adapt to the most different musical situations, to bring the transformation/sound synthesis mechanisms closer to autonomous agents.

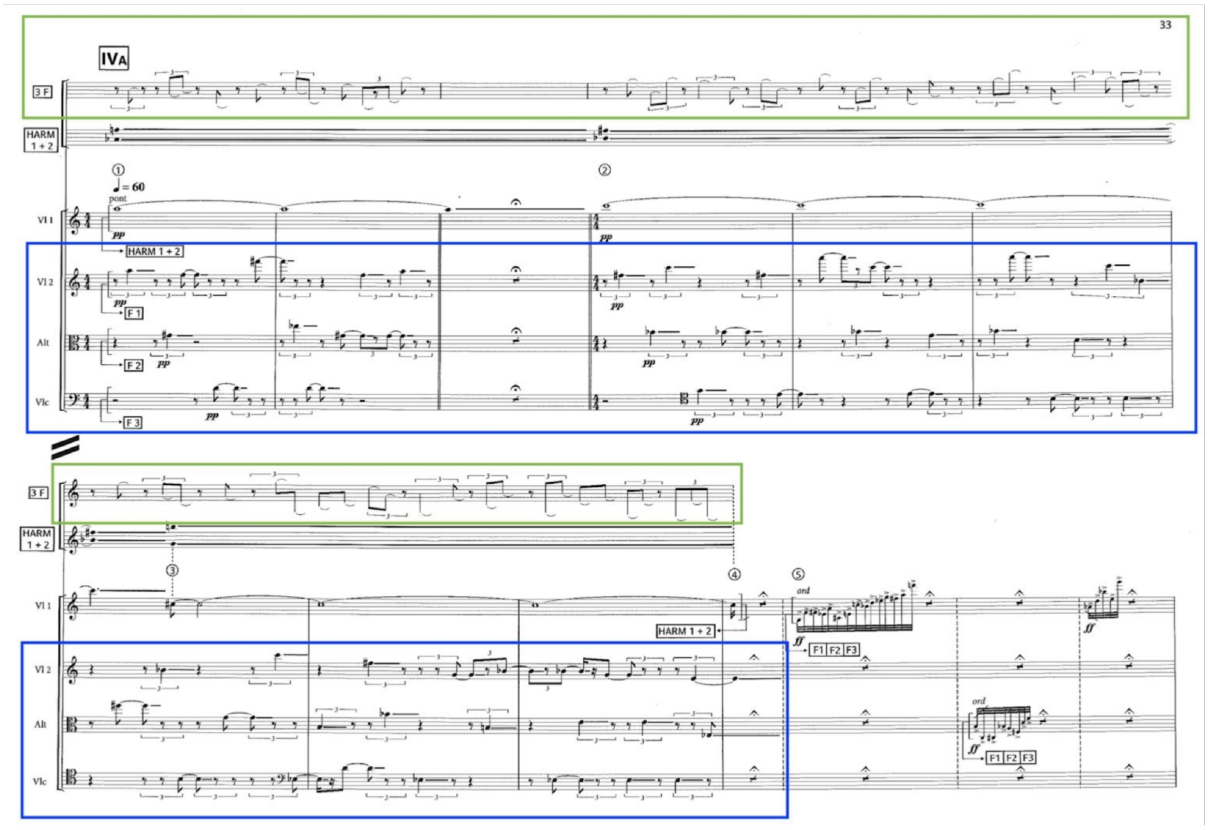
a) *Tensio* (2010), by Philippe Manoury

We can illustrate the use of the procedural paradigm through *Tensio* (2010), for string quartet and live-electronics, by Philippe Manoury. The procedural paradigm can be identified, in this piece, through the conception of an algorithm that acts just like a compositional procedure, engendering harmonic and melodic structures. The procedure relies on an additive synthesis process developed by

Miller Puckette in 2006, called *3F synthesis* (MANOURY, 2020). The *3F synthesis*, clearly influenced by the parallel frequency bands generated by modulatory processes such as *FM synthesis* and ring modulation, calculates a series of frequencies to be used in additive processes by manipulating input frequency values detected in real-time by ML processes/objects. Each of these three frequencies is subjected to all the possibilities of adding and subtracting each other. For example, given three frequencies A, B and C, we will have:

2A (A+B) (A-B) (A+C) (A-C)
2B (B+A) (B-A) (B+C) (B-C)
2C (C+A) (C-A) (C+B) (C-B)
3A (2A+B) (2A-B) (2A+C) (2A-C)
3B (2B+A) (2B-A) (2B+C) (2B-C)
3C ... etc.

Considering only the absolute numbers, we will have, as a result, a wide range of frequency values. As the spectra created in this way can become excessively dense, this cluster of frequencies undergoes an organization procedure based on a probabilistic logic, whose values can vary from 0.03 to 1. Imagining that we want to generate a spectrum composed of eight frequencies since the probability index is 1, the algorithm will select in ascending order the first eight calculated frequencies. As the probability index decreases, the algorithm randomly selects frequencies further apart from each other until reaching a total of eight components.

FIGURE 8 – Generating harmonic structures in section IV of *Tensio*

Source: score of *Tensio*, by Philippe Manoury

Thus, a proportional relationship is established between the probability index and the degree of harmonicity that the spectrum will acquire. Although such spectra are unique from a perspective of frequency combination — once a probability index less than 1 will always generate a different frequency combination by each calculation, even if the probability index value does not change -, it is possible to recognize them as belonging to the same class (MANOURY, 2013, p. 68).

Therefore, such spectral structures can create a spectral amalgamation with instrumental sounds.

At the beginning section **IVA** in *Tensio*, we can see the application of this procedure — in conjunction with the trigger paradigm, as demonstrated before. Each of the three input frequencies are pulled from a different instrument, i.e., the first frequency is retrieved from violin 2; the second from the viola, and the third from the cello. However, the process is performed only when short sounds are detected — longer sounds are ignored.

This procedure results in a harmonic structure that behaves as a kind of "virtual sub-group", whose harmonic/spectral and rhythmic content derives from the instrumental material itself. Since the number of frequency components calculated in this part of the piece is twelve, and the probability index is 0.19, the 3F synthesis calculation ends up generating a sound structure with a relatively high degree of inharmonicity. Figure 8 shows the layer resulting from this process, in green, superimposed on the string quartet, in blue.

In *Tensio*, another compositional process that implies the employment of the procedural paradigm is the retrieval of musical information related to instrumental performance to create melodic structures driven by a *Markov chain*. Manoury uses pitches and durations values retrieved by real-time analysis algorithms to create a transition matrix capable of generating a sequence of pitches and durations based on the probability of appearance of each one of them after each other. In this case, this succession of pitches generated by the Markovian process is mapped to the first of the three frequencies used in the calculation of *3F synthesis* procedure, which will ultimately engender melodic structures. A second strategy consists of mapping the sequence of values generated by the Markov process to trigger and control a database of sound samples of string instruments called *Synful Orchestra*[6] developed by Eric Lindenmann.

Hence, the composer uses the *Synful Orchestra* as a virtual performer ("virtual string quartet") that performs the parametric musical information resulting from the Markov process and transforms them into melodic structures. With this strategy, the string quartet is expanded by computational means at the idiomatic level. The *Synful Orchestra*'s mechanism consists precisely in this use of pre-recorded instrumental parts that are recombined to create new phrases and sound articulations. It explores, therefore, sound materials typical of the literature of these instruments, and since they perform melodic structures generated from a stochastic process and interval and temporal manipulations — reordering sequences of pitches, intensities and durations —, they enable the creation of rearrangements that can technically and idiomatically extend the string quartet instruments.

The algorithm responsible for carrying out the Markovian process has some control parameters that influence the characteristics that the resulting melodic structures will present: (1) tempo of the

---

[6] https://www.synful.com/

melodic lines; (2) interval contractions and expansions; (3) pitch duration control (the higher the value, the greater the uniformity of durations); (4) transposition control; (5) minimum and maximum amplitude range; (6) interval multiplier that allows overlapping streams.

b) *Traverser le réseau* (2014), by Rael Toffolo

Traverser le réseau (2014), for cello and live-electronics, by Rael Toffolo illustrates another type of strategy in the use of computational modeling of compositional procedures using ML/MIR techniques. According to Toffolo, the title of the piece [crossing the net] alludes to Henri Pousseur's creative exploration of harmonic networks, which is the central point of the poetics that guides both instrumental writing and the conception of the processes of transformation, spatialization and sound generation present in this piece (TOFFOLO, 2014).

In *Traverser le réseau*, the composer seeks to computationally modeling the harmonic networks as described by Henri Pousseur in POUSSEUR (2009), aiming at generating material for the creation of harmonic-spectral structures which will be engendered in real-time from the cello-computer interaction.

Pousseur's harmonic networks consists of a serial compositional procedure aimed at the organization, manipulation and generation of interval material (harmonic and melodic), and consists of the projection of a given set of pitches distributed in axes that are characterized each as a chain of a single interval (POUSSEUR, 2009, p. 249). Through this projection of a given interval cluster, it is possible to apply processes of "deformation" (interval transposition) on the axes to generate a new set of pitches.

In this context, to create a method for classifying the material generated by harmonic networks, Toffolo implements an algorithm responsible for calculating the harmonic voltage level from the categorization of intervals as polar and nonpolar (COSTÈRE, 1954, 1962) and from psychoacoustic models aimed at the perception of consonant and dissonant intervals (BIGAND; PARNCUTT; LERDAHL, 1996; HELMHOLTZ, 2009; PLOMP; LEVELT, 1965).

For this purpose, the composer uses the extraction of roughness to calculate the harmonic tension rate of each set of pitches. The model consists of calculating the resulting roughness rate of all partials belonging to one or more sounds and can be summarized in three steps: (1) the harmonic series of each note in the cluster is generated; (2) the roughness between each of the components of the harmonic series is calculated in a combined way; (3) finally, the roughness values between all frequency pairs are summed and divided by the general amplitude of the cluster, resulting in the harmonic voltage ratio (TOFFOLO, 2014, p. 77).

Throughout *Traverser le réseau*, this procedure is employed in several ways and operates in the creation of harmonic/melodic structures and the generation of complex sound masses and textures from the interaction of the instrumentalist with the computer.

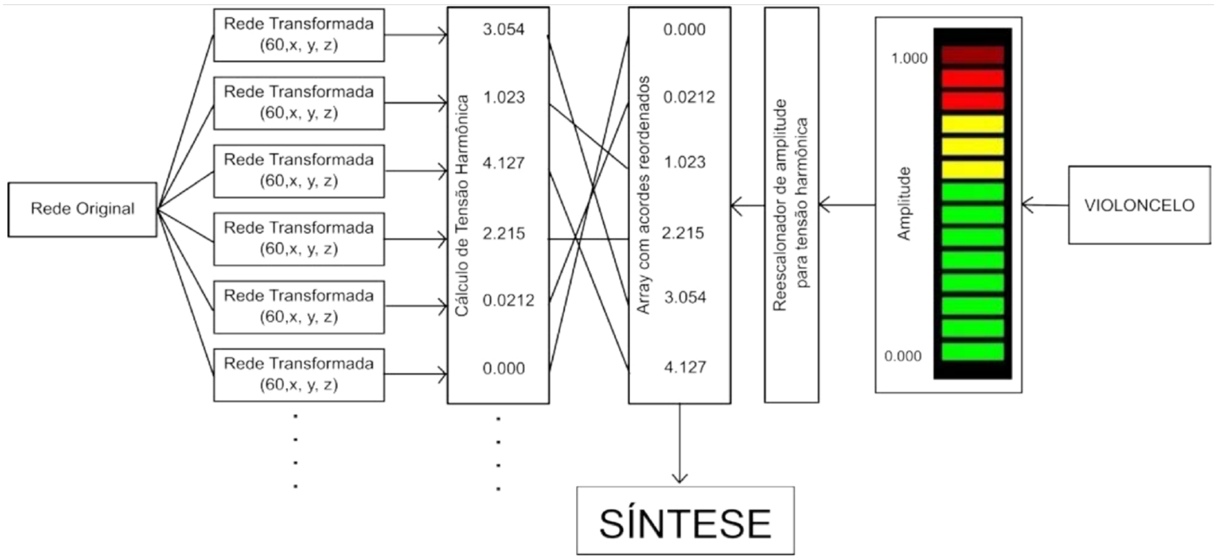
In the third section of piece, Toffolo starts from a chord taken from the *Rondes Printanières*, from *The Rite of Spring* (1913), by Igor Stravinsky, which consists of the superimposition of an F major chord on an A flat minor[7]. This chord is projected on the harmonic networks with all interval combinations of possible axes, thus generating a series of harmonic algorithms.

From the calculation of the harmonic tension ratio, these clusters are ordered gradually from the least tense to the tensest. Once classified, each cluster is associated with a range of values of the cello amplitude (extracted in real-time). While the amplitude is extracted from the instrumental performance in real-time, a cluster of pitches correlated to the extracted amplitude value is selected and sent to a bank of oscillators, which synthesizes a spectral texture. According to Toffolo, this strategy, allows the creation of sound objects/masses with a high degree of variability and high spectral density. Figure 9 shows the architectural representation of this procedure.

---

[7] The chord was transposed down a major third to adapt to the cello's texture (TOFFOLO, 2014, p. 99)

FIGURE 9 – The architecture of the implementation of harmonic networks associated with the calculation of the harmonic tension, and the mapping of the cello amplitude in *Traverser le réseau*



Source: TOFFOLO (2014, p. 106)

5. Conclusion

In two decades of development, ML/MIR have been growing exponentially as research fields, not only in the academic sphere but also in the industrial/commercial scenario. However, despite research in these fields is in constant expansion, transforming the possibilities placed to musical practices to boost research with very diversified creative biases, there is still a lack of works concerned with carrying out reflections on the compositional exploration of these tools.

We aim to contribute with understanding the creative exploration of these tools in contemporary creative practices, outlining and scrutinizing how composers have been applying these resources in musical creation in the context of the composition of pieces that combine instrumental/vocal and sound processing in real-time. By identifying the previously listed paradigms, it was possible to glimpse not only a series of technical frameworks but also to point out compositional approaches used by composers as strategies to solving technical-musical issues that aim to achieve certain poetic conceptions and aesthetic results.

As already highlighted, the generalization of these application strategies does not aim to exhaust or to rigidly categorize all possibilities of explorations of MIR/ML methods in the context of live-electronics/interactive music composition. As stated, the objective is to circumscribe in a general way, ways of thinking and conceiving the creative use of ML/MIR, seeking to offer a broader view of the use of these resources that can prove to be useful in analytical, creative, and pedagogical contexts.

ACKNOWLEDGMENT

This work has been supported by the Brazilian Agency CAPES (Coordination for the Improvement of Higher Education Personnel) and by FAPEMIG (Minas Gerais Research Funding Foundation).

REFERENCES

BIGAND, Emmanuel; PARNCUTT, Richard; LERDAHL, Fred. Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & psychophysics*, v. 58, p. 124–41, 1 jan. 1996.

BREGMAN, Albert S. *Auditory Scene Analysis*. The MIT Press, 1990. Available at: <https://mitpress.mit.edu/books/auditory-scene-analysis>. Access: 9 jun. 2021.

BRENT, William. A timbre analysis and classification toolkit for pure data. 2010, 2010.

CHADABE, Joel. Interactive Composing: An Overview. *Computer Music Journal*, v. 8, n. 1, p. 22–27, 1984.

COLLINS, Nick. Automatic Composition of Electroacoustic Art Music Utilizing Machine Listening. *Computer Music Journal*, v. 36, n. 3, p. 8–23, 1 sept. 2012.

COLLINS, Nick. Machine Listening in SuperCollider. In: WILSON, S.; COLLINS, N.; COTTLE, D. *The SuperCollider book*. Cambridge, Mass: MIT Press, 2011. .

COLLINS, Nick. *Towards Autonomous Agents for Live Computer Music: Realtime Machine Listening and Interactive Music Systems*. 2006. Universoty of Cambridge, 2006.

CONT, Arshia. Antescofo: Anticipatory Synchronization and control of Interactive parameters in Computer Music. In: INTERNATIONAL COMPUTER MUSIC CONFERENCE, 2008, Belfast. *Proceedings*... Belfast, 2008. . Access: 1 mar. 2021.

COSTÈRE, Edmond. Entre l'harmonie classique et les harmonies contemporaines, y a-t-il rupture, ou continuité? *Revue de Musicologie*, v. 36, p. 55–65, 1954.

COSTÈRE, Edmond. *Mort: Transfiguration de l'harmonie*. Presses universitaires de France, 1962.

DE CHEVEIGNÉ, Alain; KAWAHARA, Hideki. YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, v. 111, n. 4, p. 1917–1930, apr. 2002.

EINARSSON, Anna. Experiencing Responsive Technology in a Mixed Work: Interactive music as embodied and situated activity. *Organised Sound*, v. 22, n. 3, p. 418–427, dec. 2017.

EINARSSON, Anna; FRIBERG, Anders. Using Singing Voice Vibrato as a Control Parameter in a Chamber Opera. In: ICMC, 25 sept. 2015, University of North Texas. *Proceedings*... University of North Texas, 25 sept. 2015. p. 18–21.

EMMERSON, Simon. Ring Modulation and Structure. *Contact: A Journal for Contemporary Music (1971-1990)*, n. 17, 1 jun. 1977. Available at: <https://journals.gold.ac.uk/index.php/contactjournal/article/view/contactjournal17_p14-20>. Access: 3 apr. 2022.

FERNANDEZ, José Miguel; GIAVITTO, Jean-Louis; DONAT-BOUILLUD, Pierre. AntesCollider: Control and Signal Processing in the Same Score. In: ICMC 2019 - INTERNATIONAL COMPUTER MUSIC CONFERENCE, 16 jun. 2019, 16 jun. 2019. Available at: <https://hal.inria.fr/hal-02159629>. Access: 27 apr. 2021.

FERRAZ, Silvio. *Criação musical com suporte tecnológico*. . Salvador, 1999. Available at: <http://sferraz.mus.br/anp99.htm>. Access: 3 apr. 2022.

FRIBERG, Anders; SCHOONDERWALDT, Erwin; JUSLIN, Patrik. CUEX: An Algorithm for Automatic Extraction of Expressive Tone Parameters in Music Performance from Acoustic Signals. *Acta Acustica united with Acustica*, v. 93, p. 411–420, 1 may. 2007.

GARNETT, Guy E. The Aesthetics of Interactive Computer Music. *Computer Music Journal*, v. 25, n. 1, p. 21–33, 2001.

GUTSCHALK, Alexander; DYKSTRA, Andrew R. Functional imaging of auditory scene analysis. *Hearing Research*, Citation Key: gutschalk_functional_2014, v. 307, p. 98–110, 2014.

HELMHOLTZ, Hermann L. F. *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. Translated by Alexander J. Ellis. 3. ed. Cambridge: Cambridge University Press, 2009. Available at: <https://www.cambridge.org/core/books/on-the-sensations-of-tone-as-a-physiological-basis-for-the-theory-of-music/6B5630E5D439286A8946CEB9D62BE6EC>. Access: 3 apr. 2022. (Cambridge Library Collection - Music).

HUNT, Andy *et al.* The Importance of Parameter Mapping in Electronic Instrument Design. *Journal of New Music Research*, v. 32, 5 may. 2002.

IAZZETTA, Fernando. Meaning in Musical Gesture. *Trends in Gestural Control of Music [CD-ROM]*. IRCAM, 2000. p. 259–268.

IRCAM. *Raggi di stringhe, Lara Morciano*. Available at: <https://brahms.ircam.fr/works/work/29525/>. Access: 3 apr. 2022.

IRCAM. *Tensio, Philippe Manoury*. Available at: <https://brahms.ircam.fr/works/work/28236/>. Access: 3 apr. 2022.

JOHNSON, Mark. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. 1ª edição ed. University of Chicago Press, 2013.

MALT, Mikhail. Khorwa: A musical experience with “autonomous agents”. 2004, 2004.

MALT, Mikhail; GENTILUCCI, Marta. Real Time Vowel Tremolo Detection Using Low Level Audio Descriptors. *arXiv:1511.07008 [cs]*, arXiv: 1511.07008, 22 nov. 2015. Available at: <http://arxiv.org/abs/1511.07008>. Access: 25 may. 2021.

MALT, Mikhail; JOURDAN, Emmanuel. Real-Time Uses of Low Level Sound Descriptors as Event Detection Functions Using the Max/MSP Zsa.Descriptors Library. p. 12, 2009.

MANOURY, Philippe. Compositional Procedures in Tensio. *Contemporary Music Review*, v. 32, n. 1, p. 61–97, 1 feb. 2013.

MANOURY, Philippe. *La note et le son*. *Philippe Manoury» Blog Archive »*. Available at: <http://www.philippemanoury.com/?p=330>. Access: 22 feb. 2021. , 1990

MANOURY, Philippe. *Sur la synthèse 3F.* . Available at: <http://www.philippemanoury.com/?p=7084>. Access: 11 jun. 2021. , 2020

MAY, Andrew. Philippe's Manoury Jupiter. *Analytical Methods of Electroacoustic Music*. Routledge, 2005. p. 145–185.

MIRANDA, Eduardo Reck; WANDERLEY, Marcelo. *New Digital Musical Instruments: Control And Interaction Beyond the Keyboard*. 1st edition ed. Middleton, Wis: A-R Editions, Inc., 2006.

MONTEIRO, Adriano Claro. *Criação e performance musical no contexto dos instrumentos musicais digitais*. 2012. [s.n.], 2012. Available at: <http://repositorio.unicamp.br/jspui/handle/REPOSIP/284418>. Access: 20 sept. 2020.

MORCIANO, Lara. *Ecriture du son, du temps et de l'espace dans l'interaction entre instruments et dispositifs numériques synchrones*. 2018. Theses – Université Paris sciences et lettres, 2018. Available at: <https://tel.archives-ouvertes.fr/tel-02270028>. Access: 3 apr. 2022.

PADOVANI, José Henrique. Representação, intuição e contato na composição com algoritmos. Dissertação de Mestrado. Universidade Estadual de Campinas, Campinas SP, 2009.

PESTOVA, Xenia. Models of interaction: performance strategies in works for piano and live electronics.*Journal of Music, Technology and Education*, v. 2, p. 113–126, 1 dec. 2009.

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

POUSSEUR, Henri. *Apoteose de Rameau e outros ensaios*. 1. ed. 2009. Available at: <http://editoraunesp.com.br/catalogo/9788571399198,apoteose-de-rameau>. Access: 29 jun. 2021.

PUCKETTE, Miller. Max at seventeen. *Computer Music Journal*, v. 26, n. 4, p. 31–43, 2002.

PUCKETTE, Miller; APEL, Theodore; ZICARELLI, David. Real-time audio analysis tools for Pd and MSP. In: INTERNATIONAL COMPUTER MUSIC CONFERENCE, 21 oct. 1998, São Francisco. *Proceedings*... São Francisco, 21 oct. 1998.

REGIUS, Hélena; LANGER, Joakim. *Kung Kalle av Kurrekurreduttön*. Fórum Bokförlaget, 2002. Available at: <https://biblioteket.stockholm.se/titel/833558>. Access: 25 may. 2021.

RIMOLDI, Gabriel. *Metaflauta : uma abordagem ecológica aplicada ao design e processos criativos a partir de um instrumento expandido*. 2018. [s.n.], 2018. Available at: <http://repositorio.unicamp.br/jspui/handle/REPOSIP/332105>. Access: 7 apr. 2021.

ROWE, Robert. *Interactive music systems: machine listening and composing*. Cambridge: MIT press, 1993. Available at: <http://dl.acm.org/citation.cfm?id=530519>. Access: 16 jun. 2014.

ROWE, Robert. The aesthetics of interactive music systems. *Contemporary Music Review*, v. 18, n. 3, p. 83–87, jan. 1999.

TOFFOLO, Rael Bertarelli Gimenes [UNESP. *As redes harmônicas de Pousseur aplicadas à composição musical em tempo real*. 2014. Universidade Estadual Paulista, 2014. Available at: <https://repositorio.unesp.br/handle/11449/108808>. Access: 29 jun. 2021.

WANG, Song; WANDERLEY, Marcelo; SCAVONE, Gary. The Study of Mapping Strategies Between the Excitators of the Single-Reed Woodwind and the Bowed String. *Proceedings of the 7th Conference on Sound and Music Technology (CSMT)*. Springer, 2020. p. 107–119.

WINKLER, István *et al.* Multistability in auditory stream segregation: a predictive coding view. *Philosophical Transactions of the Royal Society B: Biological Sciences*, Citation Key: winkler_multistability_2012tex.pmcid: PMC3282310, v. 367, n. 1591, p. 1001–1012, apr. 2012. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3282310/>. Access: 9 jun. 2021.

ABOUT THE AUTHORS

Vinícius Cesar de Oliveira holds a bachelor's degree in music composition (2018) and a master's degree in creative processes (2022) from Unicamp. Currently, he is studying for a doctorate's degree in music at the Arts Institute (Unicamp), where he researches compositional strategies in the use of machine listening and music information retrieval associated with machine learning in the context of live-electronic music under the supervision of Prof. Dr. José Henrique Padovani. ORCID: https://orcid.org/0000-0002-7711-3949. E-mail: oviniciuscesar@gmail.com

José Henrique Padovani is Assistant Professor at the School of Music of the Federal University of Minas Gerais (UFMG), teaching Composition, Computer Music, and Music Theory. Padovani is a member of the Music Graduate Programs at the School of Music at UFMG and the Institute of Arts of the State University of Campinas (UNICAMP). More information at: http://josehenriquepadovani.com and http://musica.ufmg.br/padovani/. ORCID: https://orcid.org/0000-0002-8919-7393. E-mail: zepadovani@gmail.com