Flaubert
Revue critique et génétique

21 | 2019
Flaubert en musique

Burger Time ou les Tentations de saint Antoine, or Not Wanting to Say Anything About Mauro Lanza, Gustave Flaubert and Peter Pepper

Andrea Valle

Electronic version
URL: http://journals.openedition.org/flaubert/3650
ISSN: 1969-6191

Publisher
Institut des textes & manuscrits modernes (ITEM)

Electronic reference
Andrea Valle, « *Burger Time ou les Tentations de saint Antoine,* or Not Wanting to Say Anything About Mauro Lanza, Gustave Flaubert and Peter Pepper », *Flaubert* [Online], 21 | 2019, Online since 01 June 2019, connection on 27 September 2019. URL : http://journals.openedition.org/flaubert/3650

Burger Time ou les Tentations de saint Antoine, or Not Wanting to Say Anything About Mauro Lanza, Gustave Flaubert and Peter Pepper

Andrea Valle

1 *Burger Time ou les Tentations de Saint Antoine* is a 12-minute (12'54", to be precise) piece of electronic music composed by Mauro Lanza. Being a close friend of the composer, I would have no problem in extorting him the information that I might deem useful for the description of the piece, from the aesthetic principles at its origin down to the minute details of its technical implementation. However, this existential proximity is embarrassing for a semiologist as I am: the goal of the semiologist being the description, though certainly problematic, of how sense is constructed as immanent to the object, regardless of its genetic history. In Jules Verne's *Les enfants du Captain Grant*, Lord Glenarvan retrieves a fragmentary document in a bottle swallowed by a shark. In investigating the three-part document, the only assumption that he makes concerns the enunciation: the message has been certainly written by someone. As Glenarvan says, "les quelques mots qui ont résisté ne me laissent aucun doute à cet égard". But, on the other hand, Glenarvan also observes: "les mots tracés sur ces documents sont fort incomplets". Thus, in order to understand the message, a difficult reconstruction attempt is to be tried. These two elements, an assumption of significance and a state of incompleteness, are the starting points for the hermeneutic effort of both Lord Glenarvan and the semiologist.

2 Jacques Fontanille, in an essay dedicated to the analysis of a work by Mark Rothko (*Untitled*, 1951)[1], had to deal with a set of translucent chromatic textures, forbidding him any figurative recognition. A situation in some way analogous to contemporary music, especially in electronic music in which even the reference to instrumental recognition may be missing. In dealing with Rothko, the French semiologist thus proposed, as a working hypothesis, rather than to discuss what we already know about the author and the work, to "take the side of ignorance". In agreeing with Fontanille, ignorance will therefore be my guide, primarily in the form of *ignorantia auctoris*. In short, Mauro Lanza who? I will therefore oppose a methodological distance to a friendly proximity.

3 So, how to introduce such a distance?

4 I will start from three basic hypotheses, according to three models of distance that I will discuss later:

1. I will access the piece *Burger Time* (hence on, BT) assuming that it is a message, and that therefore it should be framed within a situation of enunciation. What is presumed about this enunciation framework is (only) that the message has an acoustic nature. Therefore, its organization cannot abstract from its perceptual access. It is an audible message and it must be analyzed through listening;
2. as a control step for the previous results, I will propose a formal analytical framework, i.e. a machine-based procedure to extract information from BT. Such a step aims at an "objective" description, where the adjective here has no ontological value, but simply a methodological one: it means "consistent with its starting assumptions". It is therefore necessary to access the piece as an acoustic signal recorded on a decodable support;
3. Finally, I will reduce the problem of context to a microcorpus consisting of the piece itself and the program notes. I will delimit the content according to such a microcorpus and verify a minimum interpretation on the data gathered from the previous steps, that is, I will try to propose an overall articulation between expression and content.

Audible industry

5 The first model of a maximal hermeneutical distance is based on space and time, but still within the boundaries of the human. Considering space, such a maximal distance could be anthropological, as it happens when a distant culture is approached. Yet, as anthropology has shown, even in the case of remote populations, if a contact happens, then a shared knowledge becomes possible, up to the inclusion of the interpreter in a participant observation setting. A second type of anthropological distance takes into account (also) time: even in this case, however, forms of partial continuity are possible within cultural communities, culture being exactly an extra-bodily memory that passes through generations. However, going back a very long way, the situation becomes methodologically complex and fascinating. For example, starting from about 40,000 years ago, during the Upper Paleolithic, populations of *Homo sapiens* (therefore genetically exactly like us) have painted caves around the world. Despite a possible genetic and perhaps cultural continuity, the meaning of these representations is typically reduced to the recognition of geometric and animal elements. The sense of these elements, of their configuration, of the overall situation is lost, and can be conceived only, and explicitly, as a reconstruction that the interpreter takes on from a textual datum that must be interrogated. So, the actual interpretations (cave hunting,

shamanic rite, or merely an unstructured desire to express themselves) strongly diverge, while still being plausible[2]. But in Lower Paleolithic, millions of years before *Homo sapiens*, the situation, now dealing with stones, can become much more complex. A first case is when it is unclear whether or not a stone is the product of a lithic industry. If it is, then the stone arises in status as an expression of a certain content (the function, for example scraping); if it is not, the stone is pushed back into nature as a sort of a pareidolia, a natural accident, as a face that appears in a cloud. In this extreme hermeneutic case, the object lies on the boundary between nature and culture. A second case involves chipped stones clearly resulting from human workmanship, but whose function is totally inaccessible. Such a lack is crucial for the recognition of the "object" itself. As an example the interpretation of certain faceted stones has been historically reversed: in older interpretations, the object was the stone, obtained from the removal of lithic flakes; in more recent ones, the lithic flakes are no more waste, rather they are considered the proper product of lithic reduction, while the chipped stone is simply the remaining core[3]. Of course, as the object changes, its possible functions vary accordingly. In both the previous cases, interpretation is supported by a minimum, that can be thought in terms of a sort of abstract ergonomic principle: the inscription of a (proto)human corporeality into the object. In the first case, if the stone has chippings, then it is an artifact, and the threshold is exceeded. Thus, the chipping traces on the stone declare the presence of a Model Author, to speak with Eco[4], a presence that converts the stone into an artifact. This authorship is described only by the way in which the body has operated on the stone. In the second case, this threshold is already surpassed: the object is an artifact. Still what escapes completely to interpretation is how the human body is co-articulated with the object. How does the object enter in relation with other objects, primarily the body of its human user? Here the problem is partially different and mainly narrative: what theater of actions can those forms of corporeality, the human body and the lithic object, set up? Is it a scraper or a blade?

6 Following the model of lithic industry, I will thus consider BT under the notion of audible industry.

7 First, the nature of BT as an object has to be questioned, and, second, the forms of corporeality that it sets up in relation to the audible domain have to be investigated.

8 The nature of BT as an artifact is undoubted: the acoustic datum shows a specific craftsmanship. Such a feature would be less obvious if the documentary evidence presented itself as a recording of a previous acoustic situation. In the context of electronic music, one may think about field recording, that is to say, a genre in which the work is conceived, in the extreme case, as a recording of a certain acoustic landscape. How can we assume that BT is not a reproduction of an existing soundscape? A fundamental clue is a set of acoustic features (for example reverberation and distribution of sound in space) that allows the listener to reconstruct an acoustically realistic scene, following the optical model of photography. Now, these elements of "aural realism" (on the model of optical realism) are absent from BT. On the contrary, these aspects are carefully avoided. BT therefore does not show a documentary nature, rather it is constructed as a sort of chiseled object, presenting itself to the ear with clear boundaries and a precise organization.

9 Once the first threshold has been exceeded, it is possible to observe how the BT object, while not proposing itself as an acoustic documentary, still refers to a clearly figurative

background. But what is a figure in the audible domain? How is it related to corporeality?

10 In relation to this, Bregman has proposed a fruitful approach, the so-called Auditory Scene Analysis[5]. This theoretical framework is based on an ecological assumption, as it aims at studying audible perception as located in an environmental context, in the human *Umwelt*, to speak with von Uexküll. Hence, auditory perception is characterized as a system that takes in input the acoustic data and organize them in "streams" according to a set of "heuristics", that is, non-deterministic -but typically effective- strategies. Heuristics share two main features: first, they can be either linked to the anthropological constitution of the human (e.g. Gestalt principles) or historically determined according to cultural competences (e.g. in case of trained listening); second, they take as a starting point the spectral information accessible to the auditory apparatus, that converts a single mechanical vibration into a set of separate information related to various frequencies. The phenomenological correlate of this process is actually quite obvious in "everyday listening"[6]. When listening to a surrounding soundscape, a single mechanical vibration involving air molecules reaches the ear. Yet, we are able to recognize many different sounds at the same time. In Bregman's hypothesis, raw data are decomposed into frequencies by the auditory apparatus and then re-aggregated by various heuristics, all pursuing the same goal: to obtain "a nice consistent story about the sound"[7], that is, to provide an ecologically plausible reconstruction of the acoustic sources at its origin. A complex soundscape, that reaches the ears as a single vibration, is thus decomposed into frequencies, these data are re-aggregated into streams by heuristics, streams are interpreted as products of a set of sources (a voice, the dog barking in the distance, a pneumatic hammer on the road, a car horn). While such a listening process is apparent in everyday listening, indeed music listening involves other processes, with a variable degree of cultural complexity, that may for example underline the relationship between certain streams rather than their binding to certain sources. Yet, also Pierre Schaeffer, in proposing his theory of the four modes of listening[8], had noted how the first listening mode to be activated is a "causal" one. For this causal, indexical mode, listening is listening to the causes at the origin of sound.

11 Apart from the plurality of possible listening heuristics, two features are relevant:

1. auditory scene analysis may be seen as a common anthropological background. In the listening scenario, the figurative dimension, i.e. the recognition of elements according to a certain encyclopedia, thus takes the form of a sort of mechanics that deals with actions to be linked to perceptions: the "form" in the visible domain has its audible equivalent in an "action"[9];
2. back to BT, it is neither safe nor necessary to assume that we are dealing with "music". More generally and cautiously, I will assume that BT is simply an audible message.

12 Therefore, my starting assumption is to reorganize the perceptual content of BT as related to a plausible auditory scene, made up of various acoustic sources. To include at least a musical reference, one could describe this approach as a rendering in terms of an abstract instrumentality. I will call "figures" these set of sound actions, as they are related to a minimal figurative, "mechanical" so to say, recognition framework.

13 With respect to the visible domain, the retrieval of figures in the audible one is complicated by a peculiar condition, that of a reciprocal transparency. A visible object that stands on top of another typically hides it; on the contrary, two audible objects can

be thought, in relation to visual perception, as transparent. Hence, the main difficulty for the audible system is how to handle mixtures of overlapping sounds[10]. BT seems to cope with this basic issue by mostly avoiding ambiguous mixtures, as it displays a catalog of clear-cut acoustic actions that are typically arranged by contrasts, so that they can be recognized also in complex sound textures. Indeed, this is another clue that conjures up the artificial nature of BT. This is apparent at the beginning of the piece: here, two of these figures are presented in an alternate way. Thanks to such a perceptual organization, it is possible, as a first approximation, to identify various figures that represent different audible actions, to be labeled accordingly. This labeling procedure obviously exploits some features of the figures, but depends on the encyclopedia of the interpreter: its main goal is simply to provide an identifier, just like a Bronze Age petroglyph in the alpine Vallée des Merveilles can be labeled -with an obvious anachronism- as "Le Christ"[11].

14 Thanks to the neatness of the contours, 11 figures can be identified in BT, divided into two groups according to the opposition between "flows" and "objects". Group I consists of a set of flows. A flow can be thought of as a coherent multiple emission from a certain source[12].

15 Flows are:

- *Trumpet*: the sound at the beginning of the piece, in the acute middle register. It recalls a toy trumpet, or rather a section of toy trumpets, since parallel, overlapping events can be heard;
- *Cymbalom*: in the beginning of BT it is immediately opposed to Trumpet. It shows a sort of percussive and plucked nature, hence the name;
- *Voice*: a vaguely vocal flow, in the middle-low register. If compared to Frog (see later), however, it does not have a creaky feature. It is very often associated with Cymbalom, but two facts favor its identification as autonomous. First, sometimes it appears without Cymbalom, and, second, when the two flows are together, they show uncorrelated behaviors;
- *Frog*: a sort of vocal and/or animal emission, with a creaky feature that vaguely recalls a croaking (hence the label). It is presented in solo as a sequence of repetitions after the first section that alternates between Trumpet and Cymbalom/Voice. It also (probably) makes a first fleeting appearance in pianissimo after the very first alternation of Trumpet and Cymbalom (see Figure 1 and 2);
- *Breath*: an eminently noisy, even if pitched, sound, as blowing into a bottle;
- *Pluck*: this flow has a plucked string-like quality. It differs from Cymbalom because it displays a sort of metallic nature that provides also a stronger and noisier attack.

16 These flows are organized into note-like sequences, that is, they are made up of short events, clearly delimited in time, i.e. provided with a beginning and an end. Thus, flows have two main features:

1. they are made up of temporally well-formed, "eumorphous" sound events[13];
2. even if displaying a complex timbre, they still shows a pitched quality.

17 The second group consists of a set of figures that are characterized by exact, even if sometimes incomplete, repetition. One of the criticisms to Schaeffer's theory of sound object is that, when applied to sound, the very notion of object somehow over-emphasizes the clarity of a certain event with respect to its background, while at the same time hiding its nature of temporal process[14]. To avoid this criticism, the previous

figures have been identified as “flows” rather than as “objects”. But in the case of Group II the notion of object captures exactly the nature of samples of its members.

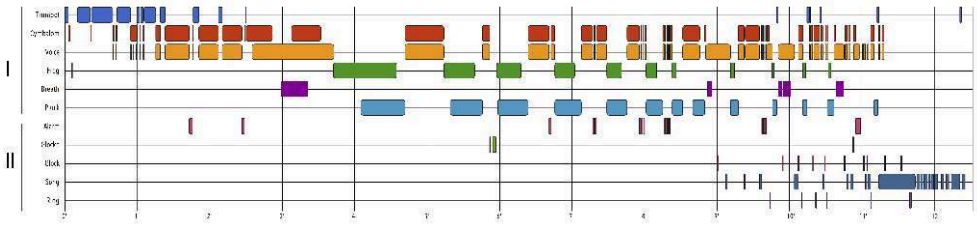
18 Objects are:

- *Alarm*: a two-tone object made up of a sequence of two pitches;
- *Glock*: a sort of glockenspiel gliding towards the acute, sometimes presented in an incomplete form;
- *GlockB*: again a glockenspiel, but with different arrangement with respect to Glock, so that it might be considered as a set of variants of the latter. Given the many identical occurrences of Glock, GlockB is used for the all the others;
- *Doorbell*: a tremolo doorbell-like sound;
- *Song*: the most peculiar object is a clearly melodic fragment that is partially or completely repeated. It is actually a composite object, as it is made up of various sources grouped by its organization.

19 Such a breakdown into two groups, “flows” and “objects” already outlines a possible difference in terms of enunciation. Group I describes a set of “actors”, each provided with a specific sound generation competence (an “instrument”). In relation to Eco’s theory of modes of sign production[15], each flow can therefore be reconstructed as the “imprint” of a certain actor. Group II instead is made up of internally organized blocks that are iterated in a substantially identical way. These objects are properly “signals” as defined by soundscape studies: short, with precise boundary, exactly repeated[16]. Again referring to modes of sign production, they are produced not by “imprint” but as “replicas” and exactly because of this they can be interpreted as signals. Now, signals are culturally established, but there is no information about the conventional meaning of these objects in BT: still, they show an opportune set of features to support a signal-like function. What is recognized is not an imprint of an actor “on stage”, rather the (ri)presentation of an already produced object, as a sample or a citation. Such a productive mode is called “ostension” by Eco[17]. In short, these objects seem to involve aspects of replica and ostension. Better, they are ostensions of replicas: pre-built pieces adhering to a type (replica) assembled like in a collage or in sampling procedures in popular music, i.e. through ostension. Thus, the grouping reveals an interesting aspect of the enunciation at work. In BT, enunciation operates according to two different regimes: on the one hand, it sets up a theater in which unknown but ecologically viable instruments are sounding, on the other hand it glues on the same surface (to use a visual analogy) a set of ready-made fragments. However, the work of enunciation is subtler, because such a layering procedure, rather than yielding a collage of heterogeneous elements, is otherwise homogenized: objects are not treated as materially heterogeneous samples (e.g. like tape excerpts pasted onto a tape assemblage), rather they seem to be realized by other actors, that is, they are presented like instrumental quotations. Therefore, the grouping reveals a sort of holographic effect set up by enunciation. On the one hand, BT is a unique surface/stage, on the other it reveals, according to grouping, two different levels.

20 Assuming the previous figurative identification as a starting point, it is possible to annotate the BT audio signal, by storing for each flow or object its attack and its duration as a textual information. Collecting the annotated data, a visual “figurative” score can be automatically generated[18], such as the one shown in **Figure 1**[19].

Figure 1. A figurative score for BT starting from annotated figures in the audio signal



21 Figures are divided into the two groups from top to bottom.

22 Starting from the figurative score, some considerations can be drawn:

- Group I is organized temporally as a sort of spring, as the flows first dilate and then shrink in duration, until the final compression.
- Cymbalom and Voice are almost strictly synchronized;
- Frog, except for a curious initial (and maybe dubious) appearance in *ppp*, is functional to dilatation, since it is presented at the peak of dilatation with a long sequence, then it shrinks;
- Group II is introduced very progressively, and, quite differently from Group I, is subdued to a process of thickening;
- Alarm works as a sort of marker for this general acceleration, while Glock, GlockB and Doorbell work mostly as a punctuation for Song;
- the sum of the two processes results in a general increasing of density with a synchronous climax around 11'20";
- repetitions of Song are presented as a final coda in solo;
- before the end, a last, isolated, Trumpet event takes place;
- with respect to this global trend in two aligning processes, Breath has an ambiguous status because, while it belongs to Group I (it is already presented at 3'), nonetheless it contributes to the thickening process of Group II. It could be seen as a mediating object, which holds the two processes together.

23 BT's overall organization is therefore firmly non-architectural, as it is not based on the presence of sections. Eventually, some figures may work as markers for certain phases of the processes (one may think about the presentation of Frog at about 3'40"), but nothing more. Thus, it is the notion of process that describes the overall configuration style of the piece. Such a procedural feature is crucial and it also emerges in details that cannot be revealed by the low temporal resolution of the annotated score of Figure 1 (but see Figure 2). Each flow in Group I is in fact a process itself, with its own internal articulation. The density of events in each flow is very high, and events are usually structured in their flow by means of a temporal grid that brings out metric regularities. This valuable and almost constant density is able to partly mask the overall configuration of the piece, as it forces the listener's perception to focus on a short time span. Moreover, BT shows some specific local configurations. The most striking is the alternation of Trumpet and Cymbalom (and/or Voice). This alternation can be seen as a sort of syntagma, still undergoing to a unique temporal manipulation. While the couple Cymbalom/Voice initially acts as a short tail to Trumpet, during the following presentations the process is reversed, so that Trumpet is shortened and Cymbalom/Voice stretched. This process is apparent at the beginning of Figure 1. Also, when Trumpet occurs again (just before 10') the same syntagmatic organization is respected, this time in the context of a general temporal compression.

24 Thus, in BT the "form" is the result of a recursive organization: a set of processes (e.g. the flows) organized into processes (e.g. the Trumpet/Cymbalom syntagma) organized into processes (the spring or the thickening in the two groups) organized into an overall process (that synchronizes Group I and II).

Sound test

25 Back to the quest for models of hermeneutical distance, one could hypothesize an ontological distance, in which the human interpreter is replaced by a machine. The model is that of blood test, in which data are extracted from a blood sample, automatically analyzed, and the results are made available to the human specialist[20]. Thus, we can think of an analytical machine that performs a "sound test" by automatically extracting data from the audio signal. By working according to its technical specifications, indeed such a machine implements hermeneutical criteria, whose definition is still due to the interpreter: but from the methodological point of view, the fundamental difference is that such a device does not inadvertently introduce ad hoc criteria during the analysis process.

26 For Bregman, streaming starts from the acoustic datum converted by the auditory apparatus into frequency information (in particular from the basilar membrane in the organ of Corti). This information is then grouped by cognitive processes into streams that represent the acoustic behavior of various sources. Physiology is not relevant here (and it is not relevant even in Bregman's theoretical framework, if not as a background): my hypothesis will be to extract relevant information in relation to phenomenological appreciation. In fact, frequency is not a perceptual entity, rather a theoretical notion defined by acoustics. Perception is related to pitches. A very general fact is that pitches are arranged along an axis (from low to high, regardless of orientation) and organized by a module (the octave). This starting point is codified in different cultures in the form of a variable punctuation (e.g. various scales or gamuts), which determines the positions of pitches and their number in the octave module. Pitch perception is rooted in the tonotopic functioning of the basilar membrane, a physiological datum that is evolutionarily very ancient, as it has been already found in fishes[21]. Very generally, it simply means that frequency detection in the ear is organized as in pitch appreciation (regardless of scales etc, of course).

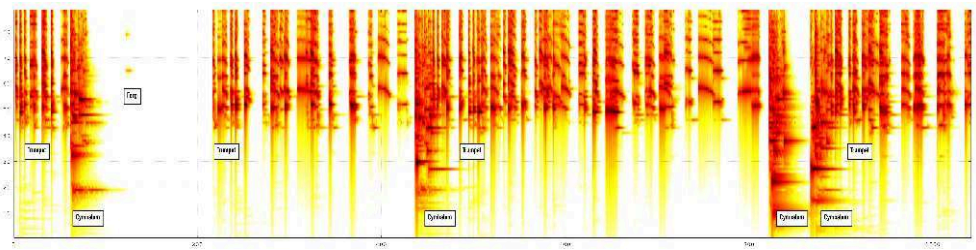
27 In short, I will start from two assumptions:

1. auditory streams are organized starting from the information converted into the frequency domain by the ear. This conversion occurs not in terms of linear frequency, but tonotopically, that is, even if very roughly, by pitch;
2. vice versa, pitch relations are perceptually basic and accessible to the subject (they are common currency for all musicians) and are based on the tonotopic organization of the ear.

28 It is thus possible to build a simple machine that converts the audio information into the pitch domain, assuming that, while very far from the complexity of the real physiological behavior, it is not inadequate, at least in very general terms, to represent listening data. Such a design is intended to provide an easy to understand analytic device, as frequency is represented in terms of musical pitch. This simple machine can be thought as a "cartoonified" model[22] of the ear with respect to the perception of pitches, abstracting from actual physiological operations while focusing on information

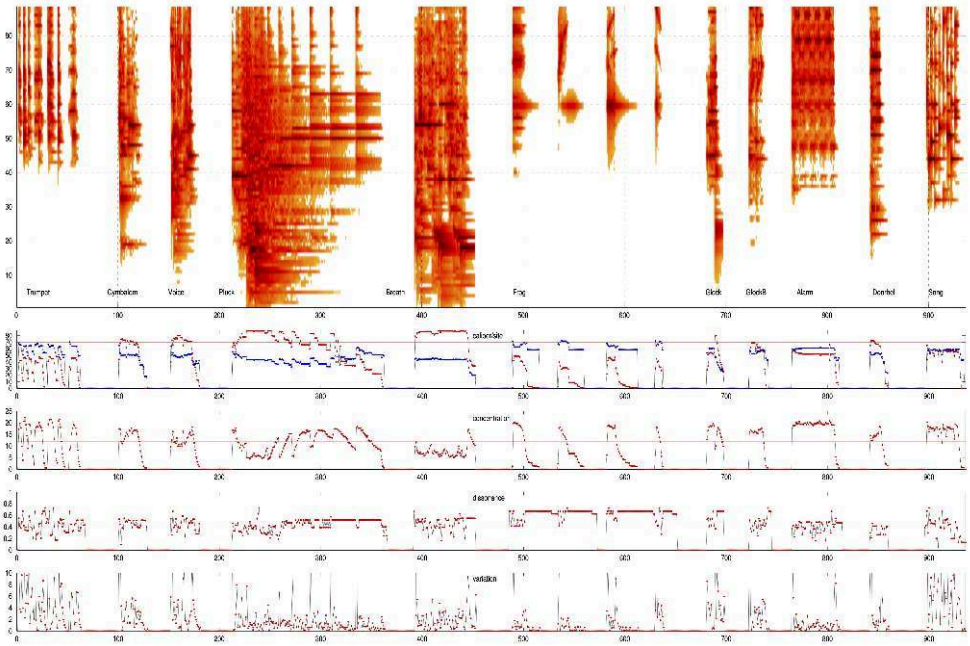
accessible outside, that is, perceptually. As a simple working hypothesis, I will then assume the Western equal temperament -and in particular the piano keyboard- as a grid to anchor pitches. Therefore, the axis of pitches will be subdivided into 88 discrete points, from the lowest A up to the highest C♯ on a piano. Such a grid offers reasonable resolution and range. In short, the analytic device describes what would happen if the sound would pass through the piano. Or, to put it in another way, as if it were to be approximated by playing on the piano a certain number of keys with a weighted pressure. During the analysis, the machine samples the sound at a regular rate. At each sampling interval, it provides in output, for each pitch, an estimation of its dynamics[23]. This vector of 88 elements can be thought as a type of spectrum. A spectral analyzer of this type therefore takes in input the signal and outputs a matrix, in which the columns represent the 88 dynamic values for the pitch grid in a sample, and the rows the sequence of dynamic values for a pitch over time. The matrix, depicting the behavior of sound as decomposed in pitches during time, is a form of sonogram, and can be plotted automatically for further explorations. Figure 2 represents a sonogram of BT's beginning, obtained with a rate of 20 samples per second. The x-axis represents time in number of samples and the y-axis the 88 pitches, starting from the lowest one (0 = 27.5 Hertz, low A on piano). Dynamic values are displayed by the so called "blackbody radiation" colormap that goes from white (lowest, i.e. absence of energy) to yellow (low), red (high), up to black (highest)[24]. The excerpt in **Figure 2** contains the first 53 seconds, and can be easily compared to the first minute of Figure 1. At the beginning (samples 0-200), it is possible to notice Trumpet, Cymbalom and the first appearance of Frog in pianissimo. Then, other presentations of the first two flows follow. Not taking into account Frog, the difference between Trumpet and Cymbalom is visually evident, in terms of average register (high vs. low), spectrum occupation (reduced vs. extended) and internal articulation (segmented vs. continuous). As a side note, the temporal resolution of Figure 2 allows to bring out in detail that continuous, internal procedural articulation of flows that I have discussed above. Starting from these analytical data, it is thus possible to propose some indices to be extracted automatically, so to verify their relevance in relation to the previous figure-based considerations.

Figure 2. Pitch-based spectral representation of BT, first 53 seconds



29 The first step consists in finding good representatives of the identified figures: here, "good" indicates a compromise between class representativeness and a technical constraint, the acoustic isolation from the other figures to obtain a "clean" extraction. A set of 11 figures are retrieved from BT's audio signal, placed in sequence and the resulting signal is then analyzed (**Figure 3**).

Figure 3. Pitch-based spectral representation of 11 figures, and relative indices



30 The second methodological hypothesis is to filter the spectrum. In Figure 2 it can be easily noticed how the energy is concentrated in certain areas of the sonogram, which obviously have a greater perceptual weight (in terms of perceived volume). If white indicates silence, then a yellow area is a sort of halo surrounding more relevant dynamic regions. After choosing a dynamic threshold, all the samples having a lower dynamic are thus reduced to zero, while those having a higher one are kept unchanged. Figure 3, top, shows the filtered sonogram of 11 labeled figures (note that Frog is represented by several emissions). This filtering operation carves out from the sonographic representation a set of visually representative shapes. How to characterize them further? In the following I am proposing various indices to be extracted automatically from the spectral matrix:

1. Pierre Schaeffer has proposed the notion of "mass" as a phenomenological generalization of pitch[25]. The mass of a sound is the way in which the pitch axis is occupied. Mass can be described by two different features. The first is "caliber". Caliber indicates how much the axis is occupied: a sound with a recognizable pitch has a reduced caliber, while a noisy sound has a relevant one. The second feature is "site", that can be thought as a position on the pitch axis. It corresponds to pitch in a sound with reduced caliber, while, with the progressive increase of the latter, it tends to indicate its mean value (noisy sounds do not allow to recognize a pitch, rather they can be characterized by a register, such as low, middle, high). It is not at all obvious to provide a formal definition of these two phenomenological parameters so that they can be implemented computationally. Without claiming a faithful translation of Schaeffer's proposals, I will therefore define the caliber as the maximum distance between pitches with non-zeroed dynamics in the filtered spectrum and subsequently the site as the average value between upper and lower values[26]. In short, in the sonogram the caliber of a sound is defined by its upper and lower dynamic boundaries, and the site is the average contour. Intuitively, given a filtered sample with energy spread between pitches 60 and 80, the caliber will be 20 and the site 70[27]. In Figure 3, the caliber/site curve plots the behavior of caliber (in red) and site (in blue) for all the figures. In the case of Breath, the caliber practically coincides with pitch axis and the site is thus simply its

midpoint. Instead, in Trumpet the caliber is limited, and the site meaningfully indicates the acute register where Trumpet is placed.

2. A second index may take into account energetic concentration. In the filtered sonogram, it is apparent that, in each figure, energy can be concentrated into certain pitches (painted in dark brown or black) or scattered all over the spectrum. As a simple measure of this feature, for each sample an average dynamics is calculated as a reference point. Then, for every pitch in the sample, the absolute difference between the dynamics and the average value is calculated. The sum of all the absolute differences divided by 88 provides the average "concentration" for the sample. The higher concentration is, the more energy is assigned to positive and negative peaks with reference to average dynamics. This means that some pitches are prominent on a lower background, while others have (almost) no energy: the sound maybe pitched (if a single pitch is dominating) or "channelled"[28] (if more components can be heard, like in bell sound). Vice versa, if concentration is low, more energy is distributed among all pitches, and the sound will be noisier. In the concentration curve, the blown nature of Breath results in a low concentration, while the presence of a clear pitch in Frog is shown in the initial peak. Alarm, with its bi-tonal nature, displays a maximal constant concentration.
3. Another possible index is based on the calculation of the internal dissonance of a figure. The term "dissonance" is far from being unequivocal, since it includes psychoacoustic measures that however are compared with historical and acoustic data. In the proposed implementation, each interval (i.e. relation between tow pitches) from unison (1:1) to major seventh (1:11) is associated with an estimation of the dissonance curve proposed by Sethares[29] in relation to 12 equally tempered intervals, normalized between 0 and 1. In short, each interval is assigned a dissonance value. In the analysis, for each sample only the six pitches with higher dynamics are taken into account. Each interval in this 6-note chord receives its dissonance value, and the overall dissonance is normalized between 0 and 1. In Figure 3, the dissonance curve shows that the dissonance index has a minimum value for Doorbell and a maximum one for Frog (a bit surprisingly, given its clear pitch). The variation of dissonance over time may also reveal interesting aspects. The maximum variation is associated to Trumpet and Song: indeed, they are in some way sequences of chords of variable composition (that is, with a variable dissonance).
4. Finally, the last index can take onto account spectral variation. The latter is calculated through the absolute difference between the dynamics of each pitch in a sample and the one relative to the same pitch in the previous sample. This means that a perfectly constant sound over time will have zero variation. In Figure 3, it can be seen that the curve is punctuated by peaks related to the attack of events: this is because between a sample representing silence (whose amplitudes are zero) and the one containing the beginning of the sound, there will be a great difference. This consideration is less obvious than what it may initially appear: rather, it indicates that all figures share a somehow percussive nature, that is, a clear attack. By contrast, if a sound starts in crescendo, there would be no variation peaks, as the difference with the preceding silence would not be so relevant. Therefore, this general behavior of all sounds underlines an important temporal feature, the relevance of "inchoativity"[30], a semantic category indicating the beginning. In relation to spectral variation, Trumpet and Song are the most variable, reflecting their nature of composite sounds including various pitches, while Pluck, but also Alarm, are the most static/stable.

31 If we consider the discussed indices, both in terms of average value and variation over time, then we can observe how the basic figures are structurally differentiated at least on the basis of an index. With two interesting exceptions, one for each Group. Among Group II (objects), Glock and GlockB share, at least in the analytical terms expressed by our machine, the same features: it has already been noted that they are, in fact,

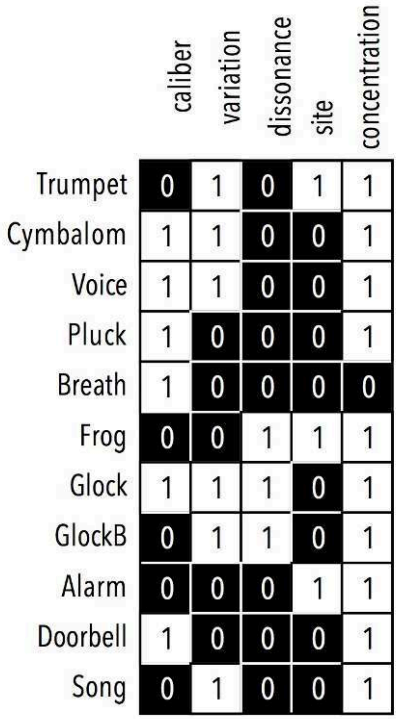
variants of a single type. In Group I (flows), instead, two examples are analytically indistinguishable, Cymbalom and Voice. Again, their close kinship was already observed while discussing Figure 1, in which they were related by parallelism or substitution.

32 If we consider the 11 figures as a microcorpus, then we can think of using indices to explicitly construct a differential structure. Identity may then be defined in a structuralist fashion, as a set of mutual differences inside a certain corpus. As a proof of concept, for each index a threshold value has been defined (see the horizontal lines in Figure 3). A binary value is assigned to each figure depending on whether the threshold is exceeded or not. This theoretical move is inspired by the classic approach by Jakobson to acoustic phonetics, that, starting from sonographic continuous representations, led him to identify phonemes as sets of binary acoustic features[31]. In short, I will thus take the previously introduced indices as the basis to construct categories, so that each category, e.g. dissonance, can receive only two values (e.g. dissonant vs. consonant). Of course, this categorical approached is tuned on the corpus but still depends on gathered data.

33 In our case, assuming that the value above the threshold is represented by 1 and the lower one by 0, we obtain a differential matrix like the one in Figure 4. The structural matrix indicates again that Cymbalom and Voice are (at least in terms of the proposed categories) structurally identical. On the other side, it is possible to individuate a difference between Glock and GlockB in relation to the caliber threshold. Thinking in terms of information theory, only 4 bits (that is: 4 categories) would be necessary to represent 11 figures ($2_{\text{values}}{}^{4\text{categories}} = 16$ values), but we need at least 5 to express all the figures in a differential way. It is like a larger structural palette ($2^5 = 32$ possible values) is prepared by the author, and only some elements are used, so that differences among the selected sound materials can be amplified.

34 Summarizing, in the selected microcorpus of 11 figures we note a common feature for all materials, the inchoative dimension that underlines their discreteness, while their identity can be effectively expressed structurally by a binary differential grid built from a set of indices.

Figure 4. Structural matrix for the figures according to index thresholds



Sounds from the outer space

35 There is a third model of hermeneutical distance, so remote that it that calls into question the aliens. The Voyager program has planned the launch of two spacecrafts, Voyager 1 and Voyager 2[32]. Each of them carried a phonograph record, the Voyager Golden Record[33], which aims to communicate the fundamental aspects of human life to an extraterrestrial intelligence. Carl Sagan, scientific director of the project, assumed that the extraterrestrial intelligence that could intercept it should be able to read the signal and correctly extract the information from the groove. The outer surface of the disk (properly, its cover) carries a set of instructions for decoding. Sagan's assumption (or hope) is therefore that there is at least a common "form of sense" between humans and aliens, in whatever way they can perceive it, a sort of truly universal community of intellect, linked to life rather than to culture. The content of the Golden Disc is made up of images, sounds and texts: in other words, a finished corpus of texts that can be used systematically by the aliens to define their interpretations. This discreteness of the corpus makes the alien hermeneutic situation a lucky one, because it clearly defines the context for all the possible messages in the record.

36 From the perspective of a now sidereal distance, and to limit the information to a minimum, one can then suppose to receive BT on a golden record, which contains only one other document. The other available document is a written text, indexed as "Burger Time Nota Programma". It declares the author ("Mauro Lanza") and the title of the piece ("Burger Time ou les temptations de saint Antoine"). These are followed by an aggregate of quotations from four sources, reported in footnotes, in three different

languages (by the way, three languages exactly as in the message recovered by Lord Glenarvan): English, French and Italian.

37 First of all, the title, *Burger Time ou les tentations de saint Antoine*, is double. By means of the conjunction, it declares two elements related by an analogy. The first is the title of an "arcade game", "Burger Time", the second is in turn double. On the one hand, it refers to a theme, a narrative and figurative *topos* typical of the "Moyen Âge fantastique", on the other it refers to a reprise of the same *topos* by Gustave Flaubert. The quotations are therefore distributed, in an interlaced and perfectly uniform way, between these two elements: four quotations related to Burger Time (from two sources) and four quotations for the temptations, two related to the medieval and gothic *topos* (from Baltrusaitis), and two from Flaubert. Thus, Burger Time, a videogame, and the temptations of Saint Anthony, a theological, but more generally cultural, *topos* are related: the analogy is developed in the reported fragments.

38 It is possible to identify a set of semantic planes that intersect the fragments. Such planes in semiotics are called "isotopies"[34] and their aim is to redundantly distribute the sense in a text to support its internal cohesion. The following isotopies are shared among the fragments.

39 The first is that of revolt. It takes two forms: one, dominant, of a figurative type, in which objects hurl themselves against the man, be them "the food that does not want to be eaten" or the "inorganic bodies of nature". The second version of revolt is more explicitly theological: now it is Satan who raises himself against God ("che si solleva contro Dio"). It can be noted that this isotopy has an aspectual dimension, that is, it is a process. This process finds its starting point in a call: such a call is developed through the isotopy of temptation ("tentation"), which shifts the presence of evil from a figurative dimension to a cognitive one. In the case of Burger Time, this temptation takes the form of a "curious dilemma" that concerns the behavior of objects, a sort of counterfactual conditional ("What if") that becomes true in what is no longer just a *Gedankenexperiment* but, rather, an actual struggle (the game). Yet this temptation is also explicitly invoked by the subject ("Je l'ai appelé"). A third isotopy, abundantly represented, is the one related to the monstrous. Monstrosity takes the form of a hyperbolic amplification: the gigantic (the antagonists of Burger Time are "gigantic foodstuffs"), the deformed ("couverts d'écailles", for example), the multiple ("les Dieux[...] qui ont plusieurs têtes). The multiple in itself is further distributed among the fragments: the monstrous is multiplied in a variety of heterogeneous elements ("di diversissima provenienza") that are opposed to the solitude of the anchorite. The isotopy of transformation is linked to the previous one but partly independent: in other words, the monstrous is the result of a sort of figurative distortion, in which the ontological frontiers between worlds collapse. The cause is attributed again partly to the subject through the "nightmare", the "fantasy", the "phantasmagoria" and the "hallucination" (as said in Italian by Baltrusaitis). Another isotopy is related to food: it structurally concerns Burger Time (the "fast food nightmare"), but it is also active when the Buddha remembers that he spent "douze ans à me nourrir de parfums". Some results of an analysis in terms of isotopies are collected in Appendix.

40 These isotopies obviously intersect constantly. But they can be integrated by relating them to an aspectual dimension: all the elements can be collected and ordered with respect to the beginning/end axis. Thus, the "game" takes the form of a "nightmare" or a "phantasmagoria" that the subject evokes ("je l'ai appelé") or which he falls into (the

case both of Peter Pepper and the anchorite): an evocation of monstrosity. Then, a hunt follows ("chasing Peter Pepper") or a fight (the monsters "spy, chase, attack") in which the subject must resist against the monstrous ("evil must be defeated") and reorder the world ("all your burgers dressed"). Subsequently, the subject undergoes a transformation through death ("Chef is sacrificed") and finally a "process of resurrection" or transfiguration ("Je devins le Bouddha!"). On closer inspection, this process is a classical narrative articulation, which Greimas has described through four phases[35]. On the cognitive side, the manipulation and sanction (respectively the evocation of the nightmarish phantasmagoria and the acknowledgment of victory, both in charge of the subject even in Burger time in which "You controlled Peter Pepper") are the beginning and ending phases. They surround a pragmatic narrative core, made up of competence and performance (respectively, the acquisition of skills, e.g. the use of the salt shaker, and the struggle).

41 In short, assuming an integrating perspective based on the redundancy of distributed isotopies, the fragmented aggregate seems to set up a classical narrative form. But if the narrative form that can be reconstructed from the program note is classical, nevertheless it is disturbed by the perturbing figurative elements (monstrosity, that turns towards grotesque in Burger Time) and by the enunciative work that deliberately collides different, heterogeneous sources in a collage without attempting any harmonization.

42 Which relationships between the two documents stored on our golden record? In other words, how does BT stand as an expression of a certain content as developed by the program notes?

43 As a conclusion of this essay on distance, I will thus propose four hypotheses for a crossover interpretation of our minimal corpus of two texts.

The monstrous

44 Far from presenting itself as an abstract sound meditation, BT proposes a sort of imaginary and somehow "teratomorphic" instrumentality: just think of Frog, halfway between human vocalization and croaking. The phantasmagoria of temptations is thus translated into a monstrous orchestra, a sort of Bachtinian grotesque sound body: it is in fact composed of strange sound sources but also of alarms and toys (think of toy trumpet), it includes stringed instruments that collapse onto themselves and sound devices that accelerate as if they contract spasmodically. The electronic medium therefore serves the composer to stage a revolt of instruments through their monstrous presences and their internal transformations.

The congeries

45 In BT each flow does not describe the sound emission of a single sound source, rather it depicts a population of actors caught in their multifarious hectic productions. BT therefore exhibits a feature that could be labeled as "congeries". I have already discussed the effect of progressive agglomeration, which thickens and concentrates the figures along the piece, up to the final diminuendo, where Song acts as an isolated closing statement. All along BT, the sequences of acoustic events have such a local density that they menace the perception of the overall process. As in the game Burger

Time, in BT the figures "would stop at nothing", like the pursuers of Peter Pepper. There is thus a sort of acoustic *horror vacui*, tempered only in short moments, according to the model of the demonic assault and the hectic chase in the grotesque fast food.

The narrative articulation

46 BT's annotated score reveals an overall process of increasing complexity and density. Without precisely individuating the proper elements of a narrative tension, such a process is clearly oriented towards an apex. After this climax, Song is presented[36]: BT's "teratophonic" theater ends with a grotesque Gestalt (the melody is obtained by a strange assembly of sounds). Yet its form -the melody- imposes itself to perception. The phantasmagorical vision comes to an end, perhaps with a transfiguration of the subject capable of winning the demons. And still, according to a circularity reminiscent of Peter Pepper's strange resurrection, a toy trumpet obnoxiously returns. Resurrection is the announcement of a cycle, of a new demonic invasion, of another match in the game.

The metalinguistic dimension

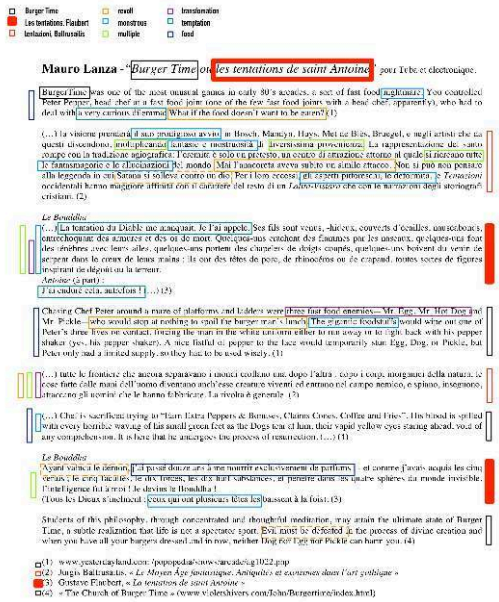
47 The program note is an aggregate that triggers a metalinguistic dimension both in its construction as a collage from multiple sources and in the direct quotation from Flaubert. I have already observed that the same logic, although partially homogenized acoustically, is active in BT through the opposition between the two groups of flows and objects. Objects, presented by ostension, work as citation signals (replicas) with respect to flows. If flows can be read as an example of the monstrous category, objects belong mostly to the grotesque, according respectively to the opposition between "les temptations de saint Antoine" and "Burger Time".

A tempting conclusion

48 There is a final element that can emerge by intertwining our two documents, and I propose it as a conclusion. It deals with temptation.

49 As the program note tells us, temptation is an invasion of the devil. Yet the Buddha demands it as necessary. The evocation of the monstrous, the trigger of the incoming nightmare, is therefore the only documentary indication concerning the subjectivity of production in our corpus. The phantasmagoric theater results from the evil-induced altered state of the hermit: he is properly the author of the nightmare, the one in charge of the hallucination. Said with a different flavor, he is the one that accepts to play Burger Time: "You controlled Peter Pepper". Thus, the program note proposes a "Model Author" that, before struggling with demonic entities, triggers himself a fantastic evocation and at the same times a game. Such a characterization of the Model Author may be projected onto BT. In that case, the authorial subjectivity (the work of the composer, "Mauro Lanza") can be thought both as an evocation of monsters and as a playful activity.

Appendix. Isotopic analysis of the program note



NOTES

1. Jacques Fontanille, "Sans titre... ou sans contenu ?", *Nouveaux Actes Sémiotiques*, 34-35-36 (*Approches sémiotiques sur Rothko*), 1994.

2. An account of the various interpretations, limited to sole notion of Naturalism, is given by Oscar Moro Abadía, Manuel R. González Morales & Eduardo Palacio Pérez, 'Naturalism' and the interpretation of cave art, *World Art*, 2, 2, 2012, pp. 219-240.

3. Ignacio de la Torre, "The origins of stone tool technology in Africa: a historical perspective", *Phil. Trans. R. Soc. B* 366, 2011, pp. 1028–1037.

4. Umberto Eco, *The Role of the Reader: Explorations in the Semiotics of Texts*, Indiana University Press, Bloomington, 1979.

5. Albert Bregman, *Auditory Scene Analysis. The Perceptual Organization of Sound*, The MIT Press, Cambridge, Mass. and London, 1990.

6. See William W. Gaver, "What in the world do we hear?: An ecological approach to auditory event perception", *Ecological psychology*, 5(1), 1993.

7. Bregman, *op. cit.*, p. 306.

8. Pierre Schaeffer, *Traité des objets musicaux,* Le Seuil, Paris, 1966. See also Michel Chion, *Guide des objets sonores. Pierre Schaeffer et la recherche musicale*, Buchet/Castel-INA, Paris, 1983. In the following discussion I will adhere to my interpretation of Schaeffer's theory from a semiotic perspective, described in "Towards a Semiotics of the Audible", *Signata*, 6, 2015, pp. 65-89.

9. From a physiological and cognitive perspective, Alain Berthoz has underlined the role of gesture in our comprehension of the world (both as an anticipation and as a program for action) (*La simplexité*, Odile Jacob, Paris, 2009, chap. 8). For a discussion of "figurative mechanics" see Valle, *op. cit.*, *passim*.

10. Bregman, *op. cit*, *passim*.

11. As an example, Jules Masson Mourey. *La figuration anthropomorphe dans l'art rupestre préhistorique et protohistorique de la région du mont Bego (Tende, Alpes-Maritimes) : modalités du schématisme, cadres chrono-culturels et pluralité des sens*", Mémoire de Master 2 en Archéologie et Histoire de l'Art spécialité Préhistoire, Aix-Marseille Université, 2016.

12. I am not using the term "stream" in order to reserve it for the general process of segregation in auditory scene analysis.

13. I am referring to the description of temporal profile in Schaeffer, *op. cit.* The term "eumorphous" for well-formed profiles is discussed in Valle, *op. cit*.

14. See Jean-Jacques Nattiez, *Musicologie générale et sémiologie*, Bourgois, Paris, 1987, chap. 4.

15. Umberto Eco, *A Theory of Semiotics*, Indiana University Press, Bloomington, 1975, chap. 3. See the discussion in Andrea Valle, "Modes of sign production", in Sarah G. Beardsworth and Randall E. Auxier (eds.), *The Philosophy of Umberto Eco,* Chicago, Open Court, 2017.

16. Barry Truax, *Acoustic Communication*, Ablex Publishing Corporation, Norwood, NJ, 1984, pp. 58-61.

17. Eco, *op. cit.*, p. 224-227.

18. Data have been annotated in the Reaper DAW https://www.reaper.fm/ exported and plotted via Nodebox https://www.nodebox.net/.

19. In the case of Song, sequences of repetitions, in particular the long event after 11', are indicated by a single graphic element.

20. Of course, in the previous model the distance was between the interpreter and the object: here it is between two types of interpreter, human and machine.

21. Arthur N. Popper and Richard R. Fray, "The Auditory Periphery in Fishes", in Richard R. Fray (ed.), Comparative hearing: Fish and Amphibians, Springer, New York, 1999, pp. 43-100. I am grateful to Prof. Geoffrey Manley for the indication.

22. The notion of cartoonification is introduced in relation to sound by Davide Rocchesso and Federico Fontana (eds.), *The Sounding Object*, Edizioni di Mondo Estremo, Firenze, 2003.

23. Technically, the implemented machine is a bank of filters centered on 88 frequencies. The amplitude is estimated in deciBel. The dB estimation is indeed a very rough approximation from the physiological point of view, but in this cartoonified framework offers a sufficient approximation for my analytical purposes. Technical details can be found here: http://www.musicaelettronica.it/spettri-fourier-e-il-geco-2/.

24. David Borland and Russell M. Taylor II, "Rainbow Color Map (Still) Considered Harmful", *IEEE Computer Graphics and Applications*, 27(2), 2007, pp. 14-17. Figure 2 and 3 have been obtained with GNU Octave: http://www.gnu.org/software/octave/.

25. Schaeffer, *op. cit.*, *passim*. See also Michel Chion, *Guide des objets sonores. Pierre Schaeffer et la recherche musicale*, Buchet/Castel-INA, Paris, 1983.

26. While the notion of caliber seems not particularly problematic, site, if so expressed, captures the notion of register but not that of exact pitch, except in the case of a sinusoidal sound.

27. Indeed, from a signal processing perspective, the notions share an analogy with bandpass characterization in terms of central frequency and bandwidth.

28. Schaeffer, *op. cit.*, *passim*., as translated by Christine North and John Dack, *Treatise on Musical Objects*, California University Press, Oakland, 2017.

29. William A. Sethares, *Tuning, Timbre, Spectrum, Scale*, Springer, London, 2005 (2nd ed), *passim* and p. 100 in particular for the curve.

30. Algirdas J. Greimas and Joseph Courtés, *Sémiotique. Dictionnaire raisonné de la théorie du langage,* Hachette, Paris, 1979, v. “Inchoativité”.
31. See Roman Jakobson, C.Gunnar M. Fant and Morris Halle, *Preliminaries to Speech Analysis. The Distinctive Features and their Correlates*, The MIT Press, Cambridge, Mass., 1952 and Roman Jakobson and Morris Halle, *Fundamentals of language*, Mouton, ‘S-Gravenhage, 1956. The same operation is at the base of Robert Cogan’s theory of sound color, *New Images of Musical Sound*, Harvard UP, Cambridge, Mass., 1984.
32. https://voyager.jpl.nasa.gov/mission/.
33. https://voyager.jpl.nasa.gov/golden-record/.
34. Greimas and Courtés, *op. cit.*, v. “Isotopie”.
35. Greimas and Courtés, *op. cit.*, v. “Narratif (schéma)”.
36. And, yes, it is a quote from the music of Burger Time, as all the objects are quotations from sound effects in the arcade game. But, no, we cannot know this information here.

ABSTRACTS

The article discusses *Burger Time ou les Tentations de Saint Antoine*, a 12-minute work of electronic music composed by Mauro Lanza. In particular, it proposes a three-layer interpretative methodology, focusing respectively on audible perception, on signal analysis and finally on a semiotic framework. In the first layer, the music work is discussed in terms of audible figures, and an annotated visual score is proposed as a result. In the second one, categories for signal analysis are introduced and the gathered data are used to build a structural matrix aiming at defining the identity of the previously identified figures. Finally, the third layer proposed an integrated semiotic interpretation by taking into account the previous layers and a textual document, in order to highlight the most relevant semantic “isotopies”.

L’article traite de *Burger Time ou les Tentations de Saint Antoine*, une œuvre de 12 minutes de musique électronique composée par Mauro Lanza. En particulier, il propose une méthodologie interprétative à trois couches, axée respectivement sur la perception audible, sur l’analyse de signal et enfin sur un cadre sémiotique. Dans la première couche, le travail musical est discuté en termes de figures audibles, et une partition visuelle annotée est proposée en conséquence. Dans la seconde, des catégories d’analyse de signal sont introduites et les données recueillies servent à construire une matrice structurelle visant à définir l’identité des figures précédemment identifiées. Enfin, la troisième couche proposait une interprétation sémiotique intégrée prenant en compte les couches précédentes et un document textuel, afin de mettre en évidence les « isotopies » sémantiques les plus pertinentes.

INDEX

motcle musique électroacoustique, sémiotique de l’audible, théorie des objets sonore, analyse de la scène auditive
Keywords: electroacoustic music, semiotics of the audible, theory of sound objects, auditory scene analysis

AUTHOR

ANDREA VALLE

Università di Torino