

FIA 2020/22

XII CONGRESSO/CONGRESO IBEROAMERICANO DE ACÚSTICA
XXIX ENCONTRO DA SOCIEDADE BRASILEIRA DE ACÚSTICA - SOBRAC

Florianópolis, SC, Brasil

The profile concept in Schaeffer's solfège: discussion of its diverse meanings and proposal of a quantitative real-time approach

Freire, S.^{1,2}; Padovani, J.H.^{1,3}; Campos, C.^{1,4}

¹ Laboratory for Performance with Interactive Systems (LaPIS), Escola de Música, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil

² sfreire@musica.ufmg.br, ³ jhp@ufmg.br, ⁴ costacaiocampos@gmail.com

Abstract

In his *Traité des Objets Musicaux* (1966), Schaeffer proposes a solfège dedicated to any sound potentially musical, built upon seven perceptual criteria: mass, dynamic, harmonic timbre, melodic profile, mass profile, grain, allure. The practice of this solfège relies on reduced listening, a deconditioning attitude that focuses on the intrinsic properties of the sounds and ignores their origins or meanings. In two recent works, we have proposed a quantitative real-time interactive approach of five perceptive criteria, excluding the profiles, based on the assumption that the use of low-level descriptors has similarities with the practice of reduced listening. Our next step was to quantitatively tackle the two remaining criteria after understanding the scope of the term profile in the Schaefferian theory. We aim to use this tool in pedagogical and creative situations. Schaeffer almost always associates the term profile with a temporal variation of at least one sonic feature. Also, he usually connects this term with a sound object bearing an optimal duration for the memorization. Moreover, this temporal variation is expected not to be perceived as cyclic, in which case he prefers the term allure. Profiles are associated not only with the main criteria (dynamic profile, harmonic profile, melodic profile, mass profile) but also with second-order variations such as the speeding up of an allure, the gradual granulation of a sound, or even a "profiled" profile. We have adapted and developed a series of time-varying low-level descriptors used for the characterization of these criteria: temporal envelope, number of grains and allures, number of partials responsible for 50% and 80% of the total energy, presence, and estimation of the fundamental frequency, spectral centroid, intrinsic dissonance, spectral distribution. In the present work, we present the correlations found between some selected descriptors and the different types of profiles.

Keywords: Treatise on Musical Objects, low-level audio descriptors, musical variation.

PACS: 43.66.Lj, 43.75.Cd, 43.75.St, 43.75.Zz.



1. INTRODUCTION

For those with a minimum acquaintance with the work of Pierre Schaeffer, it may sound odd a paper dedicated to a quantitative approach to the *profiles* and not to *sound objects*. In fact, the present work appears in tandem with a recently published article, in which we address this issue extensively [1]. Here we have the opportunity to deal in detail with a less well-known and discussed aspect of Schaeffer's theory.

The tools we have been developing from this research run in interactive real-time situations, with direct applications in musical pedagogy and creative processes involving sounds not perceptually dominated by pitch. Our focus is on simple objects and contexts with restricted sound palettes, in which different degrees of contrast and subtle differences between sounds are equally important. Therefore, our goal is to work with contextualized living data produced by musicians in their praxis.

The text begins with a summary of the *Traité des Objets Musicaux* (TOM)¹[2, 3], followed by the discussion of the use of the term profile made by Schaeffer in this book. We then introduce the framework in which this author considers profiles as musical *perceptual criteria*: the theory of variations. The next section briefly presents the computational tools used in our quantitative analysis. After depicting a few examples, we conclude with current and potential uses of this association between digital signal processing and the Schaefferian *sofège*.

2. THE TREATISE ON MUSICAL OBJECTS

In 1966, Pierre Schaeffer published the *Traité des Objets Musicaux* (TOM), in which he developed the concept of sound objects. His starting point was the observation that the traditional musical theory was no more capable of dealing with the musical reality, be it the contemporary production, the adoption of new techniques, or the contact with other musical cultures. He questions the idea of a musical note still being the source of musical structures, proposing an approach re-

lying on sound objects.

Schaeffer's definition of a sound object grows up from the phenomenology (from the concept of *époche*) and from the new possibilities of listening introduced by the recording techniques, mainly the repeated reproduction of the same event. He proposes the practice of *reduced listening*: "having ignored the source and the meaning, we perceive the sound object". [3, p. 115].

In the search for a *typology* (identification of sounds) and a *morphology* (their description), he starts from two poles of the sound object, one more abstract, the other more concrete: the poles of *values* (description) and *facture* (identification). He further develops the typology relying on the criteria pair *articulation-stress* (the same as *sustainment-intonation*), while the morphology explores the pair *matter-form*. He gives us a hint of what these latter concepts imply:

Imagine that we can 'stop' a sound to hear what it is like, at a given moment in our listening: what we can grasp now is what we will call its *matter*, complex, fixed in the tessitura and in subtle relationships with the sound contexture. Now listen to the history of the sound: we become aware of the evolution in duration of what had for an instant been fixed, *a journey in time that shapes this matter*. [3, p. 317]

The *sustainment* concept covers concrete aspects of matter and form, "closely linked to the energetic history that relates the origin of *every moment of the sound*." [3, p. 437] At last, he introduces the discussion of variations of the former criteria.

On the perceptual side, Schaeffer proposes the *criterion-dimension* pair, "the one that ultimately gives a meaning to the analysis of objects: the meaning of its musical proportions." [3, p. 300] In the Summary Diagram, he depicts the crossing of seven criteria of musical perception with three perceptive fields. The seven criteria are *mass*, *dynamic*, *harmonic timbre*, *melodic profile*, *mass profile*, *allure*, and *grain*. The *perceptual fields* are related to *pitch*, *intensity*, and *duration*. It is

¹All quotations used in this paper are from the 2017 English translation.

important to note that, in Schaeffer's theory, the *perceptual fields* are not the equivalent of the physical/acoustic parameters of sound:

Yet we must repeat that the perceptual dimensions, even if they are reduced to the three types of evaluation — pitch, duration, and intensity — do not coincide either fundamentally or numerically with the reference trihedron of acoustics: frequency, time, and level. [3, p. 402]

The dimensions can be explored in continuous or discontinuous contexts. For the discontinuous cases, the terms *site* and *module* are employed; for the continuous, *caliber* and *impact*. His utopian goal was that particular collections of sounds ("suitable objects"), perceivable through similar criteria, would permit to order ("putting them roughly into a series") or locate these sounds "in a calibration with relationships that are cardinal, not just ordinal, and even arranged in their field in the form of vectors." [3, p. 297]

In short, this seminal work depicts a clear framework (a *sofège*) for tackling different sound *typologies* (types) and *morphologies* (classes), accompanied by less developed ideas for the *characterology* (genres) and species of sounds. These concepts are well fitted to contexts where the note (and its fixed pitch) is not the most prominent musical feature or value.

3. PROFILES

Besides naming two perceptual criteria discussed in the last section, the term *profile* appears throughout the TOM in diverse contexts and with different meanings. In most cases, it is related to the dynamic curve or to the envelope of a sound object. On several occasions, the dynamic profile associates with the harmonic profile, where a distinction between linked profiles — as in the case of attack-resonance sounds — and independent profiles — for sustained sounds — is stressed. In one figure in section 31.4 there is one occurrence of this term bearing a specific link to attacks [3, p. 430].

Aside from the general idea of an envelope, we also find a more restricted use of the term profile, linked to a crucial category in Schaeffer typology: the *well-formed sounds*. These are "sounds that have an optimal duration for memorization by the ear" or "sounds with a form (or profile)." [3, p. 423]

This term is also used in the sections discussing variation. In these sections, the connotations of the term are not, however, univocal. Two of the seven main perceptual criteria bear the name profile: namely *melodic profile* and *mass profile*. In these cases, the term indicates a first-order, perceptually pregnant variation. On the other hand, the 'profile' also points to a particular pace of the variation of information ("medium variation" [3, p. 452]), among other possibilities (*progression* and *anamorphoses*).

In all these cases, it is possible to associate the term 'profile' with a temporal variation of at least one sonic feature. Moreover, this temporal variation is expected not to be perceived as cyclic, in which case Schaeffer prefers the term *allure*. The *profile* concept also covers second-order variations, such as the speeding up of an *allure* or the gradual granulation of a sound.

We still encounter a very different meaning for this term, used as the "profile of the texture of the mass". This use, contrasted to the mass profile, is not directly linked to a temporal variation, but to a "vertical" (spectral) aspect of the sound object.

The mass profile [or better, the profile of the texture of the mass] is made up of all the (perceived) intensities of the various components of the spectrum of a sound. It is a more or less overall or detailed perception of the components and consequently can only be very roughly deduced from the acoustic spectrum or Fletcher's curves. [3, p. 433]

In the *Summary Diagram* [3, p. 464-467], it is possible to find all these meanings in one place. In the first column (not numbered), dedicated to the criteria of musical perception, in rows 4 and 5, appear the *melodic* and *mass* profiles respectively.



In these rows, it is also possible to find similar uses of the term in cells 4.3, 4.4, 4.6, 4.7, 5.6, and 5.7.

In the intersection of the first column (*types*) with the fourth row (*melodic profile*), we find the term used with the meaning of a moderate variation speed (in contrast to progression and anamorphoses). In the intersection of the second column (*classes*) with the second row (*dynamic*), the term 'profile' refers to the envelope of a well-formed sound. Cells 2.6, 2.7, and 2.8 also refer to this same meaning. Finally, in cell 1.7, we encounter the expression "profile of the texture of mass".

4. THE SOLFÈGE OF VARIATIONS

In chapter 33 of the TOM, dedicated to the *Theory of Variations*, after digressing about continuous and discontinuous aspects of the musical perception, Schaeffer depicts a tentative typology of musical variations upon two axes. One axis handles the density of information of *variations* (*progression*, *profile*, and *anamorphoses*). The other one deals with the *factures* (*fluctuation*, *evolution*, *modulation*). These two axes thus grade in three general stages both the density/rate of information of a given variation, as well as what concretely gives constitutive existence (*facture*) to such a variation.

Michel Chion summarizes the concepts from which Schaeffer seeks to qualify the density/speed of information and the invoicing of variations in the following way:

A variation can be discerned on the one hand by its *facture*, i.e. the manner in which it is executed, and on the other hand by its distinctive "speed", its "density", i.e. the relationship between the variation and its duration.

a) We shall therefore consider whether the *facture* of variation is characterised:

- by a fairly noticeable instability within an essentially stable process (variation of the *Fluctuation* type);

- by a continuous and progressive transformation (*Evolution* type);
- in a "scalar" manner, i.e. by discontinuous stages (*Modulation* type).

b) On the other hand, there are three types of density:

- either the speed of variation is slow, with little density, we shall call this *progress*;
- or its speed is of medium density, and this is *profile*;
- or it is rapid and will be referred to as *anamorphosis*.

These three types of densities together with the three main types of *facture* will allow the principal types of varying objects to be classified. [4, 5, p. 76; p. 82]

Further, Schaeffer applies this classification to traditional and experimental sound objects, focusing on variations on the elementary criteria of musical perception: *matter*, *form*, and *sustainment*. The variations of *matter* deserve two new perceptual criteria: *melodic* and *mass* profiles. According to Chion, melodic profile is:

a criterion applied to varying sounds, and describing a *variation which affects the whole mass of the sound, making it describe a sort of "trajectory" in the tessitura*. As opposed to mass profile, which refers to an internal variation of mass, melodic profile is a displacement of the entire sound in the pitch-field: it is the sound itself which moves, instead of being sculpted by an internal development. [4, 5, p. 162; p. 183]

On the other hand, mass profile is:

a morphological criterion describing an *internal* variation of the sound mass which is, as it were, 'sculpted' in the

course of its development, by modifications which cause it to become thicker, thinner etc. E.g. a tonic sound developing into a thick sound of complex mass. By contrast, melodic profile is the overall trajectory of the sound mass moving in the tessitura.[4, 5, p. 164; p. 185]

The *harmonic profile* is never handled as an independent criterion in the TOM, being discussed in connection with other criteria (mostly the *dynamic*). The temporal form is already a variation on its own, and its second-order fluctuations are the subject of the criterion *allure*. Still, an *allure* variation may also distinguish a sound object through speeding up or deepening. The other *sustainment* criterion, the *grain*, may also present similar temporal variations.

In real-life sounds (or “evolving sounds”), “presenting in its texture a summation of articulations and stresses varying both in matter and in form” [3, p. 319], it is hard to isolate any criteria. This fact poses a real challenge to a quantitative (and real-time) approach like the one we describe in the next section: in which cases a detected variation should be characterized as a *melodic*, *mass*, or *harmonic* profile? How to differentiate between different dynamic profiles? How to identify and qualify allures, grains, and their variation in time? How interconnected may these variations appear? The use of thresholds and filters is practically unavoidable here.

5. ANALYZING PROFILES IN REAL TIME

In the last years, we have implemented real-time tools to approximate the Schaefferian criteria and low-level audio descriptors, using the Max programming environment². These tools are the main subject of a paper already mentioned in the Introduction[citar Vortex]. The audio descriptors divide into two main groups, linked to time and frequency representations. The time-domain descriptors are attack profile, envelope, grains (tiny and iterative), and allures. Most spectral descriptors rely on the object *sigmund*, programmed by M. Puckette, which generates a list of spectral peaks and correspondent amplitudes, besides

estimating the fundamental frequency (if one is detected). From this data, we calculate the number of spectral components necessary to express 50% and 80% of the total energy; how much energy is carried by 20 (or less) spectral peaks; the most prominent peak; fundamental frequency; intrinsic dissonance; spectral centroid; the difference between highest and lowest peak; spectral region. The spectral region returns an integer value, useful for short sounds. For longer sounds, it is preferable to use a spectral distribution in three ranges (low, mid, high), whose limits the musician can vary. This distribution depicts percentages of the total energy.

Most of these descriptors are time series sampled at every 512 audio samples (implying an interval of 10.67 ms at a 48 kHz sampling frequency). Due to its importance for percussive sounds, the attack profile (the first 400 ms or less after the onset) counts with a higher (8x) resolution. Two descriptors also generate a binary time series, where 1 means an occurrence and 0 none during the time frame: pitched/unpitched, allure/no allure. The iterative grains also produce a similar curve, with the number of occurrences (if any) in each frame. There is a configurable display to plot a selection of curves during the performance, offering a valuable visual feedback for the musical practice.

All curves are estimated between onset and offset; next they are submitted to simple statistical analysis. We calculate the average and standard deviation; temporal centroid and spread; skewness; kurtosis; peakness; flatness. We also extract values for the curve span (Δ max-min), and a 10-point average for the initial, medium and end sections.

The analysis of allures is mainly directed to the envelope curve but can also be applied to other curves when undulations or zigzagging occur. The algorithm searches for the appearance of bends surpassing a given amplitude threshold. If a bending in the opposite direction follows, we estimate and validate a depth value. In addition to the binary time series and the number of occurrences during a sound event, the descriptor for allures produces sets of values for size (depth); the time interval between peaks, peaks to valleys,

²www.cycling74.com



and vice-versa; sharpness of the bending. From this data, it is possible to infer the regularity and impact of the variation of the selected parameter.

Figure 1 depicts the pct80 (number of partials needed to express 80% of the total energy) curve of a double-bass sound produced as a transition from regular bowing to overpressure. We applied a 5-point moving average filter to this curve. Table 1 shows a selection of parameters estimated for this sound. Here we have a case of a mass profile, where a tonic (pitched) sound transforms into a complex one. In these cases, it is relevant to observe, among other descriptors, the variation of the number of partials needed to express 80% of its energy. The curve presents a high value for standard deviation, the first clue to the presence of a profile. The values estimated for the temporal centroid, skewness, span, and initial-medium-end values indicate a curve with an ascending character. In addition, the allure analysis points to a not very smooth curve that oscillates (according to the given threshold) 10 times. These oscillations are not regular regarding depth and rate, as the large standard deviations for size and interval demonstrate. Finally, the values estimated for the binary time series also indicate that the oscillations occur more likely around the middle and end portions of the sound.

Our tools are not intended to work blindly on any selection of sounds (although they could be programmed to function in this way) but, on the contrary, to support the musical practice in specific contexts. Therefore, we try to select the most significant descriptors according to the sound set chosen for each occasion. As stated by Schaeffer, “the application of the major distinctions in our typomorphology will depend in practice on particular applications and the listening criteria at the time.” [3, p. 371] These numerical values feed the input of a machine learning algorithm, which returns categorical labels when needed.

6. A FEW EXAMPLES

Musical variation, according to Schaeffer, is a nearly uncharted territory once we move away from its most common types: glissandos and dynamic profiles. One of his remarks mixes advice and warning: “the domain of combined varia-

Table 1: Data from statistical analysis of the pct80 curve shown in Figure 1.

avg / stdev		10.3 ± 5.9
temp. centr.		0.63
temp. spread		15.22
skewness		-0.01
Δ max-min		16.8
value at start		2.7
value at middle		14.9
value at end		11.6
allures	number	10
	size	5 ± 3.6
	p_p interval (ms)	288 ± 359
binary time series	temp. centr.	0.61
	skewness	-0.12

tions is intractably complex.” [3, p. 451]

In our study, we are just starting to explore this challenging and exciting realm of musical practice. Our first—and surprising—finding was that mass profiles are not very common in the percussive sounds, while melodic profiles are. On the other hand, the use of mass profiles is not rare in the contemporary repertoire for ensembles and also in some instrumental “extended” techniques.

For this paper, we have selected two melodic profiles with contrasting features in the pitch perception: a pitched rub on the skin of a pandeiro, and a non-pitched air-jet glissando played on a flute. We chose to exemplify mass profiles an “artificial” percussive object—a vibraphone cluster filtered to one note—and the already mentioned sound transformation on the double-bass³.

Figure 2 shows a Max window with the estimated curves for six different descriptors: dynamic envelope, pct80, fundamental frequency (expressed in Midicents), intrinsic dissonance, spectral centroid (also in Midicents), and spectral distribution. These curves are depicted in real time during the performance of the selected sounds. Our discussion will rely on this figure and Table 2, which shows selected results from the curves’ analysis.

³These sound files are available at: https://musica.ufmg.br/lapis/?page_id=1188

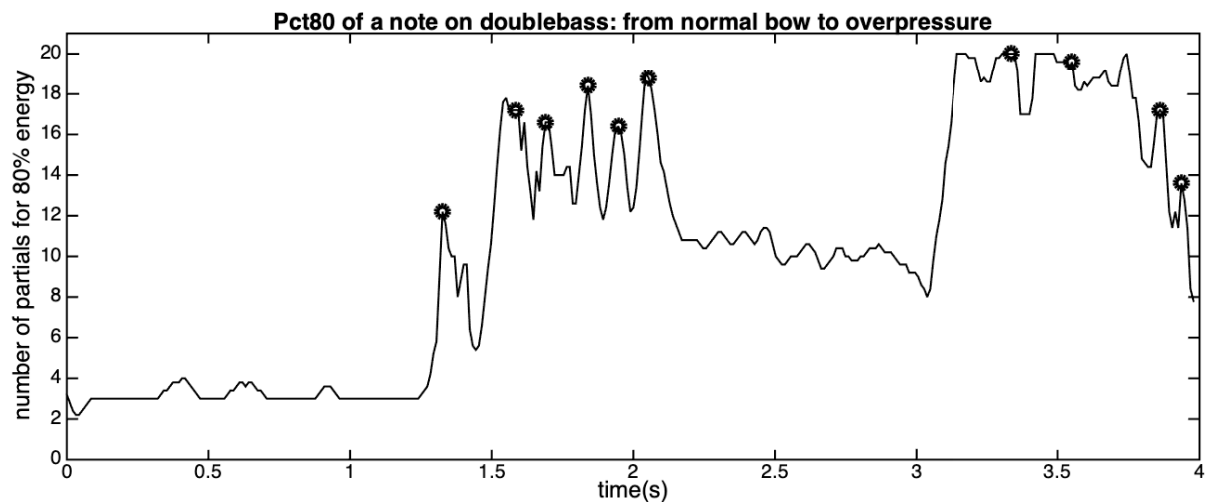


Figure 1: Pct80 curve for a double-bass sound played with a transition from regular bowing to overpressure, starting at 1.25 s. The black circles indicate the detected allures.

Melodic profiles can occur as a pitch glissando or as a noise node changing its central frequency. The first case will present a high ratio of pitched to total frames and some deviation in the pitch curve, as in the case of the pandeiro rub. In the second case, the variation of the spectral centroid will guide the analysis. The flute air-jet sound shows variation in this descriptor, corroborated by the start/medium/end values. There is also a change in the harmonic profile, indicated by variations in the descriptors pct80 and intrinsic dissonance.

We have already observed in former works that the percentage between pitched and total frames, along with the pct80 values, have a significant role in the qualification of different classes of mass. Intrinsic dissonance may also influence this task. The “cluster to note” vibraphone sound presents a profile that simplifies its mass, while the “regular bow to overpressure” double-bass sound has an opposed trajectory. This opposition appears in the skewness values estimated for pct80, dissonance, and pitch binary time series.

The spectral distribution curves suggest that melodic profiles are more stable than mass profiles, indicating that mass and harmonic timbre are actual “connecting vessels” [3, p. 412] in these cases. The four selected sounds present cyclical variation in some descriptors, which also connects to harmonic timbre and mass variations.

7. FINAL REMARKS

This paper represents the end of a preliminary stage in our attempt to approximate the Schaefferian criteria of musical perception and the interactive use of low-level audio descriptors. The next stage, already in course, is the confrontation of the implemented tools with actual musical practices. It is hard to imagine a single situation able to explore all features at once. Therefore we prefer to concentrate in specific demands from players and composers. In the practice of the contemporary flute repertoire, the use of these tools helps understanding and improving extended sonorities⁴.

8. ACKNOWLEDGMENTS

This work is supported by the Brazilian research agency CNPq (National Council for Scientific and Technological Development) and FAPEMIG (The Minas Gerais State Research Foundation).

⁴Cf. the paper “A diversidade sonora da flauta transversal: uma abordagem baseada na tipo-morfologia de Schaeffer e no uso de descritores de áudio em tempo real”, by R. Frade and S. Freire, also programmed in this FIA 2020/22 conference.

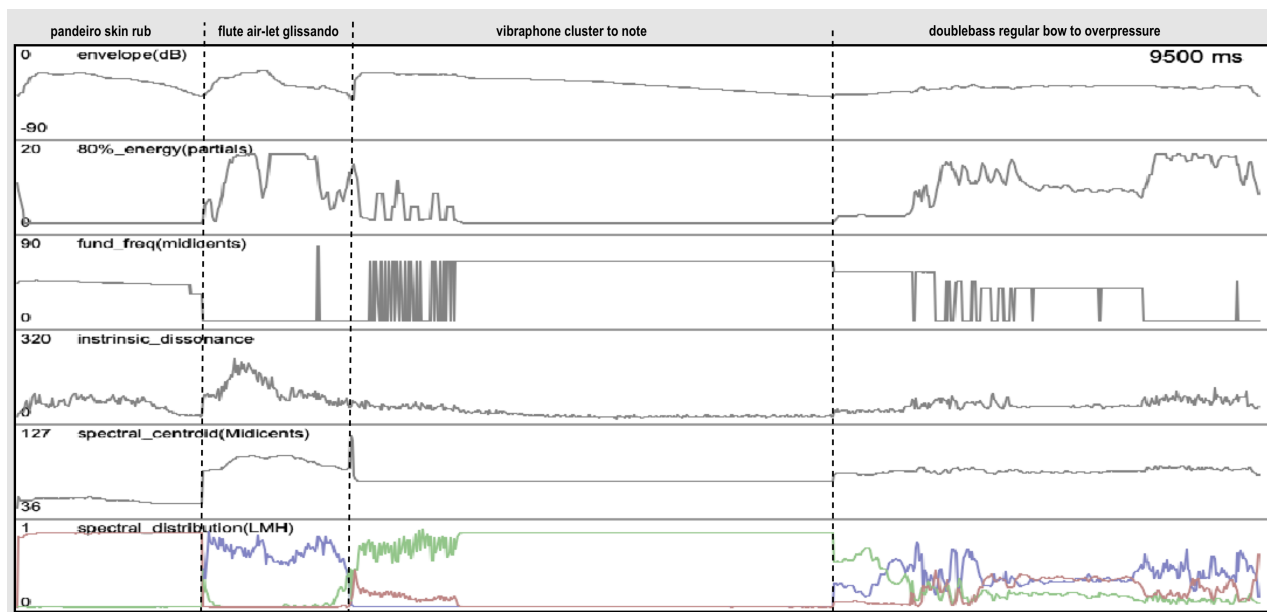


Figure 2: Max window displaying curves of 6 descriptors for 4 selected sounds. In the spectral distribution panel, the low region is depicted in red, the medium in green, and the high in blue.

REFERENCES

- [1] Sérgio Freire, José Henrique Padovani, and Caio Costa Campos. Schaeffer's Solfège, Percussion, Audio Descriptors: Towards an Interactive Musical System. *Revista Vórtex*, 10(1), April 2022. ISSN 2317-9937. doi:10.33871/23179937.2022.10.1.5. URL <https://periodicos.unespar.edu.br/index.php/vortex/article/view/4693>. Number: 1.
- [2] Pierre Schaeffer. *Traité des Objets Musicaux*. Éditions du Seuil, Paris, 1966. ISBN 978-2-02-002608-6 2-02-002608-2.
- [3] Pierre Schaeffer. *Treatise on Musical Objects: Essays Across Disciplines*. University of California Press, Oakland, California, 2017. ISBN 978-0-520-29429-5 978-0-520-29430-1.
- [4] Michel Chion. *Guide des Objets Sonores*. Buchet/Chastel, Paris, 1983. ISBN 978-2-7020-1439-4.
- [5] Michel Chion. *Guide to sound objects Pierre Schaeffer and the musical research*. London, 2009.

Table 2: Selected data from the analysis of 4 profiled sounds.

descriptor	sound	pandeiro rub	flute air gliss	vibes cluster to note	db reg. bow to overpress
fundamental frequency (midicents)	pitched/total (%)	90	4	86	61
	allure total (TS)	-	-	6 (irreg.)	3 (irreg.)
	skewness (TS)	0.5	0.37	-0.008	0.002
	pitched avg/std	47.2 ± 4.4	66 ± 7.9	73.7 ± 3.4	49.9 ± 10.2
	skewness pitched	0.001	0.37	0	0.003
	allure pitched	-	-	-	-
pct80	avg / stdev	1 ± 0	14.3 ± 5.9	1.9 ± 2.4	10.3 ± 5.9
	temp. centr.	0.5	0.5	0.31	0.63
	temp. spread	3.3	9.5	7.4	15.22
	skewness	0	0.002	0.036	-0.01
	Δ max-min	0	18.4	15.4	16.8
	values st/md/end	1 / 1 / 1	4.5 / 20 / 9.8	5.7 / 1 / 1	2.7 / 14.9 / 11.6
	presence of allures	-	3 (irreg.)	4 (irreg.)	10 (irreg.)
intrinsic dissonance	avg / stdev	54.3 ± 27	119.1 ± 51.6	39.4 ± 28.2	52.9 ± 20.5
	temp. centr.	0.42	0.43	0.31	0.57
	temp. spread	19.2	28.4	28.7	35.2
	skewness	0.001	0.004	0.008	-0.002
	Δ max-min	104	91.2	105	100.3
	values st/md/end	35 / 79 / 6	95 / 91 / 60	92 / 27 / 13	26 / 44 / 60
	presence of allures	6 (irreg.)	14 (irreg.)	10 (irreg.)	9 (irreg.)
spectral centroid (midicents)	avg / stdev	50.3 ± 2.8	97.5 ± 6.3	74.1 ± 0.9	86.7 ± 2.6
	temp. centr.	0.48	0.49	0.498	0.5
	temp. spread	23.22	28.41	46	46.6
	skewness	0.001	0.0002	0	-0.0001
	Δ max-min	8.1	18.6	13.7	11.38
	values st/md/end	51 / 50 / 47	88 / 103 / 89	77 / 74 / 74	83 / 85 / 86
	presence of allures	-	-	-	10 (irreg.)