

# A computer-based framework to analyze continuous and discontinuous textural works using psychoacoustic audio descriptors

Micael Antunes<sup>1 2</sup>, Danilo Rossetti<sup>1 2</sup>, Jônatas Manzolli<sup>1 2</sup>

<sup>1</sup> Arts Institute – Campinas University  
Rua Elis Regina, 50 – Cidade Universitária "Zeferino Vaz", Barão Geraldo, Campinas - SP

<sup>2</sup> Interdisciplinary Nucleus for Sound Communication- Campinas University  
Rua da Reitoria, 165 – Cidade Universitária "Zeferino Vaz", Barão Geraldo, Campinas, SP

micael.silva@nics.unicamp.br, danilo.rossetti@nics.unicamp.br, jonatas@nics.unicamp.br

## Abstract

This paper discusses a computer-aided musical analysis methodology anchored on psychoacoustics audio descriptors. The musicological aim is to analyze compositions centered on timbre manipulations that explore sound masses and granular synthesis as their builders. Our approach utilizes two psychoacoustics models: 1) Critical Bandwidths and 2) Loudness, and two spectral features extractors: 1) Centroid and 2) Spectral Spread. A review of the literature, contextualizing the state-of-art of audio descriptors, is followed by a definition of the musicological context guiding our analysis and discussions. Further, we present results on a comparative analysis of two acousmatic pieces: *Schall* (1995) of Horacio Vaggione and *Asperezas* (2018) of Micael Antunes. As electroacoustic works, there are no scores, therefore, segmentation and the subsequent musical analysis is an important issue to be solved. Consequently, the article ends discussing the methodological implication of the computational musicology addressed here.

## 1. Introduction

Research in the late 20th century intensified the development of computer aid analysis using computer-integrated digital signal processes. There are studies in machine listening [1], auditory scene analysis [2], and music information retrieval [3]. In line with such trajectory, we discuss here an application of so-called Audio Descriptors (AD) to a musical analysis of textural works. Audio Descriptors (AD) comprises a series of algorithms, that uses statistical and mathematical models to analyze audio data using acoustics, psychoacoustics and/or musical parameters. It allows, for example, to perform a classification of pieces and musical styles [4] and musical information retrieval and classification such as in Peeters [4], Pereira [5], and Peeters et al. [6]. Additionally, Malt & Jourdan [7] and [8] have developed tools for the analysis of contemporary music. Rossetti & Manzolli [9] studied emergent sounds in the context of electroacoustic and live electronic music. Other approaches have been developed in the Interdisciplinary Nucleus for Sound Studies (NICS-UNICAMP). Monteiro [10] developed a methodology to automatic segmentation

of electroacoustic music and Simurra & Manzolli [11] and [12] applied audio descriptors in the context of computer-assisted orchestration. In line with these previous studies, our main question here is: How can psychoacoustic descriptors contribute to an analytical methodology in the context of acousmatic compositions focused on textures?

To solve this question, we introduce and apply two audio descriptors: Volume [13] e Bark Coefficients [14]. These tools provide multidimensional graphical representations centered on two aspects: a) Psychoacoustics models like Loudness, and Critical Bandwidths and b) Spectral measures like Spectral Centroid, Spectral Spread and Loudness.

Based on these descriptors, we discuss the proposed methodology of analysis and compare two acousmatic works: *Schall* (1995) of Horacio Vaggione and *Asperezas* (2018) of Micael Antunes. Previous discussions addressing the analysis of those works are found in [15], [16], [17], and [18], here we mainly discuss the methodology and the computational tools employed in the process.

## 2. Review of literature

### 2.1 Musical Analysis and Psychoacoustics

Musical analysis based on psychoacoustic models had been particularly studied in recent years. Parncutt [19] analyzed tonal music using psychoacoustic features, based on pitch perception, fusion, and harmonicity models. Sethares [20] investigated tuning systems based on sensorial dissonance models applied in the analysis of Thai Classical Music. Sethares also developed a computer tool to generate dissonance curves aiming to calculate dissonance of sound spectra. Vassilakis [21] implemented an algorithm to calculate the roughness of audio files.

Psychoacoustic measures are useful tools regarding the study of pitch, timbre and harmony perception. In this context, Monteiro [10] implemented the Critical Bandwidth model in an external library called

*PDescriptors*<sup>1</sup> for Pure Data. Bullock [14] implemented Zwicker's model [22] in a Sonic Visualiser software plugin to generate plots on the levels of Critical Bandwidth energy of audio files (see analytical plots in topic 4).

Other psychoacoustics models are also implemented in libraries of Sonic Visualiser, Pure Data and Max/MSP, such as Sharpness, Loudness, Pitch Detection, and Inharmonicity<sup>2</sup>.

## 2.2 Sound Mass Music

The texture is a metaphor of visual and tactile perception. This concept offers a particular standpoint aiming to describe and analyze some musical phenomena [23], [24] and [25]. According to Mackay [23], "texture of music consists of its sounding components; it is conditioned in part by the number of those components sounding in simultaneity or concurrence". Ferraz [24] emphasizes that texture is "compatible with the system and typical procedures to which it belongs". Texture can be polyphonic, monophonic, harmonic, serial, pointillistic or static. He argues that textural perception is correlated with the Gestalt phenomenon [26] resulting from the interaction between the compositional sound material.

We can analyze musical texture starting with the measure of musical events that are found in the time-frequency space, giving us a measure of density [25]. We are particularly interested in sound mass texture. This phenomenon occurs when the music texture is the protagonist of the musical discourse. Further, other musical parameters (pitch, rhythm, intensity) are just parameters to create a "kind of spectromorphological hyper-instrument" [27].

Karlheinz Stockhausen stated that the perception of sound mass starts at our limit of perceiving separated musical sounds in time or frequency space [28]. This idea was explored by Ligeti with the concept of "timbre of movement" (*bewegungsfarbe*) [29]. Stockhausen's concept of sound mass could be related to the technique of micropolyphony, in which the perception of a sound texture is created by the fusion of multiple separated instrumental voices [29]. Pieces like *Atmosphères* (1961), *Continuum* (1968) or *Chamber Concerto* (1970) are examples of this concept. In the context of Ligeti's electronic music (*Glissandi*, 1957, *Artikulation*, 1958) the instrumental voices are replaced by multiple sound partials, that, juxtaposed and superimposed by similar processes to additive synthesis, form the timbre of movement.

Another theory, *Stochastic Music*, was developed by Iannis Xenakis in the 1950s [30]. Using mathematical models, Xenakis aimed to create continuous and discontinuous sound masses using a large number of *pizzicati*, *glissandi* or even electronic sounds [30]. He explored this technique in works such as *Pithoprakta* (1955-56), *Achorripsis* (1958) and *Arousa* (1971).

The granular synthesis technique is also an

important technique that generates sound masses. This process aims to generate sounds by the agglutination of thousands of 'grains', which are extremely short sound samples, belonging to the micro-time scale (in milliseconds) [31]. Perceptually, it is needed more than 20 grains per second to build a structure that is perceived as a sound event. When this fusion happens, a texture or a flux of a timbre varying in time is generated. Solomos [32] assumes that the Granular Paradigm corresponds to the corpuscular sound description, as opposed to the waveform description.

From the definitions presented above, it is possible to assume that sound masses create a kind of saturation in the auditory system. At a certain point in time, the listener starts to perceive sound events globally, i.e. a unique timbre rather than separated events [33]. Therefore, we claim here is that psychoacoustic descriptors offer a proper tool to analyze such phenomenon. Critical Bandwidths and Loudness feature extraction can be a useful tool to describe: a) how sound masses are formed and b) how they merge into a single perceived unity. Therefore, our analysis focus on these two aspects.

## 3. A Method of Computer Aid Analysis

Our method concerns the generation of plots of audio descriptors to describe spectral and psychoacoustic features from audio files (i.e. recordings of performance or acousmatic pieces). Those graphics are used to reveal important features of textural music. They are complemented by musicological discussions and further music analysis.

### 3.1 Volume Descriptors

One of our tools is a multidimensional descriptor comprising three curves in one graphic called Volume (Fig. 1 and 2, above). The volume is a concept defined by Truax [33] as the "perceived magnitude of sound or the psychological space it occupies".

Here, the Volume is a three-dimensional graphic representation created by Malt and Jourdan [13] by superposing three curves: Spectral Centroid, Spectral Spread, and Loudness. In the graphics, the Volume magnitude is given by the area between Spectral Centroid and Spectral Spread curves. Degrees of grey tonality indicates the Loudness of the perceived sound intensity. The dark is the higher level, while soft gray the is the lower.

The spectral centroid  $C(i)$  can be defined as the barycenter of the spectral energy concerning to a window of analysis [5], [10]. It is calculated as the frequency weighted mean, where  $X_i(k)$  are the magnitudes of the Discrete Fourier Transform of the  $i$  window, and  $k$  is the half of the adopted spectral components of the Transform. The spectral centroid  $C(i)$  is calculated as follows:

<sup>1</sup> <https://sites.google.com/site/pdescriptors/> Accessed in 03 May 2019.

<sup>2</sup> <https://www.vamp-plugins.org/download.html>

Accessed in 03 May 2019.

$$C(i) = \frac{\sum_{k=1}^K k * |X_i(k)|^2}{\sum_{k=1}^K |X_i(k)|^2} \quad (1)$$

where  $X_i(k)$  for  $k = 1 \dots K$  is the frequency amplitude of the window of analysis.

According to Peeters [4], the Spectral Spread  $S(i)$  describes the spread of the sound spectrum around its mean value, i.e. the variance of the defined distribution. Therefore, it is a measure of the average spread of the spectrum concerning its centroid. For noisy sounds, the Spectral Spread is expected to be high, while harmonic and tonal sounds, it is low. The Spectra Spread is defined as follows:

$$S(i) = \frac{\sum_{k=1}^K (k - C_i)^2 |X_i(k)|^2}{\sum_{k=1}^K |X_i(k)|^2} \quad (2)$$

where  $X_i(k)$  for  $k = 1 \dots K$  is the frequency amplitude of the window of analysis.

Pereira [5] and Monteiro [10] define the Loudness  $L(i)$  as the psychoacoustic measure of the sound intensity perception, which varies in frequency according to the Fletcher and Munson curves [34]. Equation 3 defines the loudness  $L(i)$  of a spectral window of analysis, according to Pereira's model. The Fletcher and Munson curves are included in equation 3 as the  $W(k)$  weight function described in equation 4.

$$L(i) = \sum_{k=1}^K |X_i[k]|^2 10^{w[k]/20} \quad (3)$$

where  $X_i(k)$  for  $k = 1 \dots K$  is the frequency amplitude of the window of analysis.

$$W(k) = -0,6 * 3,64 * f(k)^{-0,8} - 6,5 * e^{-0,6 * (f(k) - 3,3)^2} + 10^{-3 * f(k)^{3,6}} \quad (4)$$

where the frequency  $f(k)$  is measured in kHz, defined as  $f(k) = k.d$ , the difference between two consecutive spectral bins in kHz.

Finally, the *Spectral Volume*  $V(i)$  of the  $i$ -nth window of the DFT. is defined by the graphic:

$$V(i) = \{C(i), S(i), L(i)\} \quad (5)$$

where  $C(i)$  is plot as the curve profile,  $S(i)$  is plot as the distance within the profile and  $L(i)$  is the colored scale.

### 3.2 Bark Coefficients Descriptor

Bark is a unit of critical bandwidths of Zwicker's model [22]. The model consists of a subdivision of the range of audible frequencies based on the physiological

properties of the basilar membrane.

Each *bark* is usually smaller than the pitch interval of one-third-octave and each slot is indexed with a number of the unit *bark*, i.e. the critical-band number zero is associated with 0 Hz. A critical-band-index smaller than one Bark indicates that two tones belong to the same critical band [22]. Consequently, masking, beating, and roughness can occur among these tones inside the same band.

It is important to say that psychoacoustic measures are based on empirical tests, therefore, a further mathematical formulation is required. An example is the so-called Fletcher & Munson's Curve of Iso-Loudness [34], equation (4). Moreover, we have already studied the Bark Scale anchored in Traunmüller [35] and others [36]. Our contribution is to present a new formula to convert frequency into barks.

We define a conversion function,  $B(f)$ , divided into three curves and a scaling factor  $b(f)$ . Nevertheless, our definition follows Traunmüller [35], clearness is improved.

Given a value  $0 \text{ Hz} \leq f \leq 20 \text{ kHz}$ , the respective Bark Scale conversion function  $B(f)$  is calculated as follows:

$$B(f) = \begin{cases} b(f) + 0.15(2 - b(f)) & \text{if } b(f) < 2 \\ b(f) + 0.22(b(f) - 20.1) & \text{if } b(f) > 20.1 \\ b(f) & \text{otherwise} \end{cases} \quad (6)$$

where the scaling factor  $b(f)$  is given by

$$b(f) = \frac{26.81f}{(1960+f)} - 0.53 \quad (7)$$

In a previous study [18], we have used Bullock's [14] implementation to calculate the *bark coefficients* (e.g. Fig. 1 and 2, below). This descriptor divides the audible frequency range into 26 bark slots, using Zwicker's<sup>3</sup> model. It provides a measure of the sound energy into 26 slots, and it is quite similar to the Chroma audio descriptor<sup>4</sup>.

Finally, the Bark coefficients and the Volume are used in three analytical tasks:

- 1) to create the piece segmentations using the Volume and Bark Coefficients;
- 2) to analyze harmonic fusion and segregation using Bark Coefficients;
- 3) examine the Sound Flux and the spectral evolution in time from the Volume representation.

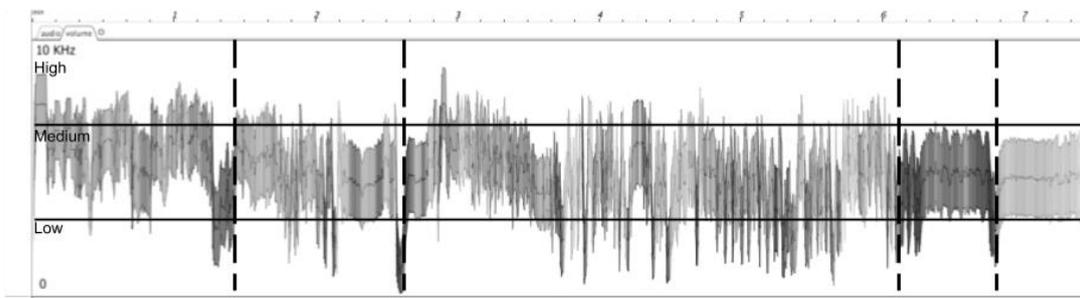
## 4. Comparative Study

Here, we present a comparative study applying our analytical methodology in to three goals: 1) segmentaion, 2) fusion and segregation and 3) sound flux and loudness, as follows.

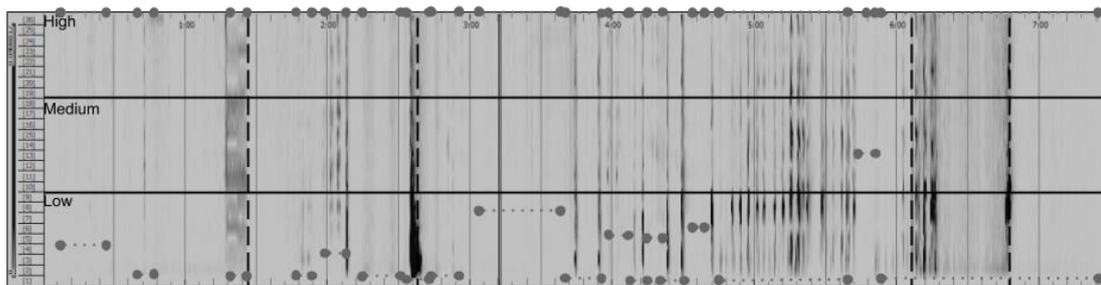
<sup>3</sup> Zwicker's [22] presents only 24 bark because the higher frequency of his table is 15500hz. But by using the equation to convert Hz in barks, we can calculate the 26

barks, containing all the range of audible frequencies.

<sup>4</sup> <https://www.vamp-plugins.org/download.html> Access 03 May 2019.



**Figure 1: Plot of Volume Schall (1995).** The vertical axis represents the frequency and the horizontal axis represents the time. The width of the line represents the spectral spread. The center line represents the spectral centroid. And the grayscale represents the loudness.



**Figure 2: Plot of the Bark Coefficients of Schall (1995).** The vertical axis represents the Bark Scale and the horizontal axis, time; grayscale, the energy of each bark bandwidth.

**Segmentation:** it is possible to segment a composition by observing peaks and valleys of energy, as well as breakpoints detected by in the Bark Scale and Volume descriptor.

**Fusion and Segregation:** it is possible to examine fusion or segregation levels of sound textures. These are related to the energy concentration of the Bark Scale.

**Sound Flux and Loudness:** it is possible to evaluated how much sound masses are not static applying the Spectral Spread and Loudness descriptors.

#### 4.1 The Two Pieces

In this article, we analyze two works: *Schall* (1995) by Horacio Vaggione, and *Asperzas* (2018) by Micael Antunes. Since they are acousmatic works, there is no traditional musical score of the pieces. Thus, the provided segmentation an important feature addressed by the computer-aid analysis.

*Schall* (1995) is an acousmatic piece by Vaggione composed using the *micromontage*, an important compositional technique of the granular paradigm. The employed sound materials are piano recordings. These samples were edited in microtime domain, intending to create granular textures. *Schall* is a representative work concerning texture and sound mass manipulation of the granular paradigm.

*Asperzas* (2018), on the other hand, is a piece in which the compositional process is based on the threshold of beats and roughness sensations. These phenomena are related with the critical bandwidths energy manipulation. Thus, the piece was created using the FM synthesis

technique, from a continuous timbre transformation in time [36]. *Asperzas* (2018) is anchored in a study of beats and roughness that is also correlated to our analytical model.

We compare both pieces under two main perspectives:

- To employ the segmentation provided by the computer-aid analysis in order to describe the macroform of the works;
- To enhance our musicological standpoint, comparing their analysis with the computer.

#### 4.2 Comparing Descriptors segmentation

The pieces' segmentation was performed by employing together the Volume and Bark Coefficients. Peaks and valleys were taken into account in order to observe possible correspondences between them. In this sense, the information coming from both descriptors can be viewed as complementary, providing consistency in the segmentation of the piece.

Figures 1 and 2 show the Volume and Bark Coefficients of *Schall* (1995). The segmentation is represented by the dashed vertical lines. By this process, the piece was divided into 5 segments, see Table 1.

Section	1	2	3	4	5
Starting	0''	1'24''	2'39''	6'07''	6'49''

**Table 1: Segmentation of Schall (1995), showing sections and its initial times.**

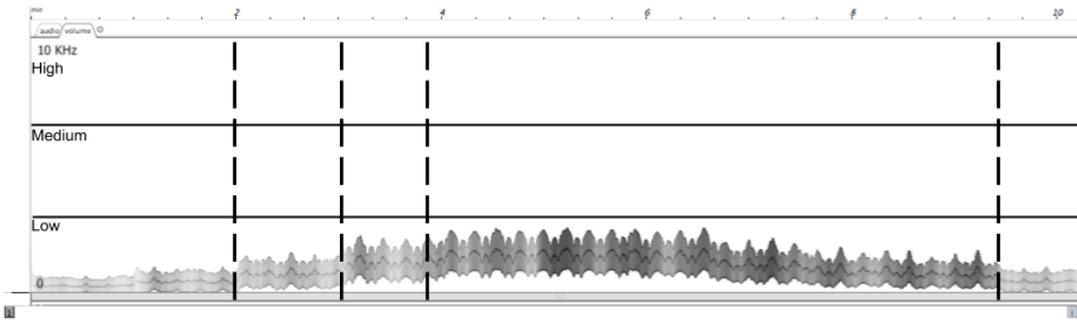


Figure 3: Volume of *Asperzas* (2018). The parameters are the same as in Figure 1.

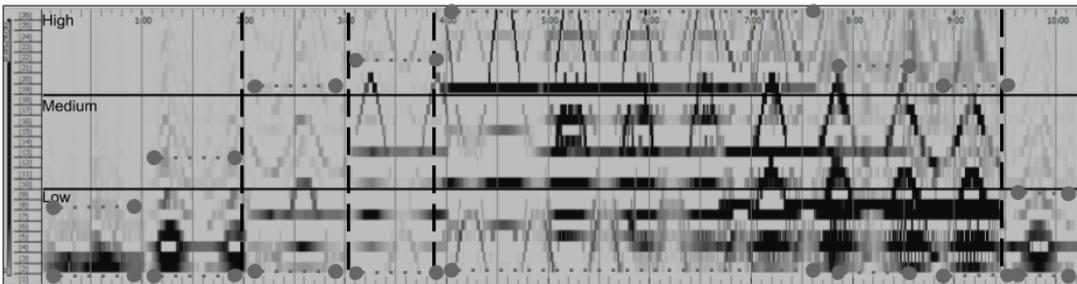


Figure 4: Plot of the Bark Coefficients of *Asperzas* (2018). The parameters are the same as in Figure 2.

Figures 3 and 4 show the Volume and Bark Coefficients of *Asperzas* (2018). The segmentation is represented by dashed lines. By this process, the piece was divided into 5 segments, see Table 2.

Section	1	2	3	4	5
Starting	0''	2'02''	3'05''	3'55''	9'22''

Table 2: Segmentation of *Asperzas* (2018), showing sections and its initial times.

### 4.3 Comparing Harmonic Fusion and Segregation

The Bark Scale graphic of *Schall* shows different density levels of grains, which mostly appear in the higher bark position. Grains with different changes in density do not allow significant harmonic fusion. For instance, in the second segment of the plot (Figure 2), the lowest level of fusion is found, while in segment 4 we have a peak of fusion.

Fusion and segregation in *Asperzas* are observed in Figure 4, within a pattern of recurrences. Therefore, there are frequencies which are present in the whole piece. We can highlight some aspects of the segment 4, in Figure 4: at the same time that some barks have constant energy (at some points, from 2 to 11), there is significant energy variation in all of the other bark slots. The consequence is a sensation of segregation because we perceive concomitant continuous frequencies and sound mass movement in their background.

Another aspect is the proximity of the barks concentrations that increases according to the evolution of the piece in time. This proximity generates several interferences. The consequence of this proximity is a high

sensation of roughness and beatings.

### 4.4 Comparing Volume (Sound Flux and Loudness)

*Schall* (1995) belongs to the granular paradigm and it presents a considerable Sound Flux, i.e. a high level of grain's movement in microtime, forming the sound texture or timbre of the work. The general level of loudness is low, except in the excerpt between 6'07'' e 6'49'', the fourth section of the piece, see Figure 1 and Table 1.

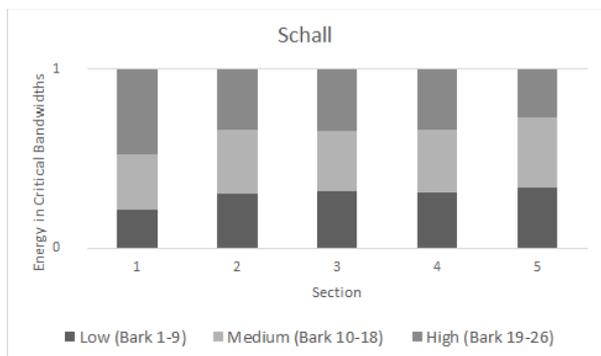
We can observe in Figure 1 that the third section (which lasts around 3'30'') presents a development focusing on the grain clouds' manipulation and variations. The flux in this section has the higher level of the piece, reached by the considerable variations of grains' intensity and frequency range. The first two sections present a considerable movement of sound clouds, but in a lower level compared to section 3. Sections 4 and 6 are more stable with lower sound flux, indicating an entropy movement inside the piece, such as the grains reaching a state of equilibrium. Therefore, in an overall view, the Spectral Volume in *Schall* (1995) is very diverse. The large Spectral Spread band indicates a preponderant timbre with noisy characteristics.

The sound flux of *Asperzas* (2018), on the other hand, has a clear direction. The development of the piece is linked with an expansion of the spectral frequencies and an increment of loudness (Section 4 has the peak of loudness). It is interesting to note that the spectral centroid of the volume representation increases at the same time that the bark concentration is higher.

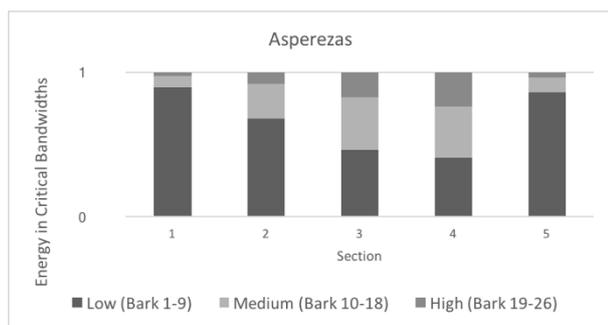
### 4.5 *Schall* (1995) versus *Asperzas* (2018)

According to the intention of comparing the two pieces, we integrate the collected data of the bark coefficients descriptor in graphics 1 and 2. Thus, it is

possible to have a global evaluation of the features extracted from the spectral development of the pieces. The graphics integrates the energy of the barks into three bands: Low (barks 1-9); Medium (barks 10-18); and High (barks 19-26). The energy of each band is plotted in Graphics 1 and 2 according to the sections of the pieces. The result is that *Schall* has its predominance in Medium and High bands while *Asperezas* has its predominance in the low bands.



**Graphic 1: The energy of each section of *Schall* into bands of Bark Scale.**



**Graphic 2: The energy of each section of *Asperezas* into bands of Bark Scale.**

We also integrated our analysis of Bark Coefficients, Spectral Fusion, Loudness and Spectral Flux in levels of variation: *Low, Medium and High*. Table 3 shows that *Schall* has a low variation on Bark; medium variation of Spectral Fusion and Loudness; and High variation of Spectral Flux. *Asperezas*, on the other hand, has a high variation in Bark and Spectral Fusion and low variation in Loudness and Spectral Flux.

	<i>Bark</i>	<i>Spectral Fusion</i>	<i>Loudness</i>	<i>Spectral Flux</i>
<i>Schall</i>	Low	Medium	Medium	High
<i>Asperezas</i>	High	High	Low	Low

**Table 3: Level of variation of the parameters of analysis.**

## 5. Discussions

The sound grains in *Schall* have some zones of concentration in each section. The montage of the grains

in microtime domain reflects points of agglutination with denser textures and moments where the grains are scattered in the frequency space, generating a subtle texture with rarefied grain clouds.

In the micromontage technique applied in *Schall*, the idea is to achieve the saturation from the manipulation of sound particles, exploring our threshold of the pitch, timbre, and loudness perception. Thus, the synthesis, from the variation of levels of density, creates masses using sound particles, causing spectral and temporal masking phenomena. Thereby, starting with an amount of discontinuity, we create a global idea of continuity in the sound.

The macroform of *Asperezas* shows movement expressed by a spectral and loudness increasing, that decreases in the final of the piece. In the microform, the constant spectral oscillation is remarkable, such as the variations on the spectral fusion, and consequently changes of roughness and beatings.

In this sense, *Asperezas*, in the opposite of *Schall*, offers us a formal process that is higher predictable. The descriptors reveal a quasi-symmetrical transformation of the sound material (see figure 4) in the microform, creating in consequence a previous expectation in listening.

A comparative analysis of the two pieces raises an interesting compositional question. The sound masses are an effect of saturation of our auditory system. However, different techniques explore different thresholds of perception in different ways.

## 6. Conclusions

We discussed a methodology of computer-aided musical analysis using psychoacoustics audio descriptors that afford tools to study textural music. The aims of this study were to understand its singularities and enhanced hypothesis on the creative process. We compared two works *Asperezas* (2018) and *Schall* (1995) under three analytical perspectives: formal segmentation, harmonic fusion/segregation, and sound flux.

These questions are relevant in the musicological context concerning the analysis of contemporary music field, as they reveal the details of the compositional practices. They also allow us to study composition, not only from the standpoint of its creation, but to perform an analysis that gives us clues about the perceptual processes related to the listening of the works. Finally, audio descriptors are considered useful tool to support these issues.

In our analysis, we observed the singularities of the sound textures explored in the pieces. In this way, we emphasize the difference between the texture explored by the granulation paradigm, supported by the *micromontage* technique and the construction of the masses through the accumulation of sound phenomena in the critical bands. In the specific case of the two pieces, it was also observed the differences in the development of the pieces, by highlighting that the expectation of listening can act from the manipulation of the sound material.

Since the analysis is performed from the audio data, it would be possible to conceive other experiments using different kinds of recordings. For example, to evaluate

different recordings of a piece using dummy head microphones in order to examine perceptual issues. The approach introduced here can also be explored in the context of live electronic and instrumental music, aiming to analyze textural music by Ligeti, Xenakis, and Penderecki, for example.

We also foresee applications in computer-aided composition, the psychoacoustic descriptors would be useful in sound designing of interactive performances and multimodal installations. They could be applied together with theories of ecological perception and dynamic cognition, creating immersive computer-controlled environments.

At least, computational musicology dialogues with experimental psychology and cognitive sciences, therefore, some of the future developments could be useful to implement algorithms of automatic classification.

## 7. Acknowledgments

Micael Antunes is supported by FAPESP in his doctoral research at Arts Institut of Unicamp/ NICS-UNICAMP, process 2019/09734-3. Danilo Rossetti is supported by FAPESP in his post-doctoral research at NICS-UNICAMP, process 2016/23433-8. Jônatas Manzolli is supported by CNPq under a Pq fellowship, process 304431/2018-4 and 429620/2018-7.

## 8. References

- [1] Rowe, R. Interactive music systems: machine listening and composing. MIT Press, 1992.
- [2] Bregman, A. S. Auditory scene analysis: The perceptual organization of sound. MIT Press, 1994.
- [3] Herrera, P. MIRages: an account of music audio extractors, semantic description and context-awareness, in the three ages of MIR. PhD Thesis, Information and Communication Technologies, Universitat Pompeu Fabra. Barcelona, 2018.
- [4] Peeters, G. A Large Set of Audio Features for Sound Description (Similarity and Classification) in the CUIDADO Project (Cuidado Project Report). Paris: Ircam, 2004.
- [5] Pereira, E. Estudos sobre uma ferramenta de classificação musical. Dissertação de Mestrado. Faculdade de Energia Elétrica e de Computação da Universidade Estadual de Campinas, 2009. Campinas: UNICAMP, 2009.
- [6] Peeters, G. et al. The Timbre Toolbox: Extracting audio descriptors from musical signals. Journal of the Acoustical Society of America, Melville, v. 130, n. 5, p. 2902-2916, November 2011.
- [7] Malt, M; JOURDAN, Emmanuel. Z. Descriptors: A Library for Real-Time Descriptors Analysis. In: SOUND AND MUSIC COMPUTING CONFERENCE, 5, 2008, Berlin. Proceedings... Berlin: Sound and Music Computing, 2008.
- [8] Malt, M; Jourdan, Emmanuel. Z. Real-Time Uses of Low Level Sound Descriptors as Event Detection Functions Using the Max/MSP Zsa.Descriptors Library. In: BRAZILIAN SYMPOSIUM ON COMPUTER MUSIC, 12., 2009, Recife. Proceedings... São Paulo: USP, SBC, 2009, p. 45-56.
- [9] Rossetti, D.; Manzolli, J. De Montserrat às ressonâncias do piano: uma análise com descritores de áudio. Opus, 23(3), 193-221. 2017.
- [10] Monteiro, A. Criação e performance musical no contexto dos instrumentos musicais digitais. Dissertação de Mestrado. Instituto de Artes da Universidade Estadual de Campinas, 2012. Campinas: UNICAMP, 2012.
- [11] Simurra, Ivan; Manzolli, Jônatas. Sound Shizuku Composition: A Computer-Aided Composition System for Extended Music Techniques. MusMat Brazilian Journal of Music and Mathematics, v. 1, n. 1, p. 86-101, 2016.
- [12] Simurra, Ivan; Manzolli, Jônatas. “O azeite, a lua e o rio”: o segundo diário de bordo de uma composição a partir de descritores de áudio. Revista Música Hodie, Goiás, v. 16, n. 1, p. 121-123, 2016.
- [13] Malt, M; Jourdan, E. Le “BSTD” – Une représentation graphique de la brillance et de l’écart type spectral, comme possible représentation de l’évolution du timbre sonore. In: L’ANALYSE MUSICALE AUJOURD’HUI, CRISE OU (R)ÉVOLUTION, 2009, Strasbourg. Proceedings... Strasbourg: STAM, 2009.
- [14] Bullock, J. Libxtract: a lightweight library for audio feature extraction; Proceedings of the 2007 ICMC; ICMA, Sweden, 2007.
- [15] Rossetti, D. Temporalidades, morfologias e granulações em *Schall* de Horacio Vaggione. Proceedings of the IV Encontro Internacional de Teoria e Análise Musical, Eitam 2017, São Paulo, SP, 2017.
- [16] Rossetti, D; Manzolli, J. Timbre emergente na música eletroacústica: Uma análise com o aporte gráfico do volume e da vivacidade espectral. Proceedings of the XXVIII Encontro da Sociedade Brasileira de Acústica. Porto Alegre, Unisinos, 2018.
- [17] Rossetti, D; Manzolli, J. Analysis of Granular Acousmatic Music: Representation of Sound Flux and Emergence. *Organised Sound* v. 24, n. 2, p. 1-12, 2019.
- [18] Silva, M. A., Botechia, T., Rossetti, D. Manzolli, J. Análise musical por descritores de áudio baseados em modelos psicoacústicos. Anais do Simpósio Internacional de Cognição e Artes Musicais 14. Campo Grande, MS, 2019. in press.
- [19] Parncutt, R. Harmony: A psychoacoustical approach. Vol. 19. Springer Science & Business Media, 2012.
- [20] Sethares, W. A. Tuning, timbre, spectrum, scale. Berlin: Springer Science & Business Media, 1998.
- [21] Vassilakis, P. N.; Fitz, K. SRA: A web-based research tool for spectral and roughness analysis of sound signals. In: Proceedings of the 4th Sound and Music Computing (SMC) Conference, 2007. p. 319-325.
- [22] Zwicker, E. Subdivision of the audible frequency range into critical bands (Frequenzgruppen). The Journal of the Acoustical Society of America, v.33, n.2, p.248-248,

1961.

[23] Mackay, J. On the perception of density and stratification in granular sonic textures: An exploratory study. *Journal of New Music Research*, v. 13, n. 4, p. 171-186, 1984.

[24] Ferraz, S. Análise e Percepção Textural: Peça VII, de 10 Peças para de Gyorgy Ligeti. *Cadernos de Estudo: Análise Musical*, vol. 3. São Paulo, pág. 68-79, 1990.

[25] Berry, W. *Structural functions in music*. Chelmsford: Courier Corporation, 1987.

[26] Cambouropoulos, E. Tsougras, C. Auditory Streams in Ligeti's Continuum: A Theoretical and Perceptual Approach. *Journal of interdisciplinary music studies*, v. 3, n. 1-2, p. 119-137, 2009.

[27] Smalley, D. Spectromorphology: explaining sound-shapes. *Organised sound*, Cambridge, v. 2, n. 2, p. 107-126, 1997.

[28] Stockhausen, K. Pontos e grupos. Im: Maconie, Robin. *Stockhausen sobre a música: palestras e entrevistas compiladas por Robin Maconie*. Tradução de Saulo Alencastre. São Paulo: Madras, 2009, pág. 45-50.

[29] Ligeti, G. *Neuf essais sur la musique*. Genebra: Éditions Contrechamps, 2010.

[30] Xenakis, I. Concerning time. *Perspectives of New Music*, Seattle, v. 27, n. 1, p. 84-92, 1989.

[31] Roads, C. *Microsound*. MIT Press. Massachusetts. 2001.

[32] Solomos, M. 2013. De la musique au son: l'émergence du son dans la musique des XXe-XXIe siècles. Rennes: Presses Universitaires de Rennes.

[33] Truax, B. Musical Creativity and Complexity at the Threshold of the 21st Century. *Interface* v. 21, p. 29-42, 1992.

[34] ISO, B. 226: 2003: Acoustics–Normal equal-loudness-level contours. *International Organization for Standardization*, v. 63, 2003.

[35] Traunmüller, H. “Analytical expressions for the tonotopic sensory scale”, in *The Journal of the Acoustical Society of America* 88.1, 1990.

[36] SILVA, M. A. *Redução da Dissonância Sensorial em uma Escala Temperada Utilizando Timbres Inarmônicos: Uma Abordagem Experimental e Aplicações Artísticas*. Master Dissertation, 2018. 96 p. Music post-graduation program- School of Arts and Communications, São Paulo University, São Paulo, 2018.