Conference Proceedings

SysMus 2018

06–08 June, 2018 | Conservatório da UFMG | Belo Horizonte

musica.ufmg.br/sysmus2018

Federal University of Minas Gerais
State University of Minas Gerais

SysMus 2018
Belo Horizonte • Brasil

SysMus18

Proceedings of the 11th International Conference of Students of Systematic Musicology

Editors:
Davi Alves Mota
Tairone Nunes Magalhães

Cover design:
Débora Line

Belo Horizonte, Brazil
June 6-8, 2018

Sponsors

Contents

Preface

SysMus (International Conference of Students of Systematic Musicology) is a series of conferences organized by students for students. In 2018 it was held for the first time in Latin America, from June 6th to 8th, in the city of Belo Horizonte, Brazil. The conference was organized by members of the laboratory CEGeME (Center for Studies on the Musical Gesture & Expression), located at the School of Music of the Federal University of Minas Gerais (UFMG), and members of the Post-graduate Program in Arts at the State University of Minas Gerais (UEMG).

In 2018, the conference received a total of 60 submissions from 85 authors, originated from 47 universities from 17 different countries. The scientific committee was responsible for 235 anonymous reviews, and was composed by researchers, postdocs and PhD students. During the event there was a total of 21 paper presentations and 12 poster presentations.

On behalf of the SysMus18 Organizing Committee, we would like to thank all those who supported the conference: our Supervisors, Luiz Naveda and Mauricio Loureiro, for the guidance and support throughout the planning process; the SysMus Council, specially Manuela M. Marin for the encouragement and confidence; our sponsors – Sempre (Society for Education, Music and Psychology Research), for the kind provision of travel grants for many students who attended to the conference; the Graduate Program in Music at UFMG, specially the program coordinator, Prof. Luciana Monteiro de Castro Silva Dutra, for the promptness in helping us to make SysMus18 a great conference; the Graduate Program in Arts at UEMG; Conservatório UFMG; SMPC (Society for Music Perception and Cognition); CNPq and CAPES, Brazilian agencies for foment of research, for the provision of the research scholarships to some members of the organizing committee. We also want to thank our invited speakers Andre Holzapfel, Juan Pablo Bello and Mauricio Loureiro for the great speeches they gave, and for accepting our invitation.

Tairone Magalhães & Arícia Ferigato
SysMus18 Conference Chairs

Organizing Committee

Conference Chairs

Arícia Ferigato, *Federal University of Minas Gerais*
Tairone Magalhães, *Federal University of Minas Gerais*

Members

Aluizio Neto, *IT Specialist – Federal University of Minas Gerais*
Débora Line, *Graphic Designer – Federal University of Minas Gerais*
Leandro Souza – *Federal University of Minas Gerais*
Lívia Itaborahy – *State University of Minas Gerais*
Marina Magalhães – *State University of Minas Gerais*
Michel Brasil – *State University of Minas Gerais*
Rodrigo Borges – *University of São Paulo*
Thais Santos – *Center for Studies on the Musical Gesture & Expression*

Supervisors

Prof. Dr. Luiz Naveda – *State University of Minas Gerais*
Prof. Dr. Maurício Loureiro – *Federal University of Minas Gerais*

Editors

Davi Mota, *Federal University of Minas Gerais*
Tairone Magalhães, *Federal University of Minas Gerais*

Scientific Committee

Chair of the Scientific Committee

Davi Mota, *Scientific director – Federal University of Minas Gerais*

Reviewers

Adam Ockelford, *University of Roehampton*
Adrien Ycart, *Queen Mary University of London*
Aydin Anic, *Macquarie University*
Christine Ahrends, *Aarhus University Hospital*
Claudia Stirnat, *University of Hamburg*
Dawn Rose, *University of Hertfordshire*
Daniel Fiedler, *University of Education Freiburg*
Diana Kayser, *University of Oslo*
Ekaterina Pavlova, *King's College London*
Eva Schurig, *University of Exeter*
Fabia Bertram, *University of Cologne*
Fabian Moss, *University of Music and Dance Cologne*
Fernando Iazzetta, *University of São Paulo*
Georgios Diapoulis, *University of Jyväskylä*
Geraint Wiggins, *Queen Mary University of London*
Gesine Wermke, *Julius-Maximilian University of Würzburg*
Hayley Trower, *University of Roehampton*
Igor Reyner, *King's College London*
Jaco Meyer, *North-West University*
Jan Stupacher, *University of Graz*
Jay Appaji, *Southern Methodist University*
Johanna Wilson, *University of Jyväskylä*
John Chun-fai, *The Chinese University of Hong Kong*
Joshua Bamford, *University of Jyväskylä*
Juhan Nam, *Korea Advanced Institute of Science and Technology*
Julian Céspedes-Guevara, *University of Sheffield*
Julie Delisle, *Université de Montréal*
Katerina Drakoulaki, *National and Kapodistrian University of Athens*
Katharina Schäfer, *University of Jyväskylä*
Kirk Olsen, *Macquarie Univeristy*
Klaus Frieler, *The Liszt School of Music Weimar*
Li-Ching Wang, *University of Cambridge*
Luiz Naveda, *State University of Minas Gerais*
Makarand Velankar, *Savitribai Phule Pune University*
Manuela Marin, *University of Vienna*

Mari Tervaniemi, *University of Helsinki*
Maria Witek, *University of Birmingham*
Marta Rizzonelli, *Humboldt University of Berlin*
Marvin Heimerich, *University of Cologne*
Michelle Ulor, *University of Leeds*
Nicola Dibben, *University of Sheffield*
Nicola Pennill, *Royal College of Music*
Niels Chr. Hansen, *Ohio State University*
Nina Politimou, *Middlesex University London*
Pablo Cuevas, *Universität zu Köln*
Pablo Mendoza-Halliday, *UNAM*
Peter Kranenburg, *Meertens Insitute*
Pui Yin Kwan, *Naxos Digital Services Limited*
Rodolfo Caesar, *Federal University of Rio de Janeiro*
Scott Bannister, *Durham University*
Stefanie Bräuml, *University of Vienna*
Steven Schiemann, *University of Education Freiburg*
Suzanne Ross, *Aarhus University*
Sérgio Freire Garcia, *Federal University of Minas Gerais*
Tairone Magalhães, *Federal University of Minas Gerais*
Thais Santos, *Center for Studies on the Musical Gesture & Expression*
Theresa Schallmoser, *University of Graz*
Yu-Fen Huang, *Edinburgh University*
Zachary Wallmark, *Southern Methodist University*

EVOLUTION OF TIMBRE DIVERSITY IN A DATASET OF BRAZILIAN POPULAR MUSIC: 1950-2000

Rodrigo Borges
University of São Paulo
rcborges@ime.usp.br

Marcelo Queiroz
University of São Paulo
mqz@ime.usp.br

ABSTRACT

In this paper we discuss a method for assessing the temporal evolution of timbre diversity in an annotated dataset, and apply it to a collection of Brazilian music from the 1950's to the 2000's. Previous work have explored audio analysis for measuring the variety of acoustic features or the stylistic evolution in American Popular Music in the period 1950-2010. We aim in this study to verify up to what point a similar methodology could be applied to a considerably different dataset (Brazilian popular music) in a comparably long period (1950-2000). The measure of timbre diversity, based on Shannon's entropy function, displays its lowest value for 1950-1955, abrupt decaying from 1975 to 1990 and an increasing trend from this point until 2000.

1. INTRODUCTION

Empirical musicologists relate the recent availability of large collections of digital music with a more scientific approach to music history [1, 2]. Digital audio allows acoustic descriptors to be extracted automatically and to be applied in many tasks in the field of Music Information Retrieval (MIR) [3]. We propose using acoustic descriptors to analyze the evolution of timbre diversity in a dataset of Brazilian popular music.

Among all music dimensions, timbre is the one which resists most formalization attempts, being frequently defined by opposition, as the sound quality which is not pitch, not intensity and not duration; it is informally referred to as the sound color, or more objectively as associated to the spectral composition and its dynamic variations [4]. For practical purposes and Music Information Retrieval tasks, the acoustic descriptor named Mel Frequency Cepstral Coefficients (MFCCs) is the most frequent timbre-related characteristic extracted from music sound signals, and will be addressed in the following sections.

We attempt to express timbre diversity as a measure of how these acoustic descriptors are distributed through all possible regions in an MFCC representation space, for each group of songs belonging to the same period. Periods with higher timbre diversity should display MFCC distributions that span many regions, whereas periods of low timbre diversity should display highly concentrated MFCC distributions.

We use Serrà et al [1] and Mauch et al. [2] as references for our work, and aimed to verify to what extent their methodology can be reproduced in a different dataset distributed in a period of comparable length, namely the dataset "100 greatest Brazilian music records" [5] comprising Brazilian popular music from the 1950's to the 2000's, compiled by the specialized music magazine Rolling Stone in 2007.

The text is structured as follows. In the next section we present the two studies taken here as a basis, focusing on how they deal with the timbre dimension. The methodology for measuring timbre diversity evolution is then detailed, comprising the dataset description as number of records, artists, years and songs; feature extraction; codeword representation and diversity measurement. Results are presented, discussed and compared to the ones in the literature. Conclusions and future work are presented in Section 5.

2. RELATED WORK

Serrà et al. [1] analyze harmony, timbre and loudness descriptors extracted from 464,411 distinct music recordings from a public collection known as the "million song dataset", using recordings from 1955 to 2010. They calculated *codeword* representations for harmony and timbre, and use a power law model for expressing the diversity of the distribution of these features over the years. The idea behind this method is to assume codewords as representatives of particular harmonic or timbre structures, and to associate higher degrees of diversity to samples with a more balanced distribution: e.g. if songs from a specific year use diversified harmonic and timbre combinations, this distribution should be more balanced, but if these songs use only relatively few of them, then the distribution will be more concentrated towards fewer codewords, i.e. it would be less balanced.

In order to take feature successions into account, the authors also proposed modeling each song using transition networks, where each node represents a codeword and each link represents a temporal transition. The measures of average shortest path length, clustering coefficient and assortativity with respect to a random network, were interpreted in terms of higher or lower diversity of harmonic and timbre elements. For the specific case of timbre, the diversity rate reached its peak in the year 1965, and started to decrease from there. Despite interesting evolutionary observations, such as the "loudness race", corresponding to a constant increase in the loudness level over the years, or the timbre diversity peak value happening in the year of 1965, the authors point out to a general lack of significant statistical trends in the evolution of harmonic or timbre elements in contemporary western popular music in the period considered.

Mauch et al. [2] investigated the "US Billboard Hot 100" between 1960 and 2010, aiming to measure musical diversity and evolution of disparities, as well as demonstrating quantitative trends of harmonic and timbre properties. As motivation cues, the authors asked three questions, to be answered during the analysis: (1) did North American popular music variety increase or decrease over time?; (2) were evolutionary changes continuous or discontinuous?; and (3) if they were discontinuous, when did discontinuities occur?

They chose to represent the acoustic properties in a fashion similar to the previous authors, but using the term *topics* instead of codewords. 16 topics were calculated, 8 based on MFCC (for timbre-related aspects) and 8 based on Chroma (for harmony-related aspects). Topics were calculated with a hierarchical generative model named Latent Dirichlet Allocation (LDA).

Having calculated timbre and harmonic topics for each song, it was possible to study the evolution of topics over the years. Authors also had access to expert-based annotations, that made possible the association of semantic information to each topic, for example "drums, agressive, percussive" in the case of a particular timbre topic, and "natural minor" for a harmonic topic. Temporal evolution in the frequency of some topics revealed clear trends, as for example the topic named "energetic, speech and bright", that starts increasing in occurrence from 1980 on.

Four measures of diversity are presented by [2]: the first measure is simply the number of songs in each time period, used to verify that other diversity measures are not affected by the size of a subsample. The second measure accounts for the year-wise diversity of acoustic style clusters in the data. The third is the effective number of music topics for each year, averaged across the harmonic and timbre topics. The fourth corresponds to disparity, or the variety of measurements in the matrix of principal components derived from the topics.

By using the Kmeans clustering algorithm, authors estimated that 13 clusters would better separate data in terms of the distribution of topics. These clusters are associated to musical styles and their evolution over the years is discussed. A Self Similarity Matrix was also calculated to assess topic distribution over the years, by computing the similarity between topic distribution of different time periods. This matrix was used to detect discontinuities, that according to these authors took place specially in 1983 and 1991.

3. METHODOLOGY

In order to verify to what extent the methodology applied in [2] can be reproduced in a different dataset we adapted it to our music collection as explained in the following subsections.

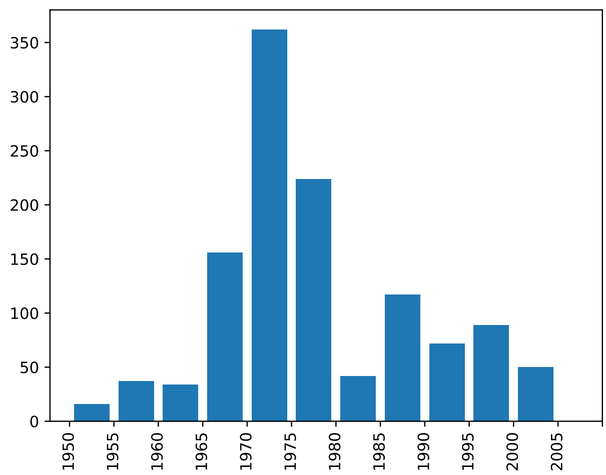


Figure 1. Distribution of songs over the 5 year periods .

3.1 Dataset

The dataset consists of 100 records released from 1950 to 2000, by 60 artists, summing up 1199 songs elected by the specialized music magazine Rolling Stone as the "100 greatest Brazilian music records list" [5]. This collection was published as representative of the opinions of 60 music researchers, producers and journalists, based on how influential they thought these records were to others artists.

Since the number of songs in each 5-year period is very unbalanced (see Figure 1), we selected random subsamples based on the period with the least number of songs, in order to allow for a more stable comparison. A table with all artist names, number of records, number of songs and number of years spanned by each artist in the dataset is presented in Table 1.

3.2 Timbre Feature Extraction

Mel frequency cepstral coefficients (MFCCs) were originally developed for automatic speech recognition and were later found to be useful for music information retrieval [3]. Even though timbre as a concept is very hard to define, since it encompasses many acoustic dimensions, MFCCs captures relevant timbre-related acoustic characteristics of the signal spectrum, and were also used in our reference work [2].

MFCC data were extracted with the Librosa library [1] using 13 coefficients, windows of 2048 samples and 75% overlap between windows. That sums up to 10.844.508 frames extracted for the whole dataset.

3.3 Codeword Representation

Codeword Representation is a technique for representing high cardinality data, allowing data to be clustered in fewer groups of similar elements, and representing each sample as a histogram. The first idea behind this technique is to

[1] http://github.com/librosa/librosa

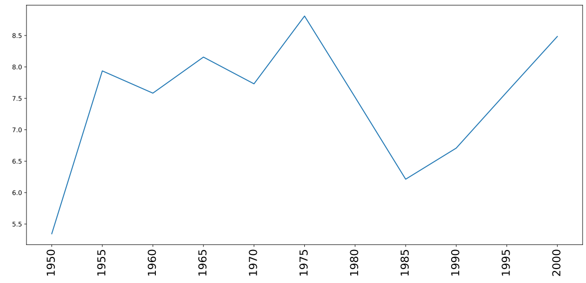


Figure 2. The entropy calculated in bits for each 5-year period.

apply unsupervised clustering techniques (e.g. Kmeans) to estimate how many clusters would ensure that the data can be well separated. In our case we used 10 clusters to represent all MFCC arrays extracted from all songs. At this point each MFCC is identified as belonging to one of the 10 clusters, and each song can be seen as a temporal succession of transitions betweens clusters. These transitions are counted for each song, ending up with a histogram indicating its distribution of MFCCs over the 10 clusters. The histograms are normalized with respect to time and become the Codeword Representation of each song.

3.4 Diversity metric

According to Mauch et al. [2], maximum diversity is achieved when frequencies are uniformly distributed in the histogram, and minimum diversity corresponds to all MFCCs belonging to a single cluster. As suggested by these authors, we take Shannon's entropy function as a measure of diversity. The average proportion of frames over each cluster $\bar{q}$ for a given 5-year period is given by

$$\bar{\mathbf{q}} = (\bar{q_1}, \bar{q_2}, \ldots, \bar{q_{10}}). \quad (1)$$

We calculate the diversity defined as

$$D = exp\left(-\sum_{i=1}^{10} \bar{q_i} \ln \bar{q_i} \right) \quad (2)$$

The maximum entropy value is attained when all $\bar{q_i}$ are equal and $D = 10$. The minimum value occurs when only one cluster is represented, and $D = 1$.

4. RESULTS

In Figure 3 it is possible to notice how the probabilities are distributed over the decades (horizontal axis) and over the clusters (vertical axis). There are periods when these probabilities are more uniformly distributed through all clusters (1970, 1975,1980), in contrast to periods when they are much more concentrated in fewer clusters (1950,1985,1990). It also possible to see clusters that present almost constant proportion over the decades, as the case of the cluster number 8, in opposition to cluster 7 that presents a peak in 1950 and then decays with time.

Another visualization in the right side of Figure 3, with the same values from the previous matrix but with columns and rows sorted by similarity. Two dendrograms are presented as representing the similarity of periods (columns) and clusters (rows), the tree-like grouping is built from euclidean distances between elements of the array. The closest periods, as indicated in the upper dendrogram, are 1970 and 2000, followed by the next most similar pair, 1965 and 1990. 1950 is considerably different when compared to all other periods.

The entropy is then calculated for each group of songs of the set and the results are shown in Figure 2. 1975 presented highest degree of diversity when compared to the other periods, 1950 presents the lowest one.

Mauch et al. [2] applies a similar measurement of diversity in a similar period of time, but mixing timbre and harmonic features, and finds the lowest value in 1985. After then it starts to increase and reach its highest value around 2000 for the case of American Popular Music, similarly to what was presented here.

5. CONCLUSIONS

We applied a part of the methodology from Mauch et al. [2] while trying to transpose the same analysis to a different music dataset. The results of the entropy-based diversity metric bring interesting trends for discussion, which might lead to interesting musicological interpretations and insights.

6. REFERENCES

[1] J. Serrà, A. Corral, M. Boguñá, M. Haro, and J. L. Arcos, "Measuring the evolution of contemporary western popular music," *Scientific Reports*, vol. 2, 07/2012 2012. [Online]. Available: http://www.nature.com/srep/2012/120726/srep00521/full/srep00521.html

[2] M. Mauch, R. M. MacCallum, M. Levy, and A. M. Leroi, "The evolution of popular music: Usa 1960–2010," *Royal Society Open Science*, vol. 2, no. 5, 2015. [Online]. Available: http://rsos.royalsocietypublishing.org/content/2/5/150081

[3] E. Pampalk, "Computational models of music similarity and their application in music information retrieval," Ph.D. dissertation, Vienna University of Technology, Vienna, Austria, March 2006. [Online]. Available: http://www.ofai.at/~elias.pampalk/publications/pampalk06thesis.pdf

[4] J. M. Grey, *An exploration of musical timbre*. Dept. of Music, Stanford University, 1975, no. 2.

[5] R. S. Brazil, ""Os 100 maiores discos da Música Brasileira" (100 greatest Brazilian music records list)," http://rollingstone.uol.com.br/listas/os-100-maiores-discos-da-musica-brasileira/, 2018, [Online; accessed 27-April-2018].

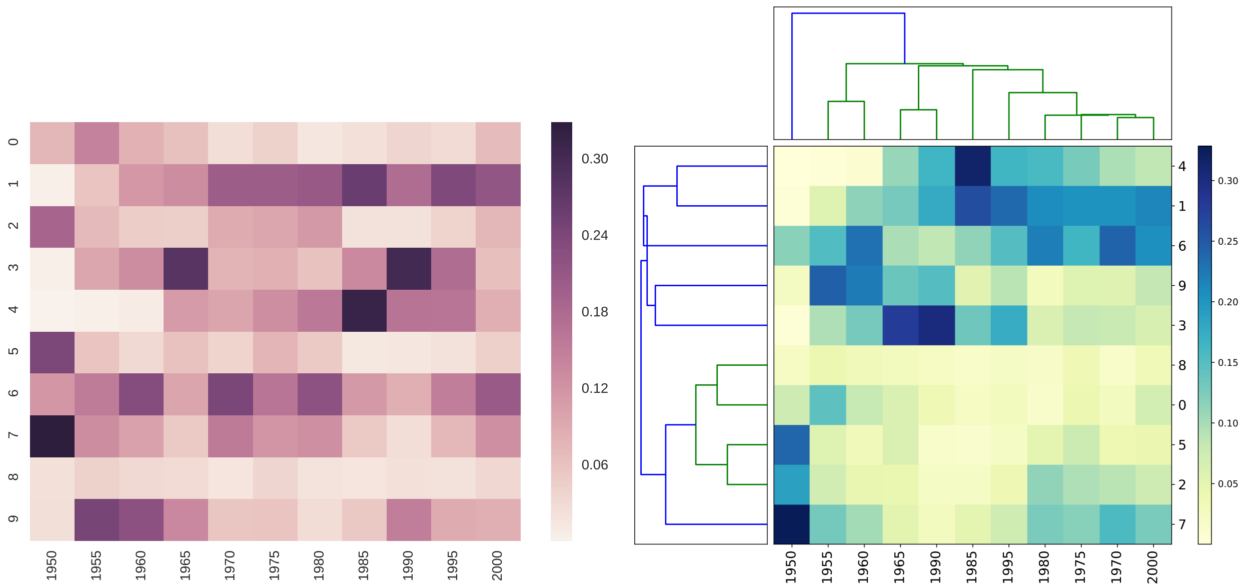


Figure 3. Left: Probabilities of each cluster summed for each 5-year period. Right: Matrix with the probabilities distributed over the clusters and over the 5-year periods, sorted by similarity between columns and rows. The two dendrograms (top and left) indicate the euclidean distances between periods and clusters, respectively.

ID	Artist	# records	# songs	ID	Artist	# records	# songs
0	Caetano Veloso	6	55	31	Júpiter Maçã	1	14
1	Gilberto Gil	5	48	32	Nélson Cavaquinho	1	13
2	Os Mutantes	5	44	33	Secos e Molhados	1	13
3	Roberto Carlos	4	48	34	Itamar Assumpção	1	13
4	Jorge Ben	4	45	35	Blitz	1	13
5	Gal Costa	4	38	36	Elizeth Cardoso	1	13
6	Tim Maia	3	42	37	Tom Zé	1	12
7	Racionais Mc's	3	32	38	Ângela Rorô	1	12
8	João Gilberto	3	30	39	O Rappa	1	11
9	Tom Jobim	3	30	40	RPM	1	11
10	Chico Science/Nação Zumbi	2	37	41	Erasmo Carlos	1	11
11	Sepultura	2	32	42	Os Paralamas Do Sucesso	1	11
12	Milton Nascimento	2	32	43	Ultraje a Rigor	1	11
13	Los Hermanos	2	29	44	Maria Bethânia	1	11
14	Raul Seixas	2	27	45	Luiz Melodia	1	10
15	Mundo Livre S/A	2	27	46	Moacir Santos	1	10
16	Titãs	2	26	47	Arnaldo Baptista	1	10
17	Marisa Monte	2	25	48	Banda Black Rio	1	10
18	Cartola	2	24	49	Novos Baianos	1	9
19	Legião Urbana	2	23	50	Rita Lee & Tutti Frutti	1	9
20	João Donato	2	22	51	Gilberto Gil; Jorge Ben	1	9
21	Paulinho da Viola	2	22	52	Arrigo Barnabé	1	8
22	Elis Regina	2	22	53	Aracy de Almeida	1	8
23	Ira!	2	20	54	João Gilberto; Stan Getz	1	8
24	Chico Buarque	2	20	55	B. Powell; V. de Moraes	1	8
25	Dorival Caymmi	2	14	56	Egberto Gismonti	1	8
26	N. Leão; Z. Kéti; J. do Vale	1	23	57	Caetano; Gal; Gil; Mutantes	1	2
27	Doces Bárbaros	1	17	58	Caetano Veloso;Gilberto Gil	1	1
28	Raimundos	1	16	59	Caetano Veloso; Gal Costa	1	1
29	Walter Franco	1	14	60	Nara Leão	1	1
30	Elis Regina; Tom Jobim	1	14				

Table 1. ID, artists, number of records and songs present in the database. We have gathered different names used by the same artists: e.g. Tom Jobim = Antônio Carlos Jobim; Os Mutante = Mutantes; Jorge Ben Jor = Jorge Ben

EFFORT IN GESTURAL INTERACTIONS WITH IMAGINARY OBJECTS IN THE CONTEXT OF DHRUPAD VOCAL IMPROVISATION

Stella Paschalidou
Dept. of Music Technology & Acoustics Eng., TEI of Crete, Greece
pashalidou@staff.teicrete.gr

Martin Clayton
Dept. of Music, Durham University, UK
martin.clayton@durham.ac.uk

Tuomas Eerola
Dept. of Music, Durham University, UK
tuomas.eerola@durham.ac.uk

ABSTRACT

This paper examines relationships between the voice and the ecological knowledge of how the human body moves when interacting with objects of the environment in the context of Hindustani vocal music. In the Dhrupad genre of singing improvisation vocalists often appear to engage with melodic ideas by manipulating intangible, imaginary objects with their hands, such as through stretching, pulling, pushing etc. Our main focus is on how much effort each of such gestures is perceived to require and whether and how this is linked to its melodic counterpart.

The work makes use of both qualitative and quantitative methods in a sequential order on original recordings of interviews, audio-visual material and 3D-movement data of Dhrupad vocal improvisations. Findings indicate that there is a certain level of consistency in gesture-sound links despite flexibility caused by idiosyncratic movement patterns of individuals. Different schemes of cross-modal associations were revealed for the singers analysed, that depend on the pitch space organisation of melodic modes (rāgas), the mechanical requirements of voice production, the macro-structure of the (ālāp) improvisation and morphological cross-domain analogies. Results further suggest that a good part of the variance in both physical effort and gesture type can be explained through a small set of sound and movement features.

1. INTRODUCTION

1.1 Background

The current paper takes an embodied approach in the analysis of Dhrupad vocal improvisation. Dhrupad is a sub-genre of Hindustani music that relies heavily on improvisation, which is rule-based and conforms to the 'rāga' (modal) system. The notion of a rāga as a movement in a melodic pitch 'space' [1] is also accompanied by smooth hand movements in the real space (deliberate or unconscious).

It is often the case that listeners and performers alike will report that they experience virtual worlds of forces in motion in relation to music and sound [2, 3]. In fact, in Hindustani music (and the Dhrupad genre in particular) singers make a frequent use of linguistic expressions for motor-based metaphors alluding to the sensation of a resistive force that the agent (the performer) needs to fight against. Additionally, during vocal performances singers seem to engage with the melody and its intricate qualities by employing and manipulating imaginary objects. They stretch, pull, push, collect, throw and execute other movements, whereby they appear to be fighting against or yielding to some imaginary resistive force. Although the object is not real, we observe them executing repeated patterns of bi-manual effortful gestures, comprising gripping (closing the hand), action and releasing phases [4].

We consider such voluntary imitations of interactions with real objects (that would not necessarily produce any sound) of particular interest to embodied music cognition research, as it has been previously argued that musical thinking is grounded in the ubiquitous patterns of actions we possess through our ecological knowledge of interacting with objects of the real world [5, 6]. Such Manual Interactions with Imaginary Objects (that we will call MIIOs) offer a special case where the hands—although free to move (see previous studies on free-hand sound tracings [7, 8]—are deliberately constrained by the conception of an object. On the other hand, due to the absence of a real mediator, they may allow significant cognitive processes to be revealed, that are associated with more fundamental concepts than unequivocal mechanical cross-modal couplings of a particular instrument.

1.2 Aim

The current paper offers a systematic examination of whether and how the kinaesthetic sensation of effort that is conveyed through movement during MIIOs relates to the voice in Dhrupad vocal performances. In other words, the paper explores whether the link between gestures and sound seems to reside in the interaction possibilities and the levels of effort that these virtual objects afford [9] according to their physical properties, such as viscosity, elasticity, weight, friction etc. Effort is understood here in its common usage, which reflects our understanding of how hard a person must work, either mentally or physically, in order to achieve an intended goal. Although effort in music has often been regarded as an important aspect of both listening and performing [10], systematic approaches to its role have been limited [11].

Singing gestures have drawn little attention in the field of embodied music cognition, although there is a growing body of work on Hindustani music (e.g. [12, 13, 4, 14, 15]). None of the existing works has however studied the Dhrupad genre and none has dealt with MIIOs and their perceived levels of effort. Dhrupad singing offers a distinctive case for studying underlying links between sound and effort, as it exhibits easily identifiable imitations of real manual interactions with objects and a plethora of smooth melodic glides (mīṇḍ) where notes are approached with a sense of pitch continuum, which makes it a good case for studying the non-discrete nature of the mechanics of gestures. It is also a suitable case of music making due to the strikingly slow melodic progression and the rigour and precision of intonation by singers. Finally, the 'oral' way of music transmission means that students do not rely on written notation, which allows the transmission of embodied knowledge through direct visual engagement of disciples with their teachers during teaching sessions. However, physical metaphors in musical thinking are not a Dhrupad exclusivity [1] and therefore, despite the specificity of the genre, the paper aims to address concerns in the study of gesture-sound relationships that are of interest to the wider research community of embodied music cognition and thus outcomes may be extended to other music lineages.

2. METHODOLOGY

The work combines qualitative and quantitative methods applied on original recordings of interviews, audio-visual material and 3D-movement data (using a 10-camera Optitrack passive marker system) of vocal improvisations by fourteen vocalists of the same music lineage (all students of Zia Fariduddin Dagar, including the maestro himself). Real recordings of performances were chosen rather than designed laboratory experiments of subjects responding to stimuli, in order to reveal robust gesture-sound links that vocalists may have established over years of practice rather than spontaneous responses to stimuli by listeners. These performances were collected for the specific study in domestic spaces in India (2010-2011). In order to avoid the metrical structure and the lyrical content of the later stages of the performance and only concentrate on melodic factors, only the slow non-metered section of the improvisation (ālāp) was used, which is sung to a repertoire of non-lexical syllables. The qualitative part of the analysis includes a thematic analysis of interview material by six vocalists and a video observation analysis of four vocalists. The quantitative part of the analysis involves the development of mathematical descriptions for the classification of MIIOs and the inference of effort levels by fitting (linear) models to a small set of movement and sonic features for two of the performers.

The thematic analysis of interview material aimed to reveal cross-modal concepts embedded in explicit and implicit knowledge by the interviewees and also to ground the coding scheme that was later developed during the video analysis stage. The analysis qlso intended to highlight the importance of visual and motor imagery in the conception of music as movement by musicians and to concentrate in particular on how the extension of the body through imaginary objects and materials may be facilitating engagement with melodic ideas. The interview material was therefore collected and thematic analysis was applied to recurrent sensorial and pictorial descriptors (adjectives, verbs and nouns) of motor-based metaphors and performer-object interactions when talking about sound and music.

The video observation analysis was performed in the ANVIL annotation environment and it aimed to use third-person observations in order to identify and classify types of MIIOs and melodic motives, to find recurrent gesture-sound associations and to draw some first conclusions about the level of consistency in their co-appearance for a single or across performances. The coding scheme was informed by findings of the previously conducted thematic analysis but mainly emerged in ANVIL progressively during multiple viewings of the video footage for each individual performer/performance. This was done by visually identifying, segmenting, labeling and classifying repeated patterns of (90) manual gesture events that allude to MIIOs. These gestures were also annotated in terms of melodic movement type, pitch interval, octave range, sung syllable and melodic context (intention of the melody in moving towards the tonic immediately after the annotated phrase) of their melodic counterpart, as well as in terms of amount of effort (on a scale between 0-10, 0 being the highest) that each gesture was perceived to require. The most prominent gesture types varied between stretching an elastic object and pulling or pushing away a rigid object. An inter-coder validation of gesture class annotations was also carried out by two professional dancers/choreographers for one of the performances, in order to assess whether the annotations of the main coder could be considered reliable. Finally, a gesture-sound association analysis was conducted in order to identify recurrent associations between categorical aspects of the hand movements and the voice.

The quantitative part of the analysis aimed at exploring whether it would be possible to computationally infer the visually annotated effort levels and gesture classes (in terms of interactions with elastic (stretching, compressing) versus rigid (pulling, pushing-away, collecting, throwing) objects) based on a small set of movement and sound features extracted from the captured material. This stage of the analysis also aimed to examine to what extent models might mostly reflect idiosyncratic elements of gesturing or more generic gesture-sound links shared across vocalists. Therefore, two sets of linear models were devised, those which best fit to the data of each individual performer (describing rather idiosyncratic aspects of gesture-sound links for each vocalist) and those that—despite their lower goodness of fit—overlapped to a greater extent across performers (displaying a more generic power over performers). Models considered as best fitting were those that displayed the best trade-off between model accuracy, compactness and simplicity in interpretation. The models were developed for the performance of Afzal Hussain in rāga Jaunpurī (scale: 1,2,*b*3,4,5,*b*6,*b*7, 18 minutes) and that of Lakhan Lal Sahu in rāga Mālkauns̒ (scale: 1,*b*3,4,*b*6,*b*7, 23 minutes). The explorative character of this process meant that a considerable part of the process involved finding a set of most important descriptors to fit the models to. Initially, the (15) most relevant gesture and audio features reported in [16] on sound-tracing experiments were used, but then a number of alternative motion/acoustic variables was progressively added as to raise the goodness of fit while retaining the compactness of the model. These were high-level statistical global measures (such as mean, SD, min, max) computed from a number of time-varying audio and movement features that derived from the raw data. Although far more sophisticated features may be probed, derivatives of position coordinates were chosen as they have proven to be robust and pertinent in other music performance contexts [17] and in sound tracing experiments [18, 19, 20].

3. RESULTS

Interview analysis revealed a high visual element in the conceptualisation of music and a heavy use of motor-based metaphors expressed through physically inspired linguistic descriptors that allude to the sensation of resistive forces a performer needs to fight against, such as elasticity/stretching, pressure/pushing, weight/lifting and friction/scratching. These recurrent descriptors were organised in meaningful ways in order to inform the coding scheme of the video analysis in later stages. The transmission of musical knowledge through visual engagement and imitation of gestural habits between teachers and disciples was also acknowledged, not perhaps as exact replicas but in terms of shared movement qualities. Some vocalists made even explicit statements about categorical gesture-sound associations.

The association analysis carried out between categorical aspects of gesture and voice during the video analysis stage resulted in a consistent link that is shared between performers and mainly reflects shared cross-modal morphologies. However, a few other modes of association where also revealed, such as a strong link between interaction type and effort level, with the stretching gesture of an elastic object requiring higher levels of effort than pulling/pushing away a rigid object, as well as a strong link between specific types of gestures with specific pitch areas of the raga and with the melodic context (the intention of the melody to ascend to or towards the tonic) in the case of Afzal Hussain.

By using the manual annotations of the preceding video analysis as ground-truth responses—or other words the correct output values—a number of compact linear models were devised in the quantitative part of the analysis, according to which the null-hypothesis that voice and gesture are unrelated was rejected. By using a small set (of four or five) non-collinear movement and sound features it was possible to estimate a good part of the variance in gesture classification and effort inference.

3.1 Idiosyncratic schemes

3.1.1 Effort level inference

Differing idiosyncratic schemes of associating effort with acoustic and movement features were identified, that are based on the pitch space organisation of the rāga as well as the mechanical requirement for producing the voice.

Hussain: The use of 5 non-collinear audio and movement features yielded a good fit of about 60% (R^2adj). According to this model, higher effort levels are required when the hands move slower and further apart and exhibit a larger speed variation. These are accompanied by melodic glides that start from lower degrees and ascend to higher degrees of the rāga scale in the range of an individual octave, thus they are considered to be associated with characteristic qualities of the specific rāga.

Sahu: The use of 4 non-collinear audio and movement features yielded an adequately good fit of about 44% (R^2adj). According to this model, higher bodily effort is required for hand movements that exhibit a larger variation of hand divergence (speed in moving the hands further apart) and a strong onset acceleration. They are accompanied by larger melodic glides that ascend to higher maximum pitches, thus reflecting the increased mechanical requirement of voice production. This may also be reflecting the gradual ascent towards the pitch climax of the ālāp macro-structure.

3.1.2 Gesture classification

Different modes of gesture class association with acoustic and movement features were revealed, especially with pitch regions of particular interest for the specific rāga and analogous cross-domain morphologies.

Hussain: The use of 5 non-collinear audio and movement features in the logistic models that were developed yielded a high classification rate of about 95% (AUC). According to this, it is more likely that interactions with elastic objects (rather than rigid) are performed by hand gestures that exhibit a low absolute mean acceleration and a large variation in hands' divergence. They are accompanied by slower and larger melodic movements that ascend to a higher degree of the scale.

Sahu: The use of 4 non-collinear audio and movement features in the logistic model yielded an adequately good fit of about 80% (AUC). According to this model, interactions with elastic objects are more likely performed with pitch movements of a larger interval and larger duration and with the hands moving faster and remaining bound to each other.

3.2 Generic scheme

3.2.1 Effort level inference

Two almost identical linear models were developed, yielding a good fit of about 53% (R^2adj) with 5 features for Hussain and 42% with 4 features for Sahu respectively. According to these models, higher bodily effort levels are required by both singers for melodic movements that start from lower and reach up to higher pitches, reflecting the mechanical strain of voice production. They are accompanied by movements which are slow on average but with a large variation of speed, and in the specific case of Hussain with the hands moving further apart.

3.2.2 Gesture classification

Two almost identical general logistic models were developed, yielding a good fit of about 86% (AUC) with 3 features for Hussain and 78% with 4 features for Sahu respectively. According to these models, interactions with elastic objects are more likely to be performed at lower pitches during larger melodic movements, and with the hands moving further apart for Hussain and less apart but faster in the case of Sahu.

4. CONCLUSIONS

The current paper has reported on the first study of gesture-sound associations during MIIOs in the Dhrupad genre of Hindustani vocal music. It has contributed to a better understanding of the role of voluntary imitations of interactions with objects of the real world by using original material of interviews and recordings of real performances captured in the field in India and by proposing a novel sequential methodological approach that comprises both qualitative and quantitative methods. MIIOs offer a particular case where motor imagery is "materialised" through physical actions directed towards an imagined object.

Findings indicate that despite the flexibility in the way a Dhrupad vocalist might use his hands while singing, the high degree of association between classes of virtual interactions and the voice provides good evidence for non-arbitrariness and generic associations that are not necessarily performer-specific or stylistic. Such association may be grounded in some shared cross-modal morphologies, but in some cases also in the melodic organization of the ālāp improvisation. Regardless of the idiosyncratic element in gesturing noticed by individual performers, it could be still argued that the type of imagined object and the nature of the imagined opposing forces reflect qualities describing melodic aspects of the voice.

Combining ethnographic data with exact measurements of real performances has brought about the advantages of ecological validity, however it has also posed a number of important challenges and limitations, such as the small dataset. Collecting a larger dataset of multiple performers, performances and rāgas for each performer will allow a more systematic comparison between performers, performances and rāgas.

5. REFERENCES

[1] Fatone, G. A., Clayton, M., Leante, L., & Rahaim, M. (2011). Imagery, melody and gesture in cross-cultural perspective. In Gritten, A. & King, E. F. (Eds.), *New Perspectives on Music and Gesture* (pp. 203-220), Surrey, Ashgate.

[2] Stern, D. (2010). *Forms of vitality : Exploring dynamic experience in psychology, the arts, psychotherapy, and development*. Oxford ; New York: Oxford University Press.

[3] Eitan, Z., & Granot, R. Y. (2006). How Music Moves. *Music Perception, 23,* 3, pp. 221-248.

[4] Rahaim, M. (2009). *Gesture, Melody, and the Paramparic Body in Hindustani Vocal Music* (Doctoral thesis, University of California, Berkeley, USA). Retrieved from http://gradworks.umi.com/33/69/3369126.html (last accessed December 15, 2016).

[5] Noë, A. (2004). *Action in perception* (Representation and mind). Massachusetts ; London: MIT Press.

[6] Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind : Cognitive science and human experience*. Cambridge, Mass.: MIT Press.

[7] Godøy, R. I., Song, M., Nymoen, K., Haugen, M. R., & Jensenius, A. R. (2016). Exploring Sound-Motion Similarity in Musical Experience. *Journal of New Music Research, 45,* 3, pp. 210-222.

[8] Caramiaux, B., Bevilacqua, F., & Schnell, N. (2010b). *Mimicking sound with gesture as interaction paradigm*. Tech. rep., IRCAM-Centre Pompidou.

[9] Gibson, J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: toward an ecological psychology* (pp. 67–82). Hillsdale, NJ: Lawrence Erlbaum.

[10] Krefeld, V., & Waisvisz, M. (1990). The Hand in the Web: An Interview with Michel Waisvisz. *Computer Music Journal, 14,* 2, pp. 28-33.

[11] Bennett, P., Ward, N., O'Modhrain, S., & Rebelo, P. (2007). DAMPER: a platform for effortful interface development. In *Proceedings of the 7th International Conference on New Interfaces for Musical Expression* (pp. 273-276), New York City, USA.

[12] Clayton, M. (2007). Time, gesture and attention in a Khyāl performance. *Asian Music*, *38*(2), 71-96.

[13] Leante, L. (2009, November). The lotus and the king: imagery, gesture and meaning in a Hindustani rāg. In *Ethnomusicology Forum* (Vol. 18, No. 2, pp. 185-206). Taylor & Francis Group.

[14] Pearson, L. (2013). Gesture and the sonic event in Karnatak music. *Empirical Musicology Review*, *8*(1), 2-14.

[15] Kelkar, T., & Jensenius, A. R. (2018). Analyzing Free-Hand Sound-Tracings of Melodic Phrases. *Applied Sciences*, *8*(1), 135.

[16] Nymoen, K., Godøy, R. I., Jensenius, A. R., & Torresen, J. (2013). Analyzing correspondence between

sound objects and body motion. *Acm Transactions on Applied Perception, 10,* 2, pp. 1-22.

[17] Rasamimanana, N. H., Flety, E., & Bevilacqua, F. (2006). Gesture Analysis of Violin Bow Strokes. *Lecture Notes in Computer Science,* 3881, pp. 145-155.

[18] Nymoen, K., Torresen, J., Godoy, R. I., & Jensenius, A. R. (2012). A Statistical Approach to Analyzing Sound Tracings. *Lecture Notes in Computer Science,* 7172, pp. 120-145.

[19] Nymoen, K., Caramiaux, B., Kozak, M., & Torresen, J. (2011). Analyzing sound tracings: a multimodal approach to music information retrieval. In *Proceedings of the 1st International ACM workshop on Music information retrieval with user-centered and multimodal strategies* (pp. 39-44), Scottsdale, AZ, USA.

[20] Glette, K. H., Jensenius, A. R., & Godøy R. I. (2010). *Extracting action-sound features from a sound-tracing study*. Tech report, University of Oslo, pp. 63-66.

ANALYSIS OF THE SONORITY: AN APPROACH BASED UPON THE PERFORMANCE

Bibiana Bragagnolo

UFPB
bibi_bragagnolo@hotmail.com

Didier Guigue

UFPB/UNICAMP/CNPQ
didierguigue@gmail.com

ABSTRACT

This research proposes a methodology of analysis of the sonority in which the performance becomes a source of information, focusing on piano pieces written by Brazilian composers. For that, our methodology includes the methodology of analysis of the sonority developed by the second author of this paper [1] and englobes two main elements: the analysis of recorded data with computational support, and the approach of the Artistic Research [2]. The first element encompasses recordings of the selected pieces played by the pianist and first author of this paper, recorded in audio and MIDI and analyzed using Sonic Visualizer and Open Music. The interpretation of this information was based on the written and performative elements of the sonority in the pieces. The analysis of the data was filtered by the performer's view, which is the main idea of the Artistic Research. The application of this methodology has already provided some consistent results in the analysis of the sonority in *Ressonâncias* and *Contrastes*, by Marisa Rezende.

1. INTRODUCTION

Analysis and performance have been separated activities in music, and their relations were often in terms of prescription; the analysis informs the performance, and never otherwise. This research aims to propose a methodology of analysis of the sonority in which the performance becomes a source of information, focusing on piano pieces written by Brazilian composers.

This understanding is based on the acceptation that the written text is not the music, and even if it is written in a very specific way, it cannot encapsulate all musical features and the interpretative decisions [3]. The starting point of our methodology is our own model for an analysis of the music based on its sonorities [1] and it includes two main elements: the analysis of recorded data with computational support and the approach of the artistic research [2].

This paper is part of a major doctoral research, in development at the Federal University of the State of Paraíba (UFPB), Brazil. The goal of this doctoral research is to develop a methodology of musical analysis, more specifically for the analysis of the sonority in piano pieces along with the musical elements given by the score.

In this paper we aim to present an experimental method for the analysis of sonority that was applied in two piano pieces: *Ressonâncias* (1983) and *Contrastes* (2001), by the Brazilian composer Marisa Rezende (1944). These pieces were selected because of the importance of the sonority in their construction, both in the performance and in the compositional project.

2. THE MORPHOLOGICAL UNDERSTANDING OF THE MUSICAL WORK

The construction of the sonority in this research is understood as the result of the interaction between performer and text (score). This understanding is supported by the idea of music as performance [3]. When the musical text is understood as something that can represent the entire musical meaning, many musical aspects, such as the performative ones, are missed. The performer, with his or her own subjectivity, musical view, background, and personal musical decisions took upon the score, contributes as an active subject to the construction of musical meaning. There is no score capable to encapsulate all the nuances and possibilities that a musical score proposes. The methodological possibility that will be presented has an experimental nature, once the musical source of meaning is no longer the written text, but the performance, through which the score has already been interpreted and filtrated by the performative view and transformed into real sound.

One of our main theoretical references for this research is the theory that defines the morphology of the musical work [3]. This theory proposes a shift from the ontological to the morphological concept of music. The morphological question verses on the perceptual aspect of music and on the transformations suffered from performance to performance and how these transformations occur. In other words, we shift from an investigation of the musical work as an ideal represented by the score to the investigation of it as music in act, and as performance. This emphasis on the process allows a more flexible attitude towards the musical notation, which no longer defines a fixed musical object, but fixes the conditions of performance to achieve a specific sonorous result.

3. ARTISTIC RESEARCH

The Artistic Research [2] is a kind of understanding that switches the research focus to the performer's view and, most commonly, in one specific performance. The justification for this kind of research encapsulates both the possibility of the inclusion of a deeper knowledge about the musical practice and the idea that one specific performance of a piece brings in itself important elements about the work as a whole.

Since the beginning of the establishment of the research in music as a scientific knowledge in the early XX Century until little time ago, performance was always subjugated as a field of research. For a long time the phenomena related to the performance were always analyzed through the perspective of other fields, as Musicology, Psychology, Musical Analysis, etc. However, only recently the performer himself was invited to contribute as a researcher, and not as an object of study, in researches regarding the performative processes. Despite being the creator of the music as sound, the voice of the performer has been muted for a long time.

As a natural consequence of the process that is happening nowadays in the contemporary Musicology, the detachment of the understanding of music as score and an approximation to the understanding of music as process/ action, the performer sees himself brought to the center of the debate. This change allowed the development of new methodologies of research, such as the Artistic Research, which is gaining emphasis in the academic production in the field of performance in the last decades.

In this kind of research, the performer acts as a researcher, in the traditional sense of the word, without losing his role as an artist. At the same time that he observes, he is observed by himself, in a way that "the artist investigates his own practices, materials and fonts" [2]. This means that a double role is attributed to the same individual, to generate an auto-reflexive process that is not-transferable and personal. This very notion of the Artistic Research is something that defies our culture, which is habituated to stablish the research through results that can be confidently transferred to other contexts [2]. However, it is the only possibility to truly understand truly the artistic and creative processes and to allow an approach totally focused on the process. In this way, the Artistic Research involves a de-territorialization of the research culture in arts, "which is nowadays still dominated by scientific and epistemic approaches" [2].

The main goal of the proposed analysis is to add the interaction between score and performer as a prime source of musical meaning, as something that will, in fact, determine musical structures. Researches where these non-textual elements are included often use adapted theoretical background from other areas (such as Linguistics). Our goal is to propose a methodology of analysis totally developed regarding the musical phenomena.

4. A METHODOLOGICAL PROPOSITION

The analysis of the selected pieces is focused on one specific performance, the one made by the pianist and first author of this paper. The interpretation of the pieces prioritized mainly the manipulation of sonority, bringing up aspects of piano timbre that can be manipulated by the performer. As methodological support, this analysis used records of the piece in public performances and in private study sessions and also a study diary. This diary brought information about the development of the piece, since the first reading of the score until the moment of its public performances. The sonority was analyzed through the recordings made during the learning and performing phases. In a first moment, the performative decisions served to separate the piece into sonority units. Done this identification, the main aspect of the sonority in each one of these units was analyzed by adopting our method for the analysis of sonority [1].

To start the reflections about the sonority in the piano, we brought the table below [4], to better understand which are in fact the elements available for the performer interference in terms of sonority. We have, in the first column, a list of the constitutive elements of timbre and, on the right, the level of control that the performer has in these elements, and how this control is possible.

Constitutive elements of timbre	**Level of control**
Number of strings put in vibration	Medium (*una corda* pedal, with no effect in grave)
Enarmony (of string, affination)	Null
Number of partials	Null
Number of partials produced in the audible area	High – velocity of attack, pedals
Relations of frequency between the partials	Very low, UC pedal in some cases
Location of partials in relation to the sensible zone of the ear	Null
Relative level (dB) of partials (in the same intensity)	Null
Partial's level of acoustic pressure	High, velocity of attack
Order and velocity of the emergence of partials	Medium, velocity of attack, articulation

Order and velocity of the extinction of partials	Relatively high (except at the upper register) relaxation in touch and/ or sustain pedal
Relation noise (of mechanism)/ sound	Very low, kind of attack
Reverberance, resonance	High, through pedals and specific touches
Velocity of hammer	High, velocity of attack
Movement of hammer	Low, kind of attack, velocity of attack
Duration and local of hammer contact with the string	Null
Action of dampers	High, relaxation of touch, sustain pedal
Harmonic table	Null

Based upon this table, it is possible to affirm that the performer has some important tools available to manipulate the piano timbre, and this manipulation will be the result of his or her interpretative decisions. With this knowledge and the practical contact with the pieces, some considerations about the sonorous structure of the piece were made. From the point of view of performative decisions, we can state that questions as the use of the sustain and *una corda* pedals, as well as the articulation (comprehended as the choices of pianistic touch), will be important elements for the analysis, because they are decisions of the performer that will directly affect the sonority of the piece and, consequently, its morphology. These sonic characteristics are understood as the variant elements of the piece, that might change from performer to performer, but that are important elements of the sonic constitution of the piece.

Also crucial is the identification of the invariant elements, the ones secured by the score, that will remain the same from performance to performance. In the case of the pieces analyzed, the register appeared as the most important element of stability in the written aspect of the sonority.

It is important to reinforce that, in this analysis, both the division into sonority units and the identification of the main aspect of the sonority in each unit will be done based on the practical relation with the piece, taking into account this relation as a performer, not only as an analyst.

5. METHODOLOGICAL PROCEDURE

The first element of our methodological procedure encompasses the recordings of the selected pieces in audio and MIDI. These recordings were made with piano Yamaha Clavinova CVP-701 and played by the pianist and first author of this paper. The MIDI data were generated by the piano and the AUDIO data was recorded using a camera SONY HDR-MV1, both recordings done concomitantly.

The data obtained in these recordings were analyzed using the software Sonic Visualizer and Open Music. The interpretation of this information was based upon the written elements of the sonority and also on the performative elements of the sonority, such as the use of the pedals and the choices of pianistic touch. The analysis of the data was filtered by the performer's auto-reflection about the piece's interpretation.

Through the data from the MIDI files, we extracted information about Velocity and Sonic Basic Quality (more about below), while from the AUDIO files, we extracted information about spectral centroid and spectrograms.

This procedure of analysis is oriented by the concept of *Compound Sonic Unit* and its relative complexity. A Compound Sonic Unit is a synthesis, at a given point of the piece's timeline, of a certain number of secondary components that interact in complementarity (e.g. density, harmonicity, periodicity, entropy, a.o.) and whose *relative complexity* can help to build a representation of the dynamic curve of the sonorities of a work. Each component is weighted onto a simple-to-complex vector according to its inner configurations or behavior, so that a more complex configuration is expected to render a relatively more complex Sonic Unit.

It is also relevant to reinforce that the sonic units were detected in these analysis departs from the experiential contact with the piece, with a performer's positioning. This means that the interpretative choices regarding variations in dynamics, timbre, pianistic touch, pedals and timing are the main elements that shape these sonic units. In these analyses the sonic units are determining the piece's contour in homogeneous sonority sections. Each different section has specific sonorous characteristics and each change of section is linked to changes in these characteristics, which in these pieces essentially takes into account elements as: use of the pedals, register and pianistic touch.

The MIDI files allowed us to extract and analyze two main elements: velocity (V) and the sonic basic quality (Q). For this task, we used the *SonicObjectAnalysisLibrary*, an Open Music library dedicated to music analysis [1] [5]. We first used the SOAL function *velocity-per-onset*, which read the file and returned the MIDI velocity values (from 0 to 127) for each note played at each onset (e.g. at

[1] This library is developed by the research group MUS3 (NICS/UNICAMP).

each attacked note(s)), to interpret diachronically the evolution of the pianistic touch.

The other element, the Q function in the library, needed to collect three sets of data: (1) the register (related to the localization of notes into the piano gamut, obtained from the analysis of MIDI notes), (2) the relative intensity of the played notes (calling again the MIDI velocities data), and (3) the use of the pedals, including the sustain pedal and *una corda*, and every combinations between the two pedals, including the use of half pedals [5]. The model for the calculation of the Q value of a given piano note is based upon the principle of a general (but nonlinear) decline of the timbric complexity of each pitch in proportion to its fundamental frequency – the higher the fundamental, the less complex its timbric structure [5]. The three reasons for this decline are: (1) the decrease of the number of audible partials for a given fundamental, (2) the decrease of the position in the spectrum of the stronger partial/s, and (3) the decrease of the duration of the sound's extinction phase [4]. This idiomatic decline is more or less modulated by the second (velocities) and third (pedals) components. In this way, while V brings specific information about the performer's approach in the instrument, revealing the exact impact of each note played (what will affect both the dynamics and the timbre in the piano), Q brings a synthesis between a static data (pitches) and a dynamic data (interpretative decisions, represented by velocities and pedals) and generates a value that represents the resonance.

In complementation to the MIDI data, we also analyzed the AUDIO files of the works, in order to evaluate the impact of the interpretative decisions onto the resulting sound of the pieces. For this task, we used Sonic Visualizes, a software that allows the visualization of different aspects of the sonic signal [6]. At this point we used specifically two tools for analysis: the spectrogram and the spectral centroid. The visualization of these two representations of the sound can reveal important information about it. The spectrogram shows data such as the number of harmonics produced for each played note, the localization and duration of these, whereas the spectral centroid shows specific information about the localization of the center of the mass of the spectrum. The spectral centroid can also show us information about the brightness of the sound: a sound with darker quality tends to present more prominent content of low frequency, while a sound with a brighter quality tends to present more predominance of superior harmonics, which can be measured by the spectral centroid [7].

This methodological procedure generated graphs and visual images of the selected parameters, which will be shown in the oral presentation, and that are available for consultation in an already published article where a complete analysis is presented in detail [8]. The results obtained showed the relevance of the performative interferences (interpretative decisions) in the construction of the sonority of the pieces.

The data obtained by V (MIDI velocities) showed how the instrumental approach happened in a real way, confirming that the interpretative plan was effectively accomplished regarding that the choices of touch (a touch with faster velocity *versus* a touch with a lower velocity) was concretized. The Q value, on the other hand, revealed the dynamic balance between the three elements included in the calculation, varying the preponderance of one onto the others and showing how the interpretative decisions occurred at the same time as some changes in the sonority in the written aspect (mainly in the register), letting us comprehend how these different elements derived from the score and from the performance acted side by side to create the timbric characteristics of the piece.

On the other hand, the parameters extracted from Sonic Visualizer confirm the correspondence between the interpretative decisions and the audio output of the recorded performances. All the previewed changes in the sonority from a brighter to a darker sound could be seen clearly in the images of the spectral centroid. Also, the sections with the higher Q values corresponded with the time point of the spectrogram, which gets the higher intensity and quantity of audible harmonics.

To conclude, we could confirm that the use of the pedals and the choice for the pianistic touches, in combination with the most relevant written elements of the sonority, dynamics and register, contributes to create the whole sonority of the pieces.

6. CONCLUSIONS

The methodological procedure presented in this paper serves as an opening to the insertion of the performance in the context of musical analysis, where the musical text and the performance act side by side in the creation of meaning in the piece, and also make it possible its existence as a sonic object. Without the score the performance of these pieces would not exist, but, on the other hand, without the performance the piece would not exist as a real happening and neither would be any sonority to be analyzed. Therefore, the collaboration between composer and performer shows itself as a manifested factor in a non-intentional way (that is independent of the intentionality of this collaboration) and inherent to the existence of the musical work.

In this analysis the conceptual basis of the morphological understanding of the musical work and the methods of the Artistic Research actuated as a bridge between theory and practice and already provided some interesting results [8], which open way for future developments.

Finally, the analyzes highlight the crucial role of the performer in the analytical process, as someone capable to add relevant information about the piece, information that are only and exclusively obtainable trough the experiential contact with it. Thus, the significance of the performer's voice in the development of musical knowledge makes itself recognized and manifested.

7. REFERENCES

1. D. Guigue. *Estética da sonoridade: a herança de Debussy na música para piano do século XX*. São Paulo: Perspectiva; CNPQ: Brasília; João Pessoa: UFPB, 2011.
2. K. Coessens., D. Crispin and A. Douglas.*The artistic turn: a manifesto.* Ghent, Leuven University Press, 2009.
3. V. da Costa. *Morfologia da obra aberta*. Curitiba, Prismas, 2016.
4. D. Guigue, L. Noda and B. Bragagnolo. "Timbre e Escrita ao Piano: por uma Incorporação do Comportamento Acústico do Piano na Composição e Análise Musical". *Música Hodie*, v. 14, n. 1, 2014, p. 137 – 158.
5. D. Guigue. *Soal 4.0, Sonic Object Analysis Library: OpenMusic Tools for analyzing musical objects structure*. 2016. Avaiable in: <https://www.academia.edu/29796175/>. Acess in: 13 jun. de 2017.
6. N. Cook, D. Leech-Wilkinson. *A Musicologist's Guide to Sonic Visualiser, 2009*. Avaiable in: <http://www.charm.rhul.ac.uk/analysing/p9_1.html>Acess in: 10 maio 2017.
7. M. Loureiro, T. Magalhães, R. Borges, T. Campolina, D. Mota and H. de Paula. "Segmentação e extração de descritores de expressividade em sinais musicais monofônicos". In: III Seminário Música Ciência e Tecnologia. 2008, São Paulo. *Anais do III Seminário Música Ciência e Tecnologia*. São Paulo: 2008. 109-125.
8. B. Bragagnolo and D. Guigue. "Análise da sonoridade em Ressonâncias de Marisa Rezende: uma abordagem a partir da performance". *Opus*, v. 23, n. 3, p. 222-253, dez. 2017.

TEXTURAL LAYERS AND POLYPHONIC TIMBRE LINKS IN ELECTRONIC DANCE MUSIC

María Marchiano
Laboratorio para el Estudio de la Experiencia Musical (FBA-UNLP)
marchiano.maria@gmail.com

Isabel Cecilia Martínez
Laboratorio para el Estudio de la Experiencia Musical (FBA-UNLP)
isabelmartinez@fba.unlp.edu.ar

Javier Dameśon
Laboratorio para el Estudio de la Experiencia Musical (FBA-UNLP)
javier.dameson@gmail.com

ABSTRACT

Polyphonic timbre refers to the overall timbral mixture in a music signal. Some of their most salient acoustic dimensions is the Sub-Band Flux. Our hypothesis is that there is a temporal alignment between the Sub-Band Flux changes and the perceived textural layers' onsets and offsets in electronic dance music (EDM). The study had two methods: (i) an experiment in which 15 professional musicians were asked to record in Reaper every perceived textural layer's onset or offset of 11 EDM's tracks, (ii) and the estimation of the transitional temporal data between homogeneous and successive states of Sub-Band Flux from sound signals. Both data sets were correlated. We found some temporal coincidences between the Sub-Band Flux changes and the perceived textural layers' onsets and offsets. Lowest spectral flux's bands show significant correlations with perceived texture, and it could mean that timbre and texture share a mutual acoustical background.

1. BACKGROUND

We perceive timbre as a multidimensional representation [1]. Nowadays, the study of timbre tries to find the acoustical features that afford our multidimensional perception.

There are two main categories of timbre based on different ideas of timbre blends: individual or monophonic timbres, and emergent or polyphonic timbres [2]. Monophonic timbre has been the main subject of study for a long time, and some of their acoustical features with perceptual correlations are spectral centroid, spectral flux, log-attack time, roughness and the Mel-Frequency Cepstral Coefficients (MFCC), among others. However, monophonic timbres are almost non-existent in our daily auditory experience. On the contrary, our sonic and musical environment is formed by polyphonic timbre.

Polyphonic timbre is defined as the overall timbral mixture in a music signal [3]. Its acoustical dimensions are currently being explored. Alluri [2] has analyzed a great quantity of acoustical features on a set of musical stimuli, to establish their relevance in the auditory experience of polyphonic timbres. One of them is the Sub-Band Flux feature. This feature "represents the fluctuation of frequency content in ten octave-scaled bands of the spectrum. (…) For each of the ten channels the spectral flux was calculated as the Euclidean distance between successive amplitude spectra" [2]. Alluri [4] found that some bands show strong correlations with timbral semantic descriptions of Indian music. The three lower bands (0-200 Hz) are related with the sensation of "fullness". Hartmann, Lartillot and Toivianen [5, 6] also find that the Sub-Band Flux plays a role in the perceptual segmentation of music. Unlike Alluri, they do not work with each band separately, but with a blend of all spectral flux's bands.

We think that the Sub-Band Flux might be a salient acoustical feature of polyphonic timbre in other music styles and in different types of perceptual tasks.

Beyond that the concept of polyphonic timbre does not refer to the textural aspect of music, but relates to the emergence of polyphonic timbres in non-monophonic textures. According to Fessel [7], the textural types are defined by five features: layers, homogeneity, linearity, accent coincidence, and attack divergence.

In electronic dance music (EDM), the layers are the most relevant feature for the definition of textural types. Normally, there is not a lead voice in the EDM, because all layers have relatively the same importance [8]. This characteristic produces some specific textural types, as mass texture, extended heterophony, paraphony, and rhythmic or timbral polyphony, and the EDM genres are partly differentiated by their textural features [9]. According to Anzil, audio-tactile sensation is a key feature in composition and in multisensorial perception of timbre in the EDM parties. This sensation is clearer in the low frequency zone.

2. HYPOTHESIS

People are perceptually sensitive to textural layers' onsets and offsets in EDM. These perceived transitions are temporarily aligned with Sub-Band Flux changes. We anticipate that the most perceptually relevant textural layers' onsets and offsets will be temporarily aligned with the most pronounced acoustic changes.

3. AIMS

To find temporal correlations between timbral and textural changes, with the long-term perspective of developing a model about these linkages.

4. METHOD

4.1 Experiment

4.1.1 Subjects

15 professional musicians participated in the experiment (MA=31; SD=6.9; average 15.8 years of music studies in institutions). All subjects had heard of EDM, but only 4 had been or currently were users of EDM parties.

4.1.2 Stimuli

11 EDM's tracks were used on the experiment (2:45 minutes average duration per track) [10].

The selection was based in the variety of EDM's genres, with different textural and timbral qualities. The genres of selected tracks were deep and electro house, techno, trap, acid and trance. This selection was made by auditory analysis of the authors.

4.1.3 Procedure

The experiment was done on Reaper v.5 with two Mackie mk3 mr6 monitors. We provided 11 Reaper projects randomly ordered (one for each track). Participants were asked to record every perceived textural layer's onset and offset of EDM's tracks by pressing the letter "m" of the computer keyboard. In order to get temporarily accurate marks, this was not a real-time task: participants could listen to each track as often as they needed, and delete, add or change the temporary location of marks.

The task involved an analytical listening, where participants had to determine what was a textural layer in the music they were working on. We asked them to make a decision using a single criteria for each track, giving the example: "if you make a mark for a little sound that you would consider a textural layer, you must mark all layers' onsets and offsets of that level of significance. This level of significance has to emerge from the music and could change from track to track".

After the task, we asked participants if the textural layer's onsets and offsets were perceived as musically relevant.

4.2 Computational modelling

All analysis were made on Matlab R2015a. We use MirToolbox1.7 for the feature extraction and novelty detection [11].

4.2.1 Sub-Band Flux extraction

Sub-Band Flux was computed by firstly filtering the sound signal in 10 spectral bands, and secondly extracting the spectral flux for each band. The analysis window normally used for Sub-Band Flux extraction is 0.025 seconds with 50% overlapping. Given the subsequent analysis, we used a 1 second window without overlapping (see *Novelty Detection*).

We used two sets of Sub-Band Flux data. On the one hand, a single temporal series of Sub-Band Flux was computed by combining the data of all bands. This series bears the Sub-Band Flux information of the complete spectrum. We will call it *general Sub-Band Flux data*, to differentiate it from the other set. On the other hand, spectral flux data of each band was used separately.

4.2.2 Novelty Detection

The novelty function detected the dissimilarity between homogeneous and successive states, and the relative importance of these temporal transitions. The peaks of the resulting novelty curve represent the moments of structural change in the sound. These segmentation points are often correlated with perceptual tasks results [12], to evaluate the perceptual relevance of the acoustical structure of a musical signal.

There are several approaches to novelty detection [13, 14, 6]. We decided to use the multi-granular method proposed by Lartillot et al. [14], because its algorithm can detect homogeneous states of different temporal lengths. The computation of the novelty curve in this approach is attached to the previous feature window analysis. We processed the signal with different window sizes to determine which have a more adjusted temporal alignment with the perceived layer's onsets and offsets. In order to get an acoustical data to correlate with perceptual task results and considering their possible temporal imprecisions, we decided that a 1 second window without overlapping was the better choice.

5. RESULTS

5.1 Perceptual Task

All participants of the experiment showed perceptual sensitivity to textural layers' onsets and offsets, and 93.3% judged that this textural layer behavior was relevant in the music of the task.

Frequency of perceived layers' onsets and offsets were calculated by counting the number of all participants' marks made in the same second (frequency range: 1 to 15). The maximum frequency represents the 3.83% of total marks of the experiment (SD of all tracks frequency=0.28, normalized).

5.2 Perceptual and Acoustic Correlations

Initially, we searched for temporal coincidences between the higher values or peaks of the novelty curve of general Sub-Band Flux and all marks of the perceptual task (of all participants and all tracks). There was a total of 47.12% coincidence, among all the tracks. Percentage of coincidences for every track is shown on Table 1.

Track	Temporal coincidence (%)	Correlation Coefficients
1	44.4	0.58
2	24	0.5
3	51.6	0.43
4	43.6	0.07
5	59.4	0.36
6	34.8	0.42
7	40.9	0.5
8	39.4	0.39
9	52.6	0.5
10	46.5	0.54
11	47.6	0.38

Table 1. Percentage of temporal coincidences between perceptual marks of experiment and novelty peaks of general Sub-Band Flux; correlation coefficients between marks' frequency and novelty values

We ran two set of correlations. In both, we searched for the novelty values of Sub-Band Flux for the temporal locations of all marks made by every participant in the perceptual task, and correlated the mark's frequency and their corresponding novelty values.

In the first set of correlations, we used the general Sub-Band Flux data, which includes information of complete spectrum. Results show significant correlations for some tracks and low correlation coefficients for others (medium-high significance = 45.5%; medium-low significance = 45.5%; no correlation = 9%) (Table 1).

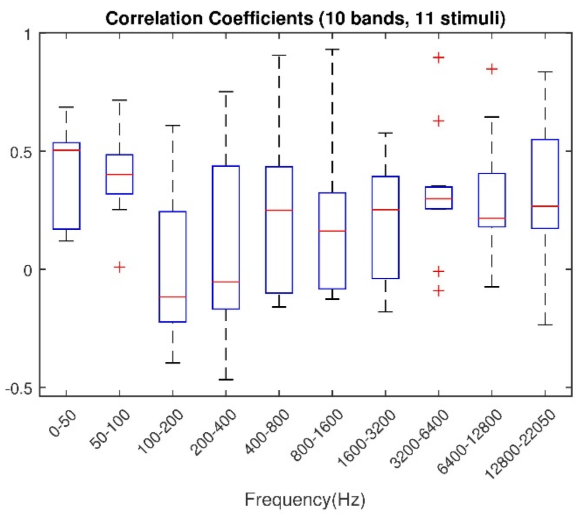


Figure 1. Box plot of correlation coefficients of each spectral flux's band between their values and the frequency of perceptual marks, in the 11 stimuli.

In the second set, we correlated the marks' frequency of the perceptual task and the values of each separately spectral flux band for each perceptual mark's temporal location. Correlation between frequency of all marks and the temporal series for the separately spectral flux's bands did not show significant results. We thought that it is possible that the perception of textural layer's onsets and offsets is linked to acoustic changes in some part of spectrum, as the Sub-Band Flux feature itself proposes. Thus, the perceptual marks not related to a high value of a spectral flux's band, could be related to another.

To confirm this, we started searching the higher peaks of the novelty curves of all spectral flux's bands. Then, we selected the perceptual marks that were temporarily coincident with any of the novelty's higher peaks, and got 10 different sets of marks (one per spectral flux's band). Finally, we correlated the selected marks' frequency and their novelty values. This method increases the correlation significance: spectral flux's bands 1 and 2 show significant correlation values (band 1's mean= 0.5; band 2's mean= 0.4) (Figure 1).

6. DISCUSSION

The present study found some links between the perceived textural layer's onsets and offsets, and Sub-Band Flux changes in EDM. However, these links are not as strong as we expected.

The analysis of a general Sub-Band Flux's novelty –i. e., novelty of the blending of all band's data in a single curve- shows a clear relationship with the perception of textural layer's changes. This connection is not so strong in a more detailed analysis. Low correlation coefficients for each spectral flux's band make us think that its changes are not so relevant to the layer's onsets and offsets perception. The links between changes in Sub-Band Flux and texture could be occurring at an acoustic and perceptual global level.

On the one hand, we think it is likely that other acoustic features are involved in the perception of texture layers. This could be explained through a joint analysis of a bigger set of features. On the other hand, it is also possible that this unstable link between Sub-Band Flux and texture is an EDM's stylistic characteristic. It would be necessary to include a more stylistically varied set of stimuli in a future study.

However, there are two spectral flux's bands which show significant correlations with the perceived layers: band 1 (0-50 Hz) and band 2 (50-100 Hz). Alluri had found that these lowest bands are related to the "fullness" descriptor of timbre sensation. This result not only contributes to the idea of the perceptual relevance of the lowest Sub-Band Flux behavior, but also allows to think of a possible shared acoustical background for timbre and texture perception.

Moreover, the lower frequencies have special treatment in the EDM and they are audio-tactilely perceived in a club with a specific kind of amplification. It is interesting to notice that even in a completely different acoustic environment, the low frequencies seem to be more perceptually relevant regarding texture than the rest of the spectrum.

7. REFERENCES

[1] McAdams, S. (1999). Perspectives on the contribution of the timbre to musical structure. *Computer Music Journal, 23*(3), 85-102.

[2] Alluri, V. (2012). *Acoustic, Neural, and Perceptual Correlates of Polyphonic Timbre* (Doctoral Thesis). University of Jyväskylä: Jyväskylä.

[3] Aucouturier, J., Pachet, F., and Sandler, M. (2005). "The Way It Sounds": Timbre Models for Analysis and Retrieval of Music Signals. *IEEE Transactions on multimedia, 7*(6), 1028-1035.

[4] Alluri, V. and Toivianen, P. (2009). Exploring Perceptual and Acoustical Correlates of Polyphonic Timbre. *Music Perception, 27*(3), 223–241

[5] Hartmann, M., Lartillot, O. and Toivianen, P. (2017a). Interaction features for prediction of perceptual segmentation: Effects of musicianship and experimental task. *Journal of New Music Research, 46*(2), 156-164.

[6] Hartmann, M., Lartillot, O. and Toivianen, P. (2017b). Musical Feature and Novelty Curve Characterizations as Predictors of Segmentation Accuracy. *SMC 2017: Proceedings of the 14th Sound and Music Computing Conference 2017,* 365-372.

[7] Fessel, P. (1997). Hacia una caracterización formal del concepto de textura. *Revista del Instituto Superior de Música, 5*, 75-94.

[8] Butler, M. (2003). *Unlocking the groove: Rhythm, meter and musical design in electronic dance music.* Indiana University: Indiana.

[9] Anzil, I. (2016). *Sensación táctil y audiotáctil en la música. El caso de las músicas electrónicas utilizadas para el baile social en locales de baile de la Ciudad Autónoma de Buenos Aires y alrededores* (Doctoral Thesis). Universidad Nacional de La Plata, La Plata.

[10] Varios. [María March]. (2017, 12, 22). EDM SysMus [Archivo de video]. Recuperado de https://www.youtube.com/playlist?list=PLWWthKwdwpnVrpumIDccjJw-utN_EYhn2

[11] Lartillot, O. and Toivianen, P. (2007). A Matlb Toolbox for Musical Feature Extraction from Audio. *Proceedings of the 10th Int. Conference on Digital Audio Effects (DAFx-07)*. Bordeaux, France.

[12] Hartmann, M. (2017). *Modelling and Prediction of Perceptual Segmentation* (Doctoral Thesis). University of Jyväskylä: Jyväskylä.

[13] Foote, J. T. (2000). Automatic Audio Segmentation Using a Measure of Audio Novelty. *Proceedings of IEEE-ICME, 1*, 452–455.

[14] Lartillot, O., Cereghetti, D., Eliard, K. amd Grandjean, D. (2013). A Simple, high-yield Method for Assessing Structural Novelty. *Proceedings of the 3rd International Conference on Music & Emotion (ICME3).* Jyväskylä, Finland.

MASCULINITY, PAGANINI, AND THE VIOLA: TOWARDS A QUEER VIOLA PERFORMANCE OF PAGANINI'S 24TH CAPRICE

Kirsty Clark

Guildhall School of Music and Drama
kirsty.clark@stu.gsmd.ac.uk

ABSTRACT

This paper explores one way in which classical music performance could be queer performance: that is, performance that challenges norms of gender and sexuality. It seeks to highlight masculine aesthetics within virtuosity, such as control, and domination of thE instrument, using Paganini's 24th caprice as a case study. Paganini was surrounded in myths of sexual promiscuity, gendered violence, and occultism (Kawabata), and this paper shows how this mythology continues to inform some performers' interpretations today. Paganini's works are included in many major violin and viola competitions: this situates Paganini and his music firmly within the musical canon and establishes virtuosity as a desirable aesthetic for the competitive violinist or violist. The inclusion of Paganini's Caprices in viola competitions is one example of violists performing transcriptions. I find a queer potential in the way the viola must make physical adjustments to 'fit' these transcriptions onto an instrument for which it was not originally written. This experience of the violist having to make physical and musical negotiations mirrors the way a queer subject must negotiate social space in which they are 'obliquely' situated in relation to the 'straight lines' of heteronormativity (Ahmed, 2006)): both the queer subject and the violist cannot comfortably fit into the social or musical space they occupy and must go 'off course' in order to 'fit'. To conspicuously 'go off course' as a queer violist performing Paganini's 24th Caprice, I use drag to exaggerate the musical and physical gestures associated with Paganini's virtuosity to highlight simultaneously the performativity of masculinity and gender in general and the history and present of hyper-masculine associations with virtuoso performances

INTRODUCTION

This paper is part of my wider doctoral thesis, an inquiry (or inqueery) into the development of a viola practice that is intentionally and legibly queer: that is, a viola practice that 'queers' the conventions that govern classical music performance practices. This paper is one way in which this 'queering' might take place

I am developing a drag king performance of Paganini's 24th violin caprice on the viola, doing so by exaggerating visual and musical signifiers of masculinity. Appropriating traditional visual signifiers of masculinity, and especially elitist masculine attire

In this paper I will illustrate the various links between Paganini's music and masculinity before situating the viola 'queerly' in relation to the canon and the social structures which enforce it.

Methods

To explore Paganini's position within the canon, I have conducted a brief survey to give an overview of violin competitions' inclusion of his works.

In order to apply the theoretical framework that I will be discussing in this paper, I will be conducting a series of performance experiments, using self-reflection to compare experiences of performing in spaces with different social codes and functions.

I have chosen Paganini's 24th Caprice as the focus for this study, as it is a piece that is relatively well-known outside the world of classical music, so should be accessible to audiences in cabaret venues who may not be familiar with Paganini's works. The theme-and-variations form of the Caprice is also useful for showcasing a wide variety of different virtuoso techniques, including those that can be difficult or risky for some violists to perform.

CONTEXTS

Paganini's music is a site where the links between virtuosity and masculinity are especially clear. The cultural context within which the music was composed and originally performed, the reception of the music and its surrounding mythology, and its status within the modern musical canon all contribute to the formation and strengthening of these links.

Contemporary mythology

As Mai Kawabata excellently sets out in Paganini: The 'Demonic' Virtuoso'[1], Paganini was surrounded by a vivid contemporary mythology which simultaneously enhanced his masculinity and positioned him as a social outsider. This mythology was constructed by written accounts and visual depictions of Paganini's onstage persona.

Paganini's innovations in violin technique were afforded, in part, by the contemporary developments in bow technology, with the introduction of the Cramer and Tourte bows, both of which share the convex stick and heavy tip of the modern bow. These innovations allowed Paganini to perform ricochet bowings, for example, with greater volume and more consistent sound than was possible with earlier Baroque bows. Peter Sheppard Skaerved [2] concludes that Paganini capitalised on the flourishing innovation in bow technology at this time by using different variations on the Baroque, Classical and transitional styles of bow to suit the music that he was writing and performing, even commissioning 'hybrid bows' to best suit his purposes. Sheppard Skaerved also observes that the 'leaping passages across all the strings' that Paganini includes in his compositions are best facilitated by an 18th-century posture where the violin is held lower than was conventional at the time. Paganini's ability to rapidly move between strings and play techniques such as ricochet and spiccato were therefore facilitated by a novel, hybrid technique that was so unfamiliar to audiences and critics that it was ascribed to magic, witchcraft, a Faustian pact with the devil, or Satanic possession [1].

The reported hysteria that Paganini's performances elicited from his audiences, especially among women [1, pp. 72-3] further strengthened this supposed link with the occult. The gendered nature of this effect contributed to the image of the violinist as masculine, sexually dominant, and violent, while the violin was cast as the helpless, feminised victim in his hands. The bows which Paganini was using were concave, unlike the convex shape of Baroque and earlier 'transitional' bows, leading to comparisons with a sword or phallus [3] which enabled the performer to 'lash'[1, pp.64-6] the violin and cause it to 'cry out' [3, p.103]. Kawabata identifies Melchior von Hugo's *Der Geiger*, which depicts a Death-figure playing a nude, armless woman's body with a phallus as a bow as exemplary of this sexualisation of the instrument itself, despite dating from a century later[1].

Contemporary images of Paganini, while not so obviously sexualised, depict him surrounded by skeletons and women who appear to be being seduced by his performance [fig. 1], or conjuring a succubus-like figure from his violin [fig. 2]. Paganini himself perpetuated his reputation as a sexually promiscuous man [1], further strengthening the associations between his performance and sexuality.

Figure 1: Johann Peter Lyser, 'Paganini der Hexenmeister'. From https://www.br-klassik.de/aktuell/news-kritik/ilya-gringolts-paganini-muenchener-kammerorchester-100.html [accessed 19/04/2018]

Figure 2: Alberto Martini, 'Paganini', from Gertrude Clarke Whittall Foundation Collection, Library of Congress, music Division

Modern reception

This association of violin virtuosity with virile sexuality is less obvious in modern performances of Paganini's violin music[1]; however, I will now show how the mythology described above has been heavily influential on the interpretations of two specific performers.

Alexander Markov is a Russian-American violinist whose performance practice has historically focused heavily on Paganini's Caprices[4]. In his performance of the 24th *Caprice*, part of his 1989 video recording of all 24 *Caprices*[5], he makes clear visual connections between himself and the composer, both in terms of his costume, with his long hair and tailcoat, and his physical gestures in performance. As Kawabata notes:

> …he flails around as if possessed, bounces the bow higher than is strictly necessary and puckers his lips with fierce concentration. [1, pp.116-17]

Markov's choice of dress suggests not only a desire to draw visual connections with Paganini but also to situate himself within a historically elitist, masculine context that further emphasises his mastery over the music and violin. Markov takes inspiration from Paganini's mythology in his performance of other repertoire: he depicts himself as a rock star violinist, wearing a purple velvet coat and an elaborately ruffled shirt [6] and even leather trousers [Fig. 3] while performing on a gold electric violin. Markov's overall image, therefore, situates him as a modern version of the 19th-cenutry violinist with a proto-rock-star reputation.

Figure 3: Alexander Markov. From: http://www.alexandermarkov.com/gallery.html

[1] It is interesting to note that heavy metal has adopted classical virtuosity, drawing on Paganini's works especially, as a signifier of masculinity. For more on this, see R. Walser, *Running with the Devil: power, gender, and madness in heavy metal music*. Hanover, NH: University Press of New England, 1993

Another modern violinist who draws heavily on Paganini's mythology and its reception is Pavel Šporcl. Šporcl, in his video of the 5th *Caprice*[7], takes on the visual language of heavy metal, including thunderclaps, fire, and an audience of women. His shadow then takes on the shape of Paganini and eventually breaks free from Šporcl to play alongside him, moving as though possessed while Šporcl remains relatively still: this, like Markov's clothing for his electric violin performances, draws on Paganini's rock-star mythology. In addition, Šporcl situates himself not only as a kind of musical 'heir' to the myth of Paganini, but also the medium through which the great virtuoso is brought back to life.

It is in this example of Paganini's reception that I find a clear reflection of the changes in performance conventions and masculinities. It might be thought that, like Markov, Šporcl would want to embody Paganini in his gestures; however, his remaining still during his performance in contrast to the wild figure of Paganini illustrates the convention within modern classical music for performers to remain show their expression primarily through musical rather than physical gestures, while Paganini exemplifies the 'Romantic idea of self-expression in performance' with his gestures. This mirrors the pressures of modern masculinity to limit emotional expression and remain stoic in all aspects of life[8, p.123]. This stoicism is evident in other performers' interpretations of this music, including Janes Ehnes, Jascha Heifetz, and William Primrose: futher

Canon

It is not only Paganini's mythology that establishes his music as an example of masculine aesthetics. His music is significant for this study because of its position within the canon. Canons, in all aspects of the art, are both a product and tool of the patriarchy, establishing a hierarchy of composers and works that has historically privileged heterosexual, cisgender white men. The canon is constructed and perpetuated within educational and cultural institutions including, but not limited to, university syllabi, literary awards and competitions. Since these institutions generally build on centuries of precedent where panels or juries are white men judging the works of white men, they inevitably replicate the structures of value and therefore power that uphold the hegemony of white patriarchy, although this is being increasingly challenged and altered as women and people of colour gain footholds within the academy.

To establish how Paganini's works fit into the violin and viola canons, I have conducted a survey of major violin and viola competitions[2]. I chose competitions as a focal point because their express purpose is to measure different musical performances in comparison to each other and reward those that best fulfil the aesthetic ideals of the competition and its panel with money, prestigious

[2] This survey was by no means exhaustive and was conducted to gain an impression of the trends and conventions among international violin and viola competitions.

concert opportunities, and/or valuable instruments. In this way, competitions uphold patriarchal capitalist ideals and are therefore a useful indicator for how certain musical works fulfil those ideals.

Competitions

My survey of violin competitions [Fig. 4] included 29 different competitions in total: of these, 18 specified at least one work by Paganini in its repertoire list, while only 3 did not specify any works by Paganini. The remaining 8 did not have any information available regarding current or previous repertoire lists.

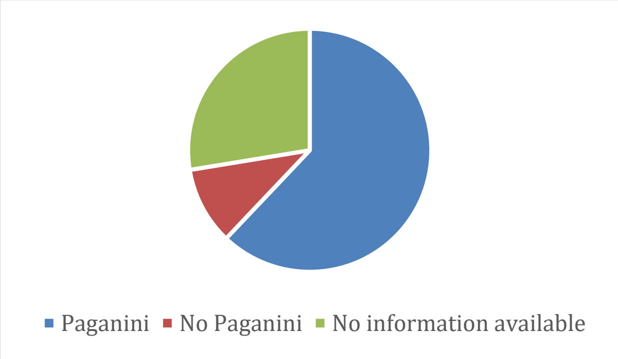


Figure 4: Paganini's works on the repertoire lists of major violin competitions

Of the twelve viola competitions I surveyed [Fig. 5], three specified works by Paganini, with only the Primrose International Viola Competition listing Paganini's *Sonata per la grand' viola*, while the other two competitions list the *Caprices*.

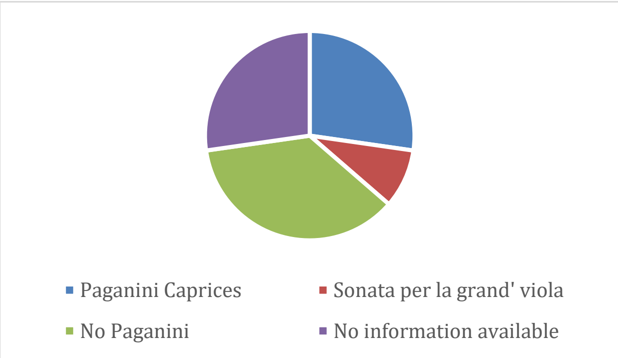


Figure 5: Paganini's works on the repertoire lists of major viola competitions

While this shows that Paganini is not very widely included in the repertoire of major viola competitions, that transcriptions of Paganini's violin works, rather than his viola work or different works originally written for solo viola, are included at all perpetuates the impression that there is insufficient repertoire of similar "difficulty" for the viola, and that Paganini's *Caprices* are a useful measure of a violist's virtuosity.

I find parallels here with the experience of the queer subject within Western society: since the *Caprices* were originally written for violin, much of the technical challenges within them remain near-impossible for many violists, even after transcription. A violist's ability to play, for example, parallel octaves or even tenths (such as the sixth variation of the 24th *Caprice*[9]), rapidly changing triple stops, or fast, delicate motifs in the highest registers, depends on their hand size, arm length, flexibility, and the size of their viola. A smaller viola, therefore, or a violist with larger hands, will therefore be closer – in terms of relative proportions – to a violin and violinist and therefore will be more likely to execute these kinds of techniques. However, since the viola is always generally larger than the violin, these motifs will always be more difficult to execute, even causing the performer significant discomfort or putting them at risk of injury. Just as the music may not 'fit' the violist's hands in performance, it also may not 'fit' the instrument: since the viola has longer, thicker strings than the violin, it is slower to speak. This makes passages such as the 4th and 10th variations of the 24th Caprice more difficult to execute on the viola than the violin.

I find parallels between the discomfort of performing this music on the viola and the discomfort of the experience of the queer subject within heteronormative society. I will now briefly sketch the theoretical background of this idea of discomfort before situating my own performance of the 24th *Caprice* within this theoretical context.

QUEER DISCOMFORT

When existing in a society which is heteronormative [10]– that is, it assumes that all people are heterosexual or cisgender until proven otherwise –the queer subject finds themselves always operating outside 'the norm'. In *Queer Phenomenology: Orientations, Objects, Others*, Sara Ahmed characterises heteronormative society as a series of 'straight lines' which form the behavioural paths that an individual is normatively expected to follow. One example of this is compulsory heterosexuality, 'an orientation towards "the other sex" ': the queer subject 'has to go "off line… to turn towards "one's own sex" [and therefore]… to leave the straight line'[11, pp.70-71]. Another line which the subject is expected to follow is that of reproduction and thus the continuation of family, which is, normatively, a consequence of the straight line of heterosexuality. Having characterised society in this way, Ahmed then frames the experience of those who follow these lines as their bodies 'extending into space', [11, 12] and therefore feel comfortable within society, whereas those who do not align with the 'straightening devices'[11] cannot extend into space to the same extent[3].

While I do not intend in any way to equate the experience of playing the viola within the elitist cultural sphere of the Western classical music world with the experience

[3] The extent to which an individual can 'extend into space' is restricted differently according to which hegemonic groups they are aligned with: for example, a white, middle-class, able-bodied queer person is less 'out of line' than a black or brown, working-class, and/or disabled queer person.

of marginalised people within the deeply entrenched, interlocking systems of white supremacy, patriarchy, heteronormativity, and capitalism, I find a 'queer potential' in the position of the viola within the classical canon and especially in the performance of transcriptions of Paganini's violin works.

TOWARDS A QUEER VIOLA PERFORMANCE

Just as Jarman identifies the voice as having queer potential in its 'genderless' 'liminality' [13], so do I identify queer potential in the viola's performance of Paganini's *Caprices*. To an audience member, it may be unclear whether they are hearing a viola or a violin perform; I plan to replicate this blurred boundary by appropriating the formal, historically elitist tailcoat and white tie favoured by Markov, binding my chest and painting my face to put on a consciously performative and exaggerated form of masculinity. In my preparation of the 24th Caprice, I have found that the combination of my physiology and viola make it almost impossible to perform the sixth variation[4], and that it is immensely difficult to replicate the agility of movement and clarity of sound that I can achieve on the violin. Just as I am unlikely to 'pass' as male when wearing drag as described above, the challenges of navigating this piece as a violist also make it unlikely that I will 'pass' as a violinist beyond the *Thema*. The viola frequently cannot 'extend into' the musical space of the *Caprice* as a result of the physical difficulties of navigating this music, just as queer bodies are limited by entrenched endemic heteronormativity and the consequent social policing of gender. I therefore am also drawing on Jack Halberstam's idea of failure as a desirable queer aesthetic [14] to develop this performance.

CONCLUSION

By 'failing' to perform masculinity or Paganini's 24th *Caprice* to hegemonic social standards, I seek simultaneously to highlight the performativity of masculinity and the many-layered signifiers of masculinity that are associated with Paganini's violin music. By doing so, I hope to offer one way in which classical music might be 'queered' and thus challenge the conventions associated with its performance.

[4] While I can physically play the 10ths in this section, it is only as a result of my joint hypermobility syndrome, and I risk injury by performing the hyperextension of my fingers and hand which is necessary to do so.

REFERENCES

[1] M. Kawabata, *Paganini: the "demonic" virtuoso*. Woodbridge: Boydell Press, 2013.

[2] P. Sheppard Skaerved, "The Paganini Project with Peter Sheppard Skaerved," Library of Congress, 06-Jun-2012.

[3] M. Kawabata, "Virtuoso Codes of Violin Performance: Power, Military Heroism, and Gender (1789-1830)," *19th-Century Music*, vol. 28, no. 2, pp. 089–107, Nov. 2004.

[4] "Violinist.com interview with Alexander Markov," 06-Sep-2014.

[5] B. Monsaingeon, *Alexander Markov plays Paganini's 24th Caprice*. Reggio Emilia.

[6] A. Markov, *Alexander Markov - Six String Violin*. .

[7] J. Herz, *Pavel Šporcl - 24 capricií pro sólové housle (oficiální video)*. 2012.

[8] R. Connell, *Masculinities*, 2. ed., reprint. Cambridge: Polity Press, 2012.

[9] N. Paganini, *24 Caprices for Viola Solo*. New York: International Music Company.

[10] A. Rich, "Compulsory Heterosexuality and Lesbian Existence," *Signs*, vol. 5, no. 4, pp. 631–660, 1980.

[11] S. Ahmed, *Queer phenomenology: orientations, objects, others*. Durham: Duke University Press, 2006.

[12] S. Ahmed, *The cultural politics of emotion*, 2. ed. Edinburgh: Edinburgh Univ. Press, 2014.

[13] F. Jarman-Ivens, *Queer voices: technologies, vocalities, and the musical flaw*, 1st ed. New York: Palgrave Macmillan, 2011.

[14] J. Halberstam, *The Queer Art of Failure*. Duke University Press, 2011.

Looking/Sounding Androgynous: An Analysis of Tomboy Body/Voice Aesthetics in East-Asia Taking Denise Ho as Example

Yin Hsuan Dai

National Taiwan University
heidi0707.taiwan@gmail.com

ABSTRACT

In contemporary popular music cultures in East Asia, tomboy-style performers have been existed as a force to defy existing gender norms and hierarchies. Yet while there have been some researches of queer performances under Euro-American contexts, few have focused on such cultures in East Asia. Taking the Hong Kong singer Denise Ho as an example, I would like to use the textual analysis of her music videos (in both visual and audio perspectives), and also the studies of audiences' commentaries based on her performance to answer the questions below: How does Ho demonstrates the tomboy/butch body/voice aesthetics on stage? How is such aesthetics distinguished from those in Euro-American contexts? How do these performances trigger the queer gaze/listening from her audience, and how does such gaze/listening help construct Ho's status as one of the most important lesbian icons in East Asia?

1. INTRODUCTION

The signification of androgyneity usually begins with a simple question when a subject is questioned by the gaze/listening of Others: is it a boy or a girl? Such straight interrogation from Others opened up the space (troubling to some people however) of how a subject blurred and permeated the gender-binary hierarchies, for those who are being looked at/listening to don't fall into either male or female categories. Yet it is exactly through the path of constant confusion and re-signification that makes androgynous characteristics emerged to the public, and through approaches that could possibly reach a greater mass such as popular cultures, the androgynous figures onstage are queered by the audience, and through the process queer themselves as well.

Among different representations of androgynous characteristics, there has been a unique style of tomboy aesthetics relating to how feminine masculinity is performed. Such tomboyish-ness contained historical roots from the past and the local contexts at the contemporary moment. However, the analysis and questions how tomboyish characteristics are represented in both visual and the equivalently important audio level in East Asia have not been widely discussed or answered. The development of queer theories and practice of gender movements have generated considerable discussions in the academic field. Yet while tomboy body aesthetics have discussed quite often in both popular music studies and culture studies in Euro-America, few scholars have put them in the East-Asian context, which is distinguishable due to different cultural scenarios. Also as a queer musicology student doing research in the academic field in East Asia, I found that the discussions of queer theories intersecting with musicological discourses are still rarely seen, or even if there have been a few, very rare discussions have focused on the specific performativity of tomboy style as an androgynous practice in local contexts, even though the identification and signification of the idea "tomboy" has long existed inside or outside queer communities for decades. To elaborate on how such tomboy aesthetics in East-Asia is represented and signified in popular music cultures, and how it triggers the homoerotic gaze/listening of the queer audience, I would like to use two songs of the famous Hong Kong singer HOCC Ho Wan-See (何韻詩), 'Rolls Royce' (勞斯萊斯) and 'Chi Ching Si' (癡情司) as examples. In this paper, I will focus on the textual analysis of the body performances, semiotics and vocality in these two songs. I want to elaborate a tomboy aesthetics influenced not only by Western gender discourses but also by East-Asian cultures. Aside from the textual analysis focusing on the singer herself, I would also analyze the texts of the audience's commentaries printed or online as supporting materials, to discover how audience look at/listen to her in a queer perspective, and moreover, how such perspective help construct Ho's status as one of the most important queer icons in East-Asia?

The notion of tomboy aesthetics in East-Asia has its own context, yet it is still more or less influenced by the queer theories and gender movements in Euro-America. Scholar Sue-Ellen Case first introduced the idea of butch-femme aesthetic in 1980s when the second-wave feminist movements have gained relatively obvious attention, yet among them lesbian discourses were still unseen among different branches of feminist theories. Though some feminists consider butch-femme roles in lesbian cultures as mere reproduction of heterosexual norms, Case suggested that such butch-femme identities were actually a "camp-style masquerade" that overflow beyond the boundaries of real/unreal identities, when lesbians are no longer the "well-adjusted women" and would take the phallic (physically or symbolically) as their own with the "mask of the castrated, or womanhood" (64, Case), and in such circumstances create alliance among lesbians.

In terms of culture studies, scholars also develop theories that focus on "female subjectivities", being both critical toward existing gender hierarchies and open to more possible queer perspectives when "looking" and "being looked at". First suggested by Laura Mulvey and also by Teresa de Lauretis, Alexander Doty…etc, the scholars implied that possible queerness was not only presented by the actresses in the films, but also by the gazes of female audiences that co-constructed a queer narrative escaping from the heterosexual normative ideology. All of such theories help prosper the queer culture studies, and affect the field of musicology as well. As Sheila Whiteley discussed in her essay "Challenging the Feminine: Annie Lennox, androgyneity and illusions of identity", for instance, queer perspectives are not only limited to visual experiences, but also listening experiences as well. By borrowing the theories of French feminist Luce Irigaray and the analyzing of melody, tonality, vocality and the performance of Annie Lennox in her music videos, Whiteley considered the performance to have "destabilized the distinctions between the natural, the artificial, depth and surface, inner and outer through which gender discourses almost always operate." (133, Whiteley).

Under the influences Euro-America gender movements and theories, queer communities in East-Asia (especially after 1990s) have also developed their own discourses in local contexts, and in this paper focusing lesbian communities as an instance, the scenarios and tensions within communities or with the general hetero-sexual-normative societies "outside" are also distinguished from the contexts mentioned above. The term to signify a gender-neutral female embodying both feminine and masculine characteristics, differing from Case's term "butch", are mostly referred to as "tomboys", which is also my approach to use this term instead of "butch" (more often used in Euro-America queer communities) to coin the figures of female masculinities in East Asia. (Chang, 1994)

Based on the Western gender theories and the developments of queer communities in East Asia, local scholars also have developed their own queer discourses in each of their specific own contexts. Yen-Ning Chao, in her PhD dissertation in 1996, traced back the history of T (tomboy) body politics in Taiwanese lesbian communities all the way to 1960s, along with the analysis of queer spaces and body performativity in a historical and ethnographic point of view. Her groundbreaking works in the field of tomboy body studies, and further opened up more space and discussions for following scholars. Chuan-Fen Chang, for instance, has operated deep and thorough fieldwork on T-Po (踢-婆) identities within Taiwanese lesbian communities in her book "Liberality of Love: A Storybook of Lesbians" in 2001. Much like a collection of interviews within lesbian communities, Chang used the conversations as examples to demonstrate how the identities T-Po culture manipulates in Taiwan. Aside from Chang's studies, recent scholars such as Yu-Yin Hu and Fan-Ting Cheng also suggested how queer theories and performances work in popular cultures and theatre, and also elaborated the tomboy body studies with East Asian cultural contexts. In one of Cheng's article in 2017 which discussed the theoretical tactics of Skin Touching, one of the most notable lesbian theatre works in Taiwan, she pointed out that the theatre work combined not only the butch-femme androgynous aesthetic, but were also playful with the crossdressing symbols of Liangzhu (梁山伯與祝英台) in the play, marking a distinctive performing style. The paper gives strong arguments on how queer theories have been appropriated and "localized" by East-Asian popular culture, and such discourses would also be my point of elaborating how Ho demonstrated a tomboy aesthetic (in both visual and audio aspects) intertwining with historical yet differential cultural backgrounds.

Before the entering into textual analysis on HOCC's works, I would first like to suggest her status in Sinophone pop music (here I choose the term instead of "Chinese" pop music due to Hong Kong's hybridized and post-colonial political background that is not entirely homogeneous from the Grand-China cultural narrative) and how her status is intertwined with local queer communities in Sinophone Asia. In Hong Kong and in East Asia, Ho was the first artist who officially came out as a lesbian in 2012, and even before her action of coming out, her musical works have long been considered to embody hidden homoerotic desires, for instance, the song Rolls & Royce that I would elaborate on later. She is well recognized also by her androgynous characteristics, in her appearance and in her unique vocality. To the public, her dress codes and styles are "tomboyish", wearing mostly jeans, sneakers, shirts or suits, with a boyish short haircut representing her androgyneity. Such "tomboy" body aesthetics are highly recognizable in East Asia, and to some extent are relevant to lesbian identities (though, I must clarify, that "tomboy" is more cling to certain exterior characteristics instead of fixed identities). To both lesbian communities and the heteronormative public, her tomboyish-ness signifies certain notion of ambiguity that permeates the binary gender boundaries, and further makes her a queer cultural icon in East Asia.

Similar gender ambiguity of Ho is also demonstrated not in her appearance but in her voice. As the only female pupil of iconic Hong Kong singer Anita Mui (梅艷芳), Ho's singing techniques have her roots from her mentor. Through critiques and comments of both Mui and Ho, many describe them to contain some smooth, mellow alto voices that are unique in Hong Kong popular music industry. Tracing back to Ho's mentor Mui's experiences of singing, we could discover that Mui's earliest access of vocal training was in Jin-Xia Singing and Dance Group (錦霞歌舞團), following the model of Yi-Xia Singing and Dance Group (藝霞歌舞團) in Taiwan back in 1960s, the latter actually followed the model of Japanese Takarazuka Revue (寶塚歌劇團), casting all-female performers onstage, some of them performing as gender-crossing roles. Such background of gender-crossing experiences in singing nurtured Mui's mellow alto vocality, and further influenced Ho's singing.

Through Mui's guidance, Ho developed her professional singing techniques. Differing from most Hong Kong popular female singers, her vocality is low, steady and warm, using mostly head voice to strengthen the resonance. The intensive usage of head voice and her mellow alto voice make her vocal lines distinguishable, yet not acceptable to all listeners in Hong Kong. While listeners hardly criticize her "techniques" on singing, some still demonstrated their uneasiness when listening to Ho's voice. Comments online include phrases such as "not sweet and feminine enough", "too (gender) neutral that it makes me feel uneasy", "most female singers contain the characteristics of either sweetness or tenderness/softness in their voices, which is why such unisex voice is not that easy to please the listeners". Yet also, there have always been comments describing Ho's vocality as "unique voices rarely heard in Hong Kong popular music" "distinguishable, steady and full of emotions, making all the covers of other's works her own music". These paradoxical comments, most importantly those who find such "gender-neutral" voice unfamiliar or uncomfortable, I would consider, is how Ho's queerness is represented at the listening level, in which her vocality blurred the lines of masculinity/femininity. Such characteristics not only challenged listeners' convention of listening, but also opens up a gender-neutral, or to put together with the body aesthetics discussed above, a "tomboyish" vocality that were unheard, or were not widely accepted.

In one of her most well-known pop hits Rolls & Royce published in the Liangzhu (梁祝下世傳奇) album in 2005, Ho (along with the lyrics writer Wei-Wen Huang) demonstrated a romantic relationship that is "hidden" and shut from the public. The romantic relationship, when we put together the lyrics, the context of Chinese traditional legend and the music video itself, could be considered as a problematic one that seems to sway between the blurred lines of homoerotic/heterosexual desires, deconstructing the real/unreal and masculine/feminine gender boundaries. Such characteristics could be interpreted in the selected lyrics below:

男子和男子　怎能親密如此
勞斯　難面對　卻跟他勾過手指
萊斯　偏偏那樣痴
終於一次　他撲過去四目對望然後
除下襯衣
迷惑中　的勞斯　此時先至知
一向沒當這好手足女子

How could a man be so intimate with another man?
Rolls dare not to face, but indeed they've held hands
While Royce were falling head over heels
Thus one day he finally seized him, looked at him and slowly take off the fabrics
The confused Rolls finally realized
Though he never thought of such a friend as a woman

In the scenario of the music video, the visual elements also conveys some certain gender-ambiguous signs, not only on Ho's androgynous look (though in this video she hasn't shown that much gender-ambiguous characteristics in her appearance), but also on the seemingly confusing characters in the video. While the storyline depicts the hidden desire between Rolls and Royce, the gender identity of Royce is actually troubling to some people throughout the whole video. In one second Royce is featured as a young boy working with Rolls, "his" best mate; on the next camera movement with the visual displacement, Royce appears to be a woman, with straight, long hair but wearing the same outfits as the "male" Royce, such as the white T-shirt and the black suits. The playfulness draws listeners/audiences into a continuous process of questioning: is it "he" or "she" I'm looking, and furthermore, is it "he" or "she" I'm listening to? When listeners/audiences are drawn into such dialectic process, the figure of Ho and her musical works are thus queered by the people, permeating the formerly unquestionable gender-binary hierarchies.

Combining the song and the video along with the traditional Chinese legend story Liangzhu, as the cultural context, we can further discuss how Ho appropriated the Liangzhu scenario to localize the notion of queer in Sinophone Asia. The legend story originated from more than a thousand years ago, already hinted the gender-crossing romances and the ambiguous sexualities of the Liang and Zhu, especially when Zhu disguised herself as male in case to get education in the preserved ancient China, and later fell in love with her male schoolmate Liang. This famous legend demonstrated a heartbreaking story of everlasting love, and most importantly, the tension, transgression and blurring of different genders, classes, and Confucian ethic rules given by the social pressures. (Liu, 2010) Such background complicates the song, distinguishing the notion of queerness from the Euro-American context. The idea of a localized and hybridized tomboy body/voice aesthetics is thus born.

Aside from the song Rolls & Royce, another song of Ho's musical works Chi Ching Su also demonstrated such localized and hybridized tomboy body/voice aesthetics. To elaborate how such specific aesthetics is performed, I would like to first introduce the background of how the song is created. The song was originally the theme song of the musical Awakening (賈寶玉), with the famous Chinese novel Dream of the Red Chambers (紅樓夢) as its background, produced by Ho herself and directed by the Hong Kong director Yi-Hwa Lin. In the musical, Ho starred as Jia Baoyu (賈寶玉), the main male character in the origin story, accompanying with twelve actresses starring the Twelve Beauties of Jinling (金陵十二金釵). In one of Lin's interviews about the play Awakening, he described Ho as his "first choice" to the character Jia Baoyu because of her unique "gender-neutral tones in her music and her personal characteristics". With the outstanding performance of Ho and her cooperation with the director, the adaptation of this Chinese classical novel gained wide reputation from audiences in Taiwan, Hong Kong and China.

Taking a gender-neutral actress as the primary choice for casting Jia Baoyu onstage is indeed a gender-crossing performance onstage. It follows the trend of experimenting queer theories into theatrical and artistic practices in Asia in the past few decades, yet such combination is not entirely contingent, but has its contextual roots from the original texts of Dream of the Red Chamber. From the original texts, Jia Baoyu has been described as a boy who "always preferred to playing with his girl mates in the garden"; instead of following a prototypically "successful" male, studying for exams in order to get a position in the court, Jia Baoyu indulged himself into "poetry, music, and even into cosmetics that were only belong to girls". Growing up in Da Guan Yuan (大觀園), a huge, prosperous family in Ching dynasty, Jia Baoyu is well protected and preserved a sense of innocence, refusing "to grow up". Not only his slender and delicate figure but also his vulnerability and sentimental hearts showed his unique characteristics unlike other male characters in the story, and also unlike a "man" that the preserved society wished him to be. To readers of all time, Jia Baoyu is a special case that demonstrated a type of feminine masculinity.

With such context, it might not be difficult to understand the relationship between the character Jia Baoyu's feminine masculinity, and Ho's interpretation of the character with certain masculine femininities. In the two versions of the music videos (and also onstage), Ho presented her well-known tomboyish characteristics throughout the scenes. In one version of music videos featuring Xi Shu, Ho and Shu were both topless, without any clothing to modify their body lines. Ho's only significant difference with Shu was her highly masculine short undercut, shaving both sides and combing back her hair to create a gender-neutral sense of body. The scenes were generally clear and simple, focusing on the two actresses' interactions, with the only element- a delicately ornamented veil- that occasionally swayed in and out of the scene, creating a sense of distance, loss and nostalgic sorrow. In the other version of music videos Ho presented with the twelve actresses in the play, the scenes were comparatively more complicated based on the visual elements as signifiers related to the play or the song. The twelve Beauties of Jinling were seated around an ornamented long table with Ho sitting in the middle, wearing formal black suits that fits her body, yet distinguished herself as a more masculine figure comparing to the twelve actresses in the scene. While performing certain feminine masculinities through her gestures and interactions with other actresses, Ho also embodied the vulnerability of Jia Baoyu through her touches and gazes with the ornaments such as Mudan flowers, fallen leaves and petals, and the butterfly in the cage, contextualizing Jia Baoyu's tenderness through such symbols that are both feminine and on the other hand "oriental". Though the adaption of the play was in a relatively modern context, the symbols still connected with the original texts that distinguished Ho's performing aesthetics different. In her mellow and relatively low vocal lines depicting Jia Baoyu's sentiments, we can even related such gender-ambiguous tones and gestures with the Xiao Sheng (小生) manner that combined the queerness with the non-Euro-American centralized text.

In this short article, I've discussed how Ho demonstrated her tomboyish body and voice aesthetics in an East Asia (or to be more specific, a Sinophone Asia) cultural context, yet there are more issues related to such hybridized queerness that can be discussed. As a Hong Kong popular music artist, Ho's performances are by and large nurtured under the hybridized, post-colonial political context of Hong Kong, and it is also important somehow to distinguish Hong Kong Sinophone pop music's gender-ambiguous "Chineseness" from the authoritative Grand-China narrative. While the totalitarian Grand-China discourse seeks to maintain and strengthen a unified, collective nationalist discourse catering to the PRC authority's policies, the notion of gender ambiguity and the negotiation (but also appropriation) of "Chineseness" in Hong Kong popular music is actually a strategy to distance themselves from the unified Grand-China discourse. There have been various discussions on how Hong Kong popular music used such strategy as never-ending identification and disidentification as ways to "become" Hong Kong citizens. Having gone through several political confrontations and protests especially in recent years of Ho's careers, I would dare to say that such appropriation of Chinesness on the one hand problematized the idea of "China as a whole", but on the other hand also demonstrated localized queerness by her personalized tomboy aesthetics. How could such tomboy aesthetics further gain alliances with various queer communities in Sinophone Asia, not only on the gender level but also on other perspectives such as in the postcolonial scenario confronting neo-liberalism in the contemporary globalized context? By coining the tomboyish style of gender-ambiguous aesthetics into these crucial issues, there could be more discussions that could be elaborated in the future.

REFERENCES

Apparudai, Arjun. Modernity at Large: Cultural Dimensions of Globalization. Minneapolis: University of Minnesota Press.

Butler, Judith. 1990. Gender Trouble: Feminism and the Subversion of Identity. New York: Routledge.

Case, Sue-Ellen. 1988. Towards a Butch-Femme Aesthetic. Discourses, Vol.11, No.1: 55-73.

Chang Chia-Hsin (張佳欣). 2016. Hong Kong Imaging China in Jin Yong's Wuxia TV Original TV Songs (香港想像「中國」：以金庸武俠電視劇中的歌曲例). Master diss., National Taiwan University.

Chang Chuan-Fen (張娟芬). 2001. Liberality of Love: A Storybook of Lesbians (愛的自由式：女同志故事書). Taipei: China Times Publishing.

Chang Chun-Mei (張君玫). 2016. The Postcolonial Cyborg: A Critical Reading of Donna Haraway and Gayatri Spivak (後殖民的賽伯格：哈洛威和史碧華克的批判書寫). Taipei: Socio Publishing.

Chao, Yen-Ning. 1996. Embodying the Invisible: Body Politics in Constructing Contemporary Taiwanese Lesbian Identities. PhD diss., Cornell University.

Chaudhuri, Shohini. 2006. Feminist Theories: Laura Mulvey, Kaja Silverman, Theresa de Lauretis, Babara Creed. New York: Routledge.

Cheng Fan-Ting (鄭芳婷). 2017. "Towards A Queer Sensory Writing: The Theatrical Tactics of Skin Touching (2015) (酷兒感官書寫：徐堰鈴《踏青去》(2015)的戲劇策略)." Journal of Women's and Gender Studies 40:153-196.

Chou Hui-Ling (周慧玲). 2009. "The Queer Stardom and Body Enactment of Leslie Cheung (不只是表演：明星過程、性別越界與身體表演——從張國 榮談起)". Journal of Theatre Studies 3: 217-248.

Chou Yao-Hui and Kao Wei-Yun (周耀輝、高偉雲). 2015. Sonic Multiplicities: Hong Kong Pop and the Global Circulation of Sound and Image (多重奏： 香港流行音樂聲像的全球流動). Hong Kong: Chinese University of Hong Kong.

Chu Yao-Wei (朱耀偉). 2009. Years Like Song: Discussing Lyrics in Hong Kong Cantonese Popular Music (歲月如歌：詞話香港粵語流行曲). Hong Kong: Sanlian Bookstore.

de Certeau, Michel. 1984. The Practice of Everyday Life. Berkeley and Los Angeles: University of California Press. PhD diss. Cornell University.

Halberstam, Judith. 1998. Female Masculinity. Duke University Press.

Hwang Hui-Jen and Tsai Bao-Chung (黃慧貞、蔡寶瓊). 2015. Gender Politics and Locals Uprising (性/別政治與本土起義). Hong Kong: Commercial Press.

Li Yen-Ju (李晏如). 2012. Interview and Critique on Awakening of Lin Yi-Hwa (一塊頑石的還魂 滄桑歷盡的釋然：看非常林奕華《賈寶玉》). Performing Arts Review 232: 108-112.

Liu Fei-Wen (劉斐玟). 2010. "Righteousness, Romance, and Transcendence: Gender, Class, and Generic Expressive Niches of Liang-Zhu Narratives (情義、性別與階級的再現與超越：梁祝敘說的文體之音)." Journal of Theatre Studies 5:27-68.

Munoz, Jose E. 1996. Disidentifications. Ph.D. diss. Duke University.

Whiteley, Sheila and Jennifer Rycenga, edit. 2006. Queering the Popular Pitch. New York: Routledge.

Yu-Ying, Hu. 2005. The Dance of Butch/Femme: The Complementarity and Autonomy of Lesbian Gender Identity. Master diss., National Taiwan University.

ELECTRONIC RESOURCES

HOCC (何韻詩). 2017. MV of HOCC Rolls & Royce ft. Guan-Shi Chen and Dean Fujioka (HOCC何韻詩《勞斯．萊斯》MV 陳冠希、藤岡靛 特別演 出). https://www.youtube.com/watch?v=lYS8QJyIpf4

HOCC (何韻詩). 2011. MV of HOCC Chi-Ching Su ft. Shu Qi (何韻詩HOCC 癡情司MV 舒淇特別演出). https://www.youtube.com/watch?v=a6d-OR2F0hM

A Music Official Channel (東亞唱片). 2011. Official MV of HOCC Chi-Ching Su (何韻詩 HOCC《癡情司》官方 MV). https://www.youtube.com/watch?v=C9E14FWilvw

Chang Yong-Hsin (張永欣). 2012. Interpretation of Hwang Wei-Wen's Rolls & Royce (黃偉文《勞斯．萊斯》的歌詞解讀). Hong Kong Lyrics Analysis Group (香港歌詞研究小組). http://cantonpopblog.blogspot.tw/2012/04/blog-post_7775.html

Yao Shi-Min (姚詩敏). 2016. Homosexuality in Hwang Wei-Wen's Lyrics (1) - Hwang and HOCC as Best Same-Sex Love Song Partners (黃偉文筆下的同性戀(上)--黃偉文與何韻詩：同志歌曲最佳拍檔). Hong Kong Lyrics Analysis Group (香港歌詞研究小組). http://hk-lyrics.blogspot.tw/2016/06/blogpost_2.html

Yao Shi-Min (姚詩敏). 2016. Homosexuality in Hwang Wei-Wen's Lyrics (2)- Hwang and HOCC as Best Same-Sex Love Song Partners (黃偉文筆下的同性戀(下)--黃偉文與何韻詩：同志歌曲最佳拍檔). Hong Kong Lyrics Analysis Group (香港歌詞研究小組). http://cantonpopblog.blogspot.tw/2012/04/blog-post_6849.html

2013. How Do You Value HOCC's Singing Techniques? (何韵诗的唱功怎么样？). ZhiHu Forum (知乎論壇). https://www.zhihu.com/question/21002276

2009，Has Anyone Find HOCC's Voice Disturbing? (有有人覺得何韻詩唱歌好難聽？). Discuss HK (香港討論區). http://www.discuss.com.hk/viewthread.php?tid=10575803&extra=&page=4

AUDIO-TO-MIDI SIMILARITY FOR MUSIC RETRIEVAL

Fábio Goródscy
University of São Paulo
fabiog@ime.usp.br

Shayenne Moura
University of São Paulo
shayenne@ime.usp.br

Marcelo Queiroz
University of São Paulo
mqz@ime.usp.br

ABSTRACT

Finding similar aspects in sound recordings is a great concern in automated music analysis applications. In this paper, we are seeking to measure how well can state-of-the-art melody-extraction algorithms be used in creating abstractions of single-voiced audio recordings for querying. Having as motivation the development of a query-by-humming application, we created experiments where we compared the performance of features automatically extracted from wav recordings, along with its ground-truth directly calculated from MIDI files. Lastly, we discuss results showing that pitch interval representations ignoring time information can preserve some discrimination capability. The aim is to explore the limits of less informative data representations that may help deciding between comparing audio-to-MIDI or audio-to-audio in query-by-humming applications.

1. INTRODUCTION

Retrieving music from a dataset is a common task nowadays. People listen to music and they usually want to remember names of songs that linger on their memory. However, much time and effort may be spent to find one specific version of a song. Frequently, a user is unable to remember the lyrics, the artist or any other common meta-data from the music she/he is looking for; in such cases the only resource available is to hum the melody that is present in their memories.

This yields an important problem in Music Information Retrieval: to search within a dataset using only hummed/sung queries. It is necessary to transform the record into a representation that could be matched with the stored data in order to retrieve efficiently the correct music to the user.

Recovering information based on sung queries has two main challenges: codification of the information, and similarity criteria. There are works as *Tararira* [1] that focus on the note's pitch and duration to construct the codification and use this information to find similar items in a dataset. One algorithm that uses this kind of symbolic information is the *SMBGT* [2], a subsequence matching framework that allows gaps on queries and targets, where you can control parameters (variance tolerance levels, maximum match length and minimum number of matched elements) to improve the retrieval.

Another approach is to use the fundamental frequency line of melody to match with the target's melodic line. In the *Follow That Tune* system [3], a modified DTW has been used in order to calculate the alignment between query and targets. Using a modified representation to summarize melodic information, Salamon [4] has built a retrieval algorithm using the Q_{max} algorithm to compute fitness values and rank targets in a dataset (by sorting them in decreasing fitness order).

This paper deals with the problem of searching hummed queries within a dataset that uses MIDI representations. For reliability of the methods and algorithms, we used different audio to MIDI transcription systems (Melodia [5] and ASyMuT [6]) and we are concerned to make the evaluation of their performances using a custom-made version of the SMBGT algorithm. In addition, we discuss the algorithm limitations using this kind of representation, comparing the human recognition of the humming record and MIDI representation.

The article is organized as follows: section 2 brings a brief introduction of relevant concepts that allows the automatic comparison of melodic information. Section 3 describes the database, the feature extraction and the matching algorithms. Section 4 describes the evaluation using two different audio to MIDI transcriptions. The results are discussed in section 5.

2. CONCEPTS

2.1 MIDI transcription

The process of automatically transcribing audio representations of hummed sounds into symbolic versions, such as MIDI, has been explored for years, and is far from being a closed problem. Nevertheless, people usually follow a similar strategy when working in this task, which we will describe in this section.

The first step for MIDI transcription is converting the audio signal into some feature space that highlights fundamental frequencies of the voiced segments in the audio signal. This is a temporal series of fundamental frequencies that can be easily converted to a pitch sequence, which is the expected output for this task. This step is less controversial and it is mostly solved for monophonic recordings, even though for polyphonic signals it is still very challenging. ASyMuT [6] is a system designed for transcribing monophonic recordings; this algorithm creates the pitch vector by analyzing the spectral representation of the

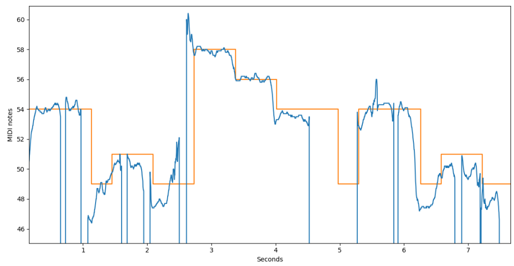


Figure 1: A pitch vector (orange) automatically extracted from a hummed query, aligned against its MIDI database version (blue).

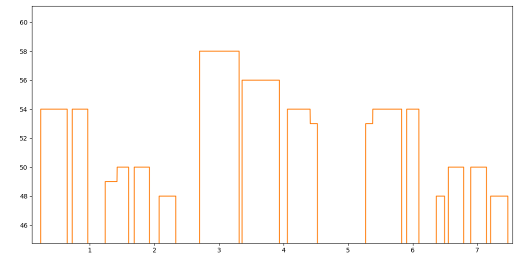


Figure 2: Final transcription after smoothing, rounding, and removing short duration notes from a pitch vector.

audio, looking for harmonic series of spectral peaks with maximal sum. Melodia [5] is designed for polyphonic signals, although it also works for monophonic signals. It has a behavior similar to ASyMuT, but it takes an extra step of evaluating the pitch candidates evolution in time, discarding extraneous pitch jumps. Figure 1 shows the results from this step.

The second step in MIDI transcription takes pitch vectors anc converts them into a sequence of discrete events, each one with specified pitch and duration, in a format similar to that of a MIDI file (that uses `note_on/note_off` events). This is the most challenging step, and there is no simple setting of the transcription algorithms that is always guaranteed to produce good results. A common pipeline for this task is to smooth the pitch vector, making jumps and fluctuations in instantaneous frequency less intense. This smoothing step is usually done by digital filters; however, choosing the right filter is hard and the most appropriate filter may vary for each recording, which is part of the challenge for this step. After obtaining the smoothed pitch vector, the pitch values can be rounded to the closest integer MIDI note. The last step is to find groups of MIDI notes that meet certain criteria (e.g. minimum duration). The final transcription is then similar to figure 2.

2.2 Interval representation

Absolute pitch values, or MIDI note values, are valuable for comparing recordings and for trying to find one specific record within a database. In order to introduce a degree of tolerance in the comparison of versions and melody contours, one solution is to consider the differences of consecutive pitch values, or *intervals*; intervals can lead to better contour matching by allowing tonal independence. One question that still has to be considered regards octave equivalence: frequently the representation of symbolic sequences can be simplified to allow only 12 pitch classes (integers between 0 and 11) instead of absolute pitch values, and also pitch class intervals instead of absolute intervals.

2.3 Time Series Matching

Ultimately, what we obtain is a sequence of timestamped events. There are various approaches for the alignment of two such sequences. One of these approaches is broadly used when building query-by-humming applications, the Dynamic Time Warping [7]. It is the default approach when comparing pitch vectors of dissimilar lenghts, and is one of the best approaches for query-by-humming when trying to maximize the recall and MRR of the results, according to Salamon et al. [8]. Even though DTW has several advantages, such as its retraceability [1] , it is not easy to adjust it to the interval representation, because the database intervals generated from the pure symbolic representation have a completely different structure than the intervals obtained by automatic transcription from the audio recording. Thus, we used another matching approach, the SMBGT algorithm [2], which is also a dynamic programming algorithm for matching sub-sequences.

The representation used and the matching algorithm have minor overall modifications compared to [2]. The SMBGT follows these steps: a sub-sequence A is compared element by element to a sub-sequence B, a matrix is calculated by placing 0 plus the last position of the matrix whenever a element is considered different (i.e. does not satisfy a pre-established condition) or 1 plus the last position of the matrix whenever it is considered equal. The algorithm has a parameter for resetting the cell score whenever it hops through a number of different cells without finding a equal element. The final score for the comparison of the two sequences is the highest value over the whole matrix, usually divided by the size of the smaller sequence, as this reflects the percentage of the longer equal sequence.

3. EXPERIMENT

3.1 Database

We considered the database that has been used in the MIREX annual competition (*Music Information Retrieval Evaluation eXchange*) for our experiments. It consists of 4431 WAV recordings of hummed queries, sampled at 8kHz. The queries were gathered through a period of 7 years, and include 195 participants and 48 songs (not all participants recorded every song in the set). This database of queries is processed and matched against a MIDI database that includes the same 48 songs among other 8474 MIDI songs from ESSEN folk music database. These 8474 extra songs were not hummed by the participants, thus they are used

[1] where we can find the exact sub-sequence that has been automatically matched to the input sequence

solely for the purpose of making the music retrieval task harder.

3.2 Feature extraction

The feature extraction algorithm works similarly to what has been described in section 2. All WAV queries are first processed for F0 determination within segmented audio frames. This gives an F0 temporal series that are later used for smoothing and approximating MIDI note values, and finally to create MIDI files. The MIDI file is used for finding intervals of two consecutive notes, yielding a sequence of intervals, which is the representation used in this experiment. We used two different implementations for this task, both of which are freely available in the Internet at the time of writing. The first one is based on Melodia [4] and its code can be found online [2]. The other implementation used was ASyMuT [6], and its code can also be found online [3]. The transcription steps are illustrated in figure 3.

3.3 Matching

Symbolic representations, such as those resulting from hummed transcriptions and MIDI re-encoding, reduce the complexity of the original melody. We need similarity measures that may be used to compare such simplified representations and the database MIDI files. For this task we used the SMBGT algorithm, which is a dynamic programming approach for the problem of comparing two sub-sequences, which was proposed within the context of a query-by-humming application [2].

MIDI transcriptions from each WAV file in the dataset are matched against every MIDI contained in the repertoire dataset. Therefore, each WAV receives 8522 similarity scores, from which a list of 10 top matches are kept. In this way, we expect to find not only the exact match, but other pieces with high melodic similarity compared to the hummed query. Even though it would be desirable, it is not reasonable to expect that the desired song is always in the first position of this list, because the tolerance added to the matching algorithm introduces ambiguities that result in high scores for many other melodies besides the ground-truth.

4. EVALUATION

Our evaluation proposal has been aimed at measuring how well the proposed method performs in trying to keep the ground-truth result among the 10-highest-scoring MIDI files. This would give us a hint on the system's ability to identify that a hummed melody matches part of a file in the dataset. For this purpose, we have taken the recall of the ground-truth belonging to the list of the 10-highest-scoring melodies. Additionally, we use the mean reciprocal rank (MRR) to find out how frequently the exact match appears in first position. Even though those measures might be seen as "quality" or "capability" of the system in giving the expected answer, we rather look at then as characteristics of the method. This way, the goal is not to achieve high recall and MRR measures, but to describe what kind of music result it is retrieving.

[2] https://github.com/justinsalamon/audio_to_midi_melodia

[3] https://github.com/adrianomitre/asymut

Smooth	Recall	MRR	Mean Position
0.05	4%	0.02	38.52
0.01	32%	0.23	4.35
1.0	6%	0.02	84.01

Table 1: Values for Melodia transcription.

Smooth	Recall	MRR	Mean Position
0.064	32%	0.19	5.13
0.128	13%	0.06	16.51

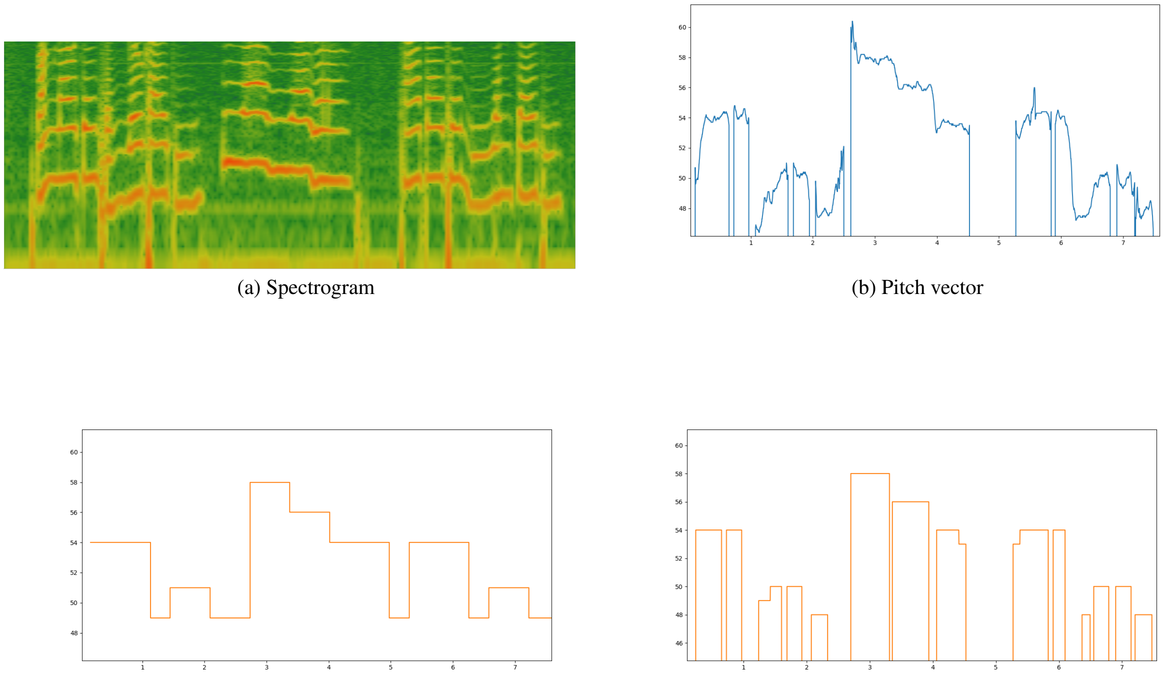
Table 2: Values for ASyMuT transcription.

Recall calculation is made by counting the number of times that ground-truth MIDI appears in the list of results, divided by the total number of queries in the dataset; i.e., recall will be 1 if every expected MIDI is returned by the system, 0.5 if half of them are retrieved, or 0 if no correct MIDI is found. MRR can be seen as the harmonic mean of the positions of the ground-truth in the retrieved list for each query; i.e. MRR will be equal to recall if every correct answer appears in first position, half the MRR if every answer is in the second position, and so on.

An important consideration regarding these metrics is their relationship to the quality of the transcription of the WAV files. The transcription algorithms have parameters for defining smoothing levels, which makes transcriptions more accurate or more tolerant to F0 deviations; setting these parameters to extreme values affects the transcription quality and the recall and MRR metrics, as can be seen in tables 1 (for MIDI transcriptions using Melodia) and 3 (for MIDI transcriptions from ASyMuT).

Furthermore, the choices of parameters for the matching algorithms also lead to changes in recall and MRR values. In the SMBGT algorithm we have explored the parameter controlling the different sizes of gaps allowed within subsequences. In our experiments, recall ranged from 34.5% to 37.0% for any combination of gaps, which means that optimization of this parameter could lead to only 2.5% improvement of recall, so that even though there is room for improvement, its impact is not considerable.

Another important point was to choose the representation that would be used for matching. We have tried to use rhythmic information from the automatic MIDI transcriptions. For comparison, we used as baseline a random classification, based on random MIDI similarity values irrespectively of the file contents. The algorithms that we used focused in keeping melodic information despite of rhythm information loss, which made our rhythmic tests always achieve very poor recall values, being in the range between 13% and 17%.



(a) Spectrogram

(b) Pitch vector

(c) Dataset

(d) Extraction

Figure 3: Steps of algorithmic transcription in our approach. a) Spectrogram of a hummed query. The strong lines in the bottom are captured by the F0-extraction algorithm. b) Result of F0 extraction. It has a similar shape to the strongest line of the spectrogram. c) MIDI file as seen in database. d) MIDI created from pitch vector.

Interval type	Recall	MRR	Mean Position
Regular	38%	0.23	4.29
Pitch classes	37%	0.23	4.33

Table 3: Recall and MRR values for different interval representations.

5. DISCUSSION

In an ideal world, a query-by-humming application would approach the behavior of the human cognition with respect to melodic similarity. As far as we have gone into this experiment, this idealized goal is not within reach by the methodology here described. One alarming behavior of the framework described is that it will frequently bring results that no human would classify as similar, for most of the queries. Furthermore, melodic similarity values will usually not indicate clearly what the correct answer is; there is always a high degree of confusion among competing MIDI files. Another problem of this methodology is that the transcription algorithms usually shorten sung notes, often because the moment of attack or release, or the period of transition between notes, makes the algorithm unsure of the pitch value in those moments and they tend to be discarded. But this note shortening makes it hard to define the parameter controlling the minimal duration for notes, because this choice would not be based on the actual problem (i.e. based on real data it would be good enough to set it to between 0.5s and 0.1s), but it would be influenced by the computer representation and the transcription errors introduced in the F0-extraction.

Regarding the choice of using absolute pitch values, pitch classes or intervals, it might be better to use the full absolute pitch vector, even though it incurs in a higher computation cost. Moreover, audio-to-audio alignment might also be an alternative to consider, if a matching strategy is defined to account for the fact that the query is tipically monophonic whereas the music database would be polyphonic (i.e. the hummed query should be aligned to the most prominent melody within the polyphonic texture).

6. CONCLUSIONS

In this paper, we considered the problem of retrieving a song from a dataset using MIDI representations calculated from hummed queries. We described the main concepts related to this task, particularly the transcription step and the SMBGT matching algorithm. We gave details of our experiments and discussed their results. Finally, we discussed some aspects relating this kind of representation to the human perception of melodic similarity, highlighting difficulties we have found. Lot of effort has been spent over this task, still it is not marked as completed. Furthermore, approaching human cognition with computers will always be complicated, and this task is a closely related to listening cognition.

Acknowledgments

The first two authors acknowledge the support by CNPq and SEMPRE: Society for Education, Music and Psychology Research.

7. REFERENCES

[1] E. López, M. Rocamora, and G. Sosa, "Búsqueda de música por tarareo," 2004.

[2] A. Kotsifakos, P. Papapetrou, J. Hollmen, and D. Gunopulos, "A subsequence matching with gaps-range-tolerances framework: A query-by-humming application," 2011.

[3] B. Stasiak, "Follow that tune – adaptive approach to dtw-based query-by-humming system," *ARCHIVES OF ACOUSTICS*, vol. 39, no. 4, p. 467–476, 2014.

[4] J. Salamon, E. Gómez, D. P. W. Ellis, and G. Richard, "Melody extraction from polyphonic music signals: Approaches, applications and challenges," *IEEE Signal Processing Magazine*, In Press (2013).

[5] J. Salamon and E. Gómez, "Melody extraction from polyphonic music signals using pitch contour characteristics," *IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING*, 2012.

[6] A. Mitre and M. Queiroz, "Um sistema automático de transcrição melódica," 2005.

[7] M. Müller, *Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications*, 2015.

[8] J. Salamon, J. Serrà, and E. Gómez, "Tonal representations for music retrieval: from version identification to query-by-humming," *Int. J. of Multimedia Info. Retrieval, special issue on Hybrid Music Info. Retrieval*, vol. 2, no. 1, pp. 45–58, Mar. 2013. [Online]. Available: http://dx.doi.org/10.1007/s13735-012-0026-0

Evaluating Melodic Encodings for Use in Cover Song Identification

David D. Wickland
University of Guelph
wickland@uoguelph.ca

David A. Calvert
University of Guelph
dcalvert@uoguelph.ca

James Harley
University of Guelph
jharley@uoguelph.ca

ABSTRACT

Cover song identification in Music Information Retrieval (MIR), and the larger task of evaluating melodic or other structural similarities in symbolic musical data, is a subject of much research today. Content-based approaches to querying melodies have been developed to identify similar song renditions based on melodic information. But there is no consensus on how to represent the symbolic melodic information in order to achieve greater classification accuracy. This paper explores five symbolic representations and evaluates the classification performance of these encodings in cover song identification using exact matching of local sequences. Results suggest the more lossy encodings can achieve better overall classification if longer melodic segments are available in the data.

1. INTRODUCTION

The landscape of todays digital music exploration paradigm has shifted greatly in recent years, and will likely continue to change. With the growth in popularity of subscription-based collections, people are discovering and consuming music in vast and varied ways on a number of devices and platforms. With such an increase in access, there is greater demand for users to explore, interact with, and share music. To this end, there is continued demand for novel and efficient ways to index, retrieve, manipulate, etc. digital music.

Symbolic melodic similarity, as a content-based approach to MIR, can be considered a sub-discipline of music similarity. The goal of melodic similarity is to compare or communicate structural elements or patterns present in the melody. Where vast efforts towards music discovery and recommender systems have historically focused on music similarity, by employing low-level feature extraction and clustering or other classification schemes, there has been comparatively less focus on melodic similarity. Many applications of similarity analysis for title retrieval, genre classification, etc., do not require the additional effort to process and interpret melodic content. Instead, relying on timbral descriptors, tags, etc. is considerably more efficient, and can often achieve equal if not better performance. However, there are numerous applications of digital musical analysis that cannot be performed without exploring the melodic content directly.

Symbolic melodic similarity research can be largely categorized into monophonic and polyphonic melodies, and into sequence similarity, or harmonic/chordal similarity. Melodies are often extracted from MIDI, MusicXML, or other digital music transcription formats, into representations such as: vectors, contours, text or numerical strings, or graphs. While considerable efforts have been made to create and evaluate melodic similarity measures with different symbolic representations, there has been little attention paid to the behaviours of these approaches with different representations.

This article explores the behaviour of local exact matching of melodies with five symbolic representations of varying information. The lengths of the local matches are used to perform cover song identification, and their classification performance is discussed.

2. LAKH MIDI DATASET

2.1 Preprocessing

The Lakh MIDI dataset was acquired for use in this research. There are many varieties of the Lakh dataset; in particular, this work employs the Clean MIDI subset, which contains MIDI files with filenames that indicate both artist and song title [1, 2]. MIDI files were scraped for track info, and any tracks titled "Melody" were parsed to acquire the melodic information. Any melody that contained two notes overlapping for greater than 50% of their duration was considered polyphonic, and was discarded. All remaining monophonic melodies were transcribed to text, including artist, song title, tempo, meter, and all melodic information (i.e. notes and durations).

Key signature data from MIDI files is unreliable. Consequently, key signatures were estimated for each melody using the Krumhansl-Schmuckler key-finding algorithm, which uses the Pearson correlation coefficient to compare the distribution of pitch classes in a musical segment to an experimental, perceptual "key-profile" to estimate which major or minor key a melody most closely belongs to [3]. The Krumhansl-Schmuckler algorithm works well for melodic segments or pieces that do not deviate from a tonal center; however, pieces that modulate or shift keys will affect the accuracy of the algorithm.

Deduplication was first handled in the original Lakh dataset, where MD5 checksums of each MIDI file were compared, and duplicates were removed. This approach is quite robust but unfortunately still requires further deduplication.

Since MIDI file or track names, or other meta data can be altered without affecting the melodic content, a further step to compare the transcribed melodies and remove duplicates was applied. This ensured that while cover songs with the same or a different artist and same song title were permitted, their transcribed melodies could not match identically.

In total, 1,259 melodies were transcribed, which gives 793,170 melodic comparisons. Of these melodies, the shortest was 14 notes long and the longest was 949 notes long. Within the 1,259 melodies, there were 106 distinct songs that had one or more corresponding cover(s) in the dataset. In total there were 202 covers present in the dataset.

2.2 Ground Truth Cover Songs Data

Using the transcribed melodies dataset, a bit map was created to annotate which melodies were covers or renditions. No consideration was given to which melody was the original work and which was the cover. The bit map was constructed such that an annotation of 1 indicated the melodic comparison was between two covers, and 0 indicated the melodies were unique (i.e. non-covers). Melodies were annotated as covers if they had the same song title and artist name, or the same song title and a different artist. Duplicate song titles by different artists were individually inspected to identify if they were genuine covers or unique songs.

3. MELODIC ENCODINGS

Melodies were encoded into five different symbolic representations of varying information loss. These encodings are: Parsons code, Pitch Class (PC), Interval, Duration, and Pitch Class + Duration (PCD). Parsons is a contour representation that ignores any intervalic or rhythmic information and only expresses the relationship between notes as $\{Up, Down, Repeat\} = \{0, 1, 2\}$. PC notation describes notes belonging to one of 12 unique pitch classes: $\{C, C\sharp, ..., B\} = \{0, 1, ..., 11\}$. The Interval representation encodes each note by its intervalic distance from the previous note (e.g. $C \uparrow G = +7,\ B \downarrow G\sharp = -3$). Interval encoding does not apply modulo operations by octave in either the positive or negative direction (i.e. intervals greater than ± 12 are permitted). Duration encoding ignores all melodic information and alphabetizes notes based on their quantized duration. Notes were quantized down to 32^{nds} using Eq. (1), where d_i is the duration of the note, tpb and met are the ticks per beat and time signature meter of the MIDI file, and $|\Sigma|$ is the size of the encoding's alphabet (i.e. $|\Sigma| = 128$ for Duration).This provides 128 possible durations up to a maximum duration of 4 bars at $\frac{4}{4}$ time. Tuples were not supported, and compound signatures were reduced to simple time signatures before quantization.

$$q_i = \left\lfloor \frac{d_i}{tpb \times met} \times \frac{|\Sigma|}{4} \right\rfloor = \left\lfloor \frac{d_i}{tpb \times met} \times 32 \right\rfloor \quad (1)$$

PCD encodes both duration and pitch class information by combining the alphabets of each encoding. Values $[0, 127]$ represent all possible durations of pitch class C, $[128, 255]$ are all possible durations of $C\sharp$, and so on. Figure 1 illustrates the PCD encoding. Both PC and PCD encodings use absolute representations of pitch values, as opposed to relative (e.g. interval). In order to compare melodies accurately, they were transposed to the same key, or the harmonic major/minor equivalent, prior to comparison.

0	1	2	3	...	128	...	256	...	1535
C	C	C	C	...	C#	...	D	...	B x4

Figure 1. Examples of the Pitch Class Duration Encoding Alphabet

4. EXACT MATCHING FOR MELODIC SIMILARITY

Evaluating melodic similarity by solving for local exact matches between musical segments often involves solving the Longest Common Substring (LCS) problem. The LCS solves for the longest string(s) that are a substring of two or more input strings. In the context of this work, melodies are encoded into strings and then compared by solving the LCS. There are two common approaches to solving the LCS: generalized suffix trees, and dynamic programming. This work employs suffix trees because of their computational efficiency.

A suffix tree is a compressed trie that represents all possible suffixes of a given input string [4]. The keys store the suffixes and the values store the positions in the input text. Constructing suffix trees was done using Ukkonen's algorithm, which constructs a suffix tree in $\mathcal{O}((n + m))$ time, where n and m are the lengths of the two input strings [4]. Similarly, the LCS can be solved in $\mathcal{O}((n + m))$ time by traversing the suffix tree.

Generalized suffix trees (GST) are created for a set of input strings as opposed to a single string. The input strings are each appended with a unique character, and then concatenated together to form one aggregate input string. In this work, each pair of melodies being compared were used to create a GST to solve for the LCS of the two melodies. Once constructed, the GST is traversed to annotate nodes as *X* for suffixes belonging to the first melody, *Y* for suffixes belonging to the second melody, and *XY* for suffixes common to both melodies. The path from root to the deepest *XY* node represents the LCS. Figure 2 shows the GST of input strings "*ABAB*" and "*BABA*", such that the concatenated input string is "*ABAB$BABA#*". Paths denoting substrings "*ABA*" and "*BAB*" are both solutions to the LCS.

5. SEQUENCE COMPLEXITY

Shannon entropy measures the average amount of information generated by a stochastic data source, and is calculated by taking the negative logarithm of the probability mass function of the character or value [5]. Shannon entropy is given by H in Eq. (2) where b is the base of the logarithm and p_i is the probability of a character number i occurring

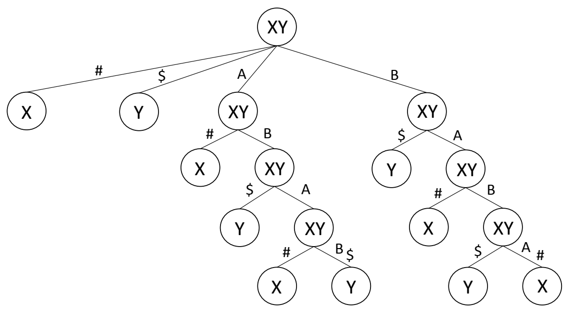


Figure 2. Annotated Generalized Suffix Tree for Input String *ABAB$BABA#* to solve for the LCS

in the input string [6]. In this work, $b = 2$, such that the units of entropy are bits.

$$H = -\sum_{i=1}^{n} p_i \log_b p_i \tag{2}$$

Shannon entropy establishes a limit on the shortest possible expected length for a lossless compression that encodes a stream of data [5]. For a given input string, when a character with a lower probability value occurs, it carries more information than a frequently occurring character. Generally, entropy reflects the disorder or uncertainty in an input, and is used in this work as an approximation to the complexity of an encoded melodic segment.

All non-cover song melodies (i.e. unique) were traversed with a sliding window to calculate the average entropy for a given window length. Figure 3 shows the average entropy, H, as a function of window length for each of the five melodic encodings. The encodings with the smallest alphabet plateau at the lowest average entropy, whereas the encoding with the largest alphabet grows toward a much larger average entropy value.

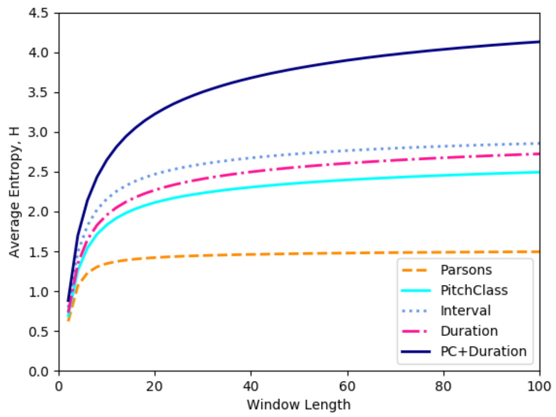


Figure 3. The average entropy, H of unique melodic segments as a function of window length

From the cover songs dataset, the exact matches for each melodic comparison were transcribed for all encodings. All match segments were categorized by their length to compute the average entropy by match length for each of the five encodings. Figure 4 shows the average entropy, H, of the exact match melodic segments as a function of their match length.

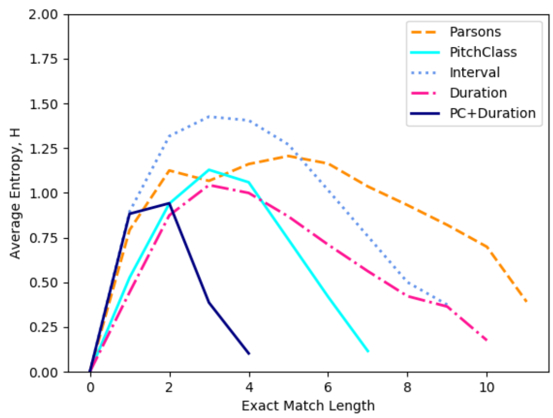


Figure 4. The average entropy, H of exact match melodic segments as a function of match length

Interval encoding achieves the greatest average entropy at a match length $l = 3$, and Parsons has greater average entropy values for the longer melodic segments (i.e. $l > 5$). PCD exhibits the lowest average entropy for nearly all match lengths. This may suggest that while larger alphabet encodings can preserve more information, exact matching techniques such as solving the LCS often discover short, repeating patterns, of comparatively low complexity.

6. COVER SONG IDENTIFICATION

6.1 Binary Classification

Binary classification is the technique of classifying the elements of a given set into two groups on the basis of a predicting or classification rule [7]. In the context of cover song identification, we are interested in identifying which melodies are unique and which are covers. With the ground truth annotated data, we can set a threshold for the length of the LCS between two melodies to predict whether they are unique or covers. Melodies with a LCS shorter than this threshold are predicted to be unique, whereas melodies with a LCS of this length or greater are predicted as covers. A confusion matrix, shown in Table 1 illustrates the four possible outcomes of these predictions: true positive (tp), false positive (fp), true negative (tn), and false negative (fn).

The Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) score are commonly used in binary classification to represent the quality of an automatic classification scheme or rule [8]. The ROC curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR) at various classification thresholds. TPR and FPR are calculated using Eq. (3) and Eq. (4) respectively. The further the curve deviates from the diagonal midline (i.e. extending from $(0, 0)$ to $(1, 1)$), the better the

		Predicted	
		p	**n**
Actual	**p′**	true positive	false negative
	n′	false positive	true negative

Table 1. Confustion Matrix for Binary Classification Scheme

quality of the classifier, assuming the positive prediction is more desired than the negative prediction.

$$TPR = \frac{tp}{tp + fn} \tag{3}$$

$$TPR = \frac{fp}{tp + tn} \tag{4}$$

The AUC score is a normalized measure of the predictive quality of a classifier. An area of 1 represents a perfect classifier, and an area of 0.5 implies the classifier is no better than random guessing.

6.2 Classification Performance

The five melodic encodings were used to compare all melodies against each other to solve for the LCS in every comparison. The lengths of the exact matches were used to predict if the two melodies being compared were covers or unique songs. Figure 5 shows the ROC curves for the five melodic encodings for all exact match length thresholds.

Parsons is the most lossy encoding (i.e. preserves the least information) but achieves the greatest AUC score of all the encodings. The PCD encoding preserves the greatest amount of information of all the encodings and achieves the lowest AUC score. It is notable that while PC is the second-most lossy encoding, its AUC score is lower than Interval and Duration, both of which have considerably larger alphabets and preserve more information. The poor performance of PC and PCD encodings may be due in part to some inaccuracy in the key-finding algorithm; however, it is unlikely these encodings would perform notably better with a perfect key-finding algorithm.

The top left corner at position $(0, 1)$ of the ROC plot represents the perfect classification scheme, with a TPR of 100% and a FPR of 0% [9]. One common approach to selecting a classifier threshold in practice is to identify the point on the curve closest to $(0, 1)$. Table 2 shows the closest point on each of the five encodings' ROC curves to $(0, 1)$, and the exact match threshold at this point. There are circumstances where a greater emphasis on TPR or FPR may be desired, and so a trade-off can be made by selecting a threshold that better suits the application of the

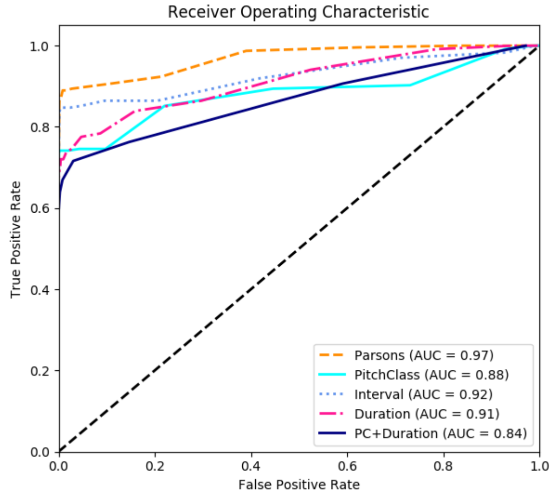


Figure 5. Receiver Operating Characteristic for the five melodic encodings using exact matching

classifier. The ability to select the classification threshold for a desired performance is an important aspect of the ROC curve.

Encoding	**FPR**	**TPR**	**Dist. to (0, 1)**	**Ex. Match Length**
Parsons	0.028	0.894	0.109	14
Pitch Class	0.043	0.746	0.258	9
Interval	0.005	0.847	0.153	13
Duration	0.159	0.839	0.226	8
PC+Duration	0.147	0.763	0.279	4

Table 2. Closest points on ROC Curves for Each Melodic Encoding and the Corresponding Exact Match Length Threshold

7. CONCLUSIONS

In this work, the behaviour of local exact matching as a measure of melodic similarity is applied to melodies encoded with five symbolic representations of varying information. Generalized suffix trees were used for each melodic comparison to solve for the longest common substring between two melodies. The lengths of these local exact matches were used to predict cover songs in a dataset of both unique and cover song melodies.

Parsons code achieves the best overall classification performance at any exact match length threshold, and it is most discriminant at an exact match length threshold of 14. Large alphabet encodings such as PCD achieve poorer classification performance. Results suggest lossy encodings such as Parsons, achieve their best classification rates with longer exact match lengths than encodings that preserve more information.

The average entropy of unique melodies in the dataset grows with the window length of the melodic segment, and with the size of the alphabet of the encoding. The average entropy results from the exact matches of cover song melodies suggests encodings that drive higher complexity exact matches are beneficial; however, ultimately the longer melodic segments are better at differentiating cover song melodies from unique song melodies.

In future work we would like to explore the effects of more granular quantization on the Duration and PCD encodings. A non-repeating contour representation should be compared to Parsons to illustrate the effects of repeating notes in exact matching and to determine if even lossier symbolic representations can achieve as good or better classification performance. It would be advantageous to compare Shannon entropy results to a practical approximation of Kolmogorov complexity such as one or more lossless compression algorithms. Lastly, an investigation of complexity and classification performance with inexact matching similarity measures, such as edit distance, could illuminate the benefits and drawbacks of the faster exact matching approach.

8. REFERENCES

[1] C. Raffel, "Lakh MIDI Dataset v0.1," http://colinraffel.com/projects/lmd.

[2] ——, *Learning-based methods for comparing sequences, with applications to audio-to-MIDI alignment and matching*. Columbia University, 2016.

[3] D. Temperley, "What's key for key? the krumhansl-schmuckler key-finding algorithm reconsidered," *Music Perception: An Interdisciplinary Journal*, vol. 17, no. 1, pp. 65–100, 1999.

[4] D. Gusfield, *Algorithms on strings, trees and sequences: computer science and computational biology*. Cambridge university press, 1997.

[5] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, p. 379?423, 1948.

[6] P. Grunwald and P. Vitányi, "Shannon information and kolmogorov complexity," *arXiv preprint cs/0410002*, 2004.

[7] A. Ng, K. Katanforoosh, and Y. Bensouda Mourri, "Neural networks and deep learning: Binary classification," https://www.coursera.org/learn/neural-networks-deep-learning/lecture/Z8j0R/binary-classification.

[8] T. Fawcett, "An introduction to ROC analysis," *Pattern recognition letters*, vol. 27, no. 8, pp. 861–874, 2006.

[9] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve." *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.

ASSESSING THE SOUND OF A WOODWIND INSTRUMENT THAT CANNOT BE PLAYED

D. Keith Bowen
Royal College of Music, London
keith.bowen@rcm.ac.uk

ABSTRACT

Historical woodwind instruments in museums or private collections often cannot be played, by virtue of their poor condition or the risk of damage. Acoustic impedance measurements may usually be performed on instruments in good condition, but only if they are in playable condition. Many museum specimens are not. However, If the bore shape and tone holes are measured accurately, we are able to compute the acoustic impedance of the instrument for all fingerings. Conclusions may then be drawn about the instrument's pitch, intonation, temperament, fingerings, effects of bore shrinkage and even the timbre of the notes. A simple linear, plane- and spherical-wave computational model, originally developed for calculating the acoustic impedance of conical-bore woodwinds, is here applied to bass clarinets for the first time. The results are assessed by experimental impedance measurements and by playing tests on an historical Heckel bass clarinet in A of 1910 that has been continuously maintained in playing condition but relatively lightly used. In all cases the lowest two to five frequency impedance peaks agreed well with the calculations. The method is shown to be a viable method for the examination of historical woodwind instruments.

1. INTRODUCTION

The aim of the investigations in this paper is to test the idea that it is possible to model the input impedance of a woodwind instrument sufficiently accurately that one may draw reliable conclusions about its behaviour purely from geometrical measurements of its bore, tone holes and keypads. This will enable the vast collections of woodwind instruments in museums to be used for primary evidence of their sounds without risk of damage.

There is a very large number of musical instruments in museum collections. However, many institutions preclude playing the instruments because of the risk of damage [1]. This is especially true for woodwind instruments where the act of playing rapidly introduces air at a much higher humidity and temperature, triggering potentially damaging reactions in the wood. Moreover, even if playing is permitted, it is fairly unlikely that a wind instrument 150 – 200 years old will be usefully playable without restoration that goes well beyond normal conservation.

However, museums will normally permit the handling and careful measurement of instruments that are not too fragile, to an accredited researcher under supervision and the guidleines of ICOM/CIMCIM [2]. This has been used to study the development of types of musical instrument and their keywork (see, for example, [3, 4, 5] for clarinets), but their sounds have so far been mostly inaccessible, apart from the small number of restored instruments.

The principle upon which the main methodology of this paper is based is that the sound of a wind instrument is largely dominated by the shape of its air column, as indicated by its resonance or 'input impedance' spectrum [6, 7]. This is not to say that the mouthpiece/reed is unimportant, but that, at least up to the middle of the clarion register, it has a much smaller effect on the intonation of each note than does the air column [8, 9].

A well-preserved instrument from 1910 was used for this trial. Standard acoustic computational methods were used to calculate the impedance spectra for each note of the instrument, and two tests of the accuracy were performed: one by measuring the input impedance directly in the laboratory, and the other by playing tests on the instrument, measuring the frequency of the note emitted at each fingering and looking at the predicted intonations produced by both 'normal' and 'alternative' fingerings.

2. COMPUTER MODELLING OF WOOD-WIND INSTRUMENTS

The development of methods of modelling woodwind instruments has taken place over more than a century, beginning with the analytical ideas of Hemholtz [10]. Major contributions were made by Bouasse [11] and especially by Benade and his collaborators [7]. The understanding of woodwind acoustics progressed through analytical expressions for lossless and then lossy systems [12, 13, 14], linear system calculations [15], analysis of the reed/mouthpiece system [e.g. 10, 16, 17, 18], impedance of the bell [19, 20] and non-linear treatment of the reed generator [21, 22]; an excellent recent treatment appears in Chaigne and Kergomard [23]. In 1979, Plitnick and Strong [24] first applied the computer modelling method to the whole instrument. They split the bore (of an oboe in this case) into short cylindrical segments approximating the conical shape of the bore (the staircase approximation), started from the calculated impedance of the bell radiating into open air and summed each complex impedance, in series for the segments and in parallel for the tone holes. A reed cavity impedance was added in parallel at the end of the sum. The result was the spectrum of impedance peaks as a function of frequency over the audible band. Note that this and most other approaches are based on linear acoustic theory and strictly only apply to small amplitudes. This suffices for the calculation of resonance peaks, but the effects of large amplitudes are critical in the understanding of the peaks actually selected, as discussed below.

This is essentially the method used today. Differences are in the expressions for tone hole impedances, for wall losses and the radiation impedance of the bell, and in the matrix formulation analogous to transmission line theory which significantly speeds up the calculation [25]. Nederveen [26] has added some valuable insight into the elements of the modelling equations. Research on simulating clarinet sound dynamically using digital formulations of the air column and reed/mouthpiece system are also reaching an interesting stage [27, 28].

The program used here was written in MatLab™ and depends largely on the equations given by Keefe [25]. This contains the main advances made in theoretical modelling since Plitnik and Strong, though we added the treatment of reed and embouchure impedance from Dalmont *et al.* [29]. We were able to use the program IMPEDPS written by Robert Cronin [30, 31] to test our program, since its source code was kindly provided to us and we could configure our program to use identical algorithms. The methodology is generally applicable to reed-driven instruments, which all share a similar non-linear generation and feedback mechanism at the reed, while the cylindrical sections of clarinets are simply cones of infinite length. We have tested the basic assertion by comparing calculated impedances to experimental measurements of impedance and to audio playing tests as described below.

Many of the investigations involving the concept of acoustic impedance so far have been to test the acoustic theory and modelling [32, 33, 34] and to control manufacture [35, 36, 37, 38, 39, 40, 41] rather than to learn about the musical behaviour of historical instruments. The main exception is the work on bassoons by Cronin and Keefe [30, 25], Dart [42] and Hichwa and Rachor [43], in which the viability of alternate fingerings, the intonation and temperament, the quality of alternate wing-joint and boot-joint designs were examined. Jeltsch, Gibiat and Forest were able to perform acoustic impedance measurements on a set of four six-key clarinets made by Joseph Baumann (fl. Paris, c. 1790 – c. 1830) [44]. The set was in very good condition, so they could compare impedance measurements with playing frequencies, and also make comparisons with a modern (Noblet) clarinet. Jean-Xavier Lefévre refers to this maker's clarinets in his famous tutor [45] and gives particular fingerings to exploit or overcome their characteristics. In their data analysis they concentrated on the harmonicity relations produced by the fingerings of the clarinets. They showed, for example, that the first register was not well tuned. Lefévre remarked on this feature in his tutor and also composed his sonatas mainly in the second register of the instrument. The modern clarinet showed much better alignment of the harmonics. Jeltsch and Shackleton have performed a similar study on early nineteenth century clarinets by Alexis Bernard et Jacques Francois Simiot [46]. Bass clarinets do not appear to have been studied so far.

The impedance spectrum shows the resonances in the tube that are capable of sustaining an oscillation in combination with the reed/mouthpiece generator. They will only make a good musical instrument if the harmonics of an oscillation based on one resonance coincide with other resonances, thus forming a 'regime of oscillation', when the non-linear generator combines with two or more resonances to form a stable tone [21, 21, 22].

3. COMPUTATIONAL METHODOLOGY

The program is an implementation of the well-established linear, small-signal plane- and spherical-wave acoustic impedance modelling equations. We shall cite sources for the key parameters and the necessary equations: the radiation impedance of a bell, the impedance of a conic segment, the impedances of tone holes and the impedance of the reed/embouchure.

3.1. Input parameters and equations

The following parameters were used: speed of sound, =347 m s^{-1}; density of air =1.19 $kg.m^{-3}$; viscosity of air =1.85 10^{-05} Pa s ; specific heat ratio C_p/C_v = 1.4; thermal conductivity of air =2.63 10^{-02} $Wm^{-1}K^{-1}$; specific heat at constant pressure C_p = 1.006 J $kg^{-1}K^{-1}$ These were chosen for appropriate playing conditions, that is, a somewhat elevated temperature and humidity and a substantially elevated CO_2 content of the exhaled air [26]. The laboratory measurements were made under normal laboratory conditions, approximately 20°C and normal atmospheric composition. Coincidentally but conveniently, the product of air density and speed of sound (which determines resonant frequencies) for these two conditions agree to better than 0.2 cents, below the limits of intonation discrimination by human ears.

The computation starts from the radiation impedance of the bell, and works up in segments to the mouthpiece. The bell formula was taken from experimental data from Benade and Murday [47], who give explicit formulae for the equivalent-length end correction due to the radiating aperture, dependent on the geometry of the aperture. This is converted into impedance by the standard formula for a lossless cylinder (e.g. [25]), since there are no walls to cause losses. As noted by Chaigne and Kergomard [2323, p. 684], there are no known formulas for the radiation impedance of a cone or flared bell, hence at present the semi-empirical formulas must suffice; however the choice does not strongly influence the end result.

The impedance of a conic section, in terms of the exit impedance of the previous section, is given in Keefe's 1990 paper on the modelling of woodwind air columns [2525]. This is a spherical wave solution, and includes viscous and thermal losses at a smooth wall. Segments end either at a tone hole, or at an output diameter within 10% of the input diameter, so that the wall losses (which depend on diameter) are calculated reasonably accurately.

Keefe's paper was also used for the tone hole corrections, with series and shunt length corrections to the segment impedance as given in his equations 5-9. Separate equations are needed for closed tone holes, open tone holes and open keyholes with a pad at a certain distance above the hole. These depend on both Keefe's theoretical models and on experiments by Benade and Murday [47] and by Cronin and Keefe [unpublished].

The reed impedance should be accounted for [48, 26]. In the initial calculations the column was terminated with an infinite impedance in order to compare closely with the experimental measurements (see below). To model the actual playing frequency, we should need the impedance as seen from the mouthpiece looking at the reed; the imaginary part of the "embouchure impedance" should be equal and opposite to that of the appropriate resonance peak to ensure no phase shift around the feedback loop to the reed; It therefore includes contributions from the reed, mouth and oral cavities. Thus the frequencies selected by the instrument will be slightly below impedance peaks of the tube alone. We have used the model of Dalmont et al [29], who show that the mouthpiece/reed combination can be taken into account by adding a small equivalent-length segment on to the top of the mouthpiece segment of the model. For soprano clarinets they found 7 ±2 mm for this correction and by considering the scaling of the equation we would expect 14 ±4 mm for bass clarinets.

4. MATERIALS AND METHODS

4.1. Description of the instrument and measurements

The instrument used for the tests was a Heckel bass clarinet in A from 1910 shown in Figure 1, owned by the author. It is a 21-key system including 5 plateau keys, and is German system with a so-called patent C#. Dated at 1910 from Heckel records [49] and formerly owned by the Kiev Symphony Orchestra, this has been kept in playing condition all its life, but lightly played (there are relatively few orchestral parts for the bass clarinet in A [50]. It is therefore a good experimental instrument for this project.

Figure 1. The Heckel bass clarinet in A used for the trials.

Bore diameters were measured with a set of graduated circular discs on the end of aluminium tubes. The bore is 23.2 mm for all its length, with a largely-conical flare beginning 153 mm from the bell. The mouthpiece was made by E. Pillinger to the dimensions of an original Heckel Bb bass clarinet mouthpiece in Nuremberg (D.N.gnm.MIR480), published by Bär [51].

Tone hole positions were measured with a calibrated tape measure to ±0.25 mm; tone hole diameters and depths and bore disc diameters were measured with a polymer caliper with accuracy ±0.1 mm. In addition to the tone hole centres and diameters, the chimney depth, diameter of the body at the tone hole position, the diameter of the tone hole keypad (where fitted) and its opening height were measured. The radius of curvature of the outer tone hole edges was estimated at 1.0 mm. These parameters all enter into the expression for the tone hole impedance when opened. Approximately 300 measurements in all were used to describe the instrument. We estimate that the parameters affecting the tuning (tone hole positions) are measured to 0.5%, corresponding to an average tuning accuracy of better than 5 cents. Since each length measurement is independent, this error applies separately to each note, and is not cumulative. The mouthpiece and crook were measured by filling with water and weighing the water, taking the average of ten measurements.

4.2. Experimental impedance measurement systems

Two systems were used to measure impedances in the laboratory: an Open University built-in-house single-microphone capillary system that has been extensively calibrated [52] (courtesy Prof. D. Sharp), and the commercial BIAS (Brass Instrument Analysis System) modified for woodwind [53, 54, 55]. One measurement (note G3) was made with the in-house system, which verified that the agreement between the methods was good. For all subsequent meas-urements the BIAS system was used. Both the BIAS and single-microphone measurement systems are capillary-based. That is, a capillary channel connects a controlled sound source to the entrance of the wind instrument to be measured. The capillary is designed to have an impedance that is frequency independent, and has a much larger magnitude than that of the air column being measured.The general principle draws from recording two characteristic signals at each end of the capillary, which allows one to obtain a good estimation of both the pressure and volume flow rate at the entrance of the measured instrument (one of which may be made constant using some active control). Provided the wavelength is sufficiently above that of the instrument's bore, the ratio of pressure over flow rate gives the plane wave component of the impedance. Phase information can also be obtained from the system through the use of a phase meter connected to the two microphones.

An adaptor was made from nylon to fit the BIAS system at one end and the mouthpiece socket of the bass clarinet at the other. The volume of the adaptor was made to be the same as that of the instrument mouthpiece at 28 cm^3, and the end fitted closely to the BIAS system.

4.3. Audio frequency measurements

In order to compare the measured and calculated impedances with the pitches actually produced, the instrument was played, and the sounds recorded over full chromatic scales. Each note was played for several seconds, without looking at a tuner and while attempting to play in the natural 'centre' of each note. The frequency was estimated by chopping the transients at the beginnings and ends of each note, and using the YIN algorithm to determine the frequency [56]. The accuracy of this method is estimated by its authors to be approximately ±1 cent.

5. RESULTS

5.1. Comparison of calculations and acoustic measurements

Waves with frequencies beyond the tone-hole cut-off limit are not reflected at the first open tone hole but transmit through to and out of the bell (which is usually designed to have a similar cut-off frequency). Such waves do not contribute to the standing waves in the instrument nor to the feedback that stabilises the oscillations of the reed, though they can contribute weakly to the sound spectrum. The tone-hole cut-off frequency for this instrument is about 1000 Hz, calculated from Benade's approximate formula [7] for an open tone-hole lattice

$$f_c = 0.11c\left(\frac{b}{a}\right)\left(\frac{1}{sl}\right)^{1/2}$$

where f_c is the cut-off frequency, c the speed of sound, a the pipe radius, b the hole radius, s the hole spacing and l the acoustic length of the holes. The result is confirmed by visual inspection of the impedance spectra. It is worth noting this value, since for bass clarinets, and also by scaling from soprano clarinets, one would normally expect a cut-off around 750 Hz [77]. This is a significant parameter to evaluate in the study of historical instruments, since it affects the musical sound and playing qualities. This is discussed by Benade [77], who notes that woodwind instruments have actually 'evolved' over the centuries so that their cut-off frequencies became approximately constant over the whole range of the instrument. We thus chose the frequency range 20 – 2000 Hz for both the measurement and calculations. The range on the instrument for analysis was chosen to be from written E2 to D5 (69.3 to 494 Hz fundamental peaks), corresponding to C#2 to B4 concert pitches). Whilst information could be obtained from higher note fingerings, it is less significant. Only one harmonic is available for generating pitches above about G4, and this can be varied widely by embouchure control in the altissimo regime. In this regime the pitch of the sound produced is more reliant on the skill of the player than on the instrument.

We first show a few notes from (written) E2 to C5 (in SPN) with experimental and calculated impedances superimposed (

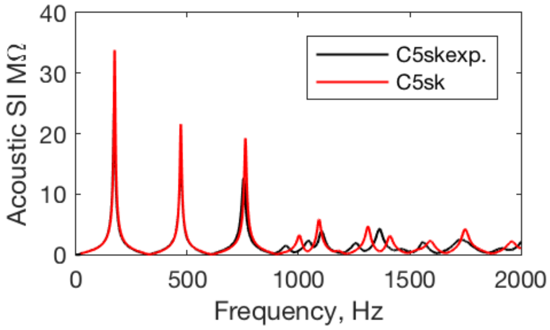
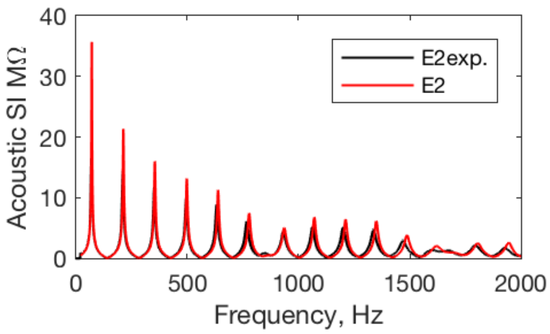
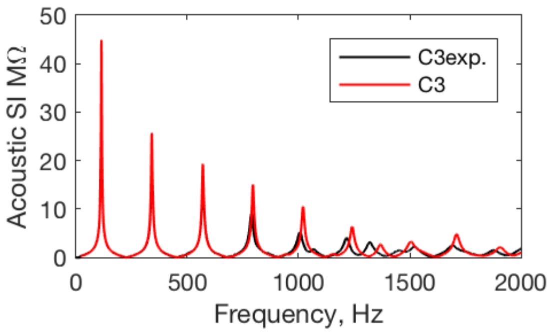
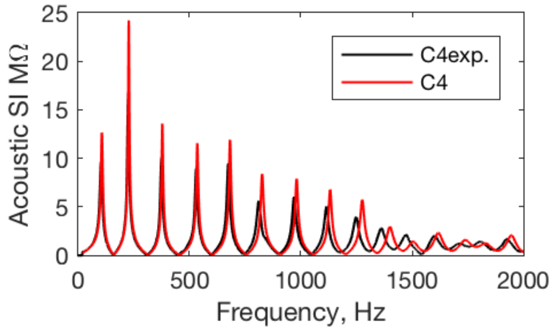


Figure 2). The experimental absolute values of the impedance peak amplitudes agree well in frequency but are up to 2× lower in amplitude. This is consistent with the results of Plitnik and Strong [24], indicating that some losses in the tube, such as fingers, pads, edges, or porosity are not taken into account.







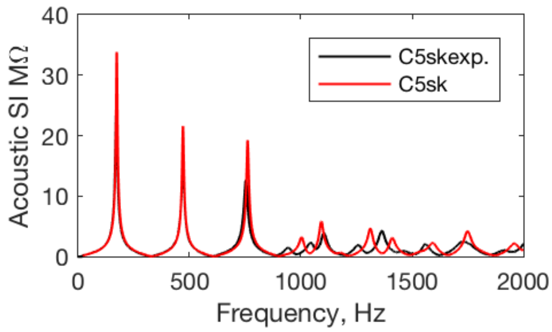


Figure 2. Four comparisons of experimental and computed results, from low written E2 up to C5.

The measured and calculated lines largely overlap for each note, but the measured amplitudes are significantly lower and the frequencies very slightly lower. Note that for C4 and above, the second impedance peak becomes the basis of the sound, through use of the speaker key, which depresses and shifts the first resonance out of a harmonic relationship with subsequent resonances.

The overall picture is shown by Figure 3, which shows the departures from equal temperament for the calculated and measured impedance values and for the frequencies shown by the playing tests. As expected, the playing frequencies are slightly below the impedance peak values. It is seen that the instrument is playing somewhat sharp, relative to equal temperament at A4=440 Hz, and becomes sharper at higher notes.

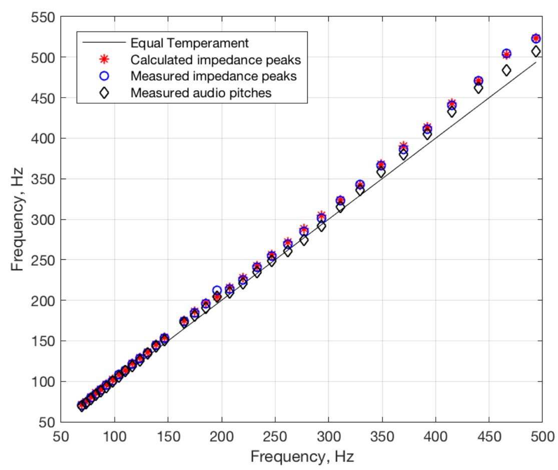


Figure 3. Graphs of calculated impedance peaks, measured impedance peaks and measured audio pitches for notes from E2 to D5. The 'break' in the instrument ranges between written Bb3 and B3 occurs at about 200 Hz and

that between C5 and C#5 at about 450 Hz. Up to the first break the first resonance frequency is plotted, between the first and second break the second resonance and above the third break, the third resonance peak.

It is useful to express the frequency differences in cents. This gives a deviation from a target pitch by an amount that is comparable over the whole range. Figure 4 shows the measured and calculated impedance peaks, and the measured audio pitches, relative to equal temperament.

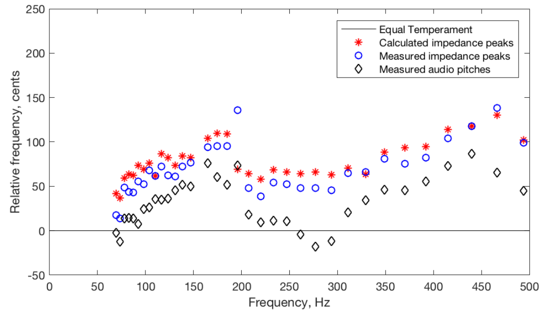
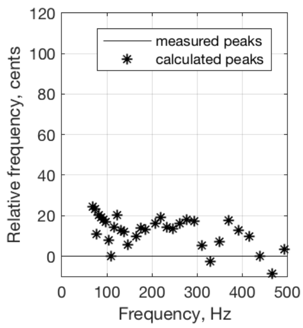


Figure 4: Deviation in cents for each note. The horizontal line at y=0 represents equal temperament at A4=440 Hz.

Whilst there is scatter, the variations in each function appear to track one another. Figure 5 therefore shows the differences between calculated and measured impedance peaks. The calculated peaks average 10 ±8 cents higher than measured peaks. Figure 6b shows the difference between the measured impedance peaks and the playing frequencies. These average at 37 ± 8 cents.



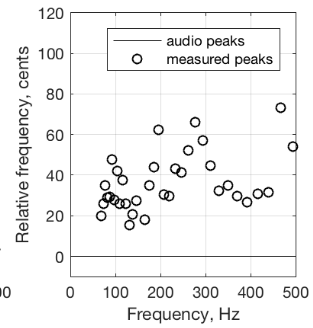


Figure 5 (left): differences between calculated and measured impedance peaks.

Figure 6 (right): differences between measured impedance peaks and audio playing frequencies at *mf* levels.

Since the impedance peak differences between calculation and experiment are reasonably consistent, they appear to be systematic and might be reduced by further development of the computation, for example to take account of other losses such as wall porosity. However, an agreement within 10 cents, which may be corrected empirically as shown below, is sufficiently accurate for the research into historical instruments.

The difference of approximately 37 cents between the measured (or corrected calculated) peaks and the playing frequencies is ascribed to the embouchure correction discussed above. The results are similar to those of Dalmont *et al.* [29] though there is more scatter, possibly because the latter used a blowing machine not a player. We therefore recalculated the impedances with a number of embouchure equivalent lengths added to the top of the column, just before the terminating impedance, simply by extending the length of the segment representing the mouthpiece volume. Our best estimate is that the equivalent length required for compensation of the small differences between calculated and experimental impedances is 3 ±1 mm and that the equivalent length of the embouchure correction should be 17 ±4 mm. The latter is consistent with the results of Dalmont *et al.* [29]. These are simply added onto the mouthpiece segment. We do not know how closely the copy of the Bb mouthpiece is to the original supplied with the A clarinet. However, its volume was accurately measured, so the results should be consistent between calculation and playing. Figure 7 shows the differences between the calculated impedance peaks and the audio frequencies for two cases, first with the mouthpiece pushed fully in and then with it pulled out by 10.8 mm. It is seen that the same correction gives consistent results in the two cases.

5.2. Investigation of alternative fingerings

Most of the application of modelling to understanding historical instruments will be comparative, for example, how in tune are the alternative fingerings? We tested this by calculating and playing several notes that have, or may have, alternative fingerings: (written) Bb2, Eb3, F3, C#4 and C5. These are referred to as 'normal' or 'fork' and are shown in Table 1. Only the calculated results are shown.

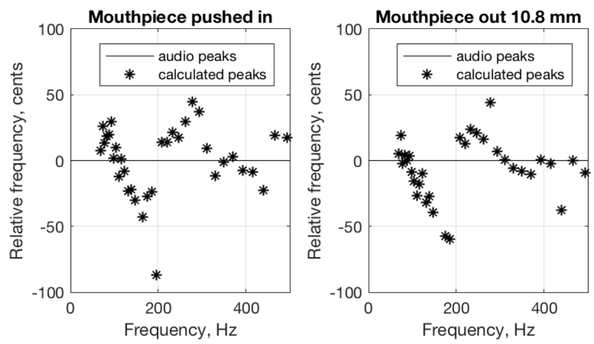


Figure 7. Comparison between calculated impedance peaks and audio playing frequencies when the overall end correction was 20 mm. (a) with mouthpiece pushed in, (b) with mouthpiece pulled out.

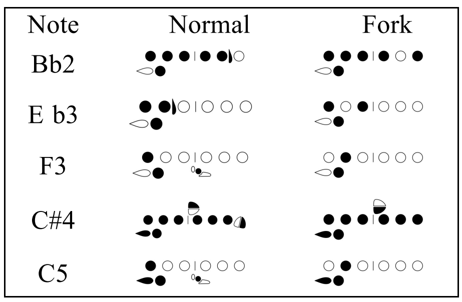


Table 1. Alternative fingerings investigated. [57].

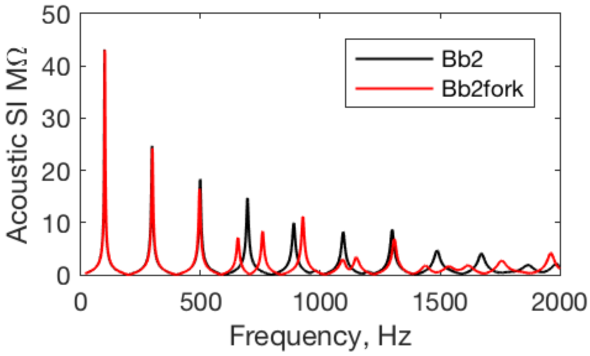


Figure 8. Calculated impedance spectra for two fingerings for the note Bb2.

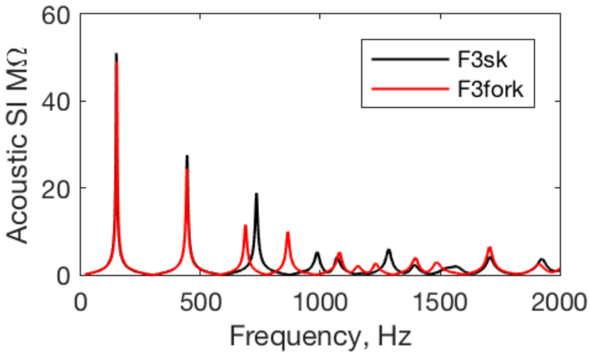


Figure 9. (right). Calculated impedance spectra for two fingerings for the note F3.

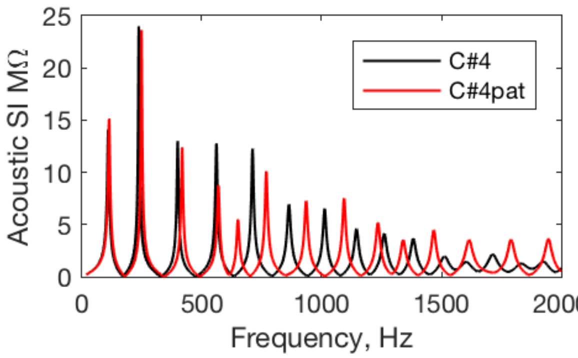


Figure 10. Calculated impedance spectra for two fingerings for the note C#4.

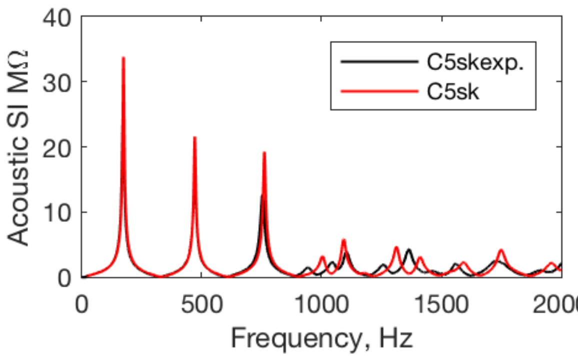


Figure 11. Calculated impedance spectra for two fingerings for the note C5.

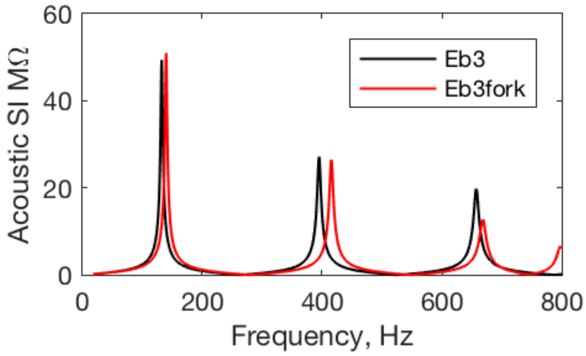


Figure 12. Calculated impedance spectra for two fingerings for the note Eb3.

The calculated impedance spectra for the notes are shown in Figure 8 Figure 12. In all except Figure 12 the fundamental and at least one other resonance aligns well between the two fingerings and these also align with fundamental and third harmonic of the designed note (not shown). For some notes, especially the "patent" C#4, the resonances are a good fit for the 5^{th} and 7^{th} harmonics also. The observation on playing was that a two-resonance match was sufficient to produce good intonation match of the fingerings, but that the timbre of the tone was better matched if more resonances were aligned.

However, the forked D#/Eb3 (Figure 12) showed no such match, and played almost a semitone sharp, just as predicted from the impedance curves. Whilst the fork fingering is often acceptable for this note on earlier German system clarinets it is clearly not the case here, and is in fact generally not the case for Albert system clarinets.

6. CONCLUSIONS AND FUTURE WORK

We recall that the model used is based on small-signal, linear, plane- and spherical-wave acoustics, with viscous and thermal wall losses. It does not take account of some loss mechanisms such as wall porosity, internal tone-hole edge turbulence and finger and pad absorption. Nevertheless, it is remarkably accurate for the absolute values of resonance frequencies and the relative heights of resonance peaks. We conclude that the method is certainly accurate enough for the purpose of reconstructing the acoustic impedance (resonance) spectra of instruments of bass clarinets. This extends the conclusion of Dalmont *et al.* [**Error! Bookmark not defined.**] from soprano clarinets, oboes and alto saxophones to bass clarinets, and provides a measurement of the embouchure equivalent length in the instrument studied.

We believe that we achieve tuning accuracy at worst within a few cents, which is entirely adequate to measure the pitch and temperament at which an instrument was designed to play. The relative accuracy within or between instruments would be much better, so we may, for example, compare the tuning of alternative fingerings for notes, determine the temperament in which the instrument was constructed or compare the overall acoustic behaviour of two different instruments. This will be performed for a number of bassoon-form bass clarinets as part of a historical and acoustic study.

The implementation in MatLab™ gives the ability to calculate a complete instrument (50 notes including alternatives) and to analyse its resonances in about one minute (on a MacBook Pro with 3 GHz Intel Core i7), and also gives the facility to introduce different models. For example, it was straightforward to introduce an embouchure equivalent length. The mouthpiece, reed and oral cavity impedances have received much theoretical and some experimental attention since IMPEDPS was written in 1994-6: for example in the second edition of Nederveen [26], Fletcher and Rossing [58] and notably Chaigne and Kegomard [23]. Some improvement could therefore eventually be made in the model by implementation of new results.

As pointed out by many others [77, 26, **Error! Bookmark not defined.**, 25, 42, 32] the knowledge of resonance peaks has utility in instrument design, restoration and modification. The effect of drilling or moving a hole, or of reaming the bore (for example, for removing the tenon compression induced by tenon lapping before cork came into use [59]) can be checked before material is removed. Playing problems with a particular instrument may also be diagnosed. Thus, it is clear that this Heckel instrument would play more in tune with a longer neck, or at a higher orchestra pitch. Examination of the neck does indicate that it might be a later replacement and not an original. The calculated impedances could also indicate how to alter a tone hole to improve the tuning, and what effect this would have on other notes. We believe, therefore, that we have quantitatively validated the computational method of acoustic impedance as a research tool for investigating and restoring both modern and historical bass clarinets and other woodwind instruments. A fuller and more detailed publication of these results is in preparation [60].

Acknowledgements

I thank Robert Cronin and Douglas Keefe for their helpful correspondence on IMPEDPS and the sources used, Colin Lawson, Gabriele Rossi Rognoni and Ingrid Pearson of the RCM for their interest and encouragement, David Sharp, Kurijn Buys and Mathew Dart for experimental cooperation and valuable discussions, the Open University for laboratory facilities and Warwick University for provision of MatLab™ software. I also thank Gary Scavone and Shi Yong for generously providing MatLab™ codes and advice based on their dissertations, and Huw Bowen for recording the audio files and for taking the picture for Figure 1. I am an RCM Doctoral Scholar supported by the Pamela Weston Award, for which I express my appreciation.

7. REFERENCES

[1] Karp, Cary. 'Restoration, Conservation, Repair and Maintenance: Some Considerations on the Care of Musical Instruments.' *Early Music* 79-84 (1979).

[2] *ICOM/CIMCIM: recommendations for regulating the access to musical instruments in public collections: 1985.* http://bit.ly/2ADcQfz

[3] Rice, Albert R. (1992). *The Baroque Clarinet*. Oxford: Clarendon Press.

[4] Rice (2003). *The Clarinet in the Classical Period*. Oxford: Oxford University Press.

[5] Rice (2009). *From the Clarinet D'amour to the Contra Bass: A History of the Large Size Clarinets, 1740-1860*. OUP USA.

[6] Campbell, Murray and Clive Greated, *The musician's guide to acoustics*. New York: Schirmer (1987) p259

[7] Benade, A.H. *Fundamentals of Musical Acoustics. New York*: Dover. Corrected reprint of 1976 edition, New York: Oxford University Press (1990).

[8] Smith, R.A. and Mercer, D.M.A., 'Possible causes of woodwind tone colour'. *Journal of sound and vibration*, 32, 347-358.

[9] Benade, Arthur H. *Horns, strings and harmony*. New York: Anchor (1960).

[10] Helmholtz, Hermann von. *Tonempfindungen.* (1863).Tr. Alexander J. Ellis. *On the sensations of tone*. (1954) New York: Dover reprint.

[11] Bouasse, H. *Instruments à Vent (Vol. I and II)*. Paris: Libraire Delagrove (1929).

[12] Lamb H., *Dynamical theory of sound*. London: Armold (1931).

[13] Olson, H.F. *Acoustical engineering*. Princeton: Van Nostrand (1957).

[14] Caussé, R., Kergomard, J. and Lurton, X. 'Input impedances of brass musical instruments'. J. Acoust. Soc. Am. 75, 241-254 (1984).

[15] Backus, J. *Small vibration theory of the clarinet*. J. Acoustic. Soc. Am. 35 305-313 (1963)

[16] Strutt, J.W. (2nd Baron Rayleigh) (1877). *The theory of sound*. London: Macmillan and Co.

[17] McGinnis, C.S and C. Gallagher. 'The mode of vibration of a clarinet reed'. . *Journal of the Acoustical Society of America*, 12, 529-531 (1941).

[18] Taillard, Pierre-André and Jean Kegomard. 'An Analytical Prediction of the Bifurcation Scheme of a Clarinet-Like Instrument: Effects of Resonator Losses'. *Acta Acustica united with Acustica* 101, 279-291 (2015).

[19] Levine, Harold and Julian Schwinger. 'On the radiation of sound from an unflanged circular pipe'. *Phys. Rev.* 73, 383-406 (1948).

[20] Dalmont, J.-P., C.J. Nederveen and N. Joly. 'Radiation impedance of tubes with different flanges: numerical and experimental investigations'. *Journal of Sound and Vibration* 244(3), 505-534 (2001).

[21] Benade, A. H. and D. J. Gans 'Sound Production in Wind Instruments', *Annals of the New York Academy of Scienc*e 155, 247-263 (1968).

[22] Worman, Walter E. *Self-Sustained Nonlinear Oscillations in Clarinet-Like Systems*. Ph.D. diss., Case Western Reserve University, Cleveland, Ohio (1971).

[23] Chaigne, Antoine and Jean Kergomard. *Acoustics of musical instruments (1st English edition)*. New York: Springer-Verlag (2016).

[24] Plitnick G.R. and W.J. Strong, `Numerical method for calculating input impedances of the oboe'. *Journal of the Acoustical Society of America*, 65, 816-825 (1979).

[25] Keefe, D.H. 'Woodwind air column models'. *Journal of the Acoustical Society of America*, 88, 35-51 (1990).

[26] Nederveen, C.J. *Acoustical aspects of woodwind instruments* (revised edition). Dekalb, IL: Northern University Illinois Press (1998).

[27] Guillemain, P., J. Kergomard, and T. Voinier, "Real-time synthesis of clarinet-like instruments using digital impedance models," *J. Acoust. Soc. Am.*, 118, 483–494, (2005).

[28] Scavone, Gary P. and Smith, Julius O. 'A stable acoustic impedance model of the clarinet using digital waveguides'. *Proc. Of the 9th International Conference on Digital Audio Effects (DAFx-06)*, Montreal, Canada, Sep. 18-20 2006.

[29] Dalmont, J.P., Gazengel, B., Gilbert, J., Kergomard, J.: 'Some aspects of tuning and clean intonation in reed instruments'. *Appl. Acoust.* 46, 19–60 (1995)

[30] Cronin, Robert H. (1996). 'Understanding the operation of auxiliary fingerings on the modern bassoon'. *Journal of the International Double Reed Society*, 24, 13-30

[31] Cronin, Robert and Douglas Keefe (1996). `Understanding the operation of auxiliary fingerings on conical doublereed instruments'. Notes for a talk presented at the 131st meeting of the Acoustical Society of America, Indianapolis, Indiana, 13-17 May 1996.

[32] Schumacher, R.T., '*Ab Initio* calculations of the oscillations of a clarinet', *Acustica* 48 71- 85 (1981).

[33] Backus, J. 'Input impedance for reed woodwind instruments'. *J. Acoust. Soc. Am.*, 56 1266-1279 (1974).

[34] Benade, A.H. and Kouzoupis, S.N. 'The clarinet spectrum: Theory and experiment'. J. Acoust. Soc. Am., 83, 292-304 (1988).

[35] Benade, Arthur H. and Douglas H. Keefe, *The Galpin Society Journal* 49 113-142 (1996)

[36] Benade, A.H. 'Benade's NX clarinet: Its genesis' (written in 1987 and published posthumously), *The Clarinet* , pp. 46-48 Jan/Feb (1994).

[37] Jameson, George. 'Benade's clarinet: Mechanical and other considerations. The Clarinet, pp. 32–33 May/June 1994

[38] Stephen Fox Clarinets. http://www.sfoxclarinets.com/Benade.html. Consulted 8 March 2018.

[39] Mamou-Mani, A. and D.B. Sharp. 'Evaluating the suitability of acoustical measurement techniques and psychophysical testing for studying the consistency of musical wind instrument manufacturing'. *Applied Acoustics*, 71, 668-674 (2010).

[40] Mamou-Mani, A., D.B. Sharp, T. Meurisse and W. Ring. "Investigating the consistency of woodwind instrument manufacturing by comparing five nominally identical oboes." *J. Acoust. Soc. Am. 131(1), Pt 2*, 728-736 (2012).

[41] Kowal, Paulina; Sharp, David and Taherzadeh, Shahram (2013). 'Analysing differences between the in- put impedances of five clarinets of different makes'. In: *Institute of Acoustics Annual Spring Conference 2013: Acoustics 2013*, 13 May 2013, Nottingham, UK.

[42] Dart, Mathew. *The Baroque Bassoon*: *form, construction, acoustics, and playing qualities.* London: London Metropolitan University. PhD thesis (2011).

[43] Hichwa, Bryant and David Rachor. 'In-depth acoustic modeling and temperament studies of 18th and early 19th century baroque bassoons comparing originals and reproductions by maker, time period, and region'. *Société Francaise d'Acoustique. Acoustics 2012*, Apr 2012, Nantes, France. 2012. Online: <hal-00810841>

[44] Jeltsch, Jean, Vincent Gibiat, and L Forest, "Acoustical Study of a Set of Six Key Baumann's Clarinets," in *Proc. International Symposium on Musical Instruments* (Dourdan, 1995), 134–40.

[45] Jean-Xavier Lefevre, *Méthode de Clarinette. Adopté Par Le Conservatoire Pur Servir Àl'étude Dans Cet Établissement* (Paris: Impr. du Conservatoire de Musique, 1802).

[46] Jeltsch, Jean and Nicholas Shackleton, "Caractéristation Acoustique de Trois Clarinettes de Facteurs Lyonnais," in *Colloque Acoustique et Instruments Anciens: Factures, Musiques et Science* (Colloque Acoustique et Instruments Anciens: Factures, Musiques et Science, Paris: Cité de la Musique, 1999), 103–24.

[47] Benade, A.H. and J.S. Murday. 'Measured end corrections for woodwind tone holes". *J. Acoust. Soc. Am.* 41, 1609 (abstract only; 1967).

[48] Thompson, Stephen C. 'The effect of the reed resonance on woodwind tone production', J. Acoust. Soc. Am. 66,1299-1307 (1979).

[49] Reiter, Edith. Wilhelm Heckel, six generations dedicated to music. Wiesbaden: Marixverlag (2014).

[50] Bowen, D. Keith. 'The Rise and Fall of the Bass Clarinet in A', *The Clarinet,* September 2011 44-51 (2011).

[51] Bär, Frank P. (2006). *Verzeichnis der Europäischen Musikinstrumente in Gemanischen Nationalmuseum Nürnberg. Band 6: Liebesklarinetten, Bassetthörner, Bassklarinetten, Metallklarinetten*. Wilmhelmshaven: Florian Noetzel.

[52] Sharp, D. Mamou-Mani, A. and Van Walstijn, M. 'A Single Microphone Capillary-Based System for Measuring the Complex Input Impedance of Musical Wind Instruments', *Acta Acust united Ac*, 97:819-829(5) (2011).

[53] Widholm, G. Pichler, H. and Ossmann, T. 'BIAS: A computer-aided test system for brass wind instruments', Paper No. 2834, Audio Engineering Society (1989).

[54] Widholm, G. Winkler, W. 'Evaluation of musical instrument quality by computer systems. Examples of

realisation', Proceedings of the SMAC93, Royal Swedish Academy of Music, ISBN: 91845289876, 560–565 (1994).

[55] Widholm, G. 'Brass wind instrument quality measured and evaluated by a new computer system', in *Proceedings of the 15th International Congress on Acoustics*, Trondheim, Norway (June 26–30, 1995), Vol. III, pp. 517–520 (1995).

[56] de Cheveigne, Alain and Hideki Kawahara, 'YIN, a fundamental frequency estimator for speech and music', *J. Acoust. Soc. Am.* 111 1917 – 1930 (2002).

[57] Pimentel, Brett, *Fingering Diagram Builder*, https://fingering.bretpimentel.com/#!/clarinet-german/albert/ . Consulted 11 March 2018.

[58] Fletcher, N.H. and Rossing, T.D. *The physics of musical instruments (2nd. Edition).* New York: Springer (2010).

[59] McGee, Terry. *'Effect of thread wrapping on flute tenons'* http://www.mcgee-flutes.com/effects_of_thread_wrapping.htm (2011). Consulted 20/11/2017.

[60] Bowen, D. Keith, David Sharp, Kurijn Buys and Mathew Dart, *Applied Acoustics,* in preparation (2018).

INTERDISCIPLINARY PERSPECTIVES ON PLAYING AN INSTRUMENT: DOES THE SHAPE OF A RECORDER MOUTHPIECE INFLUENCE THE TIMBRE?

Naomi Nordblom
Ruprecht-Karls-Universität Heidelberg
recordermouthpieces@gmail.com

ABSTRACT

This paper looks at the connection between the player and his or her instrument using the example of the recorder. Does the shape of the mouthpiece change the sound of the instrument, and if so, which frequencies are changed? In this survey, three people were recorded playing the same instrument with two different shapes of mouthpieces, the traditional shape and the other with a plastic 'hat' to change the shape of the mouthpiece. (The 'hat' had been developed by the author with the support of Stefan Kopp, professor of ortodontics.) The frequency spectrum of the recorded notes both with and without a hat was then analyzed. The spectra showed some differences. Their frequency area depends on the presence of a 'hat' and on the register of the played note. The results open new perspectives for interdisciplinary research (musicology, medicine and acoustics) to understand better the process of playing an instrument.

1. INTRODUCTION

While the recorder maker is usually responsible for the construction and maintenance of the instrument, the player is in control of its use. But in order to produce a sound, the player needs to connect with the instrument, and this area of connection is the subject of my investigation. To prevent air escaping from the sides of the mouth, it is necessary to seal the junction between the lips and the mouthpiece. This entails that the shape of the lips is determined by the shape of the recorder mouthpiece, and the lips are connected to the rest of the body, the movement of the lips is connected to the movement of other parts of the body as well.[1] Jer-Ming Chen, Dan Laurin, John Smith and Joe Wolfe have shown in their work concerning the vocal tract interactions in recorder performance, that the tensions in the mouth influence the sound of the instrument [1]. Putting all these premises together, it is highly probable that the shape of the mouthpiece has an influence on the sound of the recorder.

The idea of fabricating individualized mouthpieces for the recorder is not completely new. For example, there is the concept of the two opposite breathing types which, according to this theory, requires different shapes of mouthpieces [2]. During my experiences as a teacher, I have observed more than two different breathing and playing types. Every person has different body proportions. This matters especially for the proportions of the face and inside of the mouth – the areas that connect most with the instrument. It is quite normal for brass players, for example, to choose a mouthpiece that fits their physical condition.[2]

Figure 1. Recorder at A=440Hz made by Ralf Netsch after a model by Denner, with 'hat'. Photo by Naomi Nordblom.

With these considerations in mind, I have been developing new shapes of mouthpieces. Beginning in 2011, I have modelled them with dental silicone and have then tested the results in my daily practise routine. Since these mouthpieces sit on the top of the headjoint, I call them 'hats'. To avoid changes in the shape of the soft silicone mouthpieces, the experiments were made with 3D copies

[1] I am grateful to Professor Stefan Kopp, Johann Wolfgang Goethe-Universität Frankfurt am Main, for his advice.

[2] See for example the wide variety of brasswind mouthpieces offered for sale on the website of Bernhard W. Schmidt [3].

of the hats in ABS plastic (see Figure 1). During these experiments, I noticed not only a different feeling in the mouth, but also that some playing properties of the instrument changed as well: the timbre was more brilliant, and the notes responded more quickly. To separate opinions from facts, it is important to examine the changes in the playing properties of the hats with recorders. In this survey, I have recorded the sound of the recorders and their hats to evaluate the results with a frequency analysis and to estimate the influence of the parameter 'shape of the mouthpiece' on the playing properties of the instrument.

2. THE EXPERIMENT

In a soundproof room, the sound recorder (model ZOOM H2n) was placed with enough distance from the playing position, so that the usual playing movements of the player did not influence the recorded sound (distance 9.19ft). The apparatus for the experiment is shown in Figure 2.

Figure 2. Apparatus for the experiment, showing (A) sound recorder, model ZOOM H2n, (B) phone with metronome app bestmetronome and headphones, (C) f-alto recorder at A=415Hz by Doris Kulossa, (D) f-alto recorder at A=440Hz by Ralf Netsch. Photo by Naomi Nordblom.

Before each recording, every player receives three minutes to become accustomed to the instrument with and without a hat. First, a F major triad is played throughout the whole range of the instrument (f^1, a^1, c^2, f^2, a^2, c^3, f^3), then, two notes are chosen from this material for further analysis of the results. This relatively large amount of data helps to avoid conscious or unconscious manipulation of the sounds, as this would be more difficult over a long time span than over a short one. Additionally, the notes chosen for further interpretation are not known to the players. To make further analysis easier, every player wears headphones with a metronome marking of MM=100. Every note is played for a duration of four beats followed by two beats of rest, first with a hat and then without. Then, the recordings are converted with the software 'Audacity' to make them available for further evaluation with 'Sonic Visualiser', a program that shows the spectra of the recordings. For the evaluation of the results, the second note of each recording and pitch was selected during the steady state. The overtones shown in the spectrum and their relative amplitude to each other give some indication of the specific timbre of the played sound [4:31].

Formant	Frequency area (Hz)	Sound quality
u	200-400	large sound
o	400-600	
å	600-800	
a	800-1250	poweful
ö, ü, ä (+o,a,u)	1250-1800	nasal
e (+o)	1800-2600	bright, brilliant
i (+u)	2600-5000	

Table 1. Relationship between frequency and timbre after Meyer.

The timbre of wind instruments is mostly determined by the formants [5:401], which are fixed areas of the spectrum that are responsible for particular properties of the overall sound. These have been shown to correspond to the specific coloring of the vowels (e.g. different pronunciations of 'fork' and 'bone'). The relationship between frequency and timbre can be seen in Table 1 [4:33-34]. With a hat, a recorder sounds more powerful and brilliant. us, increased frequencies between 800Hz and 1250 Hz and above 1800Hz are expected.

This experiment requires several test series. The main test series consists of two notes (f1 and c3) played on the f-alto recorder in A=415Hz made by Doris Kulossa after a model by Jacob Denner. Three different test subjects are being recorded. Apart from the main test series, there are two control test series. The first control test series consists of the same procedure as in the main test series, but with a different instrument, a f-alto recorder in A=440Hz made by Ralf Netsch after a model by Denner. The second control test series uses the same instrument as in the main test series, but the recordings are taken on three consecutive days and always involve the same player. This series is important to examine whether the results are reproducible and to estimate the margin of error of the main test series.

3. RESULTS

The results consist of the spectra of the played notes of which an example is shown in Figure 3. The blue graph shows the spectrogram with a hat and the green graph without one. To compare the results of the recording with a hat to the recording without, both graphs were reproduced overlapping each other. As some parts of the graph below are hidden, every measurement contains two pictures: one picture with the green graph in front (so that the blue overlap shows the extra frequencies with a hat) and one with the blue graph in the foreground, so that the green overlap shows the supplementary frequencies without a hat. Full backup data may be accessed on the author's website:

<https://recordermouthpieces.wordpress.com>

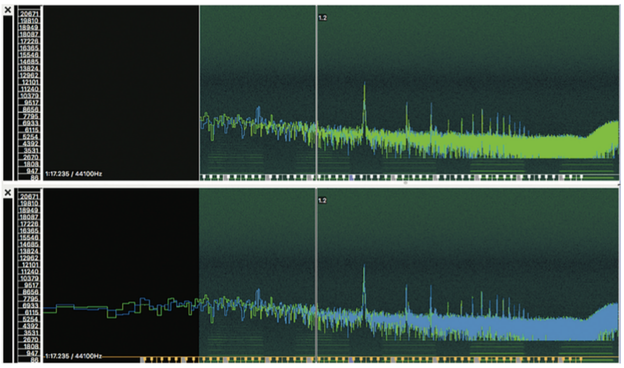


Figure 3. Spectrum of the note f^1 using a recorder at A=415Hz without (upper part) and with (lower part) a 'hat' in front.

Concerning the main test series, all three results for f^1 show a significant increase in the amplitude of the overtones between about 3000Hz and 5300Hz with a hat (see for example the upper part of Figure 3) which makes a more brilliant sound. Two of the three test subjects showed a slightly louder fundamental with the 'hat' (see Figure 3). For the note c^3, all three players had a louder fundamental with the 'hat'. Two of the three measurements with the 'hat' show amplified overtones between 4500Hz and 12700Hz, making the sound more brilliant. Two of the three players also show a louder first overtone without the 'hat'; its frequency is about 2000Hz.

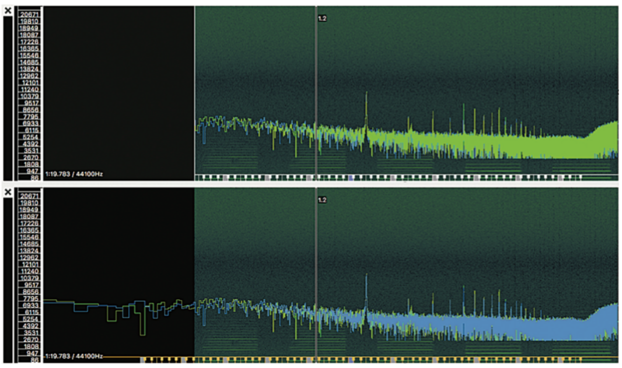


Figure 4. Spectrum of the note f^1 using a recorder at A=440Hz without (upper part) and with (lower part) a 'hat' in front.

The first control test series was carried out with a different instrument. As the hats are handmade, the shapes of the mouthpieces for the two alto recorders are not identical and neither are the two recorders. Thus, the results of the two 'hats' should not be compared to each other, but rather to the corresponding measurement without a 'hat'. All three measurements of f1 show amplified overtones between 3760Hz and 4140Hz with a 'hat' (see Figure 4 in the color section, upper part), making the sound more brilliant. The highest overtone with a 'hat' is a little bit louder than the measurement without a 'hat' (Figure 4, upper part). The place of this highest overtone depends on each individual person and lies within the range 5150–6600Hz. All three recordings of the note c^2 with 'hat' show bigger amplitudes from the second overtone upwards. is corresponds to the frequencies above about 3150Hz, and makes the sound brighter and more brilliant.

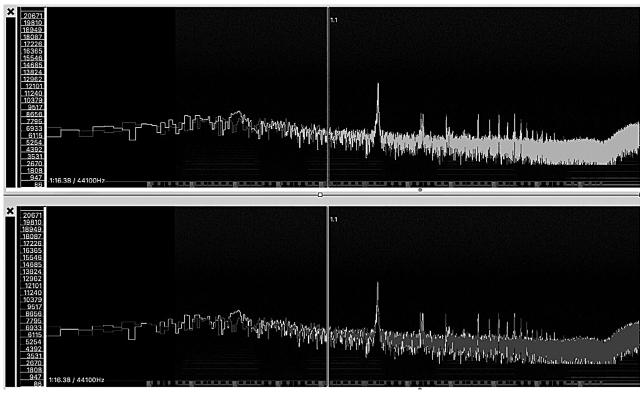


Figure 5. Spectrum of the note f^1 using a recorder at A=415Hz without (upper part) and with (lower part) a 'hat' in front, second control test series.

The second control test series was made with a 415Hz recorder on three consecutive days. This test series is particularly important to show if the results are reproducible, and to estimate the margin of error. The frequency analysis of f1 shows also amplified overtones between 3000Hz and 5300Hz with a 'hat' as in the main test series (see Figure 5, upper part). The slightly louder fundamental with a 'hat' is not present here. For c3, the louder fundamental and the weaker first fundamental with 'hat' are present on all three days.

To reduce noise from the outside, all the recordings were made in a soundproof room, with the result that some minor differences might be muffled by the walls. All the spectra show some background noise with frequencies above 13000Hz, a range that can hardly be heard and does not lie in the range that is interesting for our measurements, and so has no influence on the results of the experiment. Moreover, the number of three participants is rather small; under the available circumstances it was not possible to interpret larger amounts of data. As the control test series gives reasonable and reproducible results, the possibility for errors seems to be quite small.

4. CONCLUSIONS AND OUTLOOK

The proportions of the amplitude of the overtones determine most of the timbre. The presence of a 'hat' and the changed shape of the mouthpiece of the recorder changes the proportions between the amplitudes of the overtones. It is arguable, therefore, that the shape of the mouthpiece influences the sound of the recorder. To estimate the exact amount of this parameter and to examine its effects on other playing properties (such as articulation and the register changes) further studies with recording studio conditions and with more test subjects are needed. Instrument makers could incorporate the parameter into the design of the instrument. The results are important for music teachers: a 'bad' body position of a pupil might not be caused by a 'lack of talent', but perhaps by an instrument that does not fit his or her physical characteristics. Moreover, performers can explore new sound possibilities on their instrument. Indeed, this issue might also occur with all the other wind instruments that have a close connection between the mouth and the mouthpiece. The results suggest that there is a causal relationship between the shape of the mouthpiece and the resulting sound on the recorder. I hope to have taken a first step into examining the possibilities that the shape of the mouthpiece provides for instrument makers, performers, and teachers.

Acknowledgements

I would like to thank professor Stefan Kopp for his medical advice and Konstantinos Arapoglou, Charles Banfield and professor Iain Fenlon for their support. This article has originally appeared in the Galpin Society Journal 71 (2018) under the title „The Shape of the Mouthpiece - a New Parameter in Recorder Making?" and I am grateful for their permission to publish it here. Moreover, I would like to thank SEMPRE and the Musikwissenschaftliches Seminar der Ruprecht-Karls-Universität Heidelberg for their financial support.

5. REFERENCES

[1] D. Laurin, J.-M. Chen, J. Smith, J. Wolfe, "Vocal Tract Interactions in Recorder Performance," presented at the 19th International Conference on Acoustics, Madrid, 2-7 September 2007, Unpublished conference paper.

[2] A. Breukink. (6 April). *Mouth Pieces* [sic]. Available: http://www.adrianabreukink.com/mouthpieces.php.

[3] B. W. Schmidt. (2017). Available: http://www.schmidt-brass.de/wissenswertes-mundstuecke.htm.

[4] J. Meyer, *Akustik und musikalische Aufführungspraxis.* 4. ed., Frankfurt a.M.: Bochinsky, 1999.

[5] C. Reuter, "Zusammenklang von Musikinstrumenten," in *Akustische Grundlagen der Musik*, S. Weinzierl, Ed. (Handbuch der Systematischen Musikwissenschaft vol. 5), Regensburg: Laaber-Verl., 2014, pp. 391-408.

A Sonification Approach to Music Visualization

Roberto Piassi Passos Bodo
University of São Paulo
rppbodo@ime.usp.br

Flávio Luiz Schiavoni
Federal University of São João del-Rei
fls@ufsj.edu.br

ABSTRACT

Music visualization always helped musicologists to analyze musical pieces. Traditionally, there are a few music visual formats that are standards and broadly used. Since computers started helping music analysis, several formats arose to represent music in a digital format. In this paper we propose three forms of music representation that can create visual feedback that is different from common music visualization. Our approach can bring some discussion about how different visual feedback can help musicians to understand a musical piece. This music representation is not concerned to be a better format but it is focused in aesthetics results that can provide alternative visualizations to musicians.

1. INTRODUCTION

Musicology uses the ears and the eyes to analyze music for centuries, and since the 1960's it is possible to use computers to aid this task. The first experience to automatically extract data from music, to analyze it with a computer, started with the researcher creating a music representation to use it as computer input, and using the computational power to compare data to analyze music [1]. In this work, it is presented a brief discussion about musical data representation with punch cards and the concern about how it could be easy to share musical data between research centers using a common music notation format. Since there were no music file formats in that age, they created a symbolic music notation to define music instructions that was called MIR (Music Information Retrieval) format [1]. Since then, it started a field to extract musical information, like style analysis, aided by computers [2].

Time passed and nowadays we have several different music file formats. From symbolic music (that stores events and notations) to audio files, passing by visual scores stored as images, and also text meta-data about music stored in catalogs. All these informations and file formats can be used to represent music in the computer and consequently can be used to extract informations about music [3]. Based on these formats, different tools to musicology arose using music representation and helping musicians on musical analysis, for instance JRing [4], Open Music [5], and Music 21 [6]. Also, there are huge databases available online with music files in different formats that can be used to extract information about music using statistics and machine learning techniques.

It is important to notice that different tools for music analysis can accept different file formats. The necessity of having different music representation happens because the kind of information that can be extracted from one format is different from the information extract from other format, and they can be complementary to help understand a composition or a musical piece. Some of them uses audio analysis while other uses symbolic music, for instance.

There are also the possibility to convert one music representation format into other format changing the music information point of view. Depending on how it is stored, it is feasible to edit, transpose, move, copy, and compare musical data. Also, it can be used to perform statistical analysis and plot graphics, and so on [6].

2. DATA SONIFICATION

A field that is not directly related to musicology but is also focused in transforming virtual musical content from one format to other is the so-called sonification. Classically, a goal of sonification is to transform complex multidimensional data into intuitive audio [7] as in text-to-speech or accessibility tools. Alternatively to this classic goal, it is possible to use sonification to create purely aesthetic sounds that can be used to inspire compositional process or just to be enjoyed by users [8].

Using this alternative approach of sonification, the web became a huge repository of data like text, images and web pages to be sonified and tried out as sonic elements (for instance, to compositional purposes or to performances). The process of sonifying data (such as HTML, image, and text) has some similarities with the process of synthesize symbolic music or even play audio files. Nonetheless, while traditional music notation formats, like MIDI, ABC or MusicXML has canonical sound mappings, other data types that does not have musical semantics need some aesthetic mappings to be played as music. Thus, using appropriated mappings it is possible to convert any data file to music information and listen to it.

Since the universe of mapping can be infinite and we cannot try every possibility to map web data information into music, some rational decisions can be taken to help this mapping process. A text can be sonified like ABC files or Lylipond format, where a character (or a group of) is mapped to a set of musical commands. The HTML file format is similar to MusicXML since they are both hierarchical formats, and their structure can be mapped to music structure. Lastly, BMP images can be converted to WAVE files where the bytes of the audio samples are the bytes of the image pixels. A work presenting this sonification can be found in [9], including some sonification tools that create sounds from web pages, image files, and snippets of text, data that are not musical but that have similar structures with some digital musical formats.

3. FROM DATA SONIFICATION TO MUSIC VISUALIZATION

The set of mappings used in the sonification process has an interesting feature: they have well-defined inverses. These inverse mappings can be used to generate data files from music information as alternative representations for songs, such as a piano roll representing the list of MIDI events. One can convert two songs to text and look for the longest common subsequence between them to detect some melody similarity, or can convert one song to an image and detect repeated visual patterns in order to segment such music, and so on. Thus, these visual representations can be useful for music visualization, creating new possibilities and different approaches to music analysis.

What kind of web pages can we achieve using a MIDI score as HTML page structure? What kind of visual feedback can be reached using an audio file plotted as a bidimensional image? What kind of text we can get generating from a MusicXML score? These questions leaded the results of this research.

3.1 From score to HTML

When working with HTML sonification we did the mapping of seven CSS properties (width, height, top, left, padding, margin, border-width) to four synthesizer parameters (pitch, duration, dynamics, onset). The mapping functions were all linear: $synthProperty = a * cssProperty + b$. So, they have inverses in the following format: $cssProperty = c * synthProperty + d$, where $c = \frac{1}{a}, d = -\frac{b}{a}, a \neq 0$. Thus, also using linear equations we can convert music information to CSS properties to see different visual renditions of the same musical piece. Not only by varying the coefficients, but also by varying the mapped sound properties.

To this task we choose the MIDI file format as the input of this inverse mapping with the notes presented in Figure 1. MIDI protocol has the NOTE_ON event that contains the MIDI note number (pitch), the velocity (dynamics), and its own onset time. Lastly, with the respective NOTE_OFF event we can compute the duration of the note. With these data in hands we can convert the notes from a melodic sequence to elements in a HTML page. For instance, a simple conversion with duration to width, velocity to height, pitch to top, and onset to left, can generate a HTML page that is similar to a piano roll. In Figure 2 we have the x axis representing time and the y axis representing pitch (just like the piano roll), but we also have the height of the boxes showing different dynamics (the taller ones came from notes with higher velocities).

```
0, 0, Header, 1, 1, 120
1, 0, Start_track
1, 0, Control_c, 0, 7, 127
1, 0, Tempo, 600000
1, 0, Program_c, 0, 0
1, 0, Note_on_c, 0, 60, 120
1, 200, Note_off_c, 0, 60, 0
1, 200, Note_on_c, 0, 62, 110
1, 450, Note_off_c, 0, 62, 0
1, 450, Note_on_c, 0, 64, 100
1, 600, Note_off_c, 0, 64, 0
1, 600, Note_on_c, 0, 65, 90
1, 850, Note_off_c, 0, 65, 0
1, 850, Note_on_c, 0, 67, 80
1, 1000, Note_off_c, 0, 67, 0
1, 1000, Note_on_c, 0, 69, 70
1, 1250, Note_off_c, 0, 69, 0
1, 1250, Note_on_c, 0, 71, 60
1, 1400, Note_off_c, 0, 71, 0
1, 1400, Note_on_c, 0, 72, 50
1, 1650, Note_off_c, 0, 72, 0
1, 1650, End_track
0, 0, End_of_file
```

Figure 1. A C major scale presented as a CSV file.

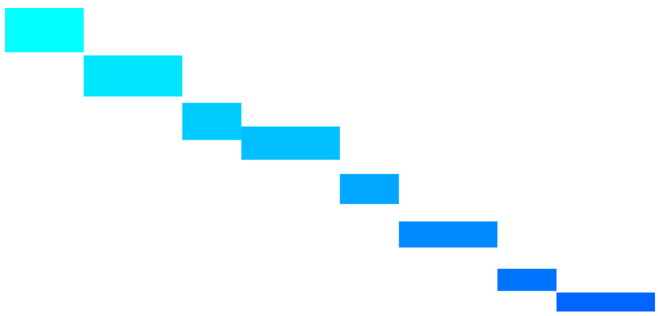


Figure 2. The visualization of a C major scale using the mapping: duration to width, velocity to height, pitch to top, and onset to left.

Other mappings can be explored, creating visualizations that are not so familiar (or traditional) as the piano roll, but that contain all the original information from the music piece, which is still easily recoverable. Figures 3, 4 and 5 show three alternative renderings of the same C major scale displayed at Figure 2 (the mapped properties are indicated in the respective captions).

The renderings displayed at the first three figures, we used width, height, top and left properties. Using only these four CSS rules, the musical attributes of the original song established the dimensions and positions of the HTML elements in a direct and explicit way. Observe that in Figure 3 higher pitches produced taller rectangles, longer durations made wider rectangles, bigger velocities placed the elements further down, and later onset times put

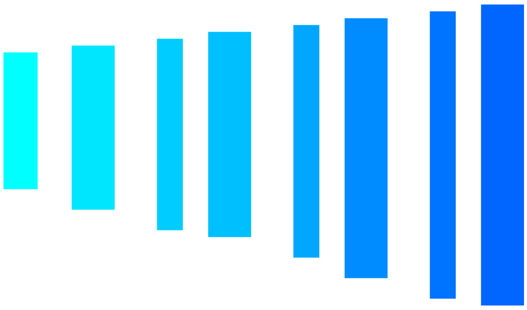


Figure 3. The visualization of a C major scale using the mapping: pitch to height, duration to width, velocity to top, and onset to left.

Figure 4. The visualization of a C major scale using the mapping: pitch to height, duration to top, velocity to left, and onset to width.

the elements farther to the right. In Figure 4 higher pitches produced taller rectangles also, longer durations put the elements further down, bigger velocities placed the elements farther to the right, and later onset times made wider rectangles.

One interesting feature of the chosen CSS properties is that we can use padding (internal gap) to control both dimensions of the HTML elements at the same time. In Figure 5 we can see that all displayed elements are squared, because we mapped duration to padding (observe that with longer durations, the squares became larger). Another feature is that we can use margin (external gap) to control together the vertical and horizontal displacement of the elements. Notice that in Figure 5 all HTML nodes are placed in a diagonal, because we mapped onset time to margin (then, with later onset times, the elements are further shifted to the lower right corner). Lastly, in Figure 5 once more, we mapped velocity to border-width, so we have thicker borders when the notes have faster attacks.

3.2 From audio to Image

When working with image sonification we did a data abstraction in which we used the image's pixels as the audio's samples. This was done using a flattened version of the image (which is normally represented by a matrix) to obtain an one-dimensional array of bytes for the audio. We only had to make some choices regarding if the image had 3 bytes per pixel (RGB images) or 1 byte per pixel (gray-scale images).

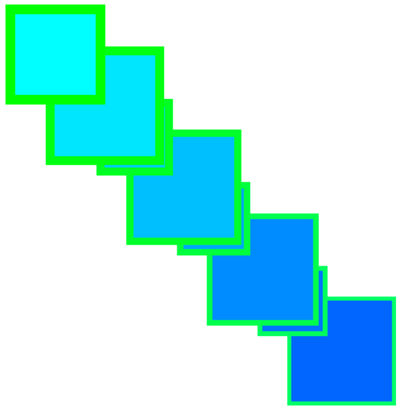


Figure 5. The visualization of a C major scale using the mapping: pitch to top, duration to padding, velocity to border-width, and onset to margin.

This abstraction works in the reverse order: we can use the audio's samples as the image's pixels. To use the bytes of the audio as the bytes of the image we need first to choose a value for image's width to be able to perform a reshape. Thus, we can convert a one-dimensional array to a 2D data structure. Again we have to be extra careful with the two basic types of images: RGB and gray-scale. Moreover, usually we have audio files with 16 bits per sample (Compact Disc standard), and this value does not match directly with 24-bit pixels, nor with 8-bit pixels.

For simplicity, we can consider an audio with 8 bits being transformed to a gray-scale image. The Figure **??** shows an image generated from an audio file containing a 210Hz sinusoid lasting 1 second (adding up 44,100 samples). The width of the image - 210 pixels - was chosen to put one complete period of the sinusoid per row of the image matrix. This choice aligned the pixels in a way that make the periodic pattern visually quite prominent.

Figure 6. A 8-bit wave file of a 210Hz sinusoid lasting 1 second in 44100Hz mapped to a image with 210 pixels of width.

A similar conversion has been made with an audio file containing a 441Hz sinusoid lasting 10 seconds (adding up 441,000 samples). The chosen width for the image was 1000 pixels (then the height is equal to 441 pixels). Thus, we get ten complete periods of the sinusoid per row. Again, there is an alignment of the pixels making the periodic pattern easily seen (as we can verify in the Figure **??**).

Figure 7. A 8-bit wave file of a 441Hz sinusoid lasting 10 seconds in 44100Hz mapped to a image with 1000 pixels of width.

Next, we tried musical pieces with richer spectrum. For instance, we concatenated two audios (a bass line with 7.5 seconds, and a guitar chord progression with 15 seconds) following the AABAAB pattern, and resulting in a 60 seconds audio (2,646,000 samples). Then, we convert it to a squared image ($width = height = 1627 \approx \sqrt{2,646,000}$) filling in the matrix with zeros when the bytes from the audio ended. In Figure 8 we can clearly see some visual patterns, and we can even recognize that they are repeated (we placed some white lines marking the start and the end of each segment).

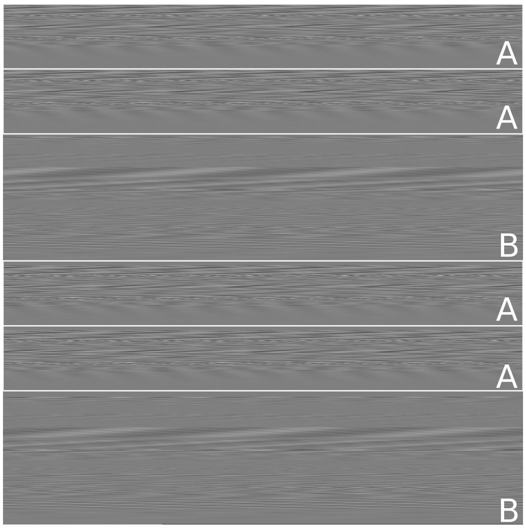


Figure 8. A 60 second audio file with subdivisions following the AABAAB pattern converted to a squared image.

In order to generate colored images, we did an alternative code that receives three 8-bit audios as input. It converts each audio file to a gray-scale image, and afterwards each image is assigned to a different color channel: the first image to Red channel, the second one to Green, and the third to Blue. Figure 9 shows four images: the first three ones are the gray-scale images generated from 882Hz, 441Hz, 220.5Hz sinusoids, and the last one is the RGB image generated from the three images above. Observe that the lighter colors in the RGB image are placed where the three gray-scale images are white, and the darker colors are placed where the three images above are black. We are also able to identify where each one of them is active alone (columns where the RGB image have red, green or blue colors), and even where they are active in pairs (columns where the RGB image have magenta, cyan or yellow colors).

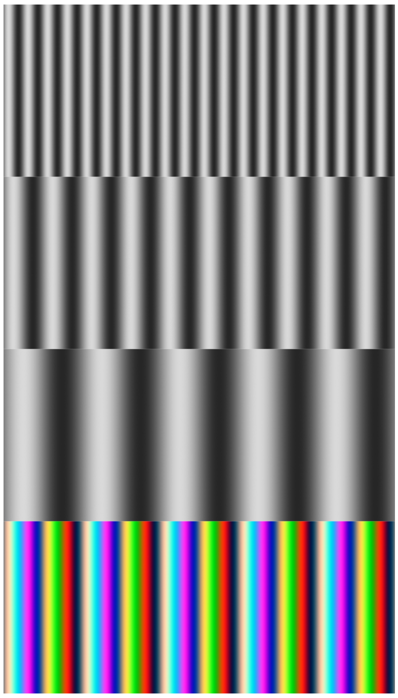


Figure 9. Three audio files being converted to a RGB image, each one being mapped to a different color channel.

These color combinations reflect the energy distributions in the original audio files. For instance, the points where the RGB image is white are the ones where the three audio files have energy peaks at the same time. The points where the pixels are black as the ones where the audios have energy valleys. Each other possible color can tell us which image was lighter at that point, thus it tells which audio had higher energy at the relative time.

3.3 From score to Text

When working with text sonification we implemented an algorithm that produces a melodic sequence mapping letters to notes. Because we were working with texts from web sites, we considered the most frequent letters in English to create the mappings. In this text to melody algorithm, we used the first twelve most frequency letters to determine the notes' pitches, the next eight ones to determine notes' durations (from 64th to double whole), one letter to produce rests and other letter to add a dot to a note [9].

In order to implement the inverse mappings for this sonifications process, we considered the MusicXML format, mainly because it has notes' durations by value. If we used MIDI, as in the audio to HTML mapping, we need to perform a quantization of the notes' durations in milliseconds to figure out their values, which is a delicate process prone to errors.

In Figure 10, we can see a typical note in MusicXML syntax, and three arrows pointing out the mappings: the

```
<note>
<pitch>
<step>C</step>⟶ E
<octave>4</octave>⟶,
</pitch>
<type>eighth</type>⟶ m
...
</note>
```

Figure 10. Example of a single note from a MusicXML score.

```
p E , R , 0 , T , S , C , D , E -
p E - R - 0 - T - S - C - D - E .
```

Figure 11. A pitch shifting performed in a C major scale from the fourth octave (represented by the char ',') to the fifth octave (represented by the char '-').

C pitch is mapped to the 'E' character; the 4th octave is mapped to ',' (whose ASCII value modulus 10 is 4); and the eighth-note is mapped to 'm'. A basic C major scale in the fourth octave - C4, D4, E4, F4, G4, A4, B4, C5 - with all durations as quarter-notes is mapped to the string: "pE,R,0,T,S,C,D,E-" (the "p" char appears once, because all notes have the same duration; if the last four notes were 8th-note the score would be "pE,R,0,T,mS,C,D,E-").

This conversion from MusicXML to text has the potential to generate a symbolic sequence that can be easily compared. Computing the longest common subsequence (LCS), for instance, we can check the differences between two melodic lines. The LCS algorithm can show us variations such as additions, deletions and substitutions of characters from one sequence to the other. In the context of the text sequences that we generate from MusicXML files, this editions can mean a change in pitch, but also a change in the octave, or duration, or rest, etc.

In Figure 11 we have an example of a pitch shifting being performed in a C major scale from C4 to C5, changing it to a C major scale from C5 to C6. In Figure 12 we have an example of a time stretching being performed in a C major scale, changing all quarter-notes to eighth-notes. And in the Figure 13, we have samples for the three types of editions that we mentioned before: the deletion of a D4 note, the addition of a C4 note, and the substitution of a B4 note by a F4 note.

```
p E , R , 0 , T , S , C , D , E -
m E , R , 0 , T , S , C , D , E -
```

Figure 12. A time stretching performed in a C major scale from quarter-notes (represented by the char 'p') to eighth-notes (represented by the char 'm').

```
p E , R , 0 , T     , S , C , D , E -
p E ,     0 , T E , , S , C , T , E -
```

Figure 13. The deletion of a D4 note (represented by "R,"); the addition of a C4 note (represented by "E,"); the substitution of a B4 note (represented by "D,") by a F4 note (represented by "T,").

4. DISCUSSION

The visualization of a score is a common way to extract data from music. Normally, visual music representation uses a common approach: x axis representing time and y axis representing pitch. This common approach is present in traditional scores, piano rolls and also in graphs plotted using symbolic music.

Our representations are probably out of this common approach and for this reason it can be weird, interesting, or useful, just because it explores different relations, representations, and mappings from the conventional approaches to display musical data.

We think these visualizations as part of a creative process, a combinational creativity [10] based on structural mapping from one parameter in sound to other parameters in visual content. Certainly, this is not a definitive form, a one-size-fits-all solution, and there are too much more to be experimented. The only certain we have is that our approach is a little bit out of formal methods to think visual music and we think it as an aesthetic form to visualize music.

The methods presented here are probably more suitable to bring some intuition about the music visually represented, or to open a discussion about other music formats, other ways to represent music and which music attribute is really intuitively represented in a graph. Furthermore, it can blur the frontier between web content and music content exposing common attributes from both sides.

If the images presented here can be useful to help someone to understand and analyze music, one can advocate that it can be possible to do the same work just listening to the music. In fact, it is, specially if you have good trained ears to to it. But if we intend to have a pervasive form to store the analysis, to teach it, to write about it, and to share it, it can be useful to think about music writing and visualization. Contemporary composers are always looking for other forms to write music than the traditional scores. Maybe we can inspire them to find new representations.

5. CONCLUSION

In 1857, the Frenchman Édouard-Léon Scott de Martinville invented the Phonautograph, the first device invented to record music. This equipment had an interesting feature: it records music but can not play it. The result of a record, called phonautograms, could be used only to analyze music in a visual form.

With the advent of the computer, it was possible to create software to music analysis that are faster and more precise than that performed by human beings. Altogether, computer analysis do not replace human sensibility to understand and interpret music.

Nowadays, we have some different formats to represent and visualize music and every each format can present a kind of information that can be complementary to other representation and it is interesting to use more than one format to reach a better comprehension of a music piece.

Thinking about web content and starting from web sonification, we created a few mapping algorithms to music visualization based on common structured formats.

We considered all of the renderings very interesting and expressive. And they are just a glimpse of what the conversions, like MIDI to HTML code, are capable of doing. In this particular case, the actual coefficients of the linear mappings are chosen by the user, so this tool have an enormous potential that even the developers have not yet explored in full.

The three conversion tools that are presented in this paper are hosted at the following links:

- `https://github.com/rppbodo/midi2html`
- `https://github.com/rppbodo/musicxml2text`
- `https://github.com/rppbodo/audio2image`

6. ACKNOWLEDGMENT

We would like to thank the Computer Music Research Group at the University of São Paulo and the ALICE (Art Lab in Interfaces, Computers, and Everything else) from the Federal University of São João del-Rei. The author Roberto Piassi Passos Bodo would like to thank the support by SEMPRE: Society for Education, Music and Psychology Research.

7. REFERENCES

[1] M. Kassler, "Toward musical information retrieval," *Perspectives of New Music*, pp. 59–67, 1966.

[2] A. Mendel, "Some preliminary attempts at computer-assisted style analysis in music," *Computers and the Humanities*, vol. 4, no. 1, pp. 41–52, 1969.

[3] J. Futrelle and J. S. Downie, "Interdisciplinary communities and research issues in music information retrieval." in *ISMIR*, vol. 2, 2002, pp. 215–221.

[4] A. Kornstädt, "The jring system for computer-assisted musicological analysis." in *ISMIR*, Bloomington, IN, USA, 2001, pp. 93–98.

[5] J. Bresson, C. Agon, and G. Assayag, "Openmusic: Visual programming environment for music composition, analysis and research," in *Proceedings of the 19th ACM International Conference on Multimedia*, ser. MM '11. New York, NY, USA: ACM, 2011, pp. 743–746. [Online]. Available: http://doi.acm.org/10.1145/2072298.2072434

[6] M. S. Cuthbert and C. Ariza, "music21: A toolkit for computer-aided musicology and symbolic music data," pp. 637–642, 2010.

[7] T. Hermann, A. Hunt, and J. G. Neuhoff, *The sonification handbook*. Logos Verlag Berlin, 2011.

[8] O. Ben-Tal and J. Berger, "Creative aspects of sonification," *Leonardo*, vol. 37, no. 3, pp. 229–233, 2004.

[9] R. P. P. Bodo and F. L. Schiavoni, "Synesthesia add-on: a tool for html sonification," in *Proceedings of the 16th Brazilian Symposium on Computer Music*, São Paulo - SP - Brazil, 2017, pp. 75–80.

[10] M. A. Boden, "Creativity and artificial intelligence," *Artificial Intelligence*, vol. 103, no. 1-2, pp. 347–356, 1998.

Musical App in Hypersensitivity to Sounds and Neurodevelopmental Disorders: Applicable Strategies

Débora Line
UFMG
deboralg@ufmg.br

Cybelle Loureiro
UFMG
cybelle@musica.ufmg.br

Raquel Prates
UFMG
rprates@dcc.ufmg.br

ABSTRACT

Sêntimus is a musical app dedicated to children diagnosed with Hypersensitivity to Sounds associated with Neurodevelopmental Disorders. The research is part of Master's Degree project in progress, conducted at the Department of Music in partnership with the Department of Computer Science at the Federal University of Minas Gerais, Brazil. The objective of this app is to stimulate expressive behaviors of pleasure or dislike, characteristic of Hypersensitivity to Sounds, through game play and music. We have proposes investigated how treatment principles and symptoms for Hypersensitivity to Sounds, Neurodevelopmental Disorders can be taken into account in the design of a game aimed at children with mild or moderate levels of both diseases. The results obtained were 12 Principles and Strategies such as “Exposure Gradually”, “Attention” and “Social Interaction” and 7 Recommendations by researchers serious game.

1. INTRODUCTION

People with Hypersensitivity to Sounds complain of excessive irritation related to some sounds and noise levels considered common [1]. People with Neurodevelopmental Disorders usually present deficits that produce impairments related to personal, social, academic, or occupational functioning [2]. These symptoms hamper the daily lives of this population. Music and games can help the treatment of these pathologies. The treatment of Hypersensitivity to Sounds is based on exposing patients to sounds, which can cause discomfort to them. In this context, games can motivate patients to be exposed to its sounds and focus, rather than trying to escape from them.

This paper is part of a Master's Degree in progress at the Department of Music in partnership with Department of Computer Science at the Federal University of Minas Gerais, Belo Horizonte, Brazil. The research project "Musical App in Hypersensitivity to Sounds and Neurodevelopmental Disorders" is aimed at children with a mild or moderate level of both disorders. The goal of the research is to stimulate modifications in expressive behaviors of pleasure or dislike, associated to Hypersensitivity to Sounds, through music and gameplay. To do so, we propose an app which is a musical game for children diagnosed with Hypersensitivity to Sounds and Neurodevelopmental Disorders, based on the treatment proposed by Jüris [3] and symptoms of the Neurodevelopmental Disorders described in the Manual DSM-5 [2]. The name of the game app is Sêntimus, which came from a connection of two Latin words: Sentimentum and Musica.

The purpose of this paper is to understand the specific needs of children diagnosed with Hypersensitivity to Sounds associated with Neurodevelopmental Disorders, and how to take them into account in designing a digital game for them. To do so, we have reviewed the literature to understand symptoms and treatment associated with these conditions, as well as the literature describing musical games and software aimed at patients. Based on these findings we propose a set of strategies of how to take into account these symptoms and treatment principles in the design of a digital game. We then describe how the strategies are used in the development of Sêntimus.

Our work brings relevant contributions to the research on the use of technology to support Hypersensitivity to Sounds associated with Neurodevelopmental Disorders. It proposes we have investigated how treatment principles and symptoms for Hypersensitivity to Sounds, Neurodevelopmental Disorders can be taken into account in the design of a game aimed at children with mild or moderate levels of both diseases.

This paper is organized as follows: Section 2 describes the Theoretical Foundation of the related topics; Section 3 describes the Methodology; Section 4 describes the Results; Section 5 describes the Applied Strategies in the Game; and Section 6 describes the Conclusion.

2. THEORETICAL FOUNDATION

This research links four areas in theoretical foundation: Hypersensitivity to Sounds, Neurodevelopmental Disorders, Music and Digital Games. In this section we present the relevant topics about each one of them, necessary to the development of the game.

2.1 Hypersensitivity to Sounds

Defined as unusual intolerance of ordinary environmental sounds, Hypersensitivity to Sounds is an alteration of the sound processing in Central Nervous System. The cochlea and hearing aid are often perfectly normal although patients complain of excessive irritation related to some sounds and noise levels considered common. In the literature we can find the following terms: *Hearing Hypersensitivity, Audiosensitivity, Phonophobia, Noise Sensitivity, Hyperacusis Dolorosa, and Reduced Tolerance to Noise*, being used in the same context [4]. In this paper, we used the term *Hypersensitivity to Sounds*, which is a disorder in the sensation of everyday sound intensities [1, 3, 5, 6]. Sounds intensities that are not perceived by most people can cause great discomfort to *hyperacusis* (people with Hypersensitivity to Sounds). Also, violent noises, such as car horns or fireworks can cause disorientation on individuals with hyperacusis, reflecting in atypical reactions. A typical feature, of the condition, is *fear of sounds* based on their meaning or association [1, 7]. Hazell presented one almost universal true example of the sound created by scratching chalk on a slate. Argues that our perception of loudness is not dictated simply by the strength or intensity of the sound arriving at the ear, but also by the association that it has some threatening qualities. *"Will the sound damage the ears?"* or *"Will it disturb sleep?"* [1, p. 3]. This example can trivialize Hypersensitivity to Sounds and anyone could think: I have hyperacusis. However, we must remember that very often the over-sensitivity for sounds begins by an irrational fear which, nevertheless, becomes an absolute belief and it also involves pain. Hazel reported cases of patients who strongly believe that their lives are ruined by environmental noise from nearby factories or generators. Patients often also complain about low-frequency sounds transmitted through the ground (considered silent because other people may be unable to hear). According to Jüris in [3] the fear/avoidance beliefs were rare, but when they are presented, nevertheless, they elevate the risk of future pain feeling episodes.

Prevalence studies of hyperacusis are not trivial, since they have a subjective nature, Jüris [3, p. 15] mentions one of the very few existing. A study in the Swedish population [8] collected data using two surveys: (1) a postal version of a random sample, which involved 589 participants; (2) an internet version, in which 595 participants' self-reported hyperacusis. The results show that, excluding participants who related hearing impairment, the point prevalence of hyperacusis rate of 7.7% (39 participants) and 5.9% (28 participants) in postal and internet, respectively. The present data suggests that hyperacusis is a common problem.

In terms of diagnostic Jüris [3, p. 13] shows that there is no consensus-based standardized procedure for hyperacusis diagnosis due to the following factors: (1) there are different types of sensitivity to sounds such as: *phonophobia, misophonia, recruitment of loudness, decreased sound tolerance* and *hypersensitivity to sounds*; (2) the concept of these terms is inaccurate, thus it is complicated to distinguish between different types of sensitivity to sounds; (3) there are still diverse terms used in audiological literature making diagnostic even more complicated, as not all people who experience unfavorable effects of sound exposure really suffer from hyperacusis; (4) sound sensitivity is diagnosed in several professions, audiology, neurology and psychiatry; (5) common procedure normally is to take the medical history, followed by audiology testing (in which pure tone audiometry is performed to detect hearing deficits); and physical and laboratory examination; (6) the main source of information diagnostic evaluation is the self-report and Questionnaire.

According to Katzenell and Segal [4], studies in etiology, in clinical guidelines, and review literature describe four main groups of causes of hyperacusis: (1) in *clinical conditions involving the peripheral auditory system*, we highlight: Bell's Palsy, Stapedectomy, Ramsay Hunt Syndrome, Recruitment and Noise-induced Hearing Loss; (2) in *clinical conditions involving the nervous system*, we highlight: Minor Head Injury, Headache, Depression, Learning Disabilities and Stuttering, Williams's Syndrome, Spinal Problems, Tinnitus Hyperacusis; (3) in *hormonal and infectious diseases*, we highlight: Lyme Disease and Addison's Disease; and (4) *unknown causes*. The authors conclude that pathogenesis of hyperacusis is not clear and it probably involves a Central Nervous System. They consider that validation of the effectiveness of the treatment is still necessary.

The symptoms of Hypersensitivity to Sounds can be variable and it can go beyond fear or uncomfortableness. Previous studies [1, 3, 8, 9] indicate that Hypersensitivity to Sounds when evident can show intermittent symptoms such as *anxiety*, *migraine*, *depression* and *severe crisis* due to sensitivity to sounds. The patients often feel *isolated* and *discouraged* considering that they cannot imagine living due to their frequent sensitivity condition.

For the treatment, the authors Hazell and Jüris [1, 3] indicate that the best solution is based on desensitization through *counseling* or *retraining*. Counseling must be undertaken by trained psychological professionals since it *"is understandably difficult to accept that sound which can be uncomfortable or even painful to the hearing, can be quite harmless to the ear"* [1, p. 4]. The retraining is designed to take away the need to plugs (e.g. earplugs) or actions which can protect the ear (e.g. avoidance to sound). The treatment is a process of *sound exposure* and can take quite a long time. Sound exposure is a procedure in which ordinary environmental sound *volume is gradually* increased. Jüris states that the aim of psychological interventions is not to "cure" the patients, but rather to reduce their distress and increase their quality of life [3, p. 21].

2.2 Neurodevelopmental Disorders

Gomes [5] states that previous research has documented that sensory-perceptual abnormalities, especially when concerning the sound, is associated to *Autism*. Hypersensitivity to Sound and Neurodevelopmental Disorders can be associated with clinical conditions involving the Nervous System [4]. Hughes [10] argues that autists often demonstrate behavior such as *holding their hands over their ears to protect or hand flapping and head shaking combined with fear of environmental situations because of sounds*.

According to Diagnostic and Statistical Manual of Mental Disorders (DSM-5) [2] the Neurodevelopmental Disorders are a *group of irregular conditions* with onset in the developmental period before the child enters grade school.

They are characterized by developmental *deficits* that produce impairments of personal, social, academic, or occupational functioning. The *limitations are very specific,* such as learning or controlling executive global functions related to social skills or intelligence impairments. The Manual (DSM-5) showed that the "Neurodevelopmental Disorders frequently co-occur in individuals with autism spectrum disorder who often have intellectual disability (intellectual developmental disorder)".

The Manual DSM-5 [2] ranked the group of conditions and classified it as Diagnostic Criteria. We next briefly present these conditions: (1) *Intellectual Disability* (Intellectual Developmental Disorder) includes both intellectual and adaptive functioning deficits in conceptual, social, and practical domains; (2) *Language Disorder* difficulties in the acquisition and use of language across modalities (i.e., spoken, written, sign language, or other); (3) *Developmental Coordination Disorder* the acquisition and execution of coordinated motor skills is substantially below the expected given the individual's chronological age and opportunity for skill learning and use; (4) *Specific Learning Disorder* presence of learning difficulties and in the use of academic skills, following deficits such as inaccurate or slow and effortful word reading (e.g., reads single words aloud incorrectly); (5) *Attention-Deficit or Hyperactivity Disorder* (ADHD) a persistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development that negatively impacts directly on social and academic/occupational activities; (6) *Autism Spectrum Disorder* (ASD) persistent deficits in social communication and social interaction across multiple contexts, including deficits in social-emotional reciprocity, nonverbal communicative behaviors used for social interaction, developing, maintaining, and understanding relationships.

2.3 Digital Games and Music

Novak [11] indicates that music, sounds, games, and entertainment are linked to communication, players and emotions, and these elements also are linked to the interaction experience. Novak argues that music role in games is very important since *music helps to set the atmosphere and feeling*. Soundtrack, sounds and sound effects can change emotions and perceptions of the players. These elements can cause a decisive impact on the player *immersive experience* [11]. As reported by Coutinho [12], a convincing atmosphere with visual elements, sounds, and screenplay can help in the immersion of the players. Fortunas [13] observed improvement in kids with Cerebral Palsy using music compositions software. The author related musicalization and experimentation progress in the users. The most interesting approach to this issue has been proposed by Cibrian [14] about music therapy and musical game (interactive surface). The results showed improvements in the autistic participants' in general experimentation. Ringland [15] developed music technology for kids with Neurodevelopmental Disorders. Ringland's studies have a theoretical foundation in therapy and argue that technology products associated with therapy can be a *complementary therapeutic tool*.

3. METHODOLOGY

Hypersensitivity Disorder patients have special needs, thus in order to better understand these needs, and propose a game that could be useful, our first step consisted of a literature review on three different topics: Hypersensitivity to Sounds [16]; Musical software and Neurodevelopmental Disorders [17]; and Musical Games for Hypersensitivity to Sounds and Neurodevelopmental Disorders [18]. The analysis of our literature review followed the methodology proposed by Laville and Dionne [19].

As a result of our review on Hypersensitivity to Sounds [16] we selected Jüris [3] as the main reference on Hypersensitivity to Sounds treatment to guide our work. In her thesis, Jüris aims to better understand the use of Cognitive Behavioural Therapy (CBT) on the clinical treatment of the condition hyperacusis. Her research focuses on *exposure to sound, psychoeducation, applied relaxation and behavior activation in patients with hyperacusis*. Thirty patients participated in the experiment. The experiment's design involved patients in a control group, as well as in a specific treatment group for 12 months. The theoretical framework in CBT is behavior, which is based on the laws of learning in both conditions: respondent and operant. Respondent conditioning applies to the establishment of a response by pairing a neutral stimulus. Operant conditioning refers to the type of learning in which an individual's behavior is modified by its consequences (e.g. positive and negative reinforcement). As a result, Jüris showed that CBT has proven to be effective for a range of psychiatric disorders and treatment for many patients with mild to moderate psychiatric problems, and specifically a promising treatment option for patients with hyperacusis.

To investigate about Neurodevelopmental Disorders we used the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) [2]. The DSM is a *taxonomic and diagnostic tool for the classification of mental disorders*, perceived as a universal reference for psychiatric diagnoses. It has been used by clinicians and researchers from different orientations (biological, psychodynamic, cognitive, behavioral, interpersonal, family/systems).

Based on the analysis of our review of musical software and games for Hypersensitivity to Sounds and Neurodevelopmental Disorders, we identified a set of recommendations [17, 18] *aimed at game and software design specific to support the special needs* of patients of Neurodevelopmental Disorders.

In this paper, we propose a set of strategies that address how the treatment principles and symptoms identified in the literature could be considered in game design aimed at hyperacusis patients. We also associate existing recommendations to each strategy. In the next section we present our results.

4. RESULTS

In this section, we present *12 Principles* (P) of clinical nature care identified in the literature. For each principle, we present a *Strategy* (S) of a technical nature, that presents considerations on how to take the principle into account in the design of a game. Finally, we associated the existing *Recommendations* (R) identified in the musical game and software literature to the principles and strategies. Notice that we have identified 7 recommendations, so they do not cover all the spectrum described by the principles.

P1 - Sound Exposure Gradually. It is an upwardly technique of graded sound exposure to patients. The patient is exposed to an initial sound volume, which is then gradually increased every week. This exposure takes place with a general purpose of environmental sound enrichment. For example repeatedly dropping a spoon, in a way that the initial height is gradually increased, and so is the sound enrichment [3, p. 31]. This technique is also presented in Jastreboff [20].

S1 - a) making a background layer of the ambient sound which gradually increases in intensity during the game; b) using graphic elements and characters emitting different sounds, adding energy for sound enrichment; c) using characters random arise in different moments.

P2 - Audiological Measures. Jüris defined a hearing threshold to identify patients with hyperacusis. To do so, a pure tone audiometry was performed in the patients using the ascending technique. Hearing thresholds were measured in dB (HL) at 125 to 8000 Hz. It was defined as the average of the frequencies 500, 1000, 2000 and 3000 Hz for each ear. For loudness discomfort levels (LDL) the measured frequencies were 250, 500, 1000, 2000, 3000 and 4000 Hz. And for at least one ear, the person was considered to meet test criteria for hyperacusis level of 90 dB or less for the frequencies of 500, 1000 and 2000 Hz [3, p. 26,27]. This technique is also presented in Anari [21].

An interesting approach to this issue has been proposed by Coelho [7, p. 171]. In her experiment with hyperacusis children it was noticed that many children did not report discomfort during the audiometry examination with measured output of: 110 dB HL at 0.25 Hz; 120 dB HL from 0.5 Hz to 6.0 kHz, and 100 dB at 8.0 kHz, adding 5 dB gradually at a time.

Anari [21] made a very important recommendation for audiological measurements. The authors started at a lower level 50–60 dB HL, different from traditional recommendations that start at 70 dB HL, as the authors considered based on their experience that 70 dB HL causes discomfort for many patients including the not hypersensitive.

S2 - a) making sound intensity model based in audiology measurements frequencies; b) equalizing audio developed; c) to setting the level of start at 50–60 dB; d) applying harmonic music in the game; e) creating configuration screen for overall game elements.

P3 - Applied Relaxation. Commonly patients with Hypersensitivity to Sounds feel tense listening to their sounds that are discomfortable for them. For this reason, Jüris applies a shortened version of the CBT applied relaxation technique. This technique is best used in treatment for stress and can be applied several times a day [3, p. 31]. This technique is also presented in Hayes [22].

S3 - a) develop mobile games that can easily be played in different situations, times and contexts; b) define few difficult situations and, for these situations.

P4 - Behavior Activation. Patients often give up different activities due to hyperacusis of sound. The goal of this part of the treatment is gradually *restart activities* that patients had given up by behavior activation. Treatment is usually used for treating depression [3, p. 32]. This technique is also presented in Schaaf [23].

Klein [24] highlights that Williams Syndrome [1] is associated with Hypersensitivity to Sounds. In a study she designs parental questionnaires to determine the prevalence of hyperacusis. The questions are: (1) *"has your child ever been unusually frightened by certain sounds?"* and the parents answered "yes" or "no". The results indicate that 83% of participants answered "yes" in the Williams group and 3% in the control group; (2) *"check the following noises that had or still do repeatedly bother your child."* aand the parents check one or more characteristics of the sounds that their child found offensive. The results were "firecracker," "power saw," "electric drill," "fire engine siren," "motorcycle," "loud auto muffler," and "blender" were selected by 60% of the parents. Features of the offensive sounds asked were: "loud" and "sudden" were chosen 73% and 74% of the time; whereas "high pitched" was chosen by 52%, and "low pitched" by 6%. The sound "thunder" was the most common self-write sound by 19%. The characteristic, benign sound, was added to the list by parents were "automatic ice maker," "television test tone," "newspaper crackling," and "church bells" [24].

S4 - a) searching and defining locals can help behavior activation; b) creating a game scene illustrating the difficult circumstances for hyperacusis; c) creating sounds related to circumstances; d) creating mode benign sound and rough sounds; e) searching and defining places can help behavioral activation; f) creating customizable game screen for input music and background screen; g) creating a noise model, various noise track based on research.

P5 - Psychoeducation. The technique applied to the education the patients, considering reasons and effects of the condition in their overall health, and the structure and contents of the treatment material. The main base of the psychoeducative method was on teaching the patient how to avoid sounds that can affect the auditory system and fear/anxiety. Jüris shows explanatory material on the treatment hyperacusia with information about the treatment of hyperacusis, the CBT model utilization, sounds applied, sound levels and assessment of riskful sounds by the participating patients [3, p. 30]. This technique is presented also in Kennerley [25].

S5 - a) creating documentation about treatment applied to the end user; b) creating audiovisual material in the game about the treatment applied; c) creating help documentation of the game; d) creating explanatory material about sound levels and model utilized.

P6 - General Mental Abilities. This condition is characterized by deficits in intellectual functions that concern reasoning, planning, abstract thinking, learning from instruction and experience, problem solving, judgment, and practical understanding. It is a typical symptom of *Intel-*

[1] Williams Syndrome is characterized by cardiac defects, physical and developmental delay, stellate eye pattern, elfin/pixie facial features and hyperacusis.

lectual Disabilities [2].

S6 - a) elaborating simple solutions for problems; b) inserting elements that represent physical objects that can be manipulated in the game.

R1 - a) the simple navigation results in game play flow; the fluidity navigation with simple tool and short cognitive processing is the secret for the fluidity; b) the graphics interface is a window for stimulating the curiosity [26]; c) inserting concrete and personal objects in game can result in an immersive experience and better creative experience to the player [27].

P7 - Communication. This condition is characterized by deficits in acquisition and use of the language, speech, and communication due to failure in the production or comprehension of vocabulary, sentence structure, and discourse. It is a classic symptom of *Communication Disorders* [2].

S7 - a) avoiding complex verbal language; b) exploring graphic language can help learning a game; c) creating navigation flow with common graphic sign; d) using methodology based in a conversation between designer and user through interface for interaction model.

R2 - a) concrete objects such as cars and dolls are considered favorite because the users do not understand how to the system works [27]; b) the use of the interaction sensors can give enjoyable and functional experiences to general children but mainly to the disabled [27].

P8 - Coordination. This condition is characterized by impairment in motor skills requiring coordination. In addition, these skills also vary according to age, and performance or participation in daily activities in family, school, social, or community life. It is a typical symptom of *Developmental Coordination Disorder* [2].

S8 - a) creating activity with simple interaction such as pressing or dragging game objects; b) creating intuitive game design when using motion sensing; c) developing the system to customize performance gradually; d) developing direct manipulation game.

R3 - a) multi-touch interaction is broader than computer possibilities [28]; b) game controllers that go beyond "button press" interaction can improve attention, expressiveness, and performance [27]; c) design project should not overwhelm the user and have limited cognitive process to a minimum necessary performance [26]; d) simple system should present user-centred experiences because depending on the mood of the user the interface can be surprising. For this reason it is important for the system to have customization options [26];

P9 - Learning. This condition is characterized by essential feature of specific learning disorder is persistent difficulties learning keystone academic skills. It is classic symptom of *Specific Learning Disorder* [2].

S9 - a) developing gradual game design with an increasing learning level while playing; b) creating clean graphic interfaces; c) creating navigation from phase to phase; d) creating a tutorial video before the phases; e) creating game scene and graphic elements corresponding to the scene.

R4 - a) stimulating visual and tactile perception can increase not only player engagement with the learning process, but also the comprehension of experiment subjects [29]. These elements are directly connected to the way one can see the objects/environments and in turn, the way we explore them [29]; b) it is not only necessary to see the object, but also to manipulate it, considering that touch is linked to the cognitive processes and also in the understanding of the concrete experiences – comprehension increases the engagement in the learning process [29]; c) discovery through action is a game plan which lets the person to add complexity in how they animate the interface through their interaction [26]; d) avoid dispensable details that requires cognitive level ability or users' interest [26].

P10 - Attention. This condition is characterized by insistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functionality or development. In Inattention the patient manifests difficulty sustaining focus. Impulsivity refers to hasty actions, without forethought. Also, Hyperactivity refers to excessive motor activity. It is a typical symptom of *Attention-Deficit/ Hyperactivity Disorder* (ADHD) [2].

S10 - a) drawing player's attention through graphics and sounds elements; b) trying to keep player's attention through animations or activities that always something have on in the game; c) constantly changing scenes; d) creating game screen: time is over; e) definging a time limit for each game phase; f) triggering curiosity through rewards.

R5 - a) color, light, texture and sound can set the experience and afford both an aesthetic platform for imagination [26]; b) trigger curiosity through repetition can create rhythms and patterns that enhance interest through experience [26]; c) experimentation can result in improvement in the children [14]; d) autistic players speak to a number of different game features that relate to their overall enjoyment such as presence of the achievement and challenge artifacts allowing creativity or autonomy. These factors and also interesting story elements and emphasizing artistic elements are important for them

P11 - Social Interaction. This condition is characterized by deficits in social communication and social interaction across multiple contexts such as, in social-emotional reciprocity; in nonverbal communicative behaviors used for social interaction; in developing, maintaining, and understanding relationships; repetitive patterns of behavior; inflexible adherence to routines, in general issues. It is a typical symptom of *Autism Spectrum Disorder* (ASD) [2].

S11 - a) developing game in which it is possible for two people to play simultaneously; b) choosing game activities that can help in player's performance such as repeating a specific movement, involving more than one player, among others; c) creating a therapist profile with phases or activities that can be sent to each individual user in the game; d) storing data on users' performance and activities in the game; e) creating activities which can be played in face-to-face groups such as in a classroom.

R6 - a) positive emotions can express positive results in building vocabulary, encouraging vocalizations, and communication forms [30]; b) sharing among the children can be promoted through multi-touch technology, passing the mobile from hand to hand [30]; c) research proper tools to aid the user in expressing himself through the system [31].

P12 - Specifiers. Severity specifiers are used to describe the current symptomatology, and may vary with the patient's context. The specifiers depending of the severity level such as: Level 1 - requiring support, Level 2 - requiring substantial support, and Level 3 - requiring very substantial support [2].

S12 - a) searching and defining the target audience of the game; b) using appropriate methodology that can help understand needs of the game audience; c) considering mild level severity specifiers of the player; d) setting age range for the user; e) use of reflexive theories about system communication; f) use cognitive process methodologies.

R7 - a) studies involving the games' production show diverse audience both in age range and in pathology support tool. Bozzi [32] created for children from 5 to 9 years old, Malinverni [33] from 4 to 6 years old, and Wrońska [30] for children from 8 to 12 years old. In terms of diagnostic Challis [27], developed a game for users with motor disabilities, Cibrian [14] for ASD, and Gehlhaar [31] for broad of disabilities spectrum.

5. APPLIED STRATEGIES IN THE GAME

Based on the principles and recommendations identified and strategies proposed we have developed the game Sêntimus. In this section we present how the strategies proposed were applied in designing the game.

The main goal in the game is popping balloons by touching them on the screen. The punctuation is associated to the number of ballonns popped, each popped balloon scores one point, and as the player makes more points, the intensity of the ambient sound in the background increases. It is divided in 4 phases: Level 1 - *Learning Mechanics*; Level 2 - *Birthday Party*; Level 3 - *Ghosts*; and Level 4 - *Relaxation*, as depicted in Figure 1. For each level, there is delimited time, and this can be modified in the configurations. The user plays first the Level 1, second Level 2, third Level 3, and finally the Level 4. For each level there are different graphic interface, sound elements and characters. The system records the player's performance data, and this data is taken into consideration in the following level.

Figure 1. The four levels of the game Sêntimus.

E1 - Sound Exposure Gradually. We created the game level with ambient background sound and graphic elements of a birthday party. If the player pops balloons, it increases the intensity of the ambient sound layer. Each balloon popped scores one point in the game. We developed an *Intensity Model* and the sound dynamic is: in Level 1 and 2 the sound is enriched gradually, increasing tension sensation/intensity levels; in Level 3 there is sound impoverishment, decreasing this pressure sensation; in Level 4 the is a leveling sound, stabilizing these sensations.

E2 - Audiological Measures. Audiology test has been conducted using pure tone (sine wave) [3]. Borges [34, p.93] studied music noise, synthesized noise and showed that these elements are associated to the uncomfortable feeling, discomfort or displeasure sensation. On the other hand, harmonic music is normally associated with the feeling of happiness and familiarity. In this work we used the complex tone, in which fundamental frequency presents energy equivalent to stipulate pure audiological tone. We developed a *Noise Model* in Brazilian context. To Level 1 and 2: the little monster characters emit breaking glass sounds and drum plate; the balloons emit popping sounds; the environment sound with people talking; and music composition made specifically for this game. In Level 3 the Model Noise is: the little ghost characters emits thunder sound; the environment sound with people talking; the environment sound with people talking; and sinister music made specifically for this game. In Level 4 we do not apply the Model Noise. The *Frequency Model* is not applied in Levels 1 and 2. In Level 3 the little ghost characters emit this frequency sound: 1047, 1174, 1396 and 523 Hz. In Level 4 there is rhythm and melody random accompaniment sound associated to each balloon. We created 9 tracks: (1) 349.23, 415.30, 698.46; (2) 293.66, 349.23; (3) 329.63, 392.00; (4) 293.66, 349.23; (5) 220.00, 261.63, 311.13, 349.23; (6) 174.61; (7) 349.23, 311.13, 261.63, 233.08, 207.65; (8) 155.56, 207.65; (9) 349.23, 415.30 Hz.

E3 - Applied Relaxation. In this work, we interpret Relaxation with exit de-stress situation, for example, when someone walks out of a noisy environment, this action can give an immediate sense of relief. In this game the Relaxation is removing tension provoked by uncomfortable sounds. In Level 4 we used the music *Happy*[2] and used the scenery of the Minions movie[3].

E4 - Behavioral Activation. Often birthday parties are one of the first collective experiences children have, and this involves different sound types. We create a birthday party environment with music and graphic elements thinking that this environment can be uncomfortable to the hyperacusis, encouraging the player not to avoid it.

E5 - Psychoeducation. We created textual documentation about the game Sêntimus and the treatment inserted in the game such as Sound Exposure Gradually, Audiological Measures among others. We still plan to create audiovisual material on the treatment, for the user's understanding.

E6 - General Mental Abilities. The interaction is simple: a game for popping balloons with concrete elements of a typical birthday party with cake, balloons, music, conversation and unpredictable sounds.

E7 - Communication. The game mechanics occur by experience, direct manipulation, trial, and error. To improve the navigation through the game we used MOLIC [35], an interaction model based on communication among de-

[2] all musical rights reserved for Pharrell Williams.
[3] all rights reserved to Illumination Entertainment.

signer and user through the interface.

E8 - Coordination. The interaction is simple: press the balloon. We created a customization screen that allows the user to change (increase or decrease) the velocity of the balloons trajectory. This configuration can help users touch (and pop) the balloons, according to their coordination needs.

E9 - Learning. In Level 1 we built the game mechanic for the player to learn to gameplay in progress, gradually between phases. The interaction is based on trial and error and individual data storage of the player. Each level has a time limit, changing the level of the game when it ends. Everyday objects are presented in the game, such as ballons and cake.

E10 - Attention. We created the graphic art with real characters mixed with 2D characters to call the attention of the player. In each level there are characters to present an attractive game, there always is some unexpected movement in the screen. The configurations screen can modify the time limit of each level or of the game. If the player does not interact for 20 seconds (does not touch the screen), a message of inactivity appears. We built the game in 4 stages to give dynamism to the game.

E11 - Social Interaction. The storage system stores the following data: login, date, time of login, time of logout, number of phase played, total time of the phase played, number of balloons generated in the phase, number balloons popped in the phase, number of characters in the phase, maximum intensity reached in the phase, punctuation in the phase, click position and timestamp of the popped balloons.

E12 - Specifiers. Loureiro [36] highlights that children with mild severity Level 1 and Level 2 Neurodevelopmental Disorders show on average 2 to 3 delay years when compared of the typical children. We defined the users for this system as children diagnosed with Hypersensitivity to Sounds associated with Neurodevelopmental Disorders, with mild or moderate levels of both diseases. We understand that each user may have different special needs, and thus, created a configurations screen. The settings that can be customized are: the speed of appearance of the balloon; the time interval in which the balloons appear; the total time of each levels; the volume of the noise model, the volume of the background music; the percentage of sound intensity for each popped balloon, and the interval of creation of the characters.

6. CONCLUSION

In this paper we have investigated how treatment principles and symptoms for Hypersensitivity to Sounds, Neurodevelopmental Disorders can be taken into account in the design of a game aimed at children with mild or moderate levels of both diseases. As a result we have organized the relevant principles, symptoms and associated recommendations, as well as proposed strategies of how to make considerations in game design that can be associated to them. We have presented Sêntimus, a musical game that illustrates how the strategies have guided design decisions. Although we have associated each strategy to a single principle, we have noticed that strategies may be associated to more than one principle, especially those of Neurodevelopmental Disorders, due to comorbidity[4]. For instance, strategy S10 associated to Attention can collaborate with strategy S9 associated to Learning. The future steps in our research involve evaluating Sêntimus with hyperacusis children as part of musictherapy session. We also intend to further evaluate Sêntimus and the strategies proposed. Interesting future directions include analyzing different relations between the strategies and principles, as well as to analyze strategies S5 e S9 according to Coutinho's classification of sound signs and define how different sound signs could be used in different contexts or as part of an increasing exposure to sound.

Acknowledgments

The authors gratefully acknowledge the financial support of CAPES.

7. REFERENCES

[1] J. Hazell, "Hypersensitivity of Hearing (Hyperacusis, misophonia, phonophobia and recruitment - decreased sound tolerance)," *Tinnitus and Hyperacusis Centre, London*, 2002.

[2] American Psychiatric Association, *Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-5)*. American: American Psychiatric Pub, 2013.

[3] L. Jüris, "Hyperacusis Clinical Studies and Effect of Cognitive Behaviour Therapy," Ph.D. dissertation, University Hospital, Akademiska, 2013.

[4] U. Katzenell and S. Segal, "Hyperacusis: review and clinical guidelines," *Otol Neurotol*, 2001.

[5] E. Gomes, F. S. Pedroso, and M. B. Wagner, "Auditory hypersensitivity in the autistic spectrum disorder," *Pró-Fono Revista de Atualização Científica*, 2008.

[6] M. S. Gonçalves and T. M. Tochetto, "Hiperacusia: Uma Abordagem Teórica," *Revista CEFAC*, 2005.

[7] C. B. Coelho, T. G. Sanchez, and R. S. Tyler, "Hyperacusis, sound annoyance, and loudness hypersensitivity in children," *Progress in brain research*, 2007.

[8] G. Andersson, N. Lindvall, T. Hursti, and P. Carlbring, "Hypersensitivity to sound (hyperacusis): a prevalence study conducted via the internet and post: Hipersensibilidad al sonido (hiperacusia): un estudio de prevalencia realizado por internet y por correo," *International journal of audiology*, 2002.

[9] J. Marriage and N. M. Barnes, "Is central hyperacusis a symptom of 5-hydroxytryptamine (5-HT) dysfunction?" *J.Laryngol.Otol.*, 1995.

[4] The presence of one or more additional diseases or disorders co-occurring.

[10] J. M. Hughes, "Therapeutic Listening Communication in Children with Autism and Hyperacusis," Ph.D. dissertation, San Jose State University SJSU ScholarWorks, 2016.

[11] J. Novak, *Game Development Essentials Third Edition*. NY: Delmar Cengage Learning; 3 edition, 2012.

[12] F. Coutinho, "Revisiting Game Accessibility for Deaf and Hard of Hearing Players," Master's thesis, UFMG, 2012.

[13] M. L. F. Fortunas, "O Desenvolvimento Musical de Crianças com Paralisia Cerebral," Ph.D. dissertation, Universidade Católica Portuguesa, 2015.

[14] F. L. Cibrian, M. Tentori, and N. Weibel, "A Musical Interactive Surface to Support the Multi-Sensory Stimulation of Children," *10th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 2016.

[15] K. E. Ringland, R. Zalapa, M. Neal, L. Escobedo, M. Tentori, and G. R. Hayes, "SensoryPaint: A Multimodal Sensory Intervention for Children with Neurodevelopmental Disorders," *UBICOMP*, 2014.

[16] D. L. Gomes and C. M. V. Loureiro, "Proposta de um Aplicativo na Hipersensibilidade Sonora no Autismo," *XIII SIMCAM - Simpósio Internacional de Cognição e Artes Musicais*, 2017.

[17] D. L. Gomes, A. M. André, and C. M. V. Loureiro, "Softwares Musicais e Transtorno do Neurodesenvolvimento: Um Estudo de Revisão de Literatura," *2º Nas Nuvens... Congresso de Música*, 2016.

[18] D. L. Gomes and C. Loureiro, "Aplicativo Musical na Hipersensibilidade Sonora nos Transtornos do Neurodesenvolvimento: Um Estudo de Revisão," *SBGames*, 2017.

[19] C. Laville and J. Dionne, *A Construção do Saber: Manual de Metodologia da Pesquisa em Ciências Humanas*. Editora UFMG, 1999.

[20] P. J. Jastreboff and J. W. Hazell, *Tinnitus retraining therapy: Implementing the neurophysiological model*. Inglaterra: Cambridge University Press, 2008.

[21] M. Anari, A. Axelsson, A. Eliasson, and L. Magnusson, "Hypersensitivity to sound: questionnaire data, audiometry and classification," *Scandinavian Audiology*, 1999.

[22] S. A. Hayes-Skelton, L. Roemer, S. M. Orsillo, and T. D. Borkovec, "A contemporary view of applied relaxation for generalized anxiety disorder," *Cognitive behaviour therapy*, 2013.

[23] H. Schaaf, B. Klofat, and G. Hesse, "Hyperakusis, phonophobie und recruitment," *Hno*, vol. 51, no. 12, pp. 1005–1011, 2003.

[24] A. J. Klein, B. L. Armstrong, M. K. Greer, and F. R. Brown, "Hyperacusis and Otitis Media in Individuals with Williams Syndrome," *The Journal of Speech and Hearing Disorders*, 1990.

[25] H. Kennerley, J. Kirk, and D. Westbrook, *An introduction to cognitive behaviour therapy: Skills and applications*. London: Sage, 2016.

[26] W. Keay-Bright, "Designing interaction though sound and movement with children on the autistic spectrum," in *International Conference on Arts and Technology*, 2011.

[27] B. Challis, "Designing for Musical Play," in *Technologies of Inclusive Well-Being*. Springer Heidelberg, 2014.

[28] J. P. Hourcade, N. E. Bullock-Rest, and T. E. Hansen, "Multitouch tablet applications and activities to enhance the social skills of children with autism spectrum disorders," *Personal and Ubiquitous Computing*, vol. 16, no. 2, pp. 157–168, 2012.

[29] M. Ma, A. Oikonomou, and L. C. Jain, *Serious games and edutainment applications*. Springer, 2011.

[30] N. Wrońska, B. Garcia-Zapirain, and A. Mendez-Zorrilla, "An ipad-based tool for improving the skills of children with attention deficit disorder," *International journal of environmental research and public health*, 2015.

[31] R. Gehlhaar, P. M. Rodrigues, L. M. Girão, and R. Penha, "Instruments for everyone: Designing new means of musical expression for disabled creators." Springer.

[32] Y. Bozzi and E. al., "GoLIAh: A Gaming Platform for home-Based Intervention in Autism – Principles and design," *Frontiers in Psychiatry*, 2016.

[33] L. Malinverni, J. Mora-Guiard, V. Padillo, L. Valero, A. Hervás, and N. Pares, "An inclusive design approach for developing video games for children with autism spectrum disorder," *Computers in Human Behavior*, 2017.

[34] R. C. Borges, "Um estudo sobre classificaçoes fisicas e perceptivas de timbres da escuta cotidiana, de sons sintetizados e de noise music," Master's thesis, Universidade Federal de Minas Gerais, 2013.

[35] S. D. J. Barbosa and B. S. da Silva, "Design da Interação Humano-Computador com MoLIC," *IHC - Symposium on Human Factors in Computing Systems*, 2014.

[36] C. M. V. Loureiro, "Musicoterapia na educação musical especial de portadores de atraso do desenvolvimento leve e moderado na rede regular de ensino," Master's thesis, UFMG, 2006.

AUGMENTED CHARANGO: AN INSTRUMENT FOR ENRICHING THE ANDEAN MUSIC SOCIAL ROLE

Julian Jaramillo Arango
Núcleo de Sonologia
Escola de Comunicações e Artes
Universidade de São Paulo
julianjaus@yahoo.com

Jaime Daniel Rojas
Independent Researcher
rojaj292@alumni.newschool.edu

ABSTRACT

Departing from Thomas Turino's research about the context of the charango (Andean stringed instrument) in rural Peru which is highly symbolic and relates to processes of courting and love and using it as an inspiration, our aim is to create a new interface design for the charango that expands its social possibilities yet to be determined. An electronic extension of the instrument will be incorporated which enables the player to get new layers of sounds serving as a new development of the NIME (New Instruments for Musical Expression), human computer interaction and computer music. Schachter's work on live electronics and the cajon (a Peruvian percussion instrument) and Escobar's expanded marimba from the Colombian Pacific coast as well as are all proper references for this research.

INTRODUCTION

This paper reports the first stages of an ongoing research devoted to build an electronic extension for the charango. First, the project points out some technical challenges since the device is intended to be portable, embedded to the instrument's body and operated in real time. Second, the creation of the digital audio processes will be guided by theoretical insights. The traditional charango sound will be transformed according to interpretative playing techniques recognized in musicological and ethnomusicological studies. As representatives of a particular social roles of the instrument, traditional techniques such as the strummed playing one, will inform some decisions in the signal processing . By including such theoretical concerns in an instrument-building reasoning, the project aims to propose methodological directions in sound and music studies. Conscious of the unconventionality of the procedure, we take into account both artistic and scientific perspectives of music interaction, as well as links between theoretical and practical endeavours.

This work in progress research described in the paper is divided into four parts. In the first section we will present a historical context by recognizing pertinent references in our region concerned with indigenous music making and new technologies. Then we will have an explanation of the research project in the second section, here we will present the proposal of a four stage plan for the research. Afterwards we will talk about the strumming technique and justify why it is the focus of the research. Finally in the last section we will discuss some activities and procedures we envision in future stages of the project.

Julian Jaramillo Arango and Jaime Daniel Rojas Vargas are early career scholars from Colombia living and working in Brazil. While Julian is currently accomplishing a postdoctoral stay at the Núcleo de Sonología that is part of the School of Arts and Communication at University of Sao Paulo, Jaime is currently living in Curitiba doing research in various projects, he holds an MA in Media Studies. They have both managed to work from a distance.

HISTORICAL CONTEXT

Definition of the instrument

The charango is a small Andean string instrument, its origins are speculated around in the 18th century around Peru and Bolivia in the high plains of the Andes mountains. Its creation was due of the contact between local and European cultures, this is an important input since there is a group of regional instruments from different countries with a similar mixtured origin.

According to the Groove Music Online dictionary the charango is:

"It is shaped like the Spanish guitar but has a small, thin soundbox and short strings, giving a sharp, high-pitched sound. The neck has between five and 18 wooden, bone or metal frets. The soundbox may have a flat wooden back of cedar or walnut, or a round back made of armadillo shell or a single piece of carved wood; its face, which has a round soundhole, is of pine, spruce, cedar or walnut and the bridge is cedar or walnut. The total length of the instrument varies from 45 to 65 cm. The instrument also exists in other forms; it has been made from a round gourd, in a pear-shape of wood or armadillo, and ornately carved in the form of a mermaid. The strings are arranged in four or, more commonly, five single, double or triple courses and number between four and 15; they are of metal, nylon or gut (now rare)."[1]

The evolution of the instrument is highly symbolic, it has even caused tensions between the countries of Bolivia and Peru, due to the fact that both claim that the instrument was invented in their territory. This makes no sense since its creation was before the establishment of the nation states. Julio Mendivl's hypothesis upon the origin of the instrument is the following: "neither the European and American development of the string instruments was linear but dialectic, multi directional, loaded with interferences, struggles and loans that present elusive when trying to establish a historic reconstruction." [2]

Ethnomusicological studies

It is Thomas Turino's study entitled: "The Charango and the Sirena: Music, and the power of Love" [3] that serves as an inspiration for the current project. The study was published in 1983 and it explores the charangos symbolic functions in the courting cycle in the town of Canas in Cusco, Peru. In the stages of engagement of couples, a duality exists between desire and anxiety, the charango serves as an element that channels both the fear and the union between the couple. Young men decorate their instrument and rely on the Sirena for power to overcome the nervousness and to perform well. The ambiguity lies in the fact that the charango represents both good when young couples are single and court, but evil when one is married and uses it, here it is seen as *bohemio.*

Contemporary music in Latin America

Even though our inspiration comes from the studies of popular music, our area of expertise has been around contemporary music. In order to introduce our references we will first present an brief synthesis of the historic context of Latin American contemporary music and its efforts to constantly integrate with both indigenous influences and new technologies.

Historical synthesis

The genre of contemporary music in the region underwent major transformations throughout the twentieth century, the Mexican revolution that began in 1910 meant the breakdown of colonial domination on the continent. This fact was reflected in the music. According to American musicologist Deane L. Root:

"Composers in these countries motivated by the awakening of a nationalistic conscience, began the use melodies, rhythms and indigenous and creole instruments in their music." [4]

Since the emancipatory feeling towards Europe was shared even in the United States and Canada, the Pan-American Association of Composers (PAAC) was created in 1928 with the aim of promoting the musical appreciation of the entire continent: North America, Central America and South America. The board of directors was composed of: Edgar Varése, Henry Cowell, Carl Ruggles and Carlos Chávez. In this organization, the results of the musical works were characterized by incorporating native elements, however they still had the European basis. Unfortunately the PAAC lasted very little and in 1934 due to lack of funds and hit by the "Wall Street Crash" it could not continue.

The Second World War brought to Latin America a period of musical imitation, there were no great autochthonous works, later the idea of musical integration of South, Central and North America breaks down because the United States in Cold War times begins to intervene intensively in the political affairs of the Latin American nations with the pretext of fighting communism. In 1959 the Cuban Revolution opens the possibility of changing this trend. One of many emblematic works that achieves a break is: *Cantata for Magic America* by Alberto Ginastera. When using the word *magic* refers to the pre-Columbian stage, as a support, he used the poems of the first Christian priests who toured the Mayan, Aztec and Inca cultures. The composer gives an epic intention, the end of the work represents the greatness and the destruction of some fantastic cultures. It is a daring and forceful bet that really breaks the history of contemporary Latin American music in two. Now, nothing will be the same. [5]

Taking advantage of this impulse, the Latin American Center of High Musical Studies (CLAEM) of the Torcuato Di Tella Institute was created in Buenos Aires, Argentina between 1963 and 1971. The center had the latest inventions in electroacoustic music with equipment such as tape recorders, microphones, oscillators among others. This program was amazing because it managed to fund students from all over Latin America to travel to Argentina and with all stipends for two years with access to the studios for creating works. Many of the pieces done during the program involved the treatment of the indigenous and popular cultures giving continuity to Alberto Ginastera's ideals. The attendees later went to influence a new generation of composers. This was the case for Graciela Paraskevaídis from Argentina that attended the CLAEM and later went to influence Bolivian composer Cergio Prudencio who created the first native instrument orchestra in the region. He acted as both composer and director with different sets of zampoñas (Andean pan instruments) interpreted by its members. He also composed pieces for charango, voice and electroacoustic music.

Named Orquesta Experimental de Instrumentos Nativos (OEIN) it revolutionized contemporary music because it switched the paradigm of approaching music from a hierarchical European centered perspective. This is by far the starting point and the most prolific of many attempts in the region to make new music with what surrounds us, in this it is case is the indigenous cultures.

Departing from the OEIN, the Orquesta de Instrumentos Autóctonos y Nuevas Tecnologías (Orchestra of autochthonous instruments and new technologies) was created in 2004 in Buenos Aires at the Universidad 3 de Febrero. The main premise of this project is to erase the boundaries between composer and interpreter allowing all members to create and execute and even build instruments. This posture has been of great controversy with experts being for and against, mainly arguing that the quality of both the interpreters and the compositions is not good enough due to this matter. What is important is to state that it is a good reference since it involves new technologies and indigenous instruments.

Influences

So far we have looked at the evolution of contemporary Latin American music and seen a couple of examples of orchestras that incorporate the indigenous and new technologies. Now we are going to look at works from individuals that deal with indigenous instruments from a compositional and technical perspective as well as from an interactive art installation perspective.

Argentinian composer Daniel Schachter [6] does a musical composition in 2009 for the Peruvian Cajon which is a box shaped percussion instrument with real time processing and electroacustics. Based on Bregman's Gestalt principles as well as Denis Smalley's Espectromorphology, Scharchter builds the content of the real time processing engine with the software *MAX*. He uses the following processes: Reverb, Reson (Resonant Filter) and Delay. A well structured score, as well as the patches accompany this creation.

Regarding examples, with interaction and gesture with aboriginal elements, there is Daniel Escobar's *Expanded Marimba* [7]. The Marimba is an indigenous instrument from the Afro-Colombian populations in the pacific coast. Its constructions is made in an artesian way. What the Colombian artist does, is to create an interaction with electronic and digital technologies so that video and lights accompany the trigger of the instrument resulting in an multi sensory experience involving video mapping using *Arduino*.

THE EXPANDED CHARANGO PROJECT

Interaction Design

All these previous experiences have given a theoretical background to start our musical interface project. As it was mentioned earlier, our goal is supporting ourselves on musicological insights to develop a technological device. While the idea of using musicology to guide a practice based process can be seen rather unorthodox, in the field of musical instrument building it can offer an opportunity to observe the results in the context of research. As we have noted in the work of Cergio Prudencio`s OEIN or even in the Daniel Schachter experiment with the cajon, the construction of new interfaces or instrument extensions is frequently intended to the experimental music performance. The same occurs with some technological devices devoted to traditional music created under the NIME (New Interfaces for Musical Expression) umbrella where the vernacular context of the instruments in not taken into consideration. While social aspects of the instrument or even traditional techniques are usually neglected in the construction of new devices and interfaces, they are rarely incorporated in the traditional music arena, or in the folkloric music one. In this regard, Franinovic and Salter warn about some assumptions present in the NIME model by stating that:

"The main issue is the almost formulaic understanding of interaction as a series of input-output processes: a gesture or action triggers an appropriately stored or mapped series of sonic responses that may be adjusted based on the range of expression of the input. This assumes an already fixed set of relations among the user/interactor, the object/instrument/sound-making body, and the environment in which the interaction with sound takes place". [8]

Under our vision, extending the charango capabilities has to do with taking into account a set of relations where its social role can be identified. The techniques developed along the charango evolution are being examined in order to find appropriate digital audio processes for the device. Our interface is intended to operate in a twofold way, by increasing the intensity and by transforming the sonority without de-characterizing the charango`s acoustical and musical identity.

Methodological considerations

With this in mind, we payed attention to the methodological aspects of the research. In the same way that in other recent projects [9, 10, 11], we expect to include the creative process in the core of the research agenda, by integrating artistic and scientific procedures. Based on design research methodologies [12, 13, 14], we have traced a one year plan including four stages: 1. Analysis 2. Synthesis 3. Experiment, and 4. Evaluation. We prefigure a research program foreseeing periods devoted to theoretical (Analysis, Synthesis) and practical activities (Experiment and Evaluation) in order to accomplish a transdisciplinary approach to creation.

The analysis stage has been dedicated to typify the instrument by scrutinizing relevant musicological, ethnomusicological, historical and theoretical sources and recognizing previous works committed to similar objectives. In addition, we have been taking personal instruction with professional charango players, studied playing methods and exploring available online lessons as a source of oral transmission knowledge. As it was reported in the Historical Perspectives section we have identified a nationalistic bias in some musicological studies concerning Andean traditional music and particularly the charango. Turino's claim about the duality between desire and anxiety in the hearth of the charango`s social role is taken as an inspiration to flee from the nationalistic tendency.

In the next section we will discuss findings related to the synthesis stage that deal with determining the charango playing techniques in order to strengthen the interface. We have implemented spectral analysis as a procedure to determine acoustical profiles and timbrcal identity. We also associate the charango`s strumming playing technique with the duality established by Turino. In the section titled *future work* we will describe the envisaged activities for the experiment and the evaluation stage.

THE STRUMMING PLAYING TECHNIQUE

The strumming playing technique is done by swinging the whole wrist really fast while playing all the strings with either the middle or the index finger. The effort is done in the forearm and virtuosity is essential since it involves high speed while controlling the movement and being efficient in terms of motion in order to maintain the sound for some time. The translation of the technique in spanish (and portuguese) is *repique,* (as it is popularly known). According to Turino: "The overall effect is a high-pitched piercing timbre with an unclear rendering of the melody and strong rhythmic drive" [15].

It is a style attributed to the "campesinos" (peasants) and associated with dances and ceremonies. It's importance shifts in the cities, where a new technique called the *t'ipi* or *rasgueo* was developed, in english it means to pinch the strings with the thumb upwards and the rest of the fingers downwards. It is common in the charango repertoire to combine both the strumming and pinch techniques for accompaniment purposes.

The strumming is commonly played using a three, four or five rhythmic pattern and usually in the measures of: 2/4, 3/4 and 6/8. Normally, the first movement is from bottom to below, the second from below to above respectively. Off course, the first movement can start above and many combinations of 3, 4 or 5 shifts can be achieved. Figure 1 shows how a four movement strumming is written. This same pattern seen on a spectrogram is shown in figure 2.

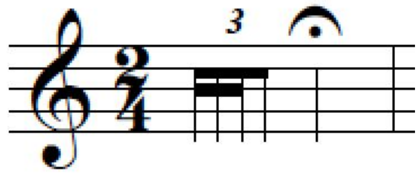


Figure 1. Strumming playing technique main rhythmic pattern.

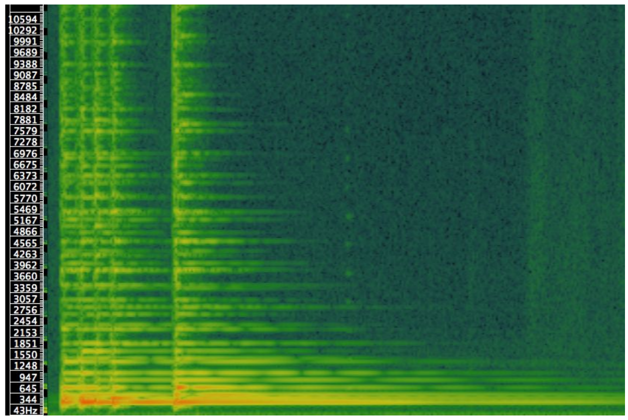


Figure 2. Rhythmic pattern spectrogram.

What can be established from the spectrogram is that the most intense frequencies are the ones around 200-700hz. Further analysis has to be done by looking at the spectrum of several strumming variations in measure, speed and changing chords in order to get valuable insights. We see a lot of potential in the rhythmic patterns that are achieved through the strumming playing techniques since they are hard to achieve and represent the greatness and uniqueness of the instrument.

FUTURE WORK

Although the practical stages of the research have not started yet, in our plan we have some previsions about the materials, the configuration and procedures we will adopt. In these phases the interface will be built and a short evaluation of it will be carried out. In order to provide a whole picture of the research we will report some technical aspects concerning the experiment and evaluation stages.

Device setup

The device will comprise modules of hardware and software. The hardware component embraces an electroacoustic system able to capture, process and reproduce the charango's acoustic signal. It is intended to be embedded in the charango's body in such a way that the player has no need of being plugged to an external audio amplification system. A user interface composed by knobs, buttons, lights and a touchscreen will allow the interpreter easily control the digital audio processes, by balancing between the acoustical and electronic signal and by increasing or decreasing other parameters` levels. While the portability of the device is one of the main challenges in the design architecture, a rechargeable direct current power supply will also be required.

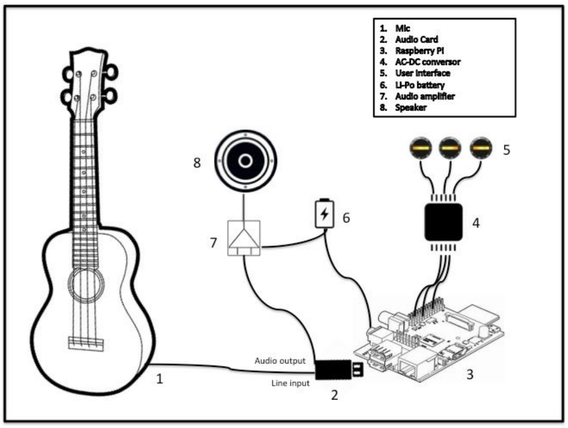


Figure 3. Components of the Charango`s interface

We chosen Raspberry pi [16] platform as the main processor since it offers an open community of coders and makers in which the project can be supported. We expect that other luthiers and tinkers, mainly those committed with South American music and its instrumental tradition, can built our interface in their laboratories or studios. In this regard we favor available and affordable materials along with open hardware and software.

The software component, in which digital audio processes will be implemented, is being developed in Pure Data [17]. We have been creating prototype patches with phase based processes, such as phaser, chorus and flanger as an strategy to deal with the rhythmic patterns of the strummed playing technique. By thickening the signal`s attack time we expect to provide alternative timbrical colours to the charango traditional sound. The implementation of Pure Data patches in Raspberry Pi is eased by libPD library [18], and the user interface is being programmed in Python.

Acknowledgments

This research is partly funded by Coordenação de aperfeiçoamento de Pessoal de Nível Superior CAPES.

REFERENCES

[1] Turino, Thomas. "Charango". in *Groove Music Online*. A. Edited by Deane Root. Accessed 20/04/2018. http://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000005445

[2] Mendivil, Julio. "Apuntes para una historia del charango andino," in *Método de Charango: Federico Tarazona*. Abril Ediciones Musicales, 2004. Translation made by the authors.

[3] Turino, Thomas. "The Charango and the "Sirena": Music, Magic and the power of love," in *Latin American Music Review*. 1983 pp 81-119

[4] Root, Dean. "The Pan American Association of Composers (1928-1934)," in *Anuario Interamericano de Investigación Musical.* 1972 pp. 49-70

[5] Paraskevaídis, Graciela. "Las venas sonoras de la otra América," in *Proceedings of the Simposium La Otra América.* Germany, 2009. Translation made by the authors.

[6] Schachter, Daniel. "Composición interactiva en un entorno de Aleatoriedad Controlada," in *Proceedings of the 12th Week of the Music and Musicology*. Buenos Aires 2015.

[7] Escobar, Daniel. "Marimba-Expandida," Accessed 20/04/2018.http://danielescobar.co/marimba-expandida/

[8] Franinović, K., Salter, C. "The Experience of Sonic Interaction," in *Sonic Interaction Design*. MIT Press, Cambridge. 2013. Pp 39-76

[9] Arango, J.J., Melán, D. "The Smartphone Ensemble. Exploring mobile computer mediation in collaborative musical performance," in *Proceedings of NIME, New Interfaces for Musical Expression*, Brisbane, Australia. 2016

[10] Gutierrez, M. P & Arango, J. J. "AirQ Jacket, design process for wearable technologies and urban ecology," in *Proceedings of the International Seminar on Electronic Arts. ISEA*, Manizales, Colombia. June 2017.

[11] Arango, J. J., Esmog Data. "Interpreting Air Quality Through Media Art and Design," in *Leonardo* April 2018, Vol. 51, No. 2, pp. 183. DOI: LEON_a_01515

[12] Findeli, A. "Research Through Design and Transdisciplinarity: A Tentative Contribution to the

Methodology of Design Research," in *Focused -- Current Design Research Projects and Methods*, Swiss Design Network, pp.67-91

[13] Jones. J. C. "Design methods". John Wiley & Sons, London (1992)

[14] Manzini, E. "Design ethics and sustainability". Dis-Indaco-Politécnico di Milano. Milan (2006)

[15] Turino, Thomas. "The urban mestizo charango tradition in southern Peru: A statement of shifting identity" in *Ethnomusicology.* 1984 pp 253-270

[16] Raspberry Pi. www.raspberrypi.org

[17] Pure Data. https://puredata.info/

[18] Brinkmann, P., Kirn, P., Lawler, R., McCormick, C., Roth, M., Steinser, H.-C.: LibPD. "Embedding pure data with libpd," in *Proc Pure Data Convention*. 2011

STOCKHAUSEN'S *STRUKTUR UND ERLEBNISZEIT*: CONCEPTUAL IMPLICATIONS IN CONTEMPORARY MUSIC ANALYSIS

Ricardo Tanganelli da Silva
Unesp - São Paulo State University
ricardo.tanganelli@gmail.com

ABSTRACT

This paper discusses the idea of music as a "shaper" of time from the analysis of the organization of sound material and its structure in the first section of Alexandre Lunsqui's *Slope Field* (2014). Definitions presented by composer Karlheiz Stockhausen in his article *Struktur und Erlebniszeit* (1955) offers the theoretical basis. Furthermore, this paper defines some processes of alterations, repetitions and perception of time flow through a brief analysis of the arrangement of musical events, putting in evidence Stockhausen's ideas concerning experiential time. Along with these concepts, the research applies definitions from Schaeffer (1966) and Mesquita (2016) in order to evaluate Stockhausen's article consequences and relevance in contemporary repertoire analysis. Finally, the research demonstrates that implications of Stockhausen's concepts about perception of musical structures offers an open field to considerations of temporal experience in contemporary music and provides a useful tool for music theory and composition. This paper is a partial result of a master dissertation developed under guidance of Prof. PhD Marcos Mesquita in the research group Cogmus.

1. INTRODUCTION

1.1 Background to Stockhausen's *Struktur und Erlebniszeit*

With the surging of post-tonal music, the possibilities of organization and unfolding the sound material in time gained an importance not seen before, becoming a subject of interest by many composers. Especially around 1950, with the integral serialism technique, the temporal aspect gained a new impetus due to its dissolution of the other formative elements of musical sound [1], becoming autonomous and self-governing. Controversial, as seen in some critical writings from non-serialist composers, integral serialism changed the perception of temporality. Although American composer Elliott Carter argues that events in constant transformation may overload the limits of human perception and eventually lead to a crisis in musical communication [2], in many aspects this technique led to the development of different possibilities of temporal perception, superposing layers with independent temporal organizations.

In order to explain the phenomena of the diversity of temporalities proposed in serial compositions, composer Karlheinz Stockhausen has shed light on the theme by calling "experiential time" the temporality experienced from the structuring of musical events arranged in the chronometric duration of music. In his article *Struktur und Erlebniszeit* (1955) Stockhausen defines a "time experienced through sound" [3] by means of processes of alteration in sound structure and durational features. Through the analysis of the first section from the second movement of Webern's *String Quartet Op. 28*, Stockhausen illustrates his concept showing a variety of ways in which Webern build strategies to provide listeners to have their time perception molded by the musical events. Figure 1 illustrates a visual schemata proposed by Stockhausen, where dots represents *pizzicati* on the strings instruments and slurs represents the legato play, in order to highlight the organization of the musical events displaced through the modes of attack.

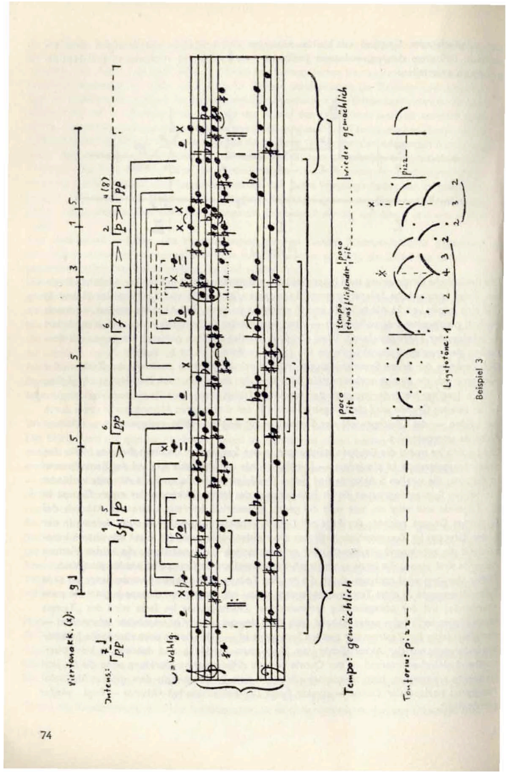


Figure 1. Stockhausen's schemata of analysis.

In this excerpt, Stockhausen argues that - although the quarter note remains consistent throughout the entire section – Anton Webern explores a distinctive temporal sensation based on perceptual data by applying changes in sound emission, modes of attack, increasing and decreasing simultaneities and intensities. Thus, even a single variation in one parameter is sufficient to articulate the temporal organization when others remains unchangeable. In this case, the bowed legato contrasted to the *pizzicati* represents the most dissimilar sound in this section. Due to the increasing density of *legato,* Stockhausen states that there is a sensation of urgency despite the unified rhythm.

Thus, Stockhausen asserts that the temporalities experienced in a musical work articulates the perception of time through differentiable phenomena, in which the distances between them link the passage of time [3]. Through these temporal interrelations, the compositional space can make the perception of temporality vary according to the level of transformation that the elements undergo. Therefore, the time of the experience is susceptible to the proper nuances of the dynamics of the sound processes of each music.

1.2 Some consequences of Stockhausen's article

As previously states, the multiplicity of experiences that music provides became a concern of a growing number of composers, giving rise to several works that deal with the subject with particular commitment. Corroborating Stockhausen's ideas, French composer Pierre Schaeffer says that the "musical duration is in direct function of the density of information" [4]. Thus, both Stockhausen and Schaeffer deals with ideas such as "information density" and "degree of change" of a given musical context as determining factors for listening to the temporal structure in the composition. In addition, his considerations come into consonance with the concept of music as a shaper of temporalities. Considerations by Boulez [5], Xenakis [6], Grisey [7], Ligeti [8] and Ferneyhough [9], for instance, deals with temporal perception and reflects compositional aesthetic concepts, influencing a generation of subsequent composers. Through this theoretical basis, one can assume that Stockhausen's proposal for analysis is competent to highlight temporal procedures in contemporary repertoire.

2. AIMS

This research aims to evaluate Stockhausen's concepts in a twenty-first century composition focusing in enlightening aspects of temporal flow and temporal motion from the musical structure though the sound material arrangement. For this, the research investigates the first section of Alexandre Lunsqui's *Slope Field* (2014), a composition written in a very continuo style, offering the appropriate material for the subject. Ultimately, this research analyzes structural devices such as repetitions, contrasts, densities of simultaneities, intensities, metric sense and temporal flow in order to relate them to Stockhausen's definitions.

3. METHODS

In order to discuss the idea of music as molding the time perception, this research assess Stockhausen's *Struktur und Erlebniszeit* in original German to assure the text integrity. To corroborate to the theoretical basis, in order to support concepts such as repetition, alteration, degree of information, degree of simultaneities, etc., the research also uses works and articles by Schaeffer [4] and Mesquita [1]. The analysis traces a correlation between Stockhausen's descriptions with its occurrences in Lunsqui's work. Thus, parameters such different degrees of repetitions, alterations, simultaneities and intensities presents in Slope Field relates to Stockhausen's ideas concerning time perception.

4. IMPLICATIONS

In accordance to the aims and the collected methods, the research resulted in an analysis attesting the applicability of Stockhausen's definitions. This analysis comprehended the first section of Lunsqui's Slope Field, work in which the composer made extensive use of repeated notes in order to create sonic "fields" with punctual and gradual changes. These repeated notes have an important structural function for the perception of the passage of time because "when we hear a piece of music, processes of alteration follow each other at varying speeds; we have now more time to grasp alteration, now less. Accordingly, anything that is immediately repeated, or that we can recollect, is grasped more rapidly than what alters" [3]. Thus, the repetition of a certain element tends to dilate the sensation of the passage of time. Still according to Stockhausen, the "experiential time is also dependent on the density of alteration: the more surprising events take place, the 'quicker' time passes; the more repetitions there are, the 'slower time passes' [3].

Gradually, the sound layers gains "deviations", with punctual alterations in the phraseological direction, giving rise to a certain internal disturbance to the homogeneous sound continuum. Figure 2 illustrates this occurrence:

Figure 2. Lunsqui's Slope Field, mm.6-12, rhythmic deviations and D-C# pitch clash.

Furthermore, Lunsqui adds "dots" in this texture, acting like a "momentum of surprise" [3] within the established context. These points acts as elements of articulation of a hitherto essentially constant and regular musical discourse, taking on the auditory interest because it presents a degree of contrast greater than the previously established context. According to Stockhausen, "there is surprise only when something unexpected occurs: on the basis of previous events we expect a particular kind of succession of alterations, and then something occurs that is quite unlike what we expected" [3]. Thus, figure 3 shows these "unexpected surprises":

Figure 3. Lunsqui's Slope Field, reliefs contrasted to continuo texture.

At the end of the first section, there is an abrupt interruption in the continuo texture, marking the higher degree of alteration until thus point. This event clearly articulates the sense of musical flow, acting like a brief caesura and marking the beginning of a new section hereafter, as shown in figure 4:

Figure 4. Continuo disrupted, marking the beginning of a new section, reorganizing the previous sound material.

Most of all, the dealing with expectancies, satisfying or frustrating them, pays an important role in this section in order to achieve the perception of temporal flow. Stockhausen wisely says that "our expectations should be aroused through a logic of structural processes, one that can be experienced at the time, in advance and […] still more in retrospect (since what has preceded reveals itself only through what follows, a reversal of causality); once our expectations are aroused, we are in a condition to assimilate information, and are thus provided with aural 'rules': only then do the ensuing displacements and effective alterations surprise us and to the corresponding degree give us information" [3]. Thus, even a very continuo process molds the experience of time perception through some highly effective devices, such as different articulations, different degrees of simultaneities, contrasts in intensities and punctual "reliefs" in a plain established texture. Moreover, some Stockhausen's concepts resembles those from Gestalt grouping principles, such as proximity, similarity, symmetry, good continuation, and common fate [10], although the author did not mention it directly.

As a multi-parametric art, music offers various articulations of time-events, and the superposition of parameters with relative independence puts a listener perception in confusion to decide what to attend. Furthermore, especially in post-tonal music, the temporal aspect pays an important role in composition, motivating an analytical approach that consider both structural procedures and auditory perception.

Is important to note that, although Stockhausen applies his analytical concept in a brief excerpt contrasted only to its following section, one could presume that his theoretical ideas could relate to formal aspects, organizing the passage of time throughout the piece as a whole.

5. CONCLUSION

Due to its technical approach based directly on the structure of sound, Stockhausen proposed a series of categories concerning time disposal and its displacement relating to perceptual characteristics. According to the analysis realized, the research concludes that the definitions in Stockhausen's article proved to be a valid tool for analysis of technical procedures even in a repertoire with continuo characteristics of distribution of musical events, highlighting aspects of temporal perception related to structural devices. Thus, these procedures can contribute to theoretical and analytical field in twenty-one century repertoire and may assist compositional techniques of temporalities.

6. AKNOWLEDGEMENTS

I thank the SysMus organizers, my counselor Prof. PhD Marcos Mesquita, the colleagues and teachers of Unesp, Cogmus research group members and Capes for financial assistance.

7. REFERENCES

[1] Mesquita, M. Serialismo integral e defasagem temporal de parâmetros sonoros. In *Subversões de protocolos: uso impróprio*. Niterói: PPGCA-UFF, 2016, pp. 7-18.

[2] Carter, E. The Time Dimension in Music (1985). In: *Collected Essays and Lectures, 1937–1995*. Rochester: University of Rochester, 1998, p. 224-228.

[3] Stockhausen, K. Struktur und Erlebniszeit. *Die Reihe II. Anton Webern*. Ed. by Herbert Eimert with the collaboration of Karlheinz Stockhausen. Viena: Universal Edition, 1955, pp. 69–79.

[4] Schaeffer, P. *Traité des objets musicaux: essay interdisciplines* (1966). New edition. Paris: Éditions du Seuil, 1998.

[5] Boulez, P. *Penser la musique aujourd'hui*. Genebra: Éditions Gonthier, 1963.

[6] Xenakis, I. Sur le temps. *Kéleütha*. Paris: L'Arche, 1988/1994, pp. 54–66.

[7] Grisey G. Tempus ex machina: a composer's reflections on musical time. *Contemporary Music Review*, v.2, 1987.

[8] Ligeti, G. & Bernard, J.W. Events, Transformations. Perspectives of New Music, Vol. 31, No. 1, 1993, pp. 164-171.

[9] Ferneyhough, B. "The Tactility of Time" (1988). *Collected Writings*. Ed. by James Boros and Richard Toop. Amsterdam: Harwood Academic Publishers, 1995, 42–50.

[10] Shepard, Roger. (1999). Cognitive psychology and music. In Perry R. Cook (ed.), *Music, Cognition, and Computerized Music*. Cambridge, MA: MIT Press, 21–35.

Combining Automatic Segmentation and Symbolic Analysis based on Timbre Features – A Case Study from Ligeti's Atmosphères

Ivan Eiji Simurra
University of São Paulo
ieysimurra@gmail.com

Rodrigo Borges
University of São Paulo
rcborges@ime.usp.br

ABSTRACT

We present an unsupervised method for detecting the timbre structures of a György Ligeti's Atmosphères (1961). The computational approach was proposed for this study in order to retrieve acoustic features to highlight sonic characteristics that may show whether the analyzed method dialogues or not with the symbolic analysis from the music score. We extracted windowed MFCC features from the phonogram, applied Principal Component Analysis to it. Next, we applied Mean Shift clustering algorithm. Our results point out to a satisfactory number of clusters to representing the extracted features statistically along with a reasonably convergence to the music symbolic segmentation. We argue that this is an empirical approach should be understood as a complement for score analysis methods and not as an alternative to it.

1. INTRODUCTION

A striking portion of the orchestral music composed since the second half of the Twentieth Century has made extensive use of non-standard music tradition techniques and/or laboured complex "sound masses". The corresponding works have extended the capacity of the written score to provide a whole conceptual processes of a composition's overall sound to its limits. When instrumental sounds are superimposed or, even more so, when they are intentionally mixture together the rift between the written score and the sounding results may become even more important. In the attempt to analyze such compositions the possibility to include information extracted from both the written score and the recording of its performance becomes a crucial issue. In this article we present and discuss an example of such an attempt based on the analysis that arises from the study of György Ligeti's Atmosphères for full orchestra (1961). The composition accrues the characteristics aforementioned. When considering the issues of identity and similarity between sound-events occurring in the piece, both acoustic features extracted from recorded sources and symbolic analysis from music score take equivalent role as a basis for investigation, a priori.

The affluence of information exists about the genesis of the composition [1–5]. They have not only supply the substratum for a preliminary analysis of the work and even so has also quite straightforwardly suggested questions of the issues just mentioned.

Additionally, the craft of musical analysis brought about directly by the music score provides considerable inputs to highlight aesthetic, conceptual and musical aspects that can be related to the composer's particular planning. However, the analysis performed only by this symbolic source may not be sufficient for the macro and microstructural understanding of the sound universe in which a composer is inserted in. This is in agreement specially on musical works that make use of compositional strategies that extend the scope of the timbre transformations. In these cases, in addition to the traditional methods of musical analysis, as highlighted by Cook [6], the computational environments allow new techniques for musical analysis.

On this particular scenario we developed an analysis study that dialogues with the area of Music Information Retrieval [7]. Hence, the computer support was brought in, first to extract acoustic features from the sound events considered and then to decide on how to classify and compare them within the context of our analysis. Many other analysis have been done so far, all of them under symbolic notation, this method is applied on the timbre features extracted from the sound recording and should provide analysts with some concrete reference.

Although the questions underlying our discussion heavily rely on information provided by the composer the application of the suggested approach to a wider context should also be viable. These possibilities will be the subject of the discussion provided in the closing sections in which we will briefly present some György Ligeti's particular compositional strategy and the Atmosphères on Section 2 and Section 2.1, respectively. On Section 3 we discuss the main stages of the method exploited on the study. On the following two Sections (4 and 5) we present the analysis and results of the study and the discussion of the results. Finally, we conclude our study in Section 6 discussing forthcoming projects.

2. GYÖRGY LIGETI

There are a multitude of technical and conceptual aspects in which it is possible to describe the compositional strategies of the hungarian composer György Ligeti. Apart from the influence of Renaissance Music (mainly by the music of Johannes Ockeghem), Serial Music (by the music of Anton Webern) and specific non-musical aspects of the creative art (by works from Paul Klee) [8], Ligeti highlighted the significant facet of his contemporary discoveries such as the electronic music developments. After a working time period on the Electronic Music Studios, in Cologne, in the last years of 1950, György Ligeti acquired new and important theoretical and technical substrates to construct his own compositional poetics. In such perspective aligning electronic music tools and instrumental music composition techniques Ligeti was able to thrive up his works by using chromaticism, polyrhythm and the he later dubbed micropolyphony [9]. Micropolyphony is a particular construct of Ligeti's own understanding of the aesthetic of sound-mass composition [4]. Accordingly, micropoliphony is the own emancipation of the surface texture, the creation of dense sonorities by chord clusters resulting from the suspension of traditional structures of musical time, form, and pitch relation [10]. Nevertheless, regardless of the fact that Ligeti did not come up with the electronic studio techniques to be the right medium for his idea, that experience in electronic studios was crucial for his composition and his stylistic development [11].

2.1 Atmosphères

As already described in Section 2, Atmosphères was composed after the György Ligeti's experiments in electronic music studios and it develops some compositional strategies that would only be possible for technological resources such as spectral filtering processes, overlapping of supersaturated sound layers and micropolyphony. By internalizing and converting these electronic techniques and their related discourses on orchestral settings, Ligeti addressed timbre as the central compositional issue in his Atmosphères. The structural form of the piece is basically based on its own textural movement. Thus, it is the timbre itself that plays the structural role. The successive structures of different sound densities established by the interlacing of the acoustic instruments convert them as the sound structural layers such as in a mixing console. The exerted audio recording was performed by New York Philharmonic and was conducted by Leonard Bernstein in 1999 (Music of Our Time. Sony Classical SMK 61845 [1]).

3. METHOD

3.1 Feature Extraction

Mel Frequency Cepstrum Coefficients are standard timbre representation in the Music Information Retrieval field [12] both because its ability to represent the spectrum in fewer coefficient and also because of its filtering process that approximates human hearing model.

We extracted windowed MFCCs from the recording using Librosa library [2] with a 2048 sized window and hop size of 512, resulting in 22443 frames extracted fot the whole piece.

[1] More details on goo.gl/Jx19ph
[2] http://github.com/librosa/librosa

Cluster	# of frames
0	9640
1	5134
2	4386
3	2624
4	660

Table 1. Number of MFFCs frames in each cluster.

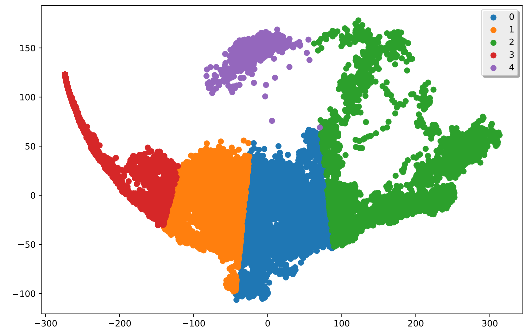


Figure 1. The two first principal components of the MFCC data extracted from the Atmospheres recording.

3.2 Clustering

In order to reduce the dimensionality of the MFCC extracted we applied Principal Component Analysis (PCA) to the data. The two first principal components contains 92% of the variance of the data.

We tested 4 clustering algorithms implemented in Scikit Learn [3] library for finding the one that separates better our timbre data logarithmically normalized. For measuring separability we calculated the silhouette coefficient that measures the distance between groups. The algorithm that presented the best silhouette (0.71) was Mean Shift.

Mean shift algorithm treats clustering problem by assuming some probability density function which generated all points. The regions of high sample density corresponding to the local maxima of this distribution.

To find these local maxima the algorithm works by allowing the points to gravitate towards areas of higher density. Those data points that converge to the same local maxima are considered to be members of the same cluster.

The number of samples located in each cluster is shown in Table 1 and in Figure 1 the two principal components are illustrated with each clusters indicated with a different colors.

3.3 Segmentation

In order to have the piece segmented according to timbre temporal changes we built a novelty curve that should indicate the points of transition between parts. We assumed clusters with fewer frames as the more informative ones. The novelty curve, then, should behave proportional to the

[3] http://scikit-learn.org

inverse of the cluster population: its highest peaks should coincide with the moment when the piece leave that cluster for going to next one successively.

4. ANALYSIS AND RESULTS

The results point to five as the most adequate number of clusters for statistically representing the extracted features and match reasonably to segmentation made manually.

From the perspective of the symbolic music analysis approach, the structural planning of the composition was based on timbre transformation by the agency of a myriad of diversified blocks of music structures built on some electronic music techniques skilled on the sixties. In each block, Ligeti conversed some techniques such as spectral filtering and time stretching saturation, to cite some, into the sound universe of an orchestra settings.

5. DISCUSSION

Atmosphères is developed by the conceptual idea of filtering processes of textural layers, represented by each musical instrumental of the orchestral settings, where each layer performs a distinct process. However, there are different transitions among these layers: sometimes the filtering is done abruptly, sometimes it is made through a gradual transformation. In some juncture it is made from the preceding material or in another moment the process is completely deserted and it presents another material.

From the perspective of the symbolic music analysis, clusters 0 and 1 stand for the first filtering process applied for Ligeti. Figure 3 highlights the moment when Woodwind instruments, first Violins and Double Bass faded out in contrast to the Violas and the Violoncellos.

In another example, cluster 4 instances higher pitches on instruments such as *Piccollo* and First Violin. As next, the lower pitches is played by Double Bass and they are exemplified once again by cluster 0. Figure 4 exhibits the transition between cluster 4 and cluster 0 on the musical score.

6. CONCLUSION

Although the questions underlying the main goal of this ongoing study rely on information provided by the composer and the literature on music analysis and musicology, the application of the suggested approach to a wider context should also be viable. The computer-assisted approach was used as basis for discussing the metaphor that inspired this particular piece but has the potential to be extended to other compositions in the repertoire. In Atmosphères, in particular, Ligeti translates electronic compositional techniques with surprising fidelity into the acoustic realm. The method proposed on the study was able to highlight the main strategies envisioned by the composer. Moreover, the segmentation layers retrieved by the novelty curve provided substantial input to spread out the musicological analysis from both acoustic and symbolic domain. Future work projections focus on the elaboration of further analytical examples, more detailed discussions of the methodological issues that may arise from the extension of the method as well as a harmonization of the computational tools involved.

Acknowledgments

The authors would like to thank the funding agency FAPESP (2015/25160-6).

7. REFERENCES

[1] V. Der Slice and J. Dewitt, "An analysis of gyorgy ligeti's" atmospheres"," Ph.D. dissertation, University of Illinois at Urbana-Champaign, 1980.

[2] J. J. Iverson, *Historical memory and György Ligeti's sound-mass music 1958-1968*. The University of Texas at Austin, 2009.

[3] J. Iverson, "The emergence of timbre: Ligeti's synthesis of electronic and acoustic music in atmosphères," *Twentieth-century music*, vol. 7, no. 1, pp. 61–89, 2010.

[4] S. Davachi, "Aesthetic appropriation of electronic sound transformations in ligeti's atmosphères," *Musicological Explorations*, vol. 12, pp. 109–110, 2011.

[5] S. Mayville, "The compositional techniques and influences behind ligeti's' atmosphères," *Intermezzo*, vol. 1, pp. 30–38, 2015.

[6] N. Cook, *A guide to musical analysis*. Oxford University Press, 1994.

[7] B. L. Sturm, "The state of the art ten years after a state of the art: Future research in music information retrieval," *Journal of New Music Research*, vol. 43, no. 2, pp. 147–172, 2014.

[8] B. R. Levy, *Metamorphosis in Music: The Compositions of György Ligeti in the 1950s and 60s*. Oxford University Press, 2017.

[9] T. Richard, *György Ligeti*. Phaidon Press; First Edition, 1999.

[10] D. Cope, *Techniques of the contemporary composer*. {Wadsworth Publishing}, 1997.

[11] B. R. Levy, "The electronic works of györgy ligeti and their influence on his later style," Ph.D. dissertation, University of Maryland, 2006.

[12] E. Pampalk, "Computational models of music similarity and their application in music information retrieval," Ph.D. dissertation, Vienna University of Technology, Vienna, Austria, March 2006. [Online]. Available: http://www.ofai.at/~elias.pampalk/publications/pampalk06thesis.pdf

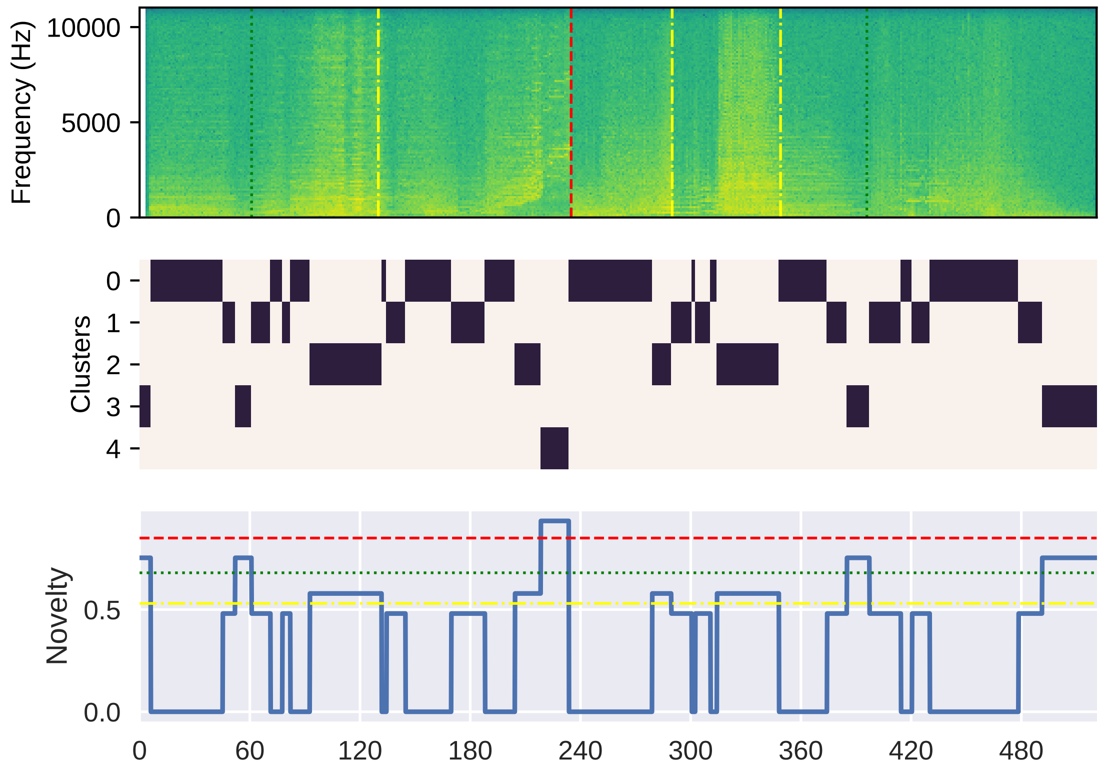


Figure 2. Top: the spectrum extracted from the piece recording; Midlle: the clusters corresponding to each frame of the piece; Bottom: the novelty curve proportionally converse to the cluster population. Three consecutive levels of segmentation are suggested (red, green and yellow) for three possibilities of resolution.

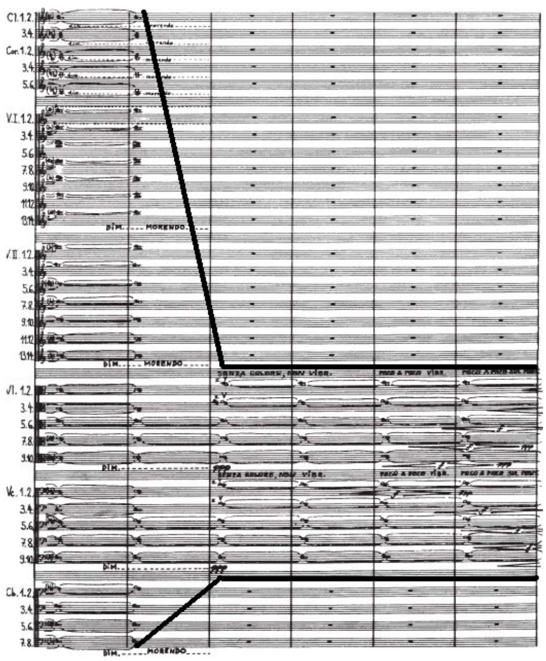


Figure 3. Filtering Process on the First Passages of Atmosphères.

Figure 4. Cluster transition from 4 to 0 on Atmosphères.

BETWEEN SOUND AND MOVEMENT IN EMBODIED TANGO PERFORMANCE. IMPLICATIONS FOR THE STUDY OF THE PERFOMATIVE STYLE IN TANGO

Demian Alimenti Bel
Comisión de investigaciones científicas de la provincia de Buenos Aires (CICba)
Laboratorio para el Estudio de la Experiencia Musical (FBA-UNLP)
demianalimentibel@gmail.com

ABSTRACT

In a recent work we studied differences that characterize styles of A. Troilo's and O. Pugliese's tango orchestras by analyzed the relationships between phenomenal accents and grouping units of the sonic melodic-rhythmic patterns. In the present work, we continue the stylistic analysis focusing on the expressiveness of the tango performer's body movements. The aim is to observe and describe expressive gestures in the movements of two *bandoneon* performers, interpreting a same tango according to Troilo and Pugliese styles. We analyzed the quality of body movements and effector gestures (Laban, 1971; Godoy and Leman, 2010) in the performance of rhythmic-melodic patterns. As to Troilo, the accentual and gestural conjunction is generally coincident with the metric scheme, and also with the grouping units. As to Pugliese, the melodic-rhythmic patterns are not performed in phase with the metric structure: the temporal distribution occurs at different levels of the metric hierarchy.

1. BACKGROUND

Traditional musicology literature approaches raised the idea of separation between composition, technique and expression in the study of music performance and in the instrumental music education. This theoretical framework left an empty epistemic in the construction of holistic categories that account for the sonic-kinetic-expressive complex in music. In another direction, the psychology of music has tried to study the musical experience from multiple approaches. In this field, some articles have investigated the musical expression (the components of timing and dynamics in instrumental performance) related to the communication of information about the musical structure [5, 13, 14]. Antecedents of the study of movement, approach to academic music, described the way structural features (changes of tempo and chronometric density) shape expressive movement during piano performance [3, 4]. So far, the study of expressive musical movement and sound has been applied more to academic music than to popular music.

1.1 Correlation between Bodily Actions, Effort-Shape and Musical Gestures in the Movement of the Bandoneon Performer.

In another way, the Laban's theory (1961) proposes and explains an analysis of simple corporal actions that can be easily observed and described in the movement of dancers, musicians, mimes and actors. In the present work, one of the simple corporal actions that interests us in the analysis of movement in the tango performance is manifested through the placement of accents and the organisation of phrases. The different combination and sequences of these simple corporal actions are composed of intentional characteristics of the person's action. In turn, each sequence can be expressed by a multiplicity of intentional actions that arise from the particular mixture of the qualities of effort (*correlation process*). It is necessary to clarify that each action produces alterations of the position of the body consuming a certain time (effort-time), it requires a specific amount of muscular energy (effort-weight), and presents a specific fluidity (effort-flow).

On the one hand, Maes and others (2014) in a recent study of musical reception in musicians with and without training, relate the Effort/Shape model that originated in Laban Movement Analysis (LMA) method with specific linguistic descriptions that are metaphorised in the participants. For example: for a sustained time with a high impulsiveness and shaking, linguistic levels such as fast, nervous, energetic and active are presented. On the other hand, Giraud and others (2016) computerised the qualities of the movement inspired by Laban's Movement Analysis (1961) in the movement expressivity during a fitness task. These authors explain that the movement can be defined from the variations of the segments of the body in space and time characterised by parameters kinematics. Following this analysis, they designed five qualities of effort-shape through a system of motion capture.

In the frame of embodied music cognition, Godoy and Leman (2010) categorised the musical gestures according to their specific function; these categories are based on compilations of books, lectures and articles developed by different authors from the same field. Particularly, they defined gestures to produce sound (*effector gestures*) whose sole purpose is to produce or modify the sound and can be subdivided into gestures of excitation and modification. Similarly, conceptualized *communicative* *gestures* (they are mainly made to communicate), *gestures to facilitate the sound* (they support several senses the effector gestures) and *gestures to accompany the sound* (they are not involved in the production of the sound, but they are produced in response to this). In the present work, the effectors gestures is the category that we use to observe and describe the movement of the bandoneon performers. The main aim is to discuss this perspective of the specific technical instrumental movement related to *low level*, that is involved in musical expression on the style of execution in tango.

For the classical manuals of the technical of the bandoneon in tango, it is suggested that the instrument should be supported on both thighs and unnecessary movements of other parts of the body should be avoided to have a better control of the sound and the fluency musical expressive. This statement has to do especially by the relation of the weight and the movement dynamics of it. But in a merely descriptive observation of tango orchestras we can affirm that the bandoneon performers show a different intentional-expressive repertoire of movement according to the particular modes of producing musical sense. Therefore, it is stated that the works surveyed here can be used as background to characterise the features of the expressive intention that is put into action in the performance of tango.

2. HYPOTHESIS

The characteristics of the expressive movement of the *bandoneonistas* during musical performance, as part of the sonic-kinetic complex of tango, correlate with the relationship between the temporal unfolding of phenomenal accents [11] and the grouping units of the sonic melodic-rhythmic-expressive patterns that were studied in a previous work [2]. Therefore, the communication of the interpretive style is related to the expressive movements (such as physical descriptions and quality of movement in the performance) that the musicians put into evidence in the practice of tango.

3. AIMS

In this study of two cases it is proposed: (i) to observe and describe the expressive gestures in the movements of two bandoneon players, performing the same tango according to the styles of Aníbal Troilo and Osvaldo Pugliese; (ii) to link the analysis of (i) with the musical discursivity manifested in the rhythmic-melodic-expressive patterns [2] and the temporal unfolding of accents in the melodic contour; and (iii) to characterise the stylistic identity based on the multimodal keys of the execution in both styles of tango.

4. METHOD

In the present work we propose a qualitative methodology to study and analysis videos recordings of historical performances in tango. This methodology pretend to address the investigation of movement in popular music starting from a perspective of descriptions of second person and exterior space of the subject [9] and no intend to measure components of movement and sound. But quantitate perspective can reinforce the initial observation and analysis of the aspects of the expressive movements in tango.

4.1 Experiment

4.1.1 Stimuli

Fragments have been selected from two video recordings of historical performances of the tango from the 20´s "Chiqué" (original Lyric and Music by R. Brignolo) performed and arranged by Aníbal Troilo's quartet and by the Osvaldo Pugliese's orchestra.

The two selected vintage videos belong, one to Aníbal Troilo's quartet recorded in 1971 in the Argentine program *"Tango Saturdays"*, and the other to the Osvaldo Pugliese's orchestra recorded in 1989 in the *"Musical Tour"* they made in Netherlands.

4.1.2 Appliances

The observation of videos and the annotation of the movement were made in Elan 5.0.0-alpha (2016) software. The software makes it possible to introduce the video in a standard format of mp4, avi and others with its respective sound signal in WAV and allows reducing the sampling speed of the video without audio for its exact annotation and segmentation of the movement.

4.1.3 Procedure

The analysis of the movement focused on the execution of the bandoneon, performed by Aníbal Troilo himself, and by Roberto Álvarez (Osvaldo Pugliese's orchestra). The analytical procedure included: (1) observation and annotation of bodily-intentional movement using Laban categories (1961) to link the descriptions of simple corporal actions and the movement quality (Time-Effort and Flow-Effort); and (2) observation and characterization of the effector gestures produced by each interpreter related to the movement of the instrument (eg: opening and closing of the air column/*fueye*).

The next table (table 1) samples the analytical correlation between: i) effector gestures and simples movements (*physical descriptions*); ii) simple corporal actions and Time-Effort (*movement quality*); and iii) temporal musical structure, phenomenal accents and grouping units of the rhythmic-melodic patterns (*units of expressive-musical sense*).

<table>
<tr><th colspan="3">PHYSICAL DESCRIPTION OF THE MOVEMENT</th></tr>
<tr><td>Head Movement
Describe head movements in two simultaneous directions
up-back down-forward</td><td>Torso Movement
Describe torso movements in one direction
back forward center</td><td>"Fueye" Movements
They describe movements of the instrument's air column, touching open or closed.
Describe effector movements in the instrument (eg: transition of opening of the bellows, recovery of air, etc.)</td></tr>
<tr><th colspan="3">MOVEMENT QUALITY</th></tr>
<tr><td colspan="2">Simple corporal actions
Velocity
the speed of movement is recorded as slow, normal or fast (Laban, 1971)
Accent in the movement
towards tension: marked accent or light accent
towards relaxation: rigorous or weak accent
=fast accent =light accent =marked
= normal accent =weak accent =rigorous
=slow</td><td>Effort categories
Time Sudden Sustained
Flow Restricted Free
Movement characteristics
Impulsiveness High Low
Smoothness Low High
Shaking High Low</td></tr>
<tr><th colspan="3">UNITS OF EXPRESSIVE-MUSICAL SENSE</th></tr>
<tr><td colspan="3">Categories generated to link the description of the physical movement and the quality of the movement with the musical structure. In each annotation it may or may not be linked to any of the three descriptions below
Metric marks/Tactus
In this type of movements, the markings are configured with some part of the body (head, arms and torso) of the tactus as of other metric levels
Grouping units
expressive movements that describe the relationship between phenomenal accents and grouping units (Lerdhal and Jackendoff, 1983)
Patterns rhythmic-melodic-expressives
expressive movement linked to the technique of variation of the rhythmic-melodic, articulatory components of the instrumentation and the control of the general progress of the tonal deployment</td></tr>
</table>

Table 1. The table shows the three categories used to describe and record the movement of the musicians in the selected fragments (*physical description of the movement, quality of movement and expressive-musical sense units*). Below each category are descriptions of their respective annotation values.

For the analysis of the textual surface, it is proceeded to: (3) Analyse the musical surface in order to describe the links between the recurrent movement patterns and the musical structure.

Particularly, the musicology analysis consisted of: listen and record melodic passages executed by the right hand of the *bandoneonists* that contained *staccato-accent-local legato* articulations (rhythmic passages), and presented rhythmic-melodic patterns. In this work the transcription was made using all available signs to account for the writing of the components of height, rhythm and articulation in the melodic plane. The notation of accents were made within the convention of the specific musical annotation in tango, in this manner usually appear in originals arrangements for typical orchestra. These were combined with annotation of the harmonic rhythm and passages of referential links to other instruments (see figures 2 and 3). Other symbols were also used as sounds with air in the bandoneons that emulate the drag of the string. The criterion of musical segmentation for the analysis presented a problem in the first instance. In the case of A. Troilo to be equal the amount of samples of videos that were identified in the execution of R. Alvarez had to take as segmentation all the A section (17 bars), while in Álvarez the first semifrase of A was selected which contains 11 measures. This was due to the fact that the repertoire of movements performed by Álvarez varies a lot in short passages. Particularity this kind of segmentation was key for the accomplishment of the conclusion of the comparative movement between both musicians.
Then it is proceeded to: (4) analyse the rhythmic-melodic surface of these passages, describe their structural features, establish the grouping units [11] and the phenomenal accents, and finally note in the score the placement of accents in the movement/phrase and the speed with which they are used (see symbols of Laban's annotation in table 1)

Finally, an analytical interpretation was elaborated based on the connections between (1), (2), (3) and (4).

Figure 2. Section A of the tango *"Chiqué"* arranged by Aníbal Troilo's quartet. This score is a transcription of the superior voice of the bandoneon (right hand) performed by A. Troilo, with the clarification of the structural harmony used and some referential passages of other instruments.

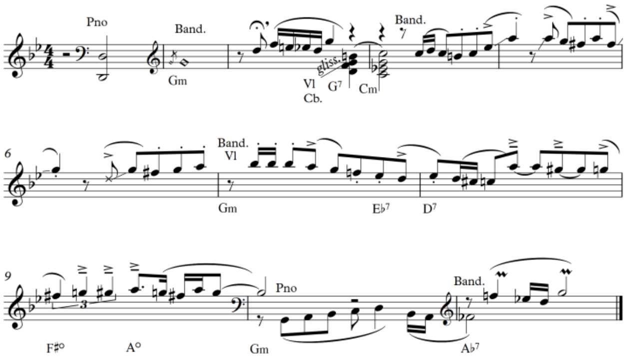


Figure 2. Section A of the tango *"Chiqué"* arranged by Osvaldo Pugliese's orchestra. This score is a transcription of the superior voice of the bandoneon (right hand) performed by R. Álvarez, with the clarification of the structural harmony used and some referential passages of other instruments.

5. RESULTS

Figure 3 shows the result for the first three grouping units in Troilo's version. It's observed in the grouping unit 1 that the initial corporal position start with the torso in the center and with the head slightly inclined forward and down (plays the group of eighth note and four sixteenth notes). The first placement of gesture accent it's directed toward the G note (red circle, quarter note), through the movement of the torso backwards and the head up and back. Continually this movement sequence it's repeated in the last two attacks of the grouping, but here we interpreted that the descent movement of torso and head accents the first triad (red circle, see figure 3). All the accentuations in the movement are executed with accents towards tension combine with a normal and fast velocity, respectively. The movement quality here its define as impulsiveness-shaking sudden, this particularly organization of the phrase demonstrate a recurrent expressive movement (given by the alternation of velocities in movement) that as well it's observed in the analysis of the following groupings.

In the grouping unit 2 it's observed that torso it keeps in the center, but the head repeat the swing down and up as it happened in the analysis of grouping unit 1. The accents are in two notes (see figure 3, red circles) product of head movement up, the first appears in the upbeat and the other in the downbeat that close the grouping. In the grouping unit 3 it's observed the same order of the movement of the head that we explains in grouping unit 2, but here its accents happens in the downbeats (see figure 3, red circles). So the movement quality its define for both groupings as a sustained impulsiveness with a sudden shaking in the close. The sustained impulsiveness is caused by slight head movement that is reiterated while the sudden shaking is produce by a net acceleration of the head movement up in the close of groupings. In the three grouping units the common characteristic is the type of flow. This is a restricted flow (the flow is contained), we suppose it's the result of the *staccato-accent-local legato* articulation, in subsequent analyses its will try to prove this interpretation.

Figure 4 shows the result for the grouping units 4, 5, 6 and 7. Here we focus on the movement of instrument. All the openings and closings of the air column (*fueye*) include moderate and little movements. In general it's observed a gradual transition of opening of the *fueye* covering each grouping units. Between groupings happens an air recovery by the *fueye*'s closing. That is, each grouping is executed with the same air flow. As for the analysis of the movement of torso and head it's the same observation as in grouping units 2 and 3, but here the up head movement is littler than in the others. This particular movement of the head is the result of a normal velocity employed in placement of accents in the phrase. That velocity is sustained along the groupings. Again the accents in the movement alternate between downbeat and upbeat. Nevertheless there is a predominance of recurrence accents in the downbeat, so the up head movement is more pronounced in the beat. Finally the movement quality is a sustained impulsiveness with sudden shaking in the close.

In figure 5 appear a global *legato* phrase and in the end of the A section resumes the *staccato-accent-local legato* articulation. So here the analysis presents new observations about the physical description of the movement and the movement quality characteristics. In the grouping units 8 and 9 the torso keeps in the center and its observed a gradual slight rolling of the head backward the close of grouping 9 (see the two red circles in measure 3, figure 5). The detail here is that this two last rhythmic cell are execute with the *fueye* closing, so it is produce a change in the instrument's timbre. All the accents annotated below the transcriptions are accents towards relaxation; this is the result of the employed slow velocity and the global *legato* articulation, which not produces movement of tension. The movement quality here its define as a sustained smoothness. The smoothness level is low because is observed in general a little head movement along the grouping units; is predominance stillness expressive (the expressive movement, perhaps, focuses on the face). Both groupings converge with a free flow; this observation is sustained for temporal fluctuation of the melodic expressive timing (the corporal attitude indicates that the movement prioritizes a concentrated relaxation). In grouping unit 10 the return of the *staccato-accent-local legato* articulation have a correlated with the expressive movement organization analysed in other grouping units. So the movement quality characteristics present a sustained impulsiveness, a sudden shaking in closing grouping and restricted flow. The accent is in the last note product of head shaking movement backward-up that conclude the A section.

In grouping unit 1 (see figure 6) it's observed a movement without sound as preparatory movement. The head and shoulders go up and return to the previous position to start playing. The employed effort is a sudden impulsiveness in rigorous accent toward relaxation. Then the grouping is executed with a quiet corporeal position and accents toward relaxation. These shoulders movements accentuate the first and the last note of the grouping unit (see red circles in figure 6). For it part grouping unit 2 start with the head movement goes forward and down that accentuated two consecutive eighth notes (see red circles in figure 6). The movement quality is a sustained impulsiveness in light accent towards tension (the movement velocity is slow). The end of the grouping show a shoulder movement that go up coincide with ascendant glissando between the E note and the A note. The grouping unit 3 repeats the same movement organization as in grouping unit 2 analysis. But here appear more accents at the level of the eighth notes (see red circles in figure 6). The close of the grouping is performed with the same shoulders movement as in grouping unit 2. Finally the grouping unit 4 started again with the head goes forward and down that accentuated the beat and no the eight notes as in the previous cases. The movement quality its define

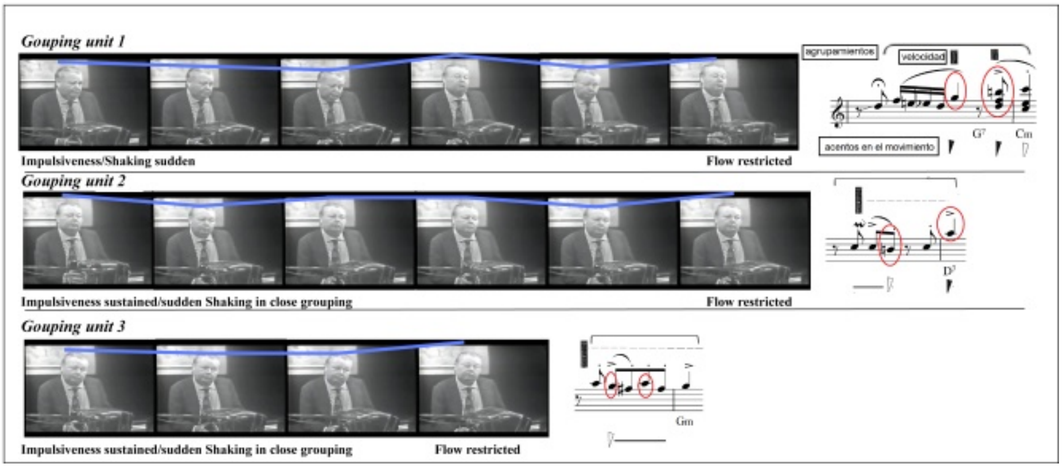


Figure 3. The grouping units 1, 2 and 3 of the tango "*Chiqué*" from A. Troilo's version. In the left side is the photogram with the frame a frame of the video image, the line blue shows the different movements of the head forward and back-down. Below each photogram is the interpretation of the movement quality (impulsiveness, shaking, etc.). On the right side is the transcription of each grouping unit with Laban's symbols notation. The movement notation represents the different types of accents in the movement and its respective velocity. The red circles indicate the organization and placement of movement accents in the phrase

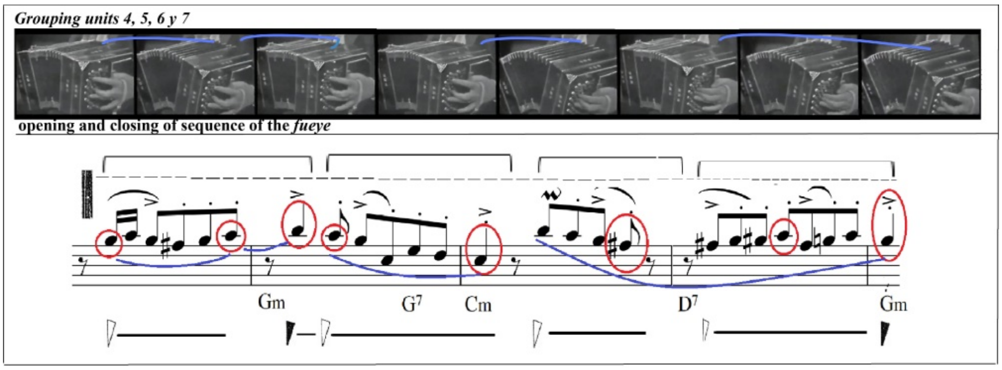


Figure 4. The grouping units 4, 5 and 6 of the tango "*Chiqué*" from A. Troilo's version. The photogram above is the result of the movement of the instrument, the sequence of opening and closing of the *fueye* (air column). Below is the transcription of each grouping unit with Laban's symbols notation. The movement notation represents the different types of accents in the movement and its respective velocity. The red circles indicate the organization and placement of movement accents in the phrase.

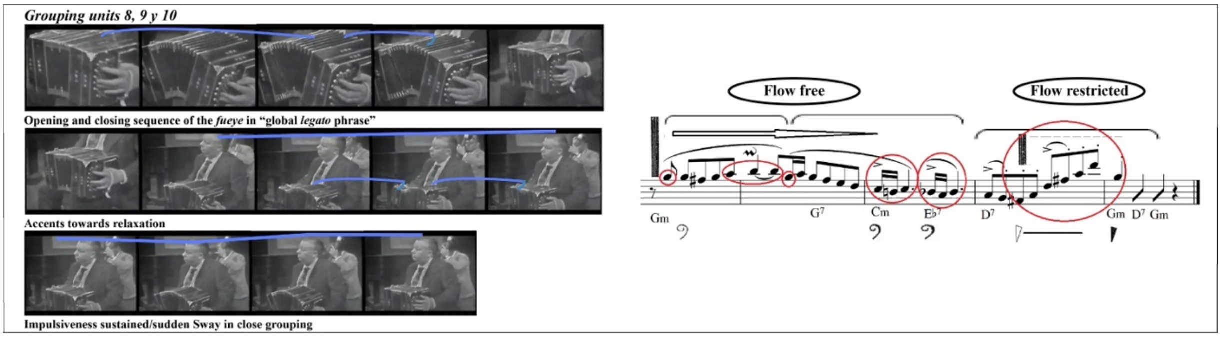


Figure 5. The grouping units 8, 9 and 10 of the tango "*Chiqué*" from A. Troilo's version. In the left side is the photogram with the frame a frame of the video image, the line blue shows the different movements of the head forward and down. Below each photogram is the interpretation of the movement quality (impulsiveness, shaking, etc.). On the right side is the transcription of each grouping unit with Laban's symbols notation. The movement notation represents the different types of accents in the movement and its respective velocity. The red circles indicate the organization and placement of movement accents in the phrase.

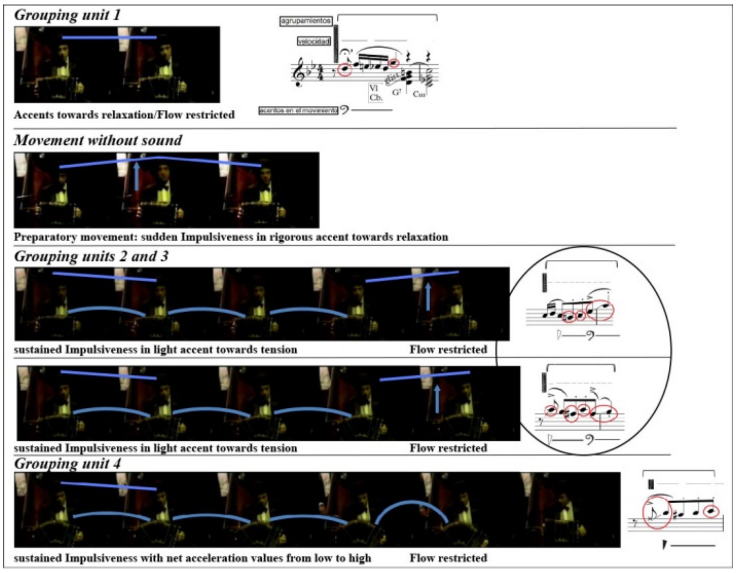


Figure 6. The grouping units 1, 2, 3 and 4 of the tango "*Chiqué*" from O. Pugliese's version. In the left side is the photogram with the frame a frame of the video image, the line blue shows the different movements of the head forward and backward, and undulate blue line display the instrument's opening and closing. Below each photogram is the interpretation of the movement quality (impulsiveness, shaking, etc.). On the right side is the transcription of each grouping unit with Laban's symbols notation. The movement notation represents the different types of accents in the movement and its respective velocity. The red circles indicate the organization and placement of movement accents in the phrase.

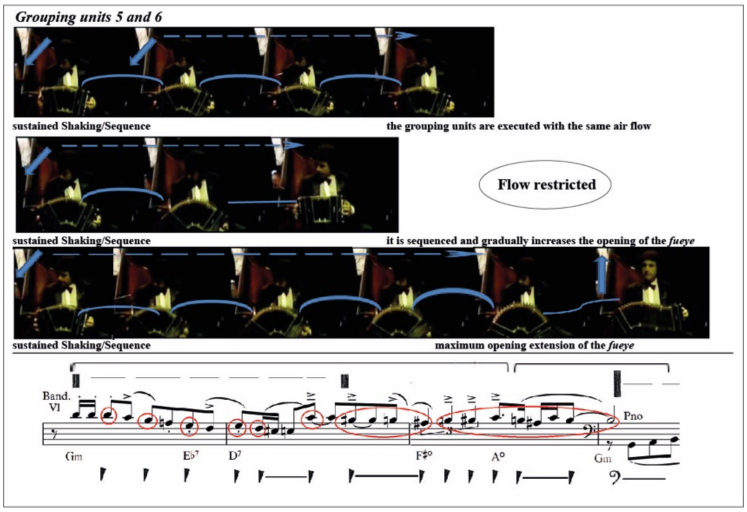


Figure 7. The grouping units 5 and 6 of the tango "*Chiqué*" from O. Pugliese's version. On the top is the photogram with the frame a frame of the video image, the line blue shows the different movements of the head forward and backward, and undulate blue line display the instrument's opening and closing. Below each photogram is the interpretation of the movement quality (impulsiveness, shaking, etc.). On the bottom is the transcription of each grouping unit with Laban's symbols notation. The movement notation represents the different types of accents in the movement and its respective velocity. The red circles indicate the organization and placement of movement accents in the phrase.

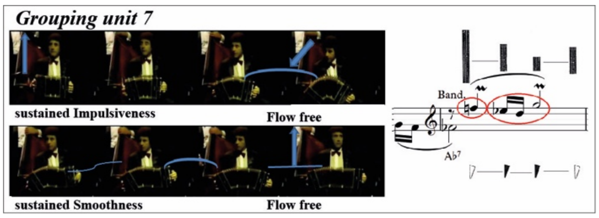


Figure 8. The grouping unit 7 of the tango "*Chiqué*" from O. Pugliese's version. In the left side is the photogram with the frame a frame of the video image, the line blue shows the different movements of the head forward and backward, and undulate blue line display the instrument's opening and closing. Below each photogram is the interpretation of the movement quality (impulsiveness, shaking, etc.). On the right side is the transcription of each grouping unit with Laban's symbols notation. The movement notation represents the different types of accents in the movement and its respective velocity. The red circles indicate the organization and placement of movement accents in the phrase.

as sustained impulsiveness whit net acceleration values from low to high. This characteristic is observed in the gradual opening of the *fueye* (see figure 6). The torso movement is irrelevant for the analysis here, most of the time stayed in the center. In all grouping units it's employed a restricted flow, as in the same way as A. Troilo's version.

The grouping units 5 and 6 (see figure 7) present the head movement goes down and forward that accentuated, at the beginning, the beat position (blue dotted line in figure 7). All accents are marked accents toward tension. This movement repeat is accompanied by the torso forward and with a fast velocity. In measure 2 and time 3 with the change of the phrasing melodic, the movement is modified in relation to this readjustment phraseological. All the accents coincide with the notes that have a kind of rhythmic syncopation. Both grouping units are executed with the same air flow and the opening of the *fueye* it's sequenced and gradually increases (achieves the maximum opening extension of the *fueye* in the last four notes, such a climax). The movement quality its define as sustained shaking sequence and with a restricted flow. All the values are high that would lead us to analyse other parts of the body.

In figure 8 appear a global legato phrase in the *bandoneonistic* performing that expand the section A to the repeat of the melody by violins string. At the beginning it's observed a preparatory movement that anticipates the execution of grouping unit 7. The head movement up and backwards is accompanied by the torso; this movement employed a slow velocity. Then when is started the first attack (whit an ornamentation) the head movement goes down and forwards, and the *fueye* opening achieves the most extension in A section (see figure 8). The movement velocity moves slow toward a fast velocity. The movement quality its define as sustained impulsiveness with a free flow. Finally the last rhythmic cell combines head movements in several directions. First goes down, forward and to the right side, then return up, backward and to the left side and finish in the center. The *fueye* present another opening form (plane aperture, see figure 8). The time-effort employed is interpreted as a sustained smoothness with a free flow. So these movements of the head, torso and the instrument seem a complementation of the initial movements. The conclusion here is that some expressive movements correspond to sound production and others accompany the phraseological and temporal modification.

6. DISCUSIONS

As to Troilo, it was observed that the movement of head, torso (sustained impulsiveness and sudden shaking) and the instrument are communicating features of the melodic phrasing that correspond to particularities of temporal and rhythmic organization of the phrase. The intentional corporal action and sound organization of the phenomenal accents are coincident within musical phrases. The placement of accents and the organisation of phrases demonstrate a recurrent expressive movement, particularly in the employ of an impulsiveness effort in accents in the upbeat and shaking sudden in the downbeat that close the grouping (with a ascendant melodic resolution). This accentual and gestural conjunction is generally coincident with the metric scheme, and also with the grouping units.

As to Pugliese (by Álvarez) it was observed that the expressive gesture of head and torso is modified by the temporal-melodic alternation between *rubato* and the regular *tactus*. When the head movement goes down and forward and it's co-articulate with the regular *tactus*, is related to the sound of *yumba*, a rhythmic pattern of the stratum of accompaniment [1]. In the grouping units was found that instrumental movement has performed a *fueye* opening that has a progressive and gradual augment in each motivic repetition. Usually achieve the most extension opening in the phrase final (such a melodic climax). However, the phenomenal accents and the expressive sound patterns are not performed in phase with the metric structure: the temporal distribution occurs at different levels of the metric hierarchy in the same phrase.

Its define on the one hand, that in Troilo, the in-phase metrical accents are used to elaborate the durational rhythm and the melodic variation at *the local phrase level*. On the other hand, in Pugliese, discursivity is elaborated based on out-of-phase expressive accents, resolved at *the global level of the grouping units*. These results coincided with the analysis of the compositional and performative components that was studied in a previous work [2]. Intrinsically the expressive movement style and the variation of rhythmic-melodic patterns contain invariant features; this means that there is a way of varying, which is more identified with one of the modes of production of musical sense rather than with the other. It's could say that there is a movement style *troiliano* and another *puglesiano*.

Finally similar expressive movements were observed in both musicians. The head movement down and forward, and up and backward are common sound production gestures and accompany the phraseological and temporal modification in the *bandoneonistic* tango performance. So these characteristic gestures are a part of the movement repertory of the musical expressive that the musicians put into evidence in the practice of tango. For it part the Flow-Effort presents a restricted flow as a result of the *staccato-accent-local legato* articulation, and a free flow for global *legato* passages. This definition has implications for the instrumental pedagogy of bandoneon in the tango execution.

To conclude the discussion it may say that the music performance provides insights for the observer to construct schemes, types or cognitive representations of the motor action displayed by the musicians. That as seen in the present study the representation is in general associated with exclusively sound cognitive material. This means, perhaps, that the music style generates in the audience cognitive representations that allow it to recognize the variations of the compositional and performative components. So the sound and visual information would be intertwined by the communication of the motor action made by the stylistic execution of the musicians. Nevertheless the observations of the movement in tango performance review the category of effector gestures [7, 10]. First it's complicate to separate the gesture in micro analytical categories such communicative gestures, gestures to facilitate the sound, gestures to accompany the sound or effector gestures. There is no an effector gesture that could be observed, there are a lot of movement that are related in simultaneous, and many of this gesture refers to an expressive manner of musical execution. For example the movement of head as alternation down, up, backward and forward are no effector gesture, is not related to the instrumental technique of the motor action or to a communicative gesture but they are mixed together to express the specific sound production. So the body movement and the sonic form combine to shape a multimodal-stylistic communicative complex in tango performance. In addition this definition of the gesture itself introduces both a material and conceptual dimension, by explicitly conceiving that motor behavior in a musical situation is a cornerstone of the mediation between mind and physical environment, leave the Cartesian dichotomy intact, limiting itself to introducing the action corporal as a mere intermediary of this opposition. Although we do not intend to provide a unique definition of *gesture*, we believe that it is useful to compare the various definitions and point out the specificity of this topic in the musical domain, especially in a concrete communication fields as is in popular music of Latin-American. These studies try to contribute definitions and debates to the embodiment frameworks and others theories of investigation of movement and sound in musical performance.

7. CONCLUSIONS

In this work we analyse the quality of body movements [8] and effector gestures [7, 10] in the execution of the rhythmic-melodic-expressive patterns of two tango *bandoneonists* in order to describe the sound-kinetic complex resulting in each performative style. Subsequent studies should incorporate analysis with mediation technologies (quantitative analysis) and in videos of the current typical tango orchestras, considered some "*troilianas*" and others "*puglesianas*". The analysis of the link between musical gestures, as carriers of meaning based on an action-oriented ontology, and the musical structure, provide a more holistic explanation of the meaning of the style construct and the modes of production of musical sense in the tango.

8. REFERENCES

[1] Alimenti Bel, D., Martínez, I.C. y Ordás, M.A. (2014b). Los mareados por la orquesta de Osvaldo Pugliese: La dinámica y la temporalidad de la yumba entendida a partir de los rasgos performativos emergentes del arreglo. En M. Escobar y L. Gasparoni (comps.). *Música Latinoamericana: Tradición e Innovación.* Congreso llevado a cabo en las III Jornadas de la Escuela de Música de la U.N.R, Rosario, Argentina.

[2] Alimenti Bel, D. y Martínez, I. C. (2018). Attributes of variation as indicators of stability in tango. Stylistic patterns in musical pieces performed by Aníbal Troilo and Osvaldo Pugliese. *Epistemus*, 5, 27-56.

[3] Davidson, J. W. (2007). Qualitative insights into the use of expressive body movement in solo piano performance: A case study approach. *Psychology of Music*, 35, 381-401.

[4] Davidson, J. W.(2012).Bodily movement and facial actions in expressive musical performance by solo and duo instrumentalists: two distinctive case studies. *Psycholgy Music* 40,595-606.

[5] Gabrielsson, A. (1999). The performance of music. *Psychology of Music*, 2, 501-602.

[6] Giraud, T et al. (2016) Impact of elicited mood on movement expressivity during a fitness task. *Human Movement Science*, 49, 9–26.

[7] Godøy, R. I. and Leman, M. (2010). (eds.), *Musical Gestures: Sound, Movement, and Meaning*. New York: Routledge.

[8] Laban, R. (1960). *The Art of Movement and Dance*. London: Macdonald and Evans.

[9] Leman, M. (2008). *Embodied Music Cognition and Mediation Technology*. Cambridge: the MIT Press.

[10] Leman, M. (2010). Music, gesture, and the formation of embodied meaning. En R. I. Godøy y M. Leman (eds.), *Musical Gestures: Sound, Movement, and Meaning*. New York.

[11] Lerdhal, F. and Jackendoff, R. (1983). *Teoría generativa de la música tonal.* Juan González-Castelao (trad.). The Massachusetts Institute of Technology.

[12] Maes, P-J and Van Dyck, E. (2014). The coupling of action and perception in musical meaning formation. *Music Perception*, 32, issue 1, 67–84.

[13] Repp, B. H. (1997c). Variability in timing in expressive piano performance increases with interval duration. *Psychon*. Bull. Rev. 4, 530–534. doi 10.3758/BF03214344.

[14] Repp, B. H. (1999). A microcosm of musical expression. II. Quantitative analysis of pianists' dynamics in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America*, 105 (3), 1972-1988.

EVALUATION OF MUSIC COGNITION IN CHILDREN AND ADOLESCENTS WITH ATTENTION-DEFICIT/ HYPERACTIVITY DISORDER

Marília Nunes-Silva
Universidade do Estado de Minas Gerais (ESMU/UEMG)
marilianunespsi@gmail.com

Aline Aparecida Rocha Caetano
Universidade do Estado de Minas Gerais (Divinópolis/ UEMG)
alinerocha-c@hotmail.com

Sara Oliveira Alves
Universidade do Estado de Minas Gerais (ESMU/ UEMG)
alvarezsaraa@gmail.com

Gabriel Telles de Mello e Silva
Universidade Federal de Minas Gerais (UFMG)
tellesmus@gmail.com

ABSTRACT

Attention Deficit Hyperactivity Disorder (ADHD) has high prevalence among children and adolescents and is characterized by attention deficits, hyperactivity, and impulsiveness. ADHD mainly compromises attention and executive functions and is associated with difficulties in phonological awareness. Considering the relationship between musical performance and the development of compromised cognitive functions in ADHD, this study aimed to evaluate the musical neuropsychological profile of individuals with ADHD and to investigate the presence of musical processing deficits. Twenty Brazilian individuals aged 7 to 13 years, were divided into two groups of 10 participants: (1) a clinical group composed by individuals diagnosed with ADHD; and (2) a control group, consisting of individuals with typical development. We assessed participants from intelligence, language and music perception tasks. The results showed that individuals with ADHD had a significantly lower performance in the intelligence test, Melodic and Meter tests, and language tasks. There were no significant differences between the groups in the performance on Rhythm and Musical Memory tests. The difficulties found in global musical temporal processing (Meter) and melodic discrimination in the clinical group could emerge as secondary characteristics to the primary deficits of attention and working memory, or present as a comorbidity of ADHD (e.g. amusias).

1. INTRODUCTION

Attention Deficit Hyperactivity Disorder (ADHD) has high prevalence among children and adolescents and is characterized mainly by attention deficits, hyperactivity, and impulsiveness [1, 2]. ADHD mainly compromises attention and executive functions (working memory, inhibitory response and self-regulation) [3, 4] and is associated with poorer performance in time perception [5, 6] and phonological skills tasks [6].

According to Barkley [1], ADHD presents as a primary deficit the lack of behavioral inhibition. This deficit could impair four executive functions: (1) working memory, (2) self-regulation of affect-motivation-arousal, (3) internalization of speech, and (d) behavioral analysis and synthesis (reconstitution). ADHD is a heterogeneous and multifactorial disorder, whose etiology has not yet been clearly identified, although evidence points to associated genetic and neurobiological aspects, mainly related to dysfunction of the fronto-subcortical pathways and imbalances in dopaminergic and noradrenergic systems [2, 7]. Individuals with ADHD often present other associated disorders. These comorbidities mainly include oppositional defiant disorder, conduct disorder, mood disorder (unipolar and bipolar), anxiety disorders and learning disorders [2, 7].

In addition to sustained attention deficits, some of the most frequent behavioral signs in ADHD diagnosis are disorganization, academic, auditory divided attention and hearing in noisy environment difficulties [8]. Studies suggest that children with ADHD present poor performance in auditory temporal tests with difficulties in perceiving frequency pattern and sound duration [6, 8, 9].

Regarding the assessment of musical memory and perception abilities in individuals with ADHD, fewer studies have been reported. Jones, Zalewski, Brewer, Lucker, and Drayna [10] have found a higher prevalence of ADHD in individuals with congenital amusia, indicating that there may be comorbidities between ADHD and musical perception deficits. Congenital amusia, also known as tone-deafness, is a developmental disorder characterized by a lifelong inability to musical processing, despite normal intelligence, memory and language abilities [11].

In a comparison group study, Carrer [12] evaluated the performance on sound and musical tasks of ADHD children with and without the use of stimulant medication, compared to typically developed children. He observed a poorer performance of the ADHD groups in temporal estimation of simple sounds tasks compared to the typically developed children group. Nevertheless, he argues that the positive average performance on most tasks observed in all groups could indicate that music modulates the symptoms of inattention in ADHD. Some studies have shown that music can be used to mediate attention and excitement in children with ADHD, and the presence of background music can reduce distraction and increase concentration on arithmetic school tasks [13]. In addition, music education is associated with better phonological awareness [14], and can contribute to the rehabilitation of children and adolescents with ADHD associated with dyslexia and other language disorders [7, 15].

There is evidence that musical training can enhance verbal, visuospatial, mathematical and reading skills [14, 16, 17, 18, 19]. Because of this association with the development of other cognitive functions, according to Särkämö, Tervaniemi and Huotilainen [20], music has been used clinically as part of the treatment of diseases involving motor, affective, attention, memory, self-regulation, and communication deficits. According to these authors, evidence suggests that music-based rehabilitation may be effective in many psychiatric, neurological, and developmental disorders, as well as chronic diseases that cause pain and anxiety. In this sense, music could also be used with children and adolescents who have ADHD, and knowing about musical neuropsychological profile of these individuals with ADHD may contribute to intervention strategies in this population. Considering this relationship between musical performance and the development of cognitive functions that may be compromised in ADHD, the present study aimed to evaluate the musical neuropsychological profile of individuals with ADHD and to investigate the presence of possible musical processing deficits.

2. METHOD

2.1. Participants

Twenty Brazilian individuals aged 7-13 years, were divided into two groups of 10 participants: (1) a clinical group composed by individuals diagnosed with ADHD, and (2) a control group, consisting of individuals with typical development. Groups were matched by sex and educational level. We recruited students from public and private schools from Divinópolis-MG. These schools do not offer formal musical instruction.

Participants on clinical group had been previously diagnosed with ADHD. Inclusion criteria of control group were: (1) have a normal academic performance; (2) have a normal intelligence quotient; (3) do not present neurological, visual or hearing diseases; (4) do not have formal music education. Table 1 shows sample characteristics regarding age, sex and institution type (public or private). There were no significant differences between groups in terms of age (U = 45.5, Z = -.34, p = .734, two-sided).

Table 1. Sample characterization.

Group	Average age (SD)	Gender (%)	Institution type (%)
ADHD	10.22 (2,06)	30% female	60% public
Contro l	10.07 (1,85)	30% female	60% public

Materials

For the data collection, we used the following tasks that evaluate musical cognition, numeric cognition, intelligence and reading:

2.1.1.*Montreal Battery of Evaluation of Musical Abilities (MBEMA):* MBEMA evaluates music cognition from three musical tests (Melody, Rhythm and Musical Memory). MBEMA follows the same principles of the Montreal Battery of Evaluation of Amusia (MBEA) [21]. However, MBEMA is an abbreviated version for children [22]. In Melody and Rhythm tests participants have to judge if the two melodies heard are the same or different from each other. In the Musical Memory test participants must say if they have heard or not the presented melody before. For evaluate beat perception we added the Meter test from the MBEA. In this test, participants should decide if the presented melody is a march (double meter) or a waltz (triple meter) [21].

2.1.2.*Brazilian School Achievement Test (Teste de Desempenho Escolar - TDE):* This test aims to evaluate the school performance of children and adolescents, and it is divided in three subtests: (1) Writing, (2) Arithmetic, and (3) Reading. In the Writing subtest participants should write 34 dictated words of increasing syllabic complexity. Arithmetic subtest consists of three oral-presented and 35 written-presented arithmetic operations. Reading subtest consists of 70 stimuli, which must be read aloud by the participant [23].

2.1.3.*Raven's progressive matrices tests (RPM):* RPM is a non-verbal test that evaluates *g*-factor of intelligence. RPM demands the ability of thinking clearly and logically. The test consists of identifying the missing element that completes a figure pattern [24, 25]. In this study, we used the Color scale for individuals under eleven year's old and Standard scale for participants over eleven years old.

2.1.4.*Phoneme elision (PE):* This task aims to evaluate participants' phonological awareness abilities. Firstly, the participants must listen to a word and repeat it. Secondly, the examiner takes a specified phoneme in the word out and individuals has to recognized and say what is the new formed word (e.g., "sola" without /a/ is "sol" [in English, it would be similar to "cup" without /k/ is "up"). The test comprises 28 items. In eight items, the child must delete a vowel, and in the other 20, a consonant. The phoneme to be suppressed could be in different positions within the words, which ranged from two to three syllables [26].

2.1.5.*Pseudowords Repetition and Reading tasks:* These tasks consist of 40 inexistent words (pseudowords) which obey to language phonological restrictions. In the Repetition task participant should repeat the pseudoword heard and in the Reading task they must read the presented pseudowords [27]. Pseudowords reading task is one of the best indicators of the ability to read by the phonological decoding mechanism (phonological route). Difficulties in this task constitute a cognitive marker for specific dyslexia. Errors in this task can also indicate attention or decoding difficulties. Pseudowords repetition task has been considered as an index of phonological working memory capacity.

2.2. Procedures

The review board of the State University of Minas Gerais (ETIC no. 1.123.506/15) approved this study. After obtaining permission from the school principals, the research project was presented in the classrooms. The parents or guardians of the interested students received an invitation letter and provided informed consent. All participants performed the tasks in one testing session of approximately 2 hours at adequate and properly prepared rooms provided by the school principals. Testing was conducted by a team of undergraduate psychology students with training in psychometrics, which was led by the first author of this article. We used Mann-Whitney U test (two-sided) with an alpha level of .05 for comparing groups results in all tasks.

3. RESULTS

The Mann-Whitney U test results for the difference between clinical (ADHD) and control groups performance are presented in Table 2 for the following tasks: (1) Montreal Battery of Evaluation of Musical Abilities (MBEMA- Melody, Rhythm and Memory tests), (2) Montreal Battery of Evaluation of Amusia (MBEA – Meter test), (3) Brazilian School Achievement Test (TDE - Writing, Arithmetic, and Reading tests), (4) Raven's progressive matrices tests (RPM), (5) Phoneme elision (PE), (6) Pseudoword Repetition task (PRP), and (7) Pseudoword Reading task (PRD).

	Mean Rank		Sum of Ranks					
Tasks	***ADHD*[a]**	***Control*[a]**	***ADHD***	***Control***	**Mann-Whitney U**	**Wilcoxon W**	**Z-score**	**p-value**
MBEMA - Melody	7.90	13.10	79.0	131.0	24.0	79.0	-1.995	.046[b]
MBEMA - Rhythm	8.40	12.60	84.0	126.0	29.0	84.0	-1.596	.111
MBEMA - Memory	9.40	11.60	94.0	116.0	39.0	94.0	-.841	.400
MBEA - Meter	7.55	13.45	75.5	134.5	20.5	75.5	-2.240	.025[b]
PRP	10.25	10.75	102.5	107.5	47.5	102.5	-.190	.849
PRD	6.85	14.15	68.5	141.5	13.5	68.5	-2.779	.005[c]
PE	7.20	13.80	72.0	138.0	17.0	72.0	-2.511	.012[b]
RPM	7.10	13.90	71.0	139.0	16.0	71.0	-2.578	.010[b]
TDE - Writing	6.45	14.55	64.5	145.5	9.5	64.5	-3.072	.002[c]
TDE - Arithmetic	8.25	12.75	82.5	127.5	27.5	82.5	-1.703	.088
TDE - Reading	7.10	13.90	71.0	139.0	16.0	71.0	-2.582	.010[b]

[a] n=20, [b] p < .05, [c] p<.01

Table 2: Mann-Whitney U test results for difference between clinical (ADHD) and control groups performance.

We found that clinical group (ADHD) had a significantly lower performance than control group in MBEMA (Melody), MBEA (Meter), language tasks (PRD and PE), RPM test and in TDE's Writing and Reading tests. There were no differences between groups concerning PRP task, Arithmetic test (from TDE), and in the MBEMA's Rhythm and Musical Memory tests.

4. DISCUSSION

From the results obtained we could observe that individuals of clinical group (ADHD) showed, in relation to control group, a lower problem-solving capacity and difficulties in tasks that evaluates phonological awareness, reading and writing. Nevertheless, individuals with ADHD showed a preserved performance in tasks related to math processing and phonological working memory capacity. They also presented a profile of musical abilities with impairment of meter and melodic perception and preservation of the rhythm and musical memory perception.

Data obtained from the language tasks may indicate the presence of associated deficits in ADHD. Dyslexia is the most frequently comorbidity associated with ADHD [28] and it is characterized by difficulties in acquiring basic reading skills such as word identification and phonological decoding [29]. The pattern of performance in ADHD group constituted of difficulties in reading, writing and phonological skills found in our study could indicate the presence of individuals with associated dyslexia in this group. However, the results did not show deficits in the phonological working memory, which is related to the accuracy of phonological representation in long-term memory. These results are controversial since reading difficulties are associated with poor performance on pseudowords repetition tasks [30].

It is worth mentioning that since there were differences between the groups regarding the general intelligence measure (Raven), with worse performance for the ADHD group, the results obtained by this group in the other tasks should be considered with caution. Despite this, studies show that there are no direct correlations between intelligence and performance in writing and reading tasks, since normal intellectual development is a condition for the diagnosis of dyslexia and other language disorders [31, 32]. In this sense, the inferior performance of the clinical group individuals in the intelligence task is not sufficient to explain their poor results in the reading and writing tasks.

In addition to the language processing difficulties, the data indicated that individuals with ADHD presented worse melodic perception (pitch), and meter perception in relation to the control group. The difficulties in melodic perception could represent a comorbidity of ADHD. The MBEMA's Melodic task aims to evaluate the melodic organization of musical perception, which is composed by contour, interval and scale dimensions [21]. The contour refers to the shape of a melody created by the patterns of the change in pitch direction, which can be ascending or descending. The interval refers to the pitch distance between two adjacent musical notes and requires a more analytical perception. The scale is a dimension related to tonal functions [21].

The difficulties found for the clinical group in the musical processing, referring to the melodic organization, may indicate that individuals with ADHD presented congenital amusia as a comorbidity, since the deficit in pitch perception is characteristic of the individuals with congenital amusia [33, 34]. The results also corroborate the study of Jones et al. [10], which verified a higher prevalence of ADHD in patients with congenital amusia. The relationship between ADHD and congenital amusia may be modulated by developmental epigenetic mechanisms [2, 35, 36]. Nevertheless, this relationship should be more investigated in studies with larger samples and specifically designed to attend to this hypothesis. We should highlight that the difficulties related to melodic processing in individuals with ADHD, instead of indicating the presence of congenital amusia, may be secondary to the attention deficit of this disorder or be related to a more spread deficit in working memory [37].

Regarding performance in tests that evaluate temporal dimensions in musical cognition, the results indicated a pattern of dissociation between meter and rhythm processing in individuals with ADHD, since they showed poor performance in the Meter but not in the Rhythm test. According to Peretz, Champod and Hyde [21], rhythm and meter dimensions are the component parts of the temporal organization of music processing. The rhythm depends on more analytical processing and refers to the temporal grouping between two events, which may vary in time proximity. The meter refers to the beat of a melody, that is, to the basic temporal regularity between a sequence of musical events, and it is related to more global aspects of musical temporal dimension.

The pattern of dissociation between meter and rhythm processing found in this work corroborates the studies showing double dissociations between these dimensions, in which one of the components is preserved, while the other is impaired [21, 38]. The difficulties of perception in beat regularity found in this study also corroborate Carrer's [12] study, in which ADHD individuals who did not use medication presented lower performance on temporal synchronization tasks. The deficits in timing processes in ADHD are evidenced on several temporal domains. According to Rubia, Halari, Christakou and Taylor [39] the neurocognitive dysfunctions in temporal processes are crucial to the impulsiveness disorder of ADHD.

Low performance on temporal processing and sound duration discrimination tasks in individuals with ADHD could be also related to the tasks demands which require attention and distinct stimulus differentiation skills. Studies indicate the relevance of executive functions and attention in temporal processing deficits presented by ADHD individuals [40, 41]. When temporal stimulus demands excessively from the sensorial capacity, other cognition functions, such as working memory and sustained attention, are recruited in order to support temporal processing [42]. This can be an explanation to the difficulties presented by ADHD individuals in the Meter test. However, individuals with ADHD did not show difficulties in Rhythm test, which also is part of the temporal dimension of music perception. This result pattern may be due to the higher demand of the Meter test compared to the Rhythm test. In the Meter test the individual must understand the melody beat regularity that was heard, and after answer if it is a waltz, or a march (trinary or binary tempo). In the Rhythm test the individual must hear the melodies and say if they are the same or different, and this difference consists of an alteration in the rhythm pattern. Besides that, the Meter task demands a higher working memory capacity and attentional processes, which as seen before, influences temporal processing in individuals with ADHD.

We must emphasize that the inferior performance of individuals of the clinical group in the intelligence test may have influenced the results in the musical tests. According to Schellenberg [43, 44], musical training may be associated with performance in intelligence tests and there seems to be a positive correlation between musical abilities and intelligence. However, it cannot be said that the inferior performance of the individuals of the clinical group in some musical tests was due to the lower intellectual capacity of these individuals in relation to the control group, since they presented a lower performance only in the Meter and Melodic tests. The results suggest, therefore, that difficulties in temporal processing (i.e. meter) and melodic discrimination (i.e. pitch) may be characteristic of ADHD or present as a comorbidity of this disorder.

Even though individuals with ADHD have deficits in music perception, musical training can help these individuals on rehabilitation of compromised functions. Learning a new skill promotes physical and chemical changes in the brain, helping to establish efficient neural network. Musical training is related to development of attentional and memory skills [20, 45]. In this sense, musical training in individuals with ADHD can constitute an alternative for developing rehabilitation or compensation strategies to their compromised cognition functions.

5. CONCLUSIONS

Though deeper investigation about music processing in individuals ADHD is necessary, the presented study results reveals that this population had primarily difficulties on understanding beat regularity in a melody, which can be related to the global temporal processing difficulties of those individuals, besides deficits on pitch processing, which could indicate congenital amusia in association to ADHD. The difficulties found in global musical temporal processing (Meter) and melodic discrimination in the clinical group could emerge as secondary characteristics to the primary deficits of attention and working memory, or to present as a comorbidity of ADHD (e.g. amusias). Knowledge of the musical neuropsychological profile of individuals with ADHD may contribute to intervention strategies in this population, since musical training is related to better performance in tasks involving verbal, visuospatial, attentional and memory skills.

Acknowledgments

We are grateful for the financial support from Programa Institucional de Apoio à Pesquisa (PaPq/UEMG-08/2015). The authors would like to thank Walquíria Lopes and Andressa Cardoso Mariano for their help in data collection.

6. REFERENCES

1. R. A. Barkley, "Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD." in *Psychological bulletin*, 121(1), 1997, pp. 65-94.

2. J. Biederman, "Attention-deficit/hyperactivity disorder: a selective overview." in *Biological psychiatry,* 57(11), 2005, pp. 1215-1220.

3. P. D. Walshaw, L. B. Alloy, and F. W. Sabb, "Executive function in pediatric bipolar disorder and attention-deficit hyperactivity disorder: In search of distinct phenotypic profiles" in *Neuropsychology Review*, 20(1), 2010, pp.103-120.

4. R. Martinussen, J. Hayden, S. Hogg-Johnson, and R. Tannock, "A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder." in *Journal of the American Academy of Child & Adolescent Psychiatry, 44(4*), 2005, pp. 377-384.

5. M. E. Toplak, C. Dockstader, and R. Tannock, "Temporal information processing in ADHD: Findings to date and new methods." in *Journal of Neuroscience Methods,* 151(1), 2006, pp. 15-29.

6. D. Gooch, M. Snowling, and C. Hulme, "Time perception, phonological skills and executive function in children with dyslexia and/or ADHD symptoms." in *Journal of Child Psychology and Psychiatry,* 52(2), 2011, pp. 195-203.

7. L. M. McGrath, C. Hutaff-Lee, A. Scott, R. Boada, L. D. Shriberg, and B. F. Pennington, "Children with comorbid speech sound disorder and specific language impairment are at increased risk for attention-deficit/hyperactivity disorder." in *Journal of abnormal child psychology,* 36(2), 2008, pp. 151-163.

8. V. R. C. Pereira, T. M. M. Santos, and M. A. G. Feitosa, "Sinais comportamentais dos Transtornos do Déficit de Atenção com Hiperatividade e do Processamento Auditivo: a impressão de profissionais brasileiros" in *Audiology Communication Research*, 18(1), 2013, pp. 1-9.

9. J. Huang, B. R. Yang, X. B. Zou, J. Jing, G. Pen, G. M. McAlonan, and R. C. Chan, "Temporal processing impairment in children with attention-deficit-hyperactivity disorder." in *Research in developmental disabilities,* 33(2), 2012, pp. 538-548.

10. J. L. Jones, C. Zalewsk, C. Brewer, J. Lucker, and D. Drayana, "Widespread auditory deficits in tune deafness." in *Ear and hearing*, 30(1), 2009, pp. 63.

11. K. Hyde, and I. Peretz, "Brains that are out of tune but in time." in *Psychological Science,* 15(5), 2004, pp. 356-360.

12. L. R. J. Carrer, "Music and sound in time processing of children with ADHD." in *Frontiers in psychiatry*, 6, 2015 pp.127

13. H. Abikoff, M. E. Courtney, P. J. Szeibel, and H. S. Koplewicz, "The effects of auditory stimulation on the arithmetic performance of children with ADHD and nondisabled children." in *Journal of Learning Disabilities* 29(3), 1996, pp. 238-246.

14. S. H. Anvari, L. J. Trainor, J. Woodside, and B. A. Levy, "Relations among musical skills, phonological processing, and early reading ability in preschool children." in *Journal of experimental child psychology,* 83(2), 2002, pp. 111-13.

15. B. F. Pennington, "From single to multiple deficit models of developmental disorders." in *Cognition,* 101(2), 2006, pp.385-413.

16. E. G. Schellenberg, "Long-Term positive associations between music lessons and IQ." in *Journal of Educational Psychology,* 98 (2), 2006, pp. 457–468.

17. K. Vaughn, "Music and mathematics: Modest support for the oft-claimed relationship." in *Journal of Aesthetic Education,* 34, 2000, pp. 149-166.

18. L. Hetland, "Learning to make music enhances spatial reasoning." in *Journal of Aesthetic Education,* 34, 2000, pp. 179-238.

19. Y .Ho, M. Cheung, and A. Chan, "Music training improves verbal but not visual memory: Cross sectional and longitudinal explorations in children." in *Neuropsychology,* 17, 2003, pp. 439-450.

20. T. Särkämö, M. Tervaniemi, and M. Huotilainen, "Music perception and cognition: development, neural basis, and rehabilitative use of music." in *Wiley Interdisciplinary Reviews: Cognitive Science,* 4(4), 2013, pp. 441-451.

21.I. Peretz, A.S. Champod, and K. Hyde. “Varieties of musical disorders: The Montreal Battery of Evaluation of Amusia.” in *Annals of the New York Academy of Sciences, 999*, 2003, pp. 58-75.

22.I. Peretz, N. Gosselin, Y. Nan, E. Caron-Caplette, S. E. Trehub, and R. Béland. “A novel tool for evaluating children's musical abilities across age and culture.” in *Frontiers in systems neuroscience*, 7, 2013, pp. 30.

23.L. M. Stein, *Teste de Desempenho Escolar: Manual Para Aplicação e Interpretação.* Casa do Psicólogo, 1994.

24.A. L. Angelini, I. C. B. Alves, E. M. Custódio, W. F. Duarte and J. L. M. Duarte. *Manual matrizes progressivas coloridas de Raven: escala especial*. Centro Editor de Testes e Pesquisas em Psicologia, 1999.

25.J. C. Raven, *Teste das matrizes Progressivas: Escala Geral.* CEPA, 2003.

26.J. B. Lopes-Silva, R. Moura, A. Júlio-Costa, V. Geraldi Haase, and G. Wood, “ Phonemic awareness as a pathway to number transcoding.” in *Frontiers in psychology,* 5, 2014, 13.

27.F. H. Santos, and O. F. A. Bueno, “Validation of the Brazilian Children's Test of Pseudoword Repetition in Portuguese speakers aged 4 to 10 years.” in *Brazilian Journal of Medical and Biological Research,* 36(11), 2003, pp. 1533-1547.

28.E. Germanò, A. Gagliano, and P. Curatolo, “Comorbidity of ADHD and dyslexia.” in *Developmental neuropsychology,* 35(5), 2010, pp. 475-493.

29.F. R. Vellutino, J. M. Fletcher, M. J. Snowling and D. M. Scanlon “Specific reading disability (dyslexia): what have we learned in the past four decades?” in Journal *of child psychology and psychiatry,* 45(1), 2004, pp. 2-40.

30.C. A. de Carvalho, A. D. S. Kida, S. A. Capellini, and C. R. de Avila, “Phonological working memory and reading in students with dyslexia.” in *Frontiers in psychology,* 5, 2014, pp.746.

31.A. C. B. Maia and M. L. Fonseca, “Quociente de inteligência e aquisição de leitura: um estudo correlacional. ” in *Psicologia: Reflexão e Crítica,* 15(2), 2002, pp. 261-270.

32.A. D'angiulli, and L. S. Siegel, “Cognitive Functioning as Measured by the WISC-R Do Children with Learning Disabilities Have Distinctive Patterns of Performance?” in *Journal of Learning Disabilities,* 36(1), 2003, pp. 48-58.

33.J, Ayotte, I. Peretz, and K. Hyde, “Congenital amusia: A group study of adults afflicted with a music-specific disorder.” in *Brain,* 125, 2002, pp. 238-251.

34.D. T. Vuvan, M. Nunes-Silva, and I. Peretz, “Meta-analytic evidence for the non-modularity of pitch processing in congenital amusia.” in *Cortex,* 69, 2015, pp. 186-200.

35.I. Peretz, S.Cummings, and M. P. Dubé, “The genetics of congenital Amusia (Tone Deafness): A family-aggregation study.” in *The American Journal of Human Genetics,* 81, 2007, pp. 582-588, 2007.

36.K. J. Mitchell, “Curiouser and curiouser: Genetic disorders of cortical specialization.” in *Current Opinion in Genetics & Development*, 21, 2011, pp. 271–277.

37. V. J. Williamson, and L. Stewart, “Memory for pitch in congenital amusia: Beyond a fine-grained pitch discrimination problem.” in *Memory*, 18(6), 2010, pp.657-669.

38.C. Liégeois-Chauvel, I. Peretz, M. Babaï, V. Laguitton, and P. Chauvel, “Contribution of different cortical areas in the temporal lobes to music processing.” in *Brain*, 121, 1998, pp. 1853-1867.

39.K. Rubia, R. Halari, A. Christakou, E. Taylor “Impulsiveness as a Timing disturbance in attention-deficit-hyperactivity disorder during temporal processes and normalization with methylphenidate.” in Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1525), 2009, pp. 1919-1931.

40.L. Valko, M. Doehnert, U. C. Müller, G. Schneider, B. Albrecht, R. Drechsler, and D. Brandeis “Differences in neurophysiological markers of inhibitory and temporal processing deficits in children and adults with ADHD.” in *Journal of Psychophysiology*, 23(4), 2009, pp. 235-246.

41. H. Y. Lee and E. L. Yang. “Exploring the Effects of Working Memory on Time Perception in Attention Deficit Hyperactivity Disorder.” in *Psychological Reports*, 2018, pp. 0033294118755674.

42.J. A. Mangels and R. B. Ivry, Time perception*. In: RAPP, B. The handbook of cognitive neuropsychology: what deficits reveal about the human mind.* Psychology Press, 2001, pp. 467-493.

43.E. G. Schellenberg, “Long-term positive associations between music lessons and IQ.” in *Journal of Educational Psychology*, 98(2), 2006, pp. 457.

44.E. G. Schellenberg, E. G. (2011). “Examining the association between music lessons and intelligence.” in *British Journal of Psychology*, 102(3), 2011, pp. 283-302.

45.D. L. Strait, and N. Kraus, “Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise.” in *Frontiers in psychology*, 2, 2011, pp.113.

EFFECTS OF ALTERED AUDITORY FEEDBACK ON THE TEMPORAL CONTROL OF DISCRETE AND CONTINUOUS MOVEMENTS

Thenille Braun Janzen
Faculty of Music,
University of Toronto, Canada
thenille.braunjanzen@utoronto.ca

William F. Thompson
Department of Psychology,
Macquarie University, Australia
bill.thompson@mq.edu.au

Ronald Ranvaud
Instituto de Ciências
Biomédicas, Universidade de
São Paulo, Brazil
ranvaud@usp.br

ABSTRACT

Research indicates that distinct mechanisms regulate the timing of discrete and continuous rhythmic movements. These mechanisms – called event and emergent timing, respectively – may respond differently to auditory feedback. In this study, we investigated whether unexpected pitches in auditory feedback would interact with the timing accuracy of discrete and continuous rhythmic movements. Participants completed two tasks: finger tapping (Experiment 1) and circle drawing (Experiment 2). A synchronization-continuation paradigm was adopted whereby each action in the continuation phase triggered a sequence of repeated piano tones; however, in some trials, the auditory feedback of one tone of the sequence was altered. Our results indicate that unexpected feedback significantly affected timing accuracy of both discrete and continuous movements, but in unique ways. For the tapping task, the perturbations led to a decrease in the intertap interval immediately following the feedback manipulations. Conversely, in the circle-drawing task, unexpected pitch changes increased the inter-response interval on the second interval after the perturbation. These results indicate that unexpected changes in feedback content induced different error correction responses in discrete and continuous rhythmic movements, and shed light on adaptation and anticipation mechanisms in the temporal control of different types of rhythmic movements.

1. INTRODUCTION

Studies have recently demonstrated that distinct mechanisms control the timing of discrete and continuous rhythmic movements [1]–[4]. Discrete rhythmic movements are periodic actions preceded and followed by a period of little or no motion (e.g., finger tapping), and are thought to rely primarily on *event timing* mechanisms. On the other hand, the temporal regularity of continuous and smooth rhythmic movements (e.g., circle drawing) emerges from the control of movement dynamics, called *emergent timing* [1], [4]. Evidence suggests that event timing requires an explicit representation of a temporal interval to be produced based on an internal clock-like mechanism, whereas emergent timing arises from the dynamic control of nontemporal parameters of the movement, such as velocity, and therefore does not require an explicit internal representation of time [1], [3], [5]. The proposal that movements are based on distinct timing mechanisms has been supported by a series of studies demonstrating, for instance, that patients with cerebellar lesion exhibited increased variability for intermittent circle drawing (event timing) but not for continuous circle drawing [5]–[7]. Neuroimaging findings also suggest that event and emergent timing may recruit different brain areas [8], [9]. Finally, mathematical models of movement control are consistent with the possibility that the motor system relies on two different timing mechanisms [4].

Studies have shown that event and emergent timing mechanisms diverge in fundamental ways with respect to motor control, particularly in relation to the role of sensory feedback. For instance, Studenka and Zelaznik [10] first reported that continuous or cyclical movements do not exhibit auditory-motor synchronization due to the lack of perceptible events demarking each temporal cycle. However, further investigations demonstrated that the introduction of auditory [2], [11] or tactile feedback [12] at the end point of each cycle provided crucial temporal references required to generate an internal representation, thus enabling sensorimotor synchronization of continuous rhythmic movements.

To further investigate the role of auditory feedback in event and emergent timing, we examined the effect of unexpected perturbations of feedback content (i.e., pitch) on the timing of self-paced finger tapping and circle drawing. Experiment 1 focused on the effect of unexpected changes in feedback content on timed tapping, whereas Experiment 2 examined the effect of unexpected changes in feedback content on timed circle drawing. Based on the assumption that expected motor and perceptual outcomes are integrated into the motor commands [13], [14], these experiments tested whether feedback would be integrated into the representation of timing not only in event timing but also in emergent timing. It was predicted that feedback perturbations would significantly disrupt the timing of both event and emergent timing.

However, if this prediction were not confirmed, there would be reason to think that continuous and discrete movements are based on distinct timing and expectancy mechanisms.

2. EXPERIMENT 1: FINGER TAPPING

2.1. Method

2.1.1.Participants

Twenty-five undergraduate students (20 females) were recruited from the Department of Psychology at Macquarie University and received partial course credit for their participation. The average age was 20.1 years (*SD* = 6.4; range 18 to 44), and formal music training of participants varied from 0 to 13 years (*M* = 4.02; *SD* = 4.5). Only one participant reported being currently involved with musical activities for at least 2 hours weekly. None of the participants reported any form of hearing or motor impairment, and none had participated previously in a tapping experiment. This study was approved by Macquarie University Human Research Ethics Committee. All participants gave informed consent and were debriefed about the goals of the experiment after their testing.

2.1.2.Stimuli and Procedure

The experimental design followed the synchronization-continuation paradigm. Participants first synchronized eight taps with metronome tones presented at a fixed inter-onset interval of 600 ms. After eight taps the metronome stopped, and the participant attempted to continue tapping at the same pace, with each tap triggering a feedback tone. All feedback tones had a piano timbre. Participants were instructed to maintain the tempo set by the pacing signal to the best of their abilities until the end of the trial.

The feedback tones in the continuation phase were organized such that the first tone in every group of four tones was a piano tone of 392 Hz frequency (G4) and the other three tones were 261.63 Hz (C4), as depicted in Figure 1. This combination of feedback tones was repeated for 20 taps in the continuation phase (5 cycles). Similar patterns of feedback tones were presented in conditions 2 and 3 (i.e., a pitch change every fourth tone). However, in Condition 2 the expected pitch change (to G4) was unexpectedly displaced upward by one semitone (to G#4) at the 9th position of the continuation phase. In Condition 3, the expected pitch change (to G4) unexpectedly did not change (remained at C4) at the 9th position of the sequence. That is, the oddball was an *unexpected non-change*. In Condition 4, all feedback sounds were C4 piano tones, and there was only one unexpected pitch change (oddball) during the sequence of feedback tones (G#4). In Conditions 1-3, the cycle of four feedback tones implied a metric structure (repeated groupings of four tones), which was reinforced in Condition 4 by the introduction of a slight increase in intensity for the first tone within each group of four tones (+10 dB).

1
2
3
4

Figure 1. Representation of the auditory feedback sequence presented in the continuation phase for each of the 4 conditions.

Participants were given four practice trials (Condition 1) to familiarize themselves with the task and to ensure that they developed expectations for the feedback tones. Each condition was randomly presented 20 times, thus totaling 80 trials per participant. Trials were discarded and re-done immediately whenever any intertap interval (ITIs) was above or below 60% of the mean ITI for the trial. With breaks offered between blocks, the task took approximately 40 minutes to be completed.

2.1.3 Data Analysis

The synchronization phase ensured that the tempo was consistent across participants; thus, only the taps in the continuation phase were subjected to statistical analysis. To account for accelerations that can occur at the transition between synchronization and continuation phases the first five taps in the continuation phase were discarded, and the remaining 15 taps were subjected to analysis. ITIs were defined as the elapsed time between taps (in milliseconds). To evaluate general interference effects of oddballs on timing we analyzed participants' coefficient of variation (CV), which is defined as the standard deviation of ITIs within a trial divided by the mean ITI (SD/Mean). The average coefficient of variation was calculated across all trials for each condition.

The analysis focused on 3 ITIs before and 3 ITIs after the oddball presentation. Initial analysis indicated that there was no significant difference among the three intervals before the oddball. Therefore, for each trial, an average of 3 ITIs preceding the oddball was calculated and labeled as ITIp (where p = *pre-oddball*). The following ITIs were coded O, O+1, O+2 (where O = oddball). For the purpose of illustration, target ITIs were aligned and averaged across positions. ITI (ms) values were averaged across trials for each participant and each condition and subjected to a 4 x 4 repeated-measures ANOVA with two factors: Condition (4 levels) and Sequence Position (4 levels). Mauchly's test indicated that the assumption of sphericity had been violated for Sequence Position ($\chi^2(5) = 14.07$, $p = .01$, $\varepsilon = 0.70$). Therefore Greenhouse-Geisser corrections were applied to *p* values where appropriate.

2.2 Results

ITI (ms) values were analyzed, and results indicated a significant main effect of Sequence Position, F (3, 72) = 7.71, p = .001. Pairwise comparisons revealed that the sequence position immediately following the oddball tone was significantly shortened in comparison to all the other positions analyzed (O < P $p < .001$; O < O+1 $p = .006$; O < O+2 $p = .02$). There was a main effect of Condition, F (3, 72) = 11.47, $p < .001$. Across the taps that were analyzed, the mean ITI was larger in Condition 4 than in the other conditions. However, this main effect is qualified by a significant interaction between Sequence Position and Condition, which revealed that the effect of the oddball was quite different in the four conditions, F (9,216) = 11.77, $p < .001$. For Condition 1 (expected pitch change), there was no significant shift in the timing of taps (see Figure 2). This finding suggests that an *expected* pitch change does not affect the timing of motor actions. On the other hand, there was a significant effect of the oddball for all other conditions.

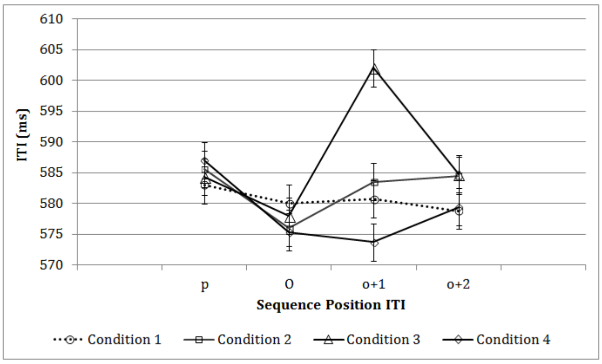


Figure 2. Intertap intervals (ms) displayed across conditions and sequence positions.

To further evaluate the effect of oddballs on timing we analyzed participants' coefficient of variation (CV). Each participant's CV values were averaged across trials for each condition and subjected to a repeated-measures analysis (ANOVA) with two factors (Condition, Sequence Position). There was a significant interaction between Condition and Sequence Position, F (3, 72) = 4.93, p = .004. Further analysis comparing the independent effect of each condition revealed that variability significantly increased after the oddball under Condition 3, suggesting that compensation was introduced after the perturbation (F = 11.001, p = .003). The coefficient of variation did not significantly change across trials in the other conditions. None of the other factors reached statistical significance.

2.3 Discussion

Experiment 1 tested the hypothesis that expectancy mechanisms mediate the interaction between feedback content and the timing of discrete motor actions, thus predicting that unexpected pitch changes in feedback would significantly interact with timing, while highly predictable pitch changes would not affect motor timing. Results confirmed this prediction by showing that only unexpected auditory feedback triggered by the discrete action significantly influenced the timing of tapping. Data revealed that unexpected feedback shortened the intertap interval immediately after the perturbation and induced a compensation in the following timing interval. The disruption of timing was observed in conditions where pitch changes were only 1-semitone, and also where an expected pitch change was omitted. These findings, therefore, suggest that expectations are intrinsic to timing and support the hypothesis that expected motor and perceptual outcomes are integrated into an internal representation [13], [14]. Further support to this finding was demonstrated by showing that expected pitch changes did not influence the timing of taps.

It was interesting to note that the effect of the unexpected pitch changes in the conditions tested affected the ITI trajectory in position o+1 significantly differently. Particularly, condition 3 - where an expected change was omitted - seems to have elicited a visibly higher overcompensation. This result suggests that expectancy mechanisms may process information on a continuous scale rather than as an all-or-none system. This suggestion is corroborated by a recent study that indicates that processing of expectancy violations is significantly modulated by the individual's action [15]. This research demonstrated that expectancy violations of one's action elicit brain responses that are significantly higher in amplitude than the passive perception of the same expectancy manipulations. In other words, expectancy violations evoke a much stronger neural response in the brain of the pianist than on the brain of the listener.

Experiment 1 supported the hypothesis of integration between feedback content and timing in event timing. However, it is not known whether unexpected changes in feedback content interact with emergent timing mechanism. This question was addressed in Experiment 2.

3. EXPERIMENT 2: CIRCLE DRAWING

3.1. Method

3.1.1.Participants

Twenty-four undergraduate students (13 females), average age 20 years (SD = 3.9 – range 18 to 36), were recruited from the Department of Psychology at Macquarie University and received partial course credit for their participation. Formal music training of participants ranged from 0 to 8 years (M = 2.45; SD = 2.5); however, only one participant was currently involved in music activities. None of the participants reported any form of hearing or motor impairment, and one had participated previously in a tapping experiment. All participants gave informed consent and were debriefed about the goals of the experiment after their testing.

3.1.2.Stimuli and Procedure

Experiment 2 was identical to Experiment 1, except for the task performed. In the circle-drawing task, participants repeatedly moved the computer mouse with the right index finger in a clockwise circular motion, tracing the circumference of a 5 cm diameter circular template displayed on the computer screen. Participants were instructed to synchronize their motion to the metronome so that the cursor would cross the mark positioned at 270 degrees on the circle at every click of the metronome.

Inter-response intervals were defined as the elapsed time between passes through the mark at 270 degrees. Participants were told that timing precision was more relevant than drawing accuracy, and they were free to choose the size of the circle they drew.

3.2. Results

To measure timing in the circle-drawing task, inter-response interval (IRI) was defined as the time elapsed between successive passes through the 270-degree mark. As in Experiment 1, in Experiment 2 IRI (ms) values were averaged across trials for each participant and each condition and subjected to a 4 x 4 repeated-measures ANOVA with two factors: Condition (4 levels) and Sequence Position (4 levels).

Results indicated a significant main effect of Sequence Position, F (3,69) = 5.79, p = .004, however pairwise comparisons revealed that unexpected pitch changes seemed to have interfered only in the position O+1 (O+1 > P, p = .005; O+1 = O p = .07; O+1 > O+2 p = .003). This result suggests that feedback changes did not interfere with the interval immediately after the oddball presentation as seen in the tapping task. It is also interesting to note that, unlike in event timing where unexpected changes tended to elicit a shortening of the interval immediately after the perturbation, in the circle drawing we see an *increase* in the interval after the perturbation.

More importantly, results indicated a main interaction of Condition and Sequence Position, F (9,207) = 2.69, p = .02, and post hoc tests revealed that Condition 1 (F (3,21) = 2.35, p = .10) and Condition 2 (F (3,21) = 2.87, p = .06) did not reach significance. Only Conditions 3 and 4 significantly interacted with intervals in the sequence position (F = 3.7 p = .02; F = 6.00 p = .004, respectively), as displayed in Figure 3.

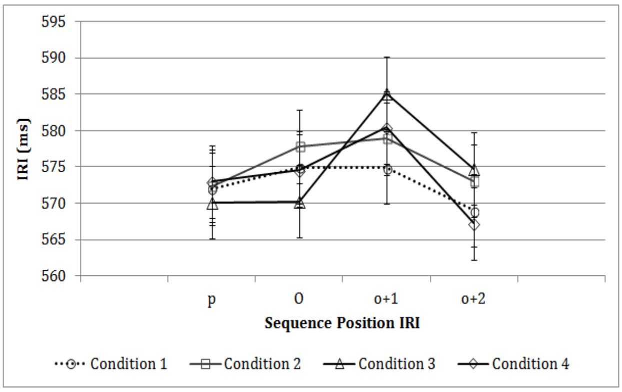


Figure 3. Inter-response intervals (ms) displayed across conditions and sequence positions.

Each participant's CV values were averaged across trials for each condition and subjected to repeated-measures ANOVA with two factors (Condition, Sequence Position). Results suggested that variability did not change significantly after the perturbation across conditions and sequence positions tested, as all factors did not reach significance (F <1). Comparing variability scores between performances on tapping (Experiment 1) and circle drawing (Experiment 2) showed that tapping was significantly more accurate than circle drawing, F (1,23) = 41.81, p < . 001. Interestingly, correlation analysis indicated that variability in circle drawing and tapping tasks were not correlated (r^2 -0.32, n = 24, p = .11).

3.3. Discussion

Experiment 2 tested whether unexpected perturbations on feedback content would disrupt the timing of continuous movements, such as circle drawing. The assumption that expected motor and perceptual outcomes of one's actions are integrated into an internal representation, such as proposed by the forward models, would predict that feedback perturbations would significantly disrupt the timing of both event and emergent timing. Results confirmed the initial prediction by showing that timing of responses in the circle-drawing task was significantly affected by unexpected feedback changes. Interestingly, it was noted that not all conditions interfered with the timing of continuous movements, as conditions 1 and 2 did not disrupt timing intervals. In these conditions, a highly expected pitch change was introduced every four tones (condition 1) and a semitone change was introduced in one of these tones in condition 2. The fact that these conditions did not interfere with the timing in the circle-drawing task suggest that continuous movements are more resistant to interference. This suggestion is corroborated by the observation that unexpected perturbations of feedback content in conditions 3 and 4 interacted with timing, but only at position O+1 and O+2.

Our findings are consistent with recent studies showing that continuous movements have longer recovery times after a phase shift perturbation [16], which indicates that movements based on emergent timing have a "large inertia in that, once the limb is moving, it is very difficult to adjust this movement pattern" [12, p. 1098], [17].

4. GENERAL DISCUSSION

These experiments aimed to investigate the role of auditory feedback in event and emergent timing by testing whether perturbations of feedback content would significantly interact with the timing of self-paced discrete and continuous movements. Experiment 1 demonstrated the crucial role of expectancy in the interaction between feedback content and timing. Results showed that unexpected changes of nontemporal aspects of the feedback such as pitch significantly interacted with the timing of tapping. It was possible to observe that expected pitch changes occurring at predictable positions did not interact with the timing of intervals, which suggests that the results reported here are associated with violation of expectations rather than mere pitch change.

It was also shown that unexpected changes in feedback content interacted with the timing of continuous movements. This result adds support to the hypothesis that expectancy mechanisms are intrinsic to the timing of motor actions. However, the effect of unexpected changes in timing differed between tapping and circle drawing. For the tapping task, the oddball led to a *decrease* in the intertap interval at the interval immediately following the change in feedback tone. On the other hand, the oddball interacted with timing by *increasing* the inter-response interval on the second position (i.e., o+1) after the perturbation in the circle drawing.

These results are consistent with studies showing that continuous movements are slower than discrete movements in adjusting following changes in the sensory input [17], [16]. Repp [18] suggested that the difference in the interference effect of unexpected events in the two tasks is related to a greater "maintenance tendency" in continuous movements. This tendency is thought to be associated with the inertia associated with the movement [12], [17], [18].

Another interpretation of these results is the association of different expectancy processes. It has been recently suggested that emergent timing is based on "strong anticipation" processes [19]. According to this hypothesis based on dynamic system approach, strong anticipations arise from the close alignment between the action and its sensory outcome. In this case, the goal of the system is to maintain smooth and uninterrupted rhythmic movements based on global and often long-term expectations. Therefore, it is possible that the decoupling between motor actions and the external environment linked to strong expectations leads to a suppression of immediate interactions between unexpected events and the motor planning of self-paced movements. On the other hand, event timing seems to be associated with weak anticipation processes. This expectancy mechanism is required in dynamic environments where unpredicted events require rapid and efficient correction in order to maintain accurate responses. Therefore, it is thought that weak anticipations entail local and short-term expectancies [14], [19], [20]. It may be possible, therefore, that weak anticipations facilitate the intervention of error correction mechanisms in event timing resulting in immediate interactions between unexpected events and timing.

5. CONCLUSIONS

Results here described corroborate the notion that timing mechanisms (i.e., event and emergent timing) diverge in fundamental ways with respect to motor control and the role of auditory feedback. Here we show that unexpected nontemporal alterations in auditory feedback content (i.e., pitch) significantly interacted with the timing of self-paced discrete and continuous movements but in distinct ways. In the tapping task perturbations led to a *decrease* in the intertap interval at the interval immediately following the feedback manipulations. Conversely, in the circle-drawing task, unexpected pitch changes *increased* the inter-response interval on the second interval after the perturbation. These findings indicate that unexpected changes in feedback content induced different error correction response in discrete and continuous rhythmic movements. Further investigations are crucial to better understand the role of adaptation/anticipation mechanisms in the temporal control of different types of rhythmic movements.

Acknowledgments

This research was supported in part by a post-graduate scholarship awarded to the first author by the National Council of Technological and Scientific Development (CNPq/Brazil) and Macquarie University (iMQRES), and a grant awarded to William F. Thompson by the Centre for Elite Performance, Expertise, and Training (CEPET), at Macquarie University, Australia.

6. REFERENCES

[1] H. N. Zelaznik *et al.*, "Timing variability in circle drawing and tapping: Probing the relationship between event and emergent timing," *J. Mot. Behav.*, vol. 37, no. 5, pp. 395–403, 2005.

[2] B. E. Studenka and H. N. Zelaznik, "Synchronization in repetitive smooth movement requires perceptible events," *Acta Psychol. (Amst).*, vol. 136, no. 3, pp. 432–441, Mar. 2011.

[3] H. N. Zelaznik, R. M. C. Spencer, and R. B. Ivry, "Dissociation of explicit and implicit timing in repetitive tapping and drawing movements," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 28, no. 3, pp. 575–588, 2002.

[4] R. Huys, B. E. Studenka, N. L. Rheaume, H. N. Zelaznik, and V. K. Jirsa, "Distinct timing mechanisms produce discrete and continuous movements," *PLoS Comput. Biol.*, vol. 4, no. 4, p. e1000061, Apr. 2008.

[5] R. B. Ivry, R. M. Spencer, H. N. Zelaznik, and J. Diedrichsen, "The cerebellum and event timing," *Ann. N. Y. Acad. Sci.*, vol. 978, pp. 302–317, Dec. 2002.

[6] R. M. C. Spencer, R. B. Ivry, and H. N. Zelaznik, "Role of the cerebellum in movements: Control of timing or movement transitions?," *Exp. Brain Res.*, vol. 161, no. 3, pp. 383–396, Mar. 2005.

[7] R. M. C. Spencer, H. N. Zelaznik, J. Diedrichsen, and R. B. Ivry, "Disrupted timing of discontinuous but not continuous movements by cerebellar lesions," *Science (80-.).*, vol. 300, no. 5624, pp. 1437–1439, May 2003.

[8] S. Schaal, D. Sternad, R. Osu, and M. Kawato, "Rhythmic arm movement is not discrete," *Nat. Neurosci.*, vol. 7, no. 10, pp. 1136–1143, Oct. 2004.

[9] R. M. C. Spencer, T. Verstynen, M. Brett, and R. Ivry, "Cerebellar activation during discrete and not continuous timed movements: An fMRI study," *Neuroimage*, vol. 36, no. 2, pp. 378–387, Jun. 2007.

[10] B. E. Studenka and H. N. Zelaznik, "Circle drawing does not exhibit auditory-motor synchronization," *J. Mot. Behav.*, vol. 43, no. 3, pp. 185–191, 2011.

[11] H. N. Zelaznik and D. A. Rosenbaum, "Timing Processes Are Correlated When Tasks Share a Salient Event," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 36, no. 6, pp. 1565–1575, 2010.

[12] B. E. Studenka, H. N. Zelaznik, and R. Balasubramaniam, "The distinction between tapping and circle drawing with and without tactile feedback: An examination of the sources of timing vari-

ance," *Q. J. Exp. Psychol.*, vol. 65, no. 6, pp. 1086–1100, 2012.

[13] D. M. Wolpert, K. Doya, and M. Kawato, "A unifying computational framework for motor control and social interaction," *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 358, no. 1431, pp. 593–602, 2003.

[14] M. C. (Marieke) van der Steen and P. E. Keller, "The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization," *Front. Hum. Neurosci.*, vol. 7, p. 253, Jun. 2013.

[15] C. Maidhof, N. Vavatzanidis, W. Prinz, M. Rieger, and S. Koelsch, "Processing expectancy violations during music performance and perception: an ERP study.," *J. Cogn. Neurosci.*, vol. 22, no. 10, pp. 2401–2413, Oct. 2010.

[16] M. T. Elliott, A. E. Welchman, and A. M. Wing, "Being discrete helps keep to the beat," *Exp. Brain Res.*, vol. 192, no. 4, pp. 731–737, 2009.

[17] B. H. Repp and S. R. Steinman, "Simultaneous Event-Based and Emergent Timing: Synchronization, Continuation, and Phase Correction," *J. Mot. Behav.*, vol. 42, no. 2, 2010.

[18] B. H. Repp, "Perfect Phase Correction in Synchronization With Slow Auditory Sequences," *J. Mot. Behav.*, vol. 40, no. 4, pp. 363–367, 2008.

[19] V. Marmelat and D. Delignières, "Strong anticipation: Complexity matching in interpersonal coordination," *Exp. Brain Res.*, vol. 222, no. 1–2, pp. 137–148, 2012.

[20] D. M. Dubois, "Mathematical Foundations of Discrete and Functional Systems with Strong and Weak Anticipations," *Anticip. Behav. ...*, pp. 110–132, 2003.

The Effect of Hand Shape Familiarity on Guitarists' Perceptions of Sonic Congruence

Keith Phillips
Royal Norther College of Music
keith.phillips@student.rncm.ac.uk

Andrew Goldman
Columbia University
ajg2232@columbia.edu

Tyreek Jackson
Columbia University
taj2118@tc.columbia.edu

ABSTRACT

Musical performance depends on the anticipation of the perceptual consequences of motor behaviour. Altered auditory feedback has previously been used to investigate auditory-motor coupling but studies to date have predominantly used MIDI piano in experimental tasks. In the present study, 21 guitarists played chords on a MIDI guitar in response to tablature diagrams. In half of the trials at random, one of the notes in the heard chord was altered and participants judged whether the feedback was altered or not, responding as quickly and accurately as possible by pressing one of two buttons on a footswitch. In a separate study, the same participants ranked the familiarity of the chord shapes. We found that participants' judgement of sonic congruence tended to be faster when the chord shape was familiar and when feedback was congruent. A follow-up study is underway to further isolate the effect of hand shape familiarity on reaction times.

RELIABILITY INTER-EXAMINERS OF THE NORDOFF ROBBINS MUSICAL COMMUNICATIVENESS SCALE BRAZILIAN VERSION

Aline Moreira André
UFMG
aline.musicasax@gmail.com

Cristiano Mauro Assis Gomes
UFMG
cristianomaurogomes@gmail.com

Cybelle Maria Veiga Loureiro
UFMG
cymbelleveigaloureiro@gmail.com

ABSTRACT

The Musical Communicativeness Scale is used to evaluate behaviors from sound and musical stimuli since the decade of 1960 in the USA. To be used also in Brazil, a validation process is required where Cultural and linguistic issues can be evaluated in the Brazilian context. For this validation process the model by Herdman, Fox-Rushby and Badia (1998) has been chosen. This model provides 6 equivalence types: conceptual, item, semantic, operational, measurement and functional. In this study the equivalence measurements has been chosen for testing reliability inter-examiners of the Musical Communicativeness Scale. The analysis of 24 music therapy methodological videos was adopted for boardline cases' people with neurodevelopmental disorders through Musical Communicativeness Scale. 1 researcher and 4 invited examiners have participated in this study stage. The inter-examiner scores have presented moderate and strong correlations (Spearman), indicating evidences of reliability for the Musical Communicativeness Scale translated and adapted to the Brazilian context.

1. INTRODUCTION

In Brazil as well as worldwide, there is a necessity for tests in order to evaluate behaviors resulting from sound and musical stimuli. Nordoff, Robbins, and Marcus (2007) state that three Scales were developed from research conducted at the University of Pennsylvania in the 1960s to analyze behaviors from musical stimuli in music therapy sessions. Among these three scales, also called Nordoff Robbins Scales, there is the Musical Communicativeness Scale. The Musical Communicativeness Scale evaluates in seven observable stages the levels of musical communicability through three variable strains: vocalizations, manipulation of musical instruments and movements with the body.

According to Nordoff, Robbins and Marcus (2007), essentially this scale is interrelated with the possibility of involving any individual in a process of awakening to musical awareness, to their perception and their pleasures, and to the experiences of communication, personal freedom, and accomplishment that the process can ensue. The musical scale also convey the stimulating impact of music, the interest musical co-activity that can be held, the releasing and uplifting enthusiasm that can be generated, and the communicative motivation that can be released. The Musical Communicativeness Scale is accomplished through three different activities: Instrumental, Vocal and Body Movement. The response of a child can arise in all three activities or in a single or two, as well as the emphasis of the reaction can be transferred to the following activity. In order to research the general means of communication associated to therapeutic activity, the scale define the successive communicativeness levels of behaviorism of each arrangement likewise they are shown in a evaluative form in such way that the developed communication activity similarly the manners remain at the same level. Yet, according to Nordoff, Robbins and Marcus (2007), the levels (1) to (4) describe the stages of activation and the differentiations of consciousness that lead to the beginning of musically intentional response — the gradual drawing out of a non-communicating child from evoked responses toward the achievement of limited communicative activity. The levels (1) to (4) can be seen as "stepping in and stepping down", which delineate the stages in a developmental process of awakening awareness as well the response to the therapists' music and their own emerging capability. Children responding on level (4) will already have, or be developing in therapy, the beginnings of a functional basis for musical participation. Levels (4) to (7) describe the stages of widening communication and personal-musical maturation that result from progressive musical engagement. These levels rate intensification of activity and advances in flexibility, quality, and breadth and strength of receptiveness and expression. These activities are manifested as a child's musical intelligence unfolds developmentally and they are progressively liberated from an emotional and/or organic dysfunction to become freely communicative and communicable.

The levels (1) to (4) examine the gradually deepening entry into a process of awakening musical awareness, and into finding and developing the receptive and expressive skills that stimulate and serve the realization of musical communicativeness. On Level (4) the beginning of an integrative grounding is taking place from which more capable and outwardly directed levels of interpersonal musical communicativeness become possible. As the response moves into Level (5) children are attaining some independence of selfhood in music as they become communicative. Level (6) co-activity and interactivity consolidates and enlivens these experiences and capabilities. Level (7) accommodates the activities of musically released, enthusiastic, and often creative of the individual. Consequently, Levels (4) to (7) can be thought of as "stepping up and stepping out."

In order for the Musical Communicativeness Scale to be used in the Brazilian context, a validation process is necessary. This process is extremely necessary for health tests in order to observe the best assurance to express each original construct and so cultural differences that can be evaluated for the new context of using. In order to carry out this process, the Universalist Validation Model developed by Herdman and colleagues (1998) was chosen to validate health tests. This model is divided into 6 (six) steps: conceptual equivalence, item equivalence, semantic equivalence, operational equivalence, measurement equivalence and functional equivalence. It is noteworthy that André, Gomes and Loureiro (2016) applied the model of Herdman et al. (1998) in order to investigate the conceptual equivalence of the Musical Communicativeness Scale. It was verified that the publications concerning to the use of the Scale increased over the years and that was used to evaluate the following populations in their original context: Autism Spectrum Disorder (ASD), described by the authors Bergmann (2015), Bell et al. (2014), Bergmann et al. (2006), which is a study of the development of neurodevelopmental disorders (Aigen et al., 1995) and the use of neurotransmitters and Mahoney (2010), healthy people, as described by the authors Australia (2008), Bunt (2003), Rahman (2008) and Wood (2006) and people with anorexia nervosa, described in Robarts's study (2000).

Subsequently, André et al. (2016) conducted a study of item equivalence, semantic equivalence and operational equivalence of the Musical Communicativeness Scale. In this process, the Musical Communicability Scale and its manual explaining English to Portuguese were translated. After the translation, the re-translation was performed for English and analysis of the translations. Specifically, four translators were invited to elaborate the Brazilian Portuguese version of the explanatory manual and the Musical Communicativeness Scale. In addition, 10 judges were invited to analyze the translated version of the explanatory manual and the Musical Communicability Scale by means of a questionnaire. All these steps presented favorable evidences to the equivalences of items, semantics and operational of the Brazilian version of the Scale of Musical Communicativeness.

The explanatory manual of the Musical Communicativeness Scale defines how the levels and domains of the Scale should be punctuated in the evaluation of the behaviors observed in a music therapy session. After verifying the explanatory manual of the Musical Communicativeness Scale presented by André et al. (2016), some adaptations through a summary of the translated manual were made. This summary did not interfere in the general sense of the Scale, but aimed to systematize its instructions. In the short version of the Handbook of Musical Communicativeness Scale, descriptions of Scale domains, items and grades were maintained, secondary information was written more objectively, and only one type of score was chosen for the Musical Communicativeness Scale. The original manual allows different types of scores. Only one of the punctuation options described in the Handbook of Musical Communicativeness Scale for the summarized manual has been used. This score is called "score by checklist," where only one item from each domain is punctuated at a time. Additional information on the summary process of the Scale manual will be described later.

In this study, the measurement equivalence of the Musical Communicativeness Scale was verified, translated and adapted to the Brazilian context as well as investigated the inter-rater reliability of this equivalence. The objective of the study was to verify the inter-rater reliability of the Musical Communicativeness Scale for the evaluation of music therapy interventions in children and adolescents with neurodevelopmental disorders.

2. METHODOLOGY

2.1. Participants

In this study, five examiners participated, among them, four were invited and one examiner was the researcher of this study. All four invited examiners are female, aged 20-25 and undergraduate students in music therapy.

In addition, two patients participated in the videos. One of them was five-year-old individual in the recording period of the videos who was diagnosed with Autism Spectrum Disorder while the other patient was 14-year-old individual during the video recording period who was diagnosed with tuberous sclerosis.

2.2. Instruments

After analyzing the translated manual of the Scale of Musical Communicability without any adaptation presented by André et al. (2016), a summary of this manual was made in order to allow greater understanding and a better targeting of data collection in this study.

All the explanations referring to all degrees, items and domains of the Scale of Musical Communicativeness were kept in the manual. Any observation considered as secondary information was discharged. This secondary information was descriptions of clinical examples from audio analysis, descriptions of modes of activities present in a music therapy session, and descriptions of different forms of Scale scores. Subsequently, the modes of activities present in the music therapy session in a more succinct way, and presented as examples of analysis and punctuation from videos of music therapy sessions were written. After excluding the other punctuation modes present in the translated manual of the Musical Communicativeness Scale, only a single way of punctuating the

Scale called “punctuation by checklist” was kept, where it allows that each domain be punctuated in a single time per analysis.

The Musical Communicativeness Scale is divided into 7 grades and 3 domains. The 7 grades assess the levels of musical communicability in 3 modes of activities: vocalizations, manipulation of musical instruments and movements with the body. The scale also allows punctuation for total patient inactivity and the sum of all modes of activity in order to verify the patient's musical communication more comprehensively.

In addition to the brief manual and the Musical Communicability Scale, 24 pre-recorded videos of Music Therapy consultations performed for people with neurodevelopmental disorders at the HC-UFMG Children's Psychiatry Outpatient Clinic and ABET (Brazilian Association of Tuberous Sclerosis) were used.

2.3. Data collected

Data from pre-filmed videos of music therapy sessions performed for people with neurodevelopmental disorders were selected and collected. Those responsible for the videos signed a Term of Free and Informed Consent authorizing the use of the same in this research. After being selected, the videos were edited in 240 time units of 30 seconds. From the 240 temporal units, a random draw of 24 stretches was carried out. After these steps, four examiners were duly trained by reading the summarized manual and were invited to hold a meeting with the researcher to clarify possible questions. Every participant has received a material for analysis and was instructed not to share information about their evaluation of the 24 videos from the Musical Communicativeness Scale. The examiners were then asked about their opinion on the understanding of the abridged version of the handbook on the Scale of Musical Communicability as part of the evaluation of the results.

This study was carried out for the Postgraduate Program in Music of the Federal University of Minas Gerais, in the area of Sonology. The same was approved and registered in the Ethics and Research Committee of UFMG, number 54578315.5.0000.5149

2.4. Data analysis

The pre-recorded videos of music therapy consultations accomplished with people suffering from neurodevelopmental disorders were analyzed from the Musical Communicability Scale. These videos were divided by 30 second units, which allowed an analysis of behaviors that occurred during each part of the music therapy sessions.

All data were stored in the Microsoft Excel 2016 spreadsheet. The Spearman correlation index was calculated to verify the inter-rater reliability of the Musical Communicability Scale. The correlation test was performed in SPSS 20.0 software, presented by Dancey and Reidy (2013).

3. RESULTS

In the field of the Musical Communicativeness Scale referring to instrumental musical communication, the mean value of the correlations was $\rho = 0.79$ and the standard deviation 0.08.

In the field of the Musical Communicativeness Scale referring to vocal musical communication, the mean value of the correlations was $\rho = 0.87$ and the standard deviation thereof was 0.06.

In the field of Musical Communicativeness Scale referring to musical communication through movements with the body, the mean value of the correlations was $\rho = 0.68$ and the standard deviation 0.19.

In addition to the individual domains of vocal, instrumental and body movement communication, the Musical Communicativeness Scale also predicts the sum of these domains in order to evaluate the patient's general musical communication. When performing the Spearman correlation test for the Total Musical Communicativeness score, it was observed that the mean value of these correlations was $\rho = 0.73$ and the standard deviation 0.06.

Moreover to the correlations among examiners, the correlation of examiner 1 (the researcher of this study) was carried out with the other examiners (collaborators). Altogether, the number of analyzes of the examiner 1 was replicated in order to compare with the 24 analyzes of each collaborating examiner. In this case, the Spearman correlation was performed with 96 analyzes (24x4).

After performing the Spearman correlation test between examiner 1 and other examiners, it was observed a moderate correlation in the body movement domain ($\rho = 0.69$) and strong correlations in the vocal domains ($\rho = 0.87$), body movement ($\rho = 0, 81$) and in total ($\rho = 0.98$).

In a preliminary study by André et al. (2016), a questionnaire was prepared for judges to classify the semantics and degree of comprehension of the translated manual of the Musical Communicativeness Scale. In this preliminary study, 60% of the judges rated the handbook as fully comprehensible and 40% rated it as partially comprehensible. According to the judges who classified the translated manual as partially comprehensible, it could be more objective and organized in a way that yields more fluency to reading. In this study, the collaborating examiners were asked about the comprehension of the summary manual of the Musical Communicativeness Scale in relation to the video analysis and in relation to the translated manual of the Scale. The whole panel of examiners rated the summarized manual as fully comprehensible and objective.

4. DISCUSSION

Hitherto, it can be acquired that Musical Communicativeness Scale has presented, in most analyzes, strong correlations, which indicates good inter-rater reliability. This information can be reinforced by the fact that there is no weak correlation in this study. These results indicate that the summary manual was indeed understandable, as the examiners stated and that they were able to understand the explanations as expected. It is also observed that according to the examiners, the summary version of the explanatory handbook of the Musical Communicativeness Scale presented better understanding than the translated explanatory manual.

The correlations made in this study that the whole panel of examiners presented similar scores to the researcher were disrupted from examiner 3 who differ from all other examiners, including the researcher, in the analysis of the domain of musical communication of corporal movement. As this difference occurred with examiner 3 only, and in a single domain, it is possible to consider this as an acceptable result because, even with some differences, the correlations were still moderate. In the others domains, examiner 3 presented strong correlations with the other examiners, including the researcher's correlations.

The Musical Communicativeness Scale has been used in Brazil in research to evaluate the behavior of children with ASD in music therapy sessions, as stated by Freire (2014), Andre e Batista (2014) and Sampaio (2015). Andre and Batista (2014) even performed the inter-rater reliability of the Nordoff Robbins Scales in this population. Nevertheless, overall the evaluation of children with ASD, the Musical Communicativeness Scale can also be used in a broader scope, for evaluating children and adolescents with neurodevelopmental disorder. According to the Diagnostic and Statistical Manual of Mental Disorders (DSM - 5), neurodevelopmental disorder can be defined as:

".... It is a set of conditions which begins in the developmental period. Disorders typically manifest early in development. In general, before children start attending school, they are characterized by deficits in development that lead to impairments in personal, social, academic or professional functioning" (AMERICAN PSYCHIATRIC ASSOCIATION, 2014).

The validation of the Musical Communicativeness Scale for the Brazilian context may contribute to several researches and to the clinical context. By using the same evaluation for assessing people with neurodevelopmental disorders may help in several contexts, since several music therapy interventions have been carried out for the Brazilian population, according to Loureiro (2006) and André et al. (2015).

Furthermore, the Musical Communicativeness Scale could also be used to evaluate behaviors in other populations, such as healthy people, as described by Australia (2008), Bunt (2003), Rahman (2008) and Wood (2006). Alternatively, in order to have these techniques being applied in Brazil, more research needs to be done. According to the Universalist Validation Model presented by Herdman and colleagues (1998), it is necessary 6 steps for the Musical Communicativeness Scale to be validated in the Brazilian context. Out of these 6 steps, 4 steps have already been accomplished. They were: conceptual equivalence, carried out by André, Gomes and Loureiro, (2016) through the study of bibliographic revision and the equivalences of items, semantic and operational carried out by André et al. (2016). In the present manuscript, the equivalence of measurement through inter-rater reliability of the Musical Communicativeness Scale was verified.

5. CONCLUSIONS

The validation of the Musical Communicativeness Scale for the Brazilian context may contribute in the future to the clinical and research context which has been used since the 1960s. For this validation process to occur, studies have already been carried out by André, Gomes and Loureiro (2016) and André et al. (2016) to verify the conceptual, item, semantic and operational equivalence of the Musical Communicativeness Scale.

In this study, the equivalence of measurement through inter-rater reliability of the Musical Communicativeness Scale was carried out. It was observed from the Spearman test, moderate and strong correlations. By same means, it was found a moderate correlation in the body movement domain (ρ = 0.68) and strong correlations in the instrumental (ρ = 0.79), vocal (ρ = 0.87) and total musical ρ = 0.73). According to collaborating examiners, the summary manual is fully understandable and the Musical Communicativeness Scale could contribute to the Brazilian context. Further studies can be carried out to validate the Musical Communicativeness Scale in the Brazilian context.

6. REFERENCES

1. AIGEN, K. Cognitive and affective processes in music therapy with individuals with developmental delays: A preliminary model for contemporary Nordoff-Robbins practice. Music Therapy, v. 13, n. 1, p. 13–46, 1995.

2. AMERICAN PSYCHIATRIC ASSOCIATION. Manual diagnóstico e estatístico de transtornos mentais DSM - 5. 5. ed. Porto Alegre: Artmed, 2014.

3. ANDRE, A. M. et al. Tecnologia e atraso do desenvolvimento: relações com a musicoterapia1o nas nuvens... congresso de música. Anais...1o Nas Nuvens... Congresso Internacional de Música, 2015

4. ANDRÉ, A. M. et al. Equivalência de itens, semântica e operacional da versão brasileira da Escala Nordoff Robbins Comunicabilidade Musical 2o Nas Nuvens... Congresso Internacional de Música. Anais...2 o Nas Nuvens... Congresso de Música, 2016

5. ANDRE, A. M.; BATISTA, D. O. Análise psicométrica das Escalas Nordoff Robbins como instrumento de avaliação no atendimento musicoterapêutico de crianças autistas atendidas no Hospital das Clínicas da UFMG. Universidade Federal de Minas Gerais, 2014.

6.ANDRÉ, A. M.; GOMES, C. M. A.; LOUREIRO, C. M. V. Escalas Nordoff Robbins: uma revisão bibliográfica. Percepta- revista de Cognição Musical, v. 3, n. 2, p. 117–131, 2016.

7.AUSTRALIA, N.-R. M. T. Music therapy program for "at risk" students Nordoff-Robbins Music Therapy Australia and James Meehan High School.

8.BELL, A. P. et al. The Music Therapy Communication and Social Interaction Scale (MTCSI): Developing a New Nordoff-Robbins Scale and Examining Interrater Reliability. Music Therapy Perspectives, v. 32, n. 1, p. 61–70, 2014.

9.BERGMANN, T. Music Therapy for People with Autism Spectrum Disorder. New York: Oxford University Press, 2015.

10. BERGMANN, T. et al. Music in diagnostics: using musical interactional settings for diagnosing autism in adults with intellectual developmental disabilities. Nordic Journal of Music Therapy, p. 1–33, 2015.

11.BUNT, L. Music therapy with children: a complementary service to music education? British journal of music education, v. 20, n. 2, p. 179–195, 2003.

12.CALTABIANO, A. The impact of music therapy on the social behaviours of children with autism in a structured outdoor inclusive setting. JOUR. University of Sydney, 2010.

13.DANCEY, C. P.; REIDY, J. Estatística sem matemática para psicologia. 5. ed. Porto Alegre: Penso Editora, 2013.

14.FREIRE, M. H. Efeitos da Musicoterapia Improvisacional no tratamento de crianças com Transtorno do Espectro do Autismo. JOUR—Universidade Federal de Minas Gerais, 2014.

15.HALLAM, S.; CROSS, I.; THAUT, M. Oxford handbook of music psychology. Oxford University Press, 2008.

16.HERDMAN, M.; FOX-RUSHBY, J.; BADIA, X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. Quality of life Research, v. 7, n. 4, p. 323–335, 1998.

17.LOUREIRO, C. M. V. Musicoterapia na educação musical especial de portadores de atraso do desenvolvimento leve e moderado na rede regular de ensino. Universidade Federal de Minas Gerais, 2006.

18.MACKINLAY, E.; FORREST, D. Making sound waves: Diversity, unity, equityMaking sound waves: Diversity, unity, equity XVIII National Conference Proceedings. Anais...2011

19.MAHONEY, J. F. Interrater agreement on the nordoff-robbins evaluation scale i: client-therapist relationship in musical activity. Music and Medicine, v. 2, n. 1, p. 23–28, 2010.

20.NORDOFF, P.; ROBBINS, C.; MARCUS, D. Creative Music Therapy:Guide to Fostering Clinical Musicianship. 2. ed. New Hampshire: Barcelona Publishers, 2007.

21.RAHMAN, Y. A. Doctoral School Summer Conference June 2008 Institute of Education. Educate, v. 8, n. 1, p. 34–64, 2008.

22.ROBARTS, J. Z. Music therapy and adolescents with anorexia nervosa. Nordic Journal of Music Therapy, v. 9, n. 1, p. 3–12, 2000.

23.SAMPAIO, R. T. Avaliação da Sincronia Rítmica em Crianças com Transtorno do Espectro do Autismo em Atendimento Musicoterapêutico. JOUR. Universidade Federal de Minas Gerais, 2015.

24.SZWEDA, S. K. The effectiveness and influence of Vocal and Instrumental Improvisation in Music Therapy on children diagnosed with autism . Pilot Study . The journal of Education Culture and Society, p. 153–166, 2015.

25.WIGRAM, A. L. Music therapy assessment: psychological assessment without words. Psyke & Logos, v. 28, n. 1, p. 25, 2007.

26.WIGRAM, T.; GOLD, C. Music therapy in the assessment and treatment of autistic spectrum disorder: clinical application and research evidence. Child: care, health and development, v. 32, n. 5, p. 535–542, 2006.

27.WOOD, S. " The Matrix": A Model of Community Music Therapy ProcessesVoices: A World Forum for Music Therapy. Anais.2006.

It's not just the music: Impact of the environment and individual differences on music medicine in pain settings.

Claire Howlin
claire.howlin@ucdconnect.ie

Darragh Lynch
darraghkevinlynch@gmail.com

Suzanne Guerin
suzzane.guerin@ucd.ie

Brendan Rooney
brendan.rooney@ucd.ie
School of Psychology, University College Dublin, Ireland

ABSTRACT

Music listening interventions (MLIs) have the capacity to reduce the experience of pain in many contexts such as cancer, procedural pain and labour. Additionally, the effect size of MLIs in pain contexts has been shown to be moderate using meta-analyses, with patient chosen music shown to predict this outcome. However, meta-analyses have not identified the cognitive mechanisms underpinning the positive effects of MLIs, which coincides with inconsistencies in terms of how MLIs are delivered (eg. duration, frequency, style, genre, music selection, sound quality) and evaluated. In order to address these inconsistencies, the current study uses a qualitative approach to gain a deeper understanding of the existing literature in terms of the overlapping characteristics of MLIs using a systematic scoping literature review. The review identifies several music related factors integral to MLIs including, musical features, music with cultural relevance, and sound quality. Additionally, the review identifies several non-music related factors such as volume control, environmental comfort, direction from the researchers and individual differences in music engagement. These additional factors should be considered or controlled for more rigorously in future music intervention studies.

1. INTRODUCTION

Music interventions are currently being implemented in pain management contexts such as geriatrics, obstetrical care, neurological dysfunction and palliative care [1]. This is supported by several meta-analyses that demonstrate the overall impact of music interventions in pain management contexts [2 – 5]. However despite the overall success of music interventions, a major problem remains in terms of the inconsistency with which music listening interventions are delivered and evaluated [4, 6]. Wide variability has been documented in relation to the duration, frequency, style, genre, equipment, preparation, choice, rationale, personnel adopted [3]. As a result MLIs currently vary from simple tape recordings in the background of a clinic room, to a full orchestral performance in hospital lobbies [7]. Furthermore this variability in how interventions are delivered is reflected by large statistical heterogeneity in effect sizes of music intervention studies [2], indicating that some interventions are more effective than others, yet it is not clear where the differences lie.

A major distinction made in music interventions is between music therapy and music medicine [4]. The key difference being that music therapy sessions involve a music therapist specially trained in music theory, and MLI sessions do not. Indeed music medicine interventions do not require delivery from a trained music therapist or necessitate additional techniques yet deliver comparable results [3]. Curiously, in the absence of a therapeutic relationship, it is not clear how the positive effects of MLIs are mediated. Currently it is deemed of paramount importance to establish whether it is the therapist or the music itself that mediate the beneficial effects of music listening [8]. A recent development in MLI research is the role of personally chosen music over experimenter chosen music [3,6]. Initially it was thought that specific musical features, such as beats per minute, tempo or valence of music may elicit the beneficial effects of music interventions in the absence of a music therapist. However using meta-analysis to examine the contribution of featural components, determined that no specific feature of the music contributes to positive outcomes witnessed in music interventions. Additionally, it is increasingly recognised that different people often respond very differently to the same piece of music [9], and reciprocally different pieces of music can be used to get the same functional effect known as functional equivalence. This has been repeatedly demonstrated with one meta-analysis including 3731 cancer patients indicating that personally chosen music has a much larger impact on pain reduction compared to researcher chosen music [4].

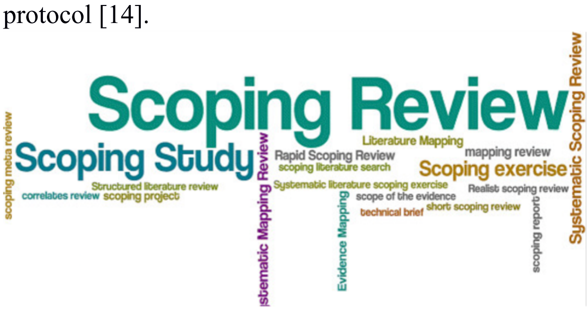
Taking these two points together this raises the broader question of what are the core components of a MLI [3,10,11]? Previous literature reviews have examined the effectiveness of music interventions in analgesic settings [3] but have failed to identify the specific mechanisms underpinning MLIs [3,10,11]. Accordingly music listening interventions need to be classified further in order to isolate the therapeutic component involved in MLI. At this point it is necessary to synthesise the factors that that have already been identified by researchers in a MLI context. This will help to structure future research aimed at identifying the cognitive mechanisms the mediate the analgesic effects of MLIs. Since meta-analyses have already examined the literature using a quantitative approach, a deeper consideration of the literature should be performed qualitatively using a systematic scoping review. This will enable a thorough examination of the literature for the current proposed models of the underlying rationale in music interventions.

Aim and Objectives

The aim of this review is to examine what is known from the existing literature about the characteristics of MLIs used in analgesic settings. The primary objective of this scoping review is to describe the characteristics of MLIs when used in different analgesic settings. This will help to identify what is known from the existing literature about the characteristics of MLIs and will help to generate testable hypotheses which can then be tested to isolate the underlying mechanisms of MLIs experimentally.

2. METHOD

This study comprises a scoping literature review using a descriptive-analytical method in line with previous guidelines [12-13]. This review involves five stages, (i) development of the research question, (ii) identifying relevant studies, (iii) study selection (iv) charting the data, and finally, (iv) summarizing and reporting the results. The research team conducted the review according to a previously published scoping review protocol [14].



Search Strategy

2.1.1 Electronic Databases

A systematic search was conducted using a subgroup of the electronic databases included in the study by Lee (2016); EBSCO Music Index and RILM, and EBSCOhost Psychology and Behavioural Sciences Collection, CINAHL Plus (EBSCO), Pubmed (Medline). These databases were chosen as they include music listening studies in a range of healthcare contexts from a multitude of health professionals, including music therapists. Additional hand searching of music therapy journals will not be performed, given that these journals are included in the Music Index RILM database and will be searched systematically. It was considered unnecessary to locate every available study, because the results of a conceptual synthesis will not change if ten rather than five studies contain the same concept but will depend on the range of concepts found in the studies, their context and whether they are in agreement or not [15]. Accordingly, this review focused on ‘conceptual saturation’ with maximum variability and in essence designing the resulting set of studies to be heterogenous rather than the homogeneity seen in meta-analyses.

2.1.2 Developing the search string

An initial search strategy was devised that was later refined in the light of early pilot searches. This included specific search parameters that allowed a targeted search without using extensive limiters that may reduce the quality of the search. Three search terms; specifically, music, listening and pain (see table 1), were used as the core terms to develop a three-pronged search string, along with their variations and MESH terms. First direct synonyms of each term known to the research team were added to the string, which was then refined by the research librarian. Second each term was used to find additional keywords or MESH terms specific to each database, and added to the string using the database search builders. Thirdly, each string was then searched for separately, before combining the searches. This procedure was followed, using identical limiters, expanders and date range for each database to ensure consistency in the resultant search string produced. Limiters included studies published in the English language, since this was the only language of the researchers, and studies published in peer review journals. Additionally, the search was limited to include articles with available abstracts, published between January 1st 2006 and June 15th 2017. This was considered necessary by the research team to ensure that the review could be completed within the planned time frame available.

Music	music OR song OR songs OR singing OR melody OR melodies OR tune OR opera OR concert OR musical
Listening	listening OR hearing
Pain	pain OR ‘Pain management’ OR neuralgia OR soreness OR ache OR tenderness OR headache OR myalgia OR analgesia

Table 1. Search String

2.2 Study Selection

2.2.1 Document Retrieval

A total of 474 documents were retrieved across the four databases (see figure). Two independent reviewers agreed to exclude 175 articles based on the abstract and title content and include 104 articles for full text review. Based on the two reviewers' decisions Krippendorf's alpha was calculated as 0.6754, indicating good agreement.

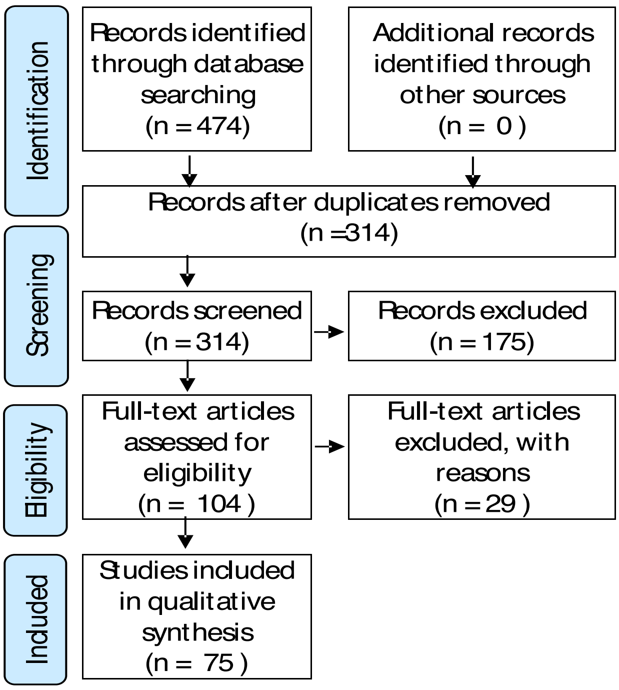


Figure 1.Prisma Flow Diagram of Article Selection

2.2.2 Quality Assessment and Full Text Review

Each article was reviewed in full by at least one author to assess the quality of the article and to deduce suitability for inclusion. A minimum quality assessment was conducted to ensure that all studies included met the appraisal criteria of the five 'fatal flaws' (see table 3; 16]. Each article was then assessed to ascertain whether it met the conditions of the inclusion and exclusion criteria outlined in table 2. Out of the 104 articles that were reviewed in full, 75 articles were included in the qualitative synthesis (See Appendix A)

Inclusion Criteria

1. Evaluate pain experience in wakeful patients in the presence of music listening
2. Are applied in a healthcare or laboratory setting
3. Use receptive music interventions where the person actively or passively observes music, either live or recorded for at least one part of the research study. This can be in conjunction with other activities, i.e. visualisation or dance performance, but this is the minimum requirement for an MLI.
4. Measure pain by either self-report measures or reduced pharmacological analgesic requirements.
5. Are empirical studies with primary data collection, including randomised, quasi-randomised, one armed trials, and qualitative accounts of MLIs

Exclusion criteria

6. Only focus on the process of making music
7. Focus on pain or injuries caused by performing music
8. Primarily examine issues of hearing loss, hearing disorders or other issues of aural health
9. Focus on the use of music as part of a patient information material

Table 2. Inclusion and Exclusion criteria

2.3 Charting the Data

2.3.1 Data Extraction

In order to extract the data systematically and consistently a data extraction form was developed in Microsoft excel by the research team in line with previous recommendations. The form consisted of a series of data points deemed necessary to answer the research question posed in the systematic review protocol. The extraction tool was piloted by two members of the research team on a random sample of 10 papers and was then reviewed to determine whether the data being extracted was consistent with the research question and purpose. The final form had a semi-structured format (see figure 2). Each variable was allocated one field in the data extraction form, and fields were expanded to accommodate text if required. Partial double data extraction was completed on a 10% random sample, and single data extraction was completed on the remainder of the papers.

2.3.2 Thematic Synthesis

A thematic synthesis strategy was developed, in line with previous guidelines [15] involving three stages (1) the reviewer coded the text, line by line according to its meaning or content, using EPPI reviewer 4 software. There was no limit imposed on the number of codes that could arise in the text. (2) Descriptive themes were developed by grouping and combining codes based on similarity to develop analytic themes. Sometimes this meant incorporating two similar codes with a slightly different meaning to produce a descriptive theme, and the overall code would be relabeled to reflect this. Codes that only applied to a very small number of studies (i.e. three or less) were re-coded into larger analytical themes where possible using the process above. (3) Descriptive Themes were then grouped together to form analytical themes. In some instances, this meant dismantling analytical themes and re-coding them into separate analytical themes for example 'cultural context' was re-coded into 'culture' or 'context' which were two themes that already existed. This helped to draw boundaries between the analytical themes in instances where descriptive themes overlapped two categories.

3. RESULTS

Inductive thematic synthesis identified the core components of MLIs identified in the existing literature. Themes were grouped in terms of music related factors; such as musical features, music with cultural relevance, and sound quality of music, and non-music related factors such as volume control, environmental comfort, direction from the researchers and individual differences in music engagement.

3.1 Music related factors

3.1.1 Musical features

Several studies focussed on the specific features of beats per minute and rhythm in the music used in the music intervention, particularly in studies where researchers selected the music on behalf of the participants. Many studies attribute the success or failure to musical features in the music such as the tempo or rhythm, however several conflicting accounts emerged in terms of the role these features play. Some studies emphasised that they used music with a slow tempo such as 60-80 beats per minute, with gentle rhythms, which was considered to be universally relaxing, and assumed to be beneficial in pain relief contexts.

"The music group listened to sedative instrumental music with a slow tempo, with 60-80 beats/minute, which had been reported to be calming and relaxing.
Björkman 2013

In contrast other studies emphasised that they intentionally used strong rhythms, with the intention that these rhythms could help to facilitate *active* listening or *active* pushing.

"*After two hours music was changed by midwife, type and volume were changed to be more rhythmic in order to help with active pushing."*
Simavli 2014

These two conflicting accounts of the role of rhythm or tempo may be reconcilable if music is considered to have different roles in different pain contexts.

"*For the active listening, the same musical pieces were available divided into two subcategories depending on the choice of patient: calm or dynamic (calm music comprised pieces with a tempo of 30-60 beats/ minute, dynamic music comprised pieces with a tempo of 60- 120 beats/minute)."*
Mercadle 2015

3.1.2 Music with cultural relevance

Another aspect of music used in MLIs, was that music of a style native to the listener was often used providing a greater opportunity for the music to have cultural relevance to the listener. This was seen both in studies where participant selected music was chosen and studies that used experimenter selected music.

"*Music was different types of the most popular and classic music in Finland, 2000 songs, from which people chose their favourite."*
Vaajoki 2012

It is likely that using traditional or folk music from a person's own culture might bear significant relevance or meaning to them. It may be a song that they've learned in school or know from a family tradition or perhaps remember hearing at a birthday or as a lullaby. From this perspective it seems that music with cultural relevance represents music with a specific sense of meaning to the listener.

"MG were allowed to listen to music of their choice from among six types, popular psalm songs based on carnatic classical ragas, classical music, devotional songs, folk songs, soft instrumental music, and bioacoustics (a soothing mixture of soft instrumental music along with nature sounds)."
Harikumar 2004

3.1.3 Sound Quality

While not mentioned in every study examined, a considerable number of studies made some reference to the type of equipment used to deliver the music intervention, and sometimes the associated sound quality.

"Music was played over MRI compatible headphones." "A soundcheck was done with a designated classical music piece to ensure that the quality and loudness were comfortable for each participant.
Dobek 2014

This equipment used across studies varied enormously across studies from high quality personalised headphones attached to patient controlled mp3 players to nearby loudspeakers. The wide disparity of equipment used would also likely introduce different levels of sound quality in each music listening experience.

"The patients listened to the stimuli using Senheiser HD 205 headphones with passive attenuation ambient noise"

Garza-Villareal 2014

"Listened to music of their choice from a selection of classical, instrumental, or rock music that was transmitted by nearby loudspeakers, as we were afraid that using headphones would cause a lack of control and increase anxiety levels."
Cakmak 2017

Additionally, it seems that there would be a big difference in attention levels required to focus on music coming from personal headphones, compared to a nearby loudspeaker, where it is still possible to hear the sounds of healthcare staff and equipment. In many cases practical concerns such as communication with staff or compatibility with the procedure were also emphasised, indicating that music interventions are often integrated with routine medical care.

3.2 Non-Music related factors

3.2.1 Volume Control

The issue of control of the volume level was identified in several studies. This is different from the actual loudness of the music and corresponds with patients or participants ability to regulate or change the volume during the music experience as they so wished.

'The patients could control the volume themselves"
Björkman 2013

There was considerable variation across studies in relation to whether subjects could control the volume during the experiment or whether it was set at a fixed level. Or whether participants were allowed to initially choose the volume level, which remained fixed for the remainder of the experiment.

"*The music volume was set at a comfortable level for subjects according to their preferred volumes. While in the POR, the music volume was controlled by the investigator."*
Chen 2015

In studies where speakers were used, it seemed that participants had little say in the volume of music played in the healthcare environment, and volume control tended to be absent in these studies. Where headphones were used volume tended to be dictated by the patients themselves depending on what they preferred or found comfortable with no specific volume advocated.

'*Volume level was self- selected by participants as appropriate to personal taste."*
Finlay 2015

This creates two very different scenarios in terms of volume control, one patient centered approach where they are put in control of the music, and another where control of the music is maintained or adjusted based on the needs or desires of the healthcare team.

3.2.2 Environmental Comfort

Many interventions described the ways in which they changed or controlled the environment during the music intervention, to enhance comfort and reduce competing noises or distractions. In many instances hospitals were also often characterized as noisy, unfamiliar and stressful environments, and a great emphasis was placed on the benefits of silence in this context as well as the benefit of music.

"This study was conducted between 3 to 6 P.M. of the day after surgery because the traffic of hospital staff was lower and patients routine care was completed and other people were less likely to be present and to interfere with the process."
Jafari 2012

In home or lab settings, emphasis was often placed in finding a comfortable position, or sitting in a comfortable chair. Additionally, emphasis was placed on avoiding other routine activities so that music was the sole focus of their engagement, without competing demands.

"These patients were asked to listen to the music for at least 25 min twice a day in available morning and evening hours while sitting on a comfortable chair in a quiet environment, being not overly hungry or full and doing no other activities."
Alparslan 2016

In hospital settings this involved reducing extraneous noise and preventing disruptions to the interventions by visitors or hospital staff.

*"In preparation for the intervention, the researcher prepared the patient and environment well (turned o*ff *cell phones, shut the door, and eliminated distractions), and tried to keep the patients from being disturbed."*
Liu 2015

This specification on finding a comfortable or quiet space, seems to emphasize the often busy or hectic nature of a hospital environment, or indeed a person's day to day routine and the need to provide a more suitable environment to aid recovery or enhance wellness.

3.2.3 Individual Differences in music engagement

Individual differences tended to be characterized in terms of demographic information, individuals cognitive style, and musicality. Demographic differences between individuals were mentioned across studies including, gender, culture, or age, and reflect the fact that these demographic differences have been shown to impact music engagement in other contexts.

"The choice of music is important, and the response to music is influenced by earlier experience of music as well as by gender, age, culture, mood, and attitude."
Björkman 2013

Musicality of participants was also assessed in many studies in terms of participants musical behavior or music engagement or musical training, however most studies did not assess musicality or music training of participants using a previously validated instrument. It was recognized in many studies that participants attitude or affinity towards music was likely to impact the effectiveness of the intervention, acknowledging that people who dislike music are less likely to benefit from an MLI.

"*It is likely that the patients who enrolled in this study already had a special affinity for music."*
Bradt 2015

The emphasis on individual cognitive style or capacity was also emphasized in several studies, with one particular study emphasizing the role of empathizing and sympathising in musical interactions.

"Empathizers are may be attracted to the emotional content of the music, whereas the Systemizers may be attracted to musicianship and performance level"
Garza-Villareal 2012

Accounting for individual differences in music interventions accounts for the fact that different people have different life experiences and intrinsic value sets that impact how they interact with music. It also suggests that these differences underpin the meaning that people derive from music and the importance music plays in their day to day life. In order to understand how people, interact with music interventions, it is important to take these individual life experiences and value sets into account.

"Providing culturally appropriate music to the mixed study sample of Malay and Chinese may have enhanced effectiveness of the music intervention."
Hook 2008

3.2.4 Directions from researcher

Instructions from the researcher team or health professional were a prominent feature of many MLIs. Participants were often given instructions on how to focus on the music, how much music to listen to or factors to consider when choosing their own music to listen to.

"*A nurse instructed patients in the intervention group on how to listen to the music when the patient was on the colonoscopy table"*
Björkman 2013

The presence or absence of such instructions are likely to affect how participants engage with the music and may also affect the overall outcome of the intervention. Additionally, participants were also given guidance on when to listen to the music, and in some cases were provided with additional reminders to encourage more regular listening of music in a home context.

"Participants were encouraged to use it as much as possible that evening and during the first 2 days, with reminders four times a day."
Good 2010

Although not every study outlined specific instructions or guidance provided, many studies did so, regardless of whether the actual music sample music was selected by the researcher, or the participant. This emphasizes differing levels of direction from the researcher, depending on the level of instructions, and differing degrees of freedom in terms of patients' choice in their self-selected music.

"*Patients listened to self-chosen music (curated by experimenter after phone call) that was highly pleasant and slower than 120 beats-per-minute."*
Garza-Villareal 2014

4. DISCUSSION

The aim of this review was to examine the characteristics of MLIs used in analgesic settings, using a qualitative approach to gain a deeper understanding of the existing literature. Using thematic synthesis this review describes the musical aspects of MLIs such as sound quality, musical features, and music with cultural relevance, as well as the less recognized non-musical features of music interventions such as volume control, environmental comfort, individual differences, directions from the researcher.

This review identified the fact that many researchers alter the environment in which the music intervention takes place, ensuring that comfort is maximized and potential disruptions or competing distractors are minimized. Thus, it seems controlling environmental noise and creating a comfortable environment during the music listening experience plays an integral part in MLI delivery. This is in line with previous studies that have demonstrated the role of social context in music listening [18]. It stands to reason that the social context of the hospital or lab environment will also impact the outcome of the music interventions. It also contextualizes music interventions within a holistic framework, emphasizing the person-music-environment interaction in mediating the outcomes of MLIs [28]. Further research should examine the role of the healthcare environment more closely to investigate the factors that either facilitate or disrupt positive outcomes MLIs.

Accounting for individual differences in music interventions accounts for the fact that different people have different life experiences and intrinsic value sets that impact how they interact with music [9]. People vary in their musical background and experience which may impact how they engage with music in MLIs. Specifically, people who don't like music, are probably less likely to benefit from a music intervention. Often people have different emotional reactions to the same piece of music; highlighting that individual cognitive factors mediate between acoustical perception and psychological response to music [19]. Cognitive interpretation in terms of memory and personality factors and learning in terms of previous cultural experience also influence emotional responses to music [20-21]. This is particularly important to consider in the context of music with cultural relevance, which may be more likely to trigger related episodic memories or evaluative associations. It also suggests that these differences underpin the meaning that people derive from music and the importance music plays in their day to day life. In order to understand how people cognitively engage with music interventions, it is important to take these individual life experiences, cultural background and value sets into account. A greater emphasis needs to be placed on the fact that aesthetic responses to music involve an interaction between the stimulus and the listener's mental state rather than treating music as a stimulus that evokes certain responses [22]. While this is well recognized more broadly in music psychology, it seems it needs to be given further consideration in an MLI for pain context, with increased focus on the role of musicality and individual musical emotions in mediating the benefits of MLIs.

Many of the studies that were reviewed, emphasized the fact that they carefully chose music or encouraged patients to choose music with a slow tempo, usually 60-80bpm. This recommendation is in line with models of entrainment and based on the idea that respiration rates vary with music tempo. Such that a relatively slow musical tempo would lead to slower respiration and subsequent relaxation coupled with a reduction in perceived pain. While logical, this recommendation was initially based on the results of only one study, and this assumption has now been repudiated by meta-analysis [3-4]. Moreover, it is completely at odds with the fact that self-selected music considered more effective than experimenter chosen music. Experimenter's tend to select music with a slow tempo and an absence of strong rhythms. This is the opposite to what would be considered high-groove music, with groove perceived more in music at tempos between 100-120 bpm [24] and depends on the rhythmic pattern of the music [25]. In contrast people tend to prefer music that is perceived as being high-groove [26] and more people may have selected high groove music when given the freedom to choose their own music. Groove is defined as wanting to move some part of the body and is associated with excitability of the motor cortex [27] indicating the physical experience involved in listening to high groove music. Since groove accounts for how the auditory experience manifests as physical motivation, more emphasis should be placed on the role of groove in MLIs in a pain context.

Previous studies emphasize the role of patient choice in the music listening process as it is considered to enhance the patients' locus of control, and subsequent ability to cope with their pain [28]. Contrary, within the context of patient selected music, some researchers maintain a certain level of control by directing patients to music of certain types, which in turn undermines the concept of truly patient selected music from an unlimited choice. This is despite the fact that the there is no evidence to support the particular types of music being advocated, or the methods of engagement being encouraged. It must also be acknowledged that while, reminders to use music interventions might be a useful way to maximize participation, directions on how to listen may not always be helpful from the participants perspective. Perhaps expert listeners already have optimal strategies for music listening that could be disrupted following instructions, whereas less experienced listeners might benefit to a greater degree from researchers' direction. At this stage the optimal level of involvement from health professionals in MLIs not clear. Additionally, it seems that the specific music choice is not the only level at which the patient can control in the music intervention, with opportunities for enhanced patient involvement by giving patients control over the volume, as well as ability to change the songs as required.

5. CONCLUSIONS

Using thematic synthesis this review describes the characteristics of musical listening interventions in pain contexts in terms of both musical factors and non-musical factors. The review highlights that it is important to consider the comfort of the patient during MLIs whether they are at home or in hospital and to minimize any potential disruptions. It also emphasizes the role of individual differences such as musicality, personal and cultural associations with music, and that should be recognized more widely in MLI literature as they are the basis of music engagement and musical emotions. The review reflects previous studies that emphasize the importance of self-selected music and identifies additional levels at which the patient can exert control in the MLI, with opportunities for enhanced patient involvement in volume control and personalized listening devices. Finally, this review suggests perceived groove as a potential difference between experimenter and participant chosen music, since experimenters tend to choose music outside of the optimal tempo range for groove to occur. These hypotheses should be explored further through experimental design to isolate the active components of MLIs.

Acknowledgments

Support to travel and attend the SYSMUS 18 conference was provided by a SEMPRE conference travel award.

This project was funded by a government of Ireland postgraduate scholarship from the Irish research council grant number R17777.

6. REFERENCES

[1] S. Mainka, R. Spingte, and & M. Thaut. Music Therapy in Medical and neurological rehabilitation settings. In: Oxford Handbook of Music Psychology, S. Hallam, I. Cross, & M. Thaut. (Eds.), 2nd edn. Oxford University Press, 2016, pp. 857-873

[2] M.S. Cepeda, D. B. Carr, J. Lau, and H. Alvarez "Music for pain relief," in *Cochrane Database of Systematic Reviews* 2006, *2*(2).

[3] J. H. Lee, J. H., "The effects of music on pain". *Journal of Music Therapy Advance Access*, 2016, pp.1-48.

[4] J., Bradt, C., Dileo, L., Magill, and A. Teague, "Music interventions for improving psychological and physical outcomes in cancer patients" in *Cochrane Database of Systematic Reviews,* (2016), 8

[5] E. A. Garza-Villareal, V. Pando, P. Vuust, P., and C. Parsons (2017). "*Music-induced analgesia in chronic pain conditions: a systematic review and meta-analysis*", Pain Physician, 2017 pp. 597-610.

[6] H.F. Tsai, Y.R. Chen, M.H. Chung, Y. M. Liao, M.J. Chi, C.C. Chang, and K.R. Chou, "Effectiveness of music intervention in ameliorating cancer patients' anxiety, depression, pain, and fatigue: A meta-analysis", *Cancer nursing*, 2015, *37*(6), E35-E50.

[7] A.C. Miu, and F. R. Balteş, "Empathy manipulation impacts music-induced emotions: a psychophysiological study on opera", *PloS one*, 2012 7(1), e30618.

[8] P. Horden, Musical Solutions: past and present in music therapy. In P. Horden (Ed) *Music as medicine: the history of music therapy since antiquity*, 2016 Routledge

[9] S. Swaminathan, and G. Schellenberg, G., "Current Emotion Research in Music Psychology." *Emotion Review, 2015 7*(2), pp. 189-197.

[10] A. Keenan, and J.K. Keithley, "Integrative Review: Effects of Music on Cancer Pain in Adults", in *Oncology nursing forum,* 2015 (Vol. 42).

[11] P., Krishnaswamy, and S. Nair, "Effect of music therapy on pain and anxiety levels of cancer patients: A pilot study", in *Indian journal of palliative care*, 2016 *22*(3), pp. 307.

[12] H., Arksey, and L. O. Malley "Scoping Studies: Towards a Methodological Framework," in *The International Journal of Social Research Methodology, 8*(1), (2005), pp.19–32

[13] D. Levac, H, Colquhoun, and K.K. O'Brien, "Scoping studies: advancing the methodology", in *Implementation Science,* 2010 5(69), 1-9.

[14] C. Howlin, S. Guerin, B. Rooney, "Investigating the cognitive mechanisms of music listening interventions in pain management settings: A Scoping Review Protocol", in *European Journal for Patient Centrered Healthcare,* In Press.

[15] J. Thomas, and A. Harden. "Methods for the thematic synthesis of qualitative research in systematic reviews", in BMC Medical Research Methodology, 2008 8(45), pp. 1-10.

[16] M. Dixon-Woods, A. Sutton, R. Shaw, T. Miller, J. Smith, B. Young, and D. Jones, "Appraising qualitative research for inclusion in systematic reviews: a quantitative and qualitative comparison of three methods" in *Journal of health services research & policy*, 2008 *12*(1), pp. 42-47.

[17] H.L. Colquhoun, D. Levac, K. K. O'Brien, S. Straus, A. C. Tricco, L. Perrier, and D. Moher, "Scoping reviews: time for clarity in definition, methods, and reporting". *Journal of clinical epidemiology*, 2014 *67*(12), pp. 1291-1294.

[18] S. Liljeström, P.N. Juslin, and D. Västfjäll, "Experimental evidence of the roles of music choice, social context, and listener personality in emotional reactions to music", in *Psychology of Music,* 2014 *41*(5), pp. 579–599.

[19] S. Koelsch *Brain and Music*. Wiley-Blackwell: 2012

[20] P.N. Juslin, and D. Västfjäll, D., "Emotional responses to music: The need to consider underlying mechanisms", *Behavioral and brain sciences*, 2008 *31*(5), pp. 559-575.

[21] P.N. Juslin, L. Harmat, and T. Eerola, T. (2013). "What makes music emotionally significant? Exploring the underlying mechanisms", in *Psychology of Music*, 2013 pp.1-25.

[22] D. Huron, *Sweet Anticipation: Music and the psychology of expectation*. MIT press 2006

[23] U. Nilsson, U. , "The anxiety and pain reducing effects of music interventions: a systematic review", in *Association of Perioperative Registered Nurses Journal*, 2008 87(4), pp. 780- 807.

[24] E. Takahide, A. Marui, S. Kawase, and P.E. Keller, "Optimal Tempo for Groove: Its Relation to

Directions of Body Movement and Japanese nori." *Frontiers in Psychology*, (2018): pp.462.

[25] M.A. Witek, E.F. Clarke, M., Wallentin, M.L. Kringelbach, and P. Vuust, P. (2014). “Syncopation, body-movement and pleasure in groove music”, in *PLoS ONE* (2014) 9:e94446.

[26] P. Janata, S.T. Tomic, J.M. Haberman, “Sensorimotor coupling in music and the psychology of the groove”, in *Journal of Experimental Psychology: General,* 2012 141: pp 54–75

[27] J. Stupacher, M. J. Hove, G. Novembre, S. Schütz-Bosbach, and P.E. Keller, “Musical groove modulates motor cortex excitability: a TMS investigation”, in *Brain Cognition,* 2013 *82*, pp. 127–136.

[28] K. A. Finlay, and J. Rogers, "Maximizing self-care through familiarity: The role of practice effects in enhancing music listening and progressive muscle relaxation for pain management", in *Psychology of Music* 43, no. 4 2015 pp. 511-529.

CHORO MARANHENSE. A BIOGRAPHICAL AND ETHNOGRAPHICAL STUDY OF AN IDENTITY IN MARANHÃO.

Raimundo João Matos Costa Neto
Universidade Estadual do Maranhão, Brazil
rjmcneto@gmail.com

Peter Ninaus
University of Music in Graz, Austria
peter.ninaus@cdx.at

ABSTRACT

The expression of a different identity within the borders of Brazil is natural. Political centralism and the picture to the exterior give a wrong idea of local cultural diversity. In the center of the cultural diversity in Maranhão next to Bumba-Meu-Boi or Reggae there is the tradition of the “Choro Maranhense” which emphasizes connections from the past to the present. This paper shows two perspectives on the practice of music of this area. The biographical part points out the work of the local composers and the independent development of another kind of Choro compared with the Tradition in Rio de Janeiro. The second part analyses the characteristics, the ideas behind and the practice of this type of music. Besides the globalization there is still a need of cultural independence and local tradition counts against internet broadcasting. “Choro Maranhense” is a good example for the power of this kind of local tradition. The aim of the work is to make a systematical research of a problematic and non-touristic region. The methods of the studies include analyses from archives and a field-study of the local practice, the social, economic and educational status of the musicians and the question of personal and local identity.

1. INTRODUCTION

“Choro” is an expression naming a small group of musicians playing with guitars, ukulele, pandeiro and solo-instrument. This paper investigates an urban musical practice with systematical, historical and ethnographical methods. The cultural complexity of the State Maranhão in the North-East of Brazil roots from its historical, commercial and geographical being. This paper shows an analytical point of view of the biography of composers and the musical practice in the capital São Luís. It combines an internal and an external view.

The expression of a different identity within the borders of Brazil is natural. Political centralism “Rio-centrism” and the picture to the exterior give a wrong idea of local cultural diversity. “Choro Maranhense” roots from a music tradition of the 19th century as a melting pot with local rhythms of Maranhão.

What is Choro (Brazilian folk music)? Where and when did the choro come? What are the first definitions and characteristics of choro? Who were the first choro musicians?

The resolutions for these questions can be found in the works written on the History of Brazilian Popular Music (EMB, 1977, KIEFER, 1976, SEVERIANO, 2008, TINHORÃO, 1998, VASCONCELOS, 1977) and Choro’s works (ARAGÃO, 2013; ARANHA, 2012; CAZES, 1998; DINIZ, 2008a-2008b; PINTO, 2009; VASCONCELOS,1984). In these works, the authors determine a place of origin, Rio de Janeiro; a chronological cut, approximately in 1870; and define the choro as first an instrumental ensemble, later as an Brazilian way of performing the songs brought from Europe and finally as a musical genre. As we can observe in the following comments:

The first choro music groups appeared around 1880 in Rio de Janeiro (...) the choro music was the feature that used the popular musician to perform, in its way, imported music, which was consumed from the half 19th century, in the halls and dances of the high society (EMB, 1977: 192).

Around 1870 (...) appears, in Rio de Janeiro, the choro music in the beginning was not really a genre of music, but the appointment of an instrumental ensemble, and soon a Brazilian way of playing European dance genres in vogue (valsas, polcas, xotes, mazurcas and cuadrilhas) (VASCONCELOS, 1977, 1984).

If I had to point out a date for the beginning of Choro's history, I would not hesitate to give the month of July, 1845, when the polka was first danced in the Teatro São Pedro (CAZES, 1998: 19).

Choro is a son of the city of São Sebastião do Rio de Janeiro. The popular instrumentalists known as chorões, appeared around 1870 (...) music that emerged from the fusion of the lundu, rhythm of accent to percussion, with European genres (DINIZ, 2001: 14).

In Rio de Janeiro at the end of the 19th century, choro musicians played in halls, parties and residences, bringing the novelties of Europe and mixing them with Brazilian rhythms of African origin that led to the new popular music (DOURADO, 2004: 79).

Artists like Joaquim Callado, Chiquinha Gonzaga and several others began sprinkling a different rhythm in the popular European genres that arrived here (Rio de Janeiro) (ARANHA, 2012: 7).

Based on the "Rio-centric" discourse observed in the historiographic accounts about the early stages of crying, it is possible to determine two fundamental factors for this genealogical construction. The first would be related to the popularization of dances with European theater companies. The European dances extrapolated the limits of the theater and the operas and invaded the halls, streets and homes of Rio de Janeiro from the middle of the century. XIX. The second would be, because of the first, a savage way of playing the musical genres homonymous to these dances, such as the waltz, the gang, the mazurka and especially the polka. Considering these factors, we can conclude that the theatrical productions and the operas brought from Europe were fundamental for the formation of the choro in Rio de Janeiro. However, the influence of European theater companies was not only limited to the Carioca scene at the time, since at the same time two theaters located in the northeast of the country also received these companies, Teatro São Luís, founded in Maranhão approximately in 1817 and Teatro Santa Isabel, founded in Pernambuco in 1850 (GOUVEIA NETO, 2009: 3). In this sense, would not it be possible to think about the occurrence of these two factors also in the scenarios of Maranhão, Pernambuco and probably in other urban centers? And, consequently, would not there be several "ways" to play the musical pieces brought from Europe? Therefore, it is proposed a discussion about the probable existence of multiple processes of formation of the choro occurred simultaneously in Brazil. The guiding thread of this discussion is the comparison between the musical and "paramusical" contexts (scenarios, actions, social habitat, etc.) of the cities São Luís (MA) and Rio de Janeiro (RJ), in the chronological cut that comprises the end of the 19th century to the middle of the 20th century. These processes of formation are related to the dynamics that involves the first definitions of the term choro in the historiography of the Brazilian music referring to the 19th century (instrumental group, meetings of musicians or festive life and way of playing the European dances) and some representations of the term choro in the 20th century (style of composition, "roda de choro" and regional group). The verification of similarities between the representations of choro in the scenarios of Carioca and Maranhão illuminate the possibilities of thinking about concomitant processes of formation of choro in different urban centers.

2. MATERIAL, QUESTIONS AND METHODS

Scores of Choro-related Pieces in the João Mohana Collection are the historical base for editing and compared analysis to music from Rio de Janeiro of the end of the 19^{th} century. As well as interviews in São Luís with musicians, teachers and scholars are giving a picture of the aim and characteristics of this music.

What is Choro (Maranhense)? Why is it important for Maranhão? Is it a new popular wave?

First, to solve the question of independent Choro-literature in Maranhão, it was necessary to collect and categorize the historic material. The second part was the participant observation to collect data about the practice and tradition. The final part consists of interviews of people of this community.

The results shown in this paper are the intermediate results combining two works. The goal is to reach systematic methods to translate the insider-view with the outsider-view. Even the categorization of musica popular in popular music sounds right, but it isn't. The anthropological sight of nomenclature will also be considered as well as the terminology for reader outside of Brazil.

3. RESULTS

A "Rio-centric" development of Choro couldn't be proofed. The popularity of the European dances of the 19th century, such as the waltz, the quadrille, the mazurka and especially the polka were provided by theater companies all over Brazil. Places like the "Teatro São Luís", founded in Maranhão approximately in 1817 and Teatro Santa Isabel, founded in Pernambuco in 1850 (GOUVEIA NETO, 2009: 3) are responsible for establishing a Choro-scene in the northeast of Brazil parallel to Rio de Janeiro. In this sense it is possible to have a simultaneous formation processes of Choro in Brazil. The term "Choro" in the historiography of the Brazilian music referring to the 19th century (instrumental group, meetings of musicians or festive life and way of playing the European dances) and some representations in the 20th century (style of composition, "roda de choro" and regional groups).
Approximately 70 pieces of the João Mohana Collection related to the choro repertoire were digitized, of which 14 were edited. There are 12 choro-composers in Mohana collection with about 900 works.
. The table below presents 12 choro composers listed from the research in this Collection:

Choro Composers	Numbers of musical pieces
Paulo Almeida	31
*Adelman Corrêa	36

Ignácio Billio	37
Onofre Fernandes	44
João de Deus Serra	45
*Pedro Gromwell	56
Antônio O. Beckman	57
Catulo da Paixão Cearense	63
*Alexandre Rayol	76
*Sebastião Pinto	142
Henrique Ciríaco F.	147
*Othon G. da Rocha	165

The composers marked with "*" also have pieces considered erudite, a fact that proves, more than the versatility of the Maranhão musicians, the fluidity of the musical contexts and the richness of the repertoire already existing at that time in Maranhão. The number of musical genres and the quantity of pieces found in the João Mohana's Collection are listed in the table below.

Musical Genres	Numbers of musical pieces
Choro	32
Maxixe	9
Polka	124
Samba	102
Schottisch	64
Tango	99
Waltz	554

In this way, we can consider an expressive number of musical works related to the repertoire of choro in the musical production of Maranhão found in the Mohana's Collection (SECMA,1997).

Some pieces were considered as classical compositions, because of the used composing technique and the richness of the repertoire already existing at that time in Maranhão (COSTA NETO, 2015).

4. ETHNOGRAPHIC RESULTS OF THE "MARANHENSE IDENTITY"

This paper shows two perspectives on this musical practice. The biographical part points out that the repertoire in Maranhão has its musical characteristics. Besides the globalization there is still a need of cultural independence and local tradition which counts against internet broadcasting. In 1997 the city of São Luís got the UNESCO title of a world cultural heritage. This year also became an impulse for the self-consciousness of the region. This movement is the expression against "Rio-centrism" and "South-North divide" in Brazil.
There is also a recognizable change in oral tradition. While many pieces of the modern repertoire still are not transcribed, there also exists the new form of learning pieces with internet steaming. The need of scores is still not so important than learning by listening.

5. DISCUSSION

This music tradition shows a development of "identity" without nationalism. In a poly-cultural society there is space for minor identities without denying the whole. A small music tradition can develop with the motor of a personal identification. Motivation is stronger than sufficient financial support.
The new Choro-tradition, like in the past, is mainly an oral tradition. But the new ways are supported by internet-portals like YouTube. In this way there is also the danger of one-way tradition without personal varieties or the risk to become captured in a new way of (cultural) colonization via internet.
This research also shows the change of the "field" in music studies and encourages to extend common views.

6. REFERENCES

ARAGÃO, Pedro. O baú do Animal: Alexandre Gonçalves Pinto e o choro. Rio de Janeiro: Folha Seca, 2013.
ARANHA, Carla. Chorinho brasileiro: como tudo começou. São Paulo: DBA Artes Gráficas, 2012.
CAZES, Henrique. Choro: do quintal ao municipal. Rio de Janeiro: Ed. 34, 1998.
COSTA NETO, Raimundo João Matos. E tem choro no Maranhão? Subsídios históricos e musicológicos para um processo de formação do choro no Maranhão entre o final do séc. XIX e meados do séc. XX. Dissertação (mestrado em Música) – Universidade Federal de Minas Gerais, Escola de Música, 2015.
DINIZ, André. Almanaque do Choro: a história do chorinho, o que ouvir, o que ler, onde curtir. Rio de Janeiro: Jorge Zahar Ed., 2008a.
______. Joaquim Callado: o pai do choro. Rio de Janeiro: Jorge Zahar, 2008b.
Enciclopédia da música brasileira: erudita, folclórica e popular (EMB). São Paulo: Art Ed., 1977. Volumes 1-2.
GOUVEIA NETO, João Costa. A importância do Teatro São Luís na efetivação do movimento cultural na São Luís da segunda metade do século XIX. Anais do XXV SIMPÓSIO NACIONAL DE HISTÓRIA-ANPUH, Fortaleza-CE, 2009.
Inventário do Acervo João Mohana: partituras. Arquivo Público do Maranhão. São Luís: Edições SECMA, 1997
KIEFER, Bruno. História da música brasileira dos primórdios ao início do século XX. Vol. 9. Porto Alegre: Editora Movimento, 1976.

PINTO, Alexandre Gonçalves. O Choro. Rio de Janeiro: FUNARTE, 2009. Primeira edição em 1936.
SEVERIANO, Jaime. Uma história da música popular brasileira. São Paulo: Ed. 34, 2008.
TINHORÃO, José Ramos. História social da música popular brasileira. São Paulo.: Ed. 34, 1998.
VASCONCELOS, Ary. Panorama da música popular brasileira na Belle Époque. Rio de Janeiro: Editora do autor, 1977.
______. Carinhoso etc (História e Inventário do Choro). Rio de Janeiro: Editora do autor, 1984.

COMIGO, O LUSTRO É OUTRO ?!
(choro)
Manoel G. Pereira

COMIGO, O LUSTRO É OUTRO?!

MARIA SAPECA
(choro)
Paulo Almeida


MARIA SAPECA
17
21
25
29
33
34
D.S. ao Coda

LISTENING TO INTERACTIONS: THE RELATIONSHIP BETWEEN *STREAMS* IN MIXED ELECTROACOUSTIC MUSIC

Davi Raubach Tuchtenhagen
Universidade Federal do Paraná
raubachdavi@gmail.com

ABSTRACT

This paper analyses interaction in mixed electroacoustic music through the perception of streams in the listening experience. The interaction between acoustical instruments and electroacoustic resources is observed in a wider theoretical scope: the interaction between streams perceived in listening. The study investigates issues related to the characterization of streams, how we perceive their interaction, and how we can describe this interaction for analytical or compositional ends. For this, two central concepts are revised: Dennis Smalley's *behavior* metaphor and Trevor Wishart's *counterpoint* idea. We present a brief analysis of selected electroacoustic mixed works by Karlheinz Stockhausen, Cort Lippe, and Tristan Murail to better illustrate listening aspects of interaction in this repertoire.

1. INTRODUCTION

The relationship between instrument and electroacoustic is the main feature of mixed electroacoustic music and is frequently called *interaction.* The term often refers to human-computer interaction (HCI) also. However, in this paper, it is not important whether the electroacoustic is fixed in support, live-electronics or interactive computer music system, but the focus is on the listening perception of interaction. This focus dislocates the interaction from an issue of technology or instrumentation to a wider scope of sound morphology. Studies [1, 2, 3] have investigated the theme with this focus by several perspectives. Menezes [1] approaches the spectral fusion and contrast between instrument and electroacoustic. Bachratá [2] explores the interaction by a multi-perspective study on musical gesture and catalogs gesture interaction in mixed music. Souza [3] presents several aspects to be considered in the study and composition of mixed works, also using several theoretical perspectives. However, these studies are little concerned with issues related to the concept of *streams*, present in Dennis Smalley's [4] and Trevor Wishart's [5] theories. This paper intends to demonstrate the analytical and creative potential in approach the interaction between instrument and electroacoustic from a perspective of interaction between *streams*. For this, the research presents (1) an investigation on the characterization of streams, (2) an explanation about the perception of streams interaction, and (3) a utilization of Dennis Smalley's [4] *behavior* metaphor and Trevor Wishart's [5] *counterpoint* idea to describe this interaction. Mixed pieces by Karlheinz Stockhausen, Cort Lippe, and Tristan Murail are briefly analyzed through these two theoretical tools (behavior and counterpoint) in order to demonstrate their applicability.

2. THE STREAMS

Stream is understood as a conceptual tool that explains textures composed of clearly distinct 'bands' of sonic activities that are part of the overall spectrum [6]. Likewise, stream is used here in the sense of *layer*, equivalent to voice in traditional theories, and refers to the grouping of successive and related sound events. A stream can be identified by differentiation in relation to another. The differentiation can be perceived in pitch, timbre, gestural configuration, space or another parameter, and can be complete or partial. As musical instruments are historically designed to maintain timbre stable and to enable the pitch to vary, often, the different fixed-timbre instruments constitute different streams [5]. In such configuration, the streams are structured in discrete values; they are in instrument x or y and are one or two, or three, etc., there is no gradation between these possibilities.

However, the discrete organization has been challenged by composers. An earlier example can be found in Schoenberg's *Klangfarbenmelodie* idea – Five Orchestral Pieces op.16, no.3 (1909) entitled *Farben* is an example. The conception of timbre (or tone-color) as a structural parameter allowed timbral transformations of a single line (stream), which is shared among instruments [7]. In this case, it is not instrument x or instrument y streams, but it is only one stream that is developed through a gradual timbre modification (almost a continuum) among instruments.

If in this case of instrumental music it is difficult to analyze streams by instruments, in discrete values, it is even more in the case of the mixed electroacoustic music. At a first glance, there is a special "instrumental" distinction: the acoustic instrument produces sound by mechanical vibrations as the electronic sounds are produced through electronic circuits, analog or digital, that are heard through loudspeakers. However, from the listening perspective, such distinction is not useful, whereas electronic sounds can project a recorded sound of instruments, and instruments can produce unconventional sounds that can be confused with electronic ones [3]. Therefore, although it seems to be easy to reduce mixed music in two streams (instrumental and electroacoustic), it is evident that it is not that simple. On the one hand both instrumental (even in a single instrument) and electroacoustic sounds can present multiple streams, and, on the other hand, one single stream can be shared between instruments and electroacoustic.

Menezes's study [1] on the issue can be helpful to attempt the distinction of streams in mixed music. According to the author, the interaction between instrument and electroacoustic happens between two poles: *fusion* and *contrast*. Fusion is the absolute similarity, characterized by spectral transferences from one side to the other, as the contrast is the absolute distinction, characterized by the absence of these transferences. Between these two poles, there are transitional stages of relative similarities and dissimilarities. In *fusion* the listeners are in a *doubt condition*, they do not know if the sound comes from the instrument or is/was processed electronically. Applying this *morphology of interaction* to distinguish streams we have that, in *fusion*, instrument and electroacoustic share the same stream. and in *contrast*, they present different streams. Besides, between these two possibilities, there are transitional stages, a gradation, between one stream and another. It is possible that a single stream splits in separated streams which could be individually developed and converged again into one. In this case, it is not one or other stream but one that engenders the other, or two (or more) that merge in one, occupying dubious regions between independence and parallelism. In his notes about the *Music for Flute and ISPW* (1994), Cort Lippe writes about a continuum between an extended-solo and a duo configuration:

> Musically, the computer part is sometimes not separate from the flute part, but serves rather to amplify the flute in multiple dimensions and directions; while at the other extreme of the continuum, the computer part has its own independent musical voice [8].

Karlheinz Stockhausen's *Kontakte* (1958-1960), for piano, percussion, and tape, is another example of the possibility of one stream split into others. At approximately 22 minutes of tape, it happens in the tape part. The score illustrates that (see Figure 1).

In the case of Kontakte, the idea is that the original sound is constituted by components. Stockhausen explains the excerpt: "The original sound is literally taken apart into its six components, and each component in turn is decomposing before our ears, into its individual rhythm of pulses" [9, p.97]. In this case, after separate, each component is developed concurrently in its own stream.

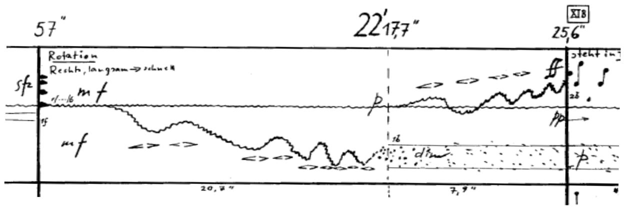


Figure 1. Splitting of one stream into two others in Stockhausen's *Kontakte*.[1]

Therefore, in mixed music, streams can be distributed in several forms among the instruments and electroacoustic including a continuum between they. In addition, there is a continuum between one stream and two independent streams. From this viewpoint, the subsequent question is how the streams interact. In the next section, the perception of this interaction is explained through the concept of *extrinsic behavioral references* [4], and the interaction between streams is analytically approached through the *behavior* metaphor [4] and the *counterpoint* idea [5].

3. THE BEHAVIOR OF THE STREAMS

3.1 Extrinsic Behavioral References

The placing of distinct sounds in a context ensures that some kind of relation must to exist between them. Smalley [4] uses the term *behavior* to conceive the relations among spectromorphologies[2] acting into a musical context. Although the author emphasizes that *behavior* may be applied at a variety of levels, here, the metaphor is used to consider the relation between streams.

The behavioral references are extrinsic, that is, they are not references to the music itself (intrinsic) but are related to a range of experiences external to the context of the music.[3] In acousmatic music[4], the possibilities in content and movement of spectromorphologies create a great and variable collection of extrinsic references [4]. This means that, in the listening experience, we associate the relationship perceived between the sounds with other lived experiences. A common example is when we listen to two instruments that intersperse their play with similar fragments and we compare this fact to a conversation. Charles Ives's *The Unanswered Question* (1908) present the extrinsic reference to "question" and "answer" explicit in the title, for example.

In acousmatic music, the behavioral relationships are perceived through the spectromorphologies alone. However, in mixed works, they are perceived with a strong in-

[1]Extracted from K. Stockhausen (composer). *Kontakte*. London: Universal Edition, 1966, p.25.

[2]Spectromorphology refers to the development of the sound spectra through the time. The term applies to musical structures of any size and level [4].

[3]Smalley approach is based on Nattiez's intrinsic and extrinsic distinction [10, pp.118-126].

[4]We consider here that acousmatic music is prerecorded and played through loudspeaker, there is no performer at stage [11]

fluence of the relationship between visual and gesture-bearing performer with the acousmatic part [4]. However important this issue is, this paper will not approach performer's visual and gestural references due to the focus on structural aspects of streams. Nevertheless, the issue could be included in future studies.

Therefore, the analysis of behavioral relationships is based on extrinsic references. From this perspective, we could speculate if it would be possible to think in other behavioral relationships based on other extrinsic references. This issue is resumed forward.

3.2 Behavior

The behavior metaphor includes any behavioral relationship. The developments of such concept are presented by Smalley in general and in specific terms. In the general terms, *behavior* is dependent on two semantic oppositions: *dominance/subordination* and *conflict/coexistence*. These oppositions represent the basis for a collection of *relationship modes*: *equality-inequality*; *reaction-interaction-reciprocity*; *activity-passivity*; *activity-inactivity*; *stability-instability*. These relationship modes are articulated in two interactive temporal dimensions. In the *horizontal* dimension, the streams pass from a context to another, it is concerned with *motion passage*. This can happen in a *voluntary* or in a *pressured* way. A key concept here is *causality*: when one event appears to be the cause of the next or the cause of a modification in a concurrent event. In the *vertical* dimension, the streams present, or not, vertical synchronization, it is concerned with *motion coordination*. That happens on a continuum between *loose* and *tight* coordination freedom [4] (See Figure 2).

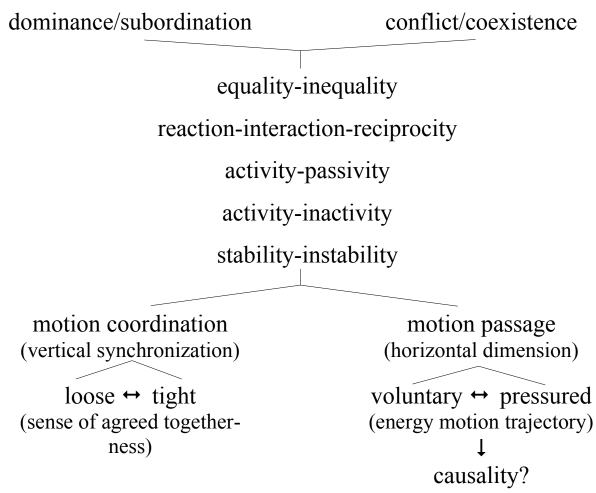


Figure 2. Behavior by Smalley.[5]

The Cort Lippe's mentioned *Music for Flute and ISPW* (1994) serves as an example. In the first section, there are two streams intercalated in the flute, one is more continuous and has the *piano* dynamic; the other consists in *forte* attacks. The *forte* stream is synchronized with the computer part, in fact, together they form a single stream. *Piano* and *forte* streams are *tight* coordinated in the vertical dimension and they alternate suddenly, in a *pressured* way. It is possible to perceive that the *forte* stream establishes an *instability* upon the more stable *piano* stream (*relationship mode*). In addition, the relationship between the two is *conflicting* and, as the music develops, we perceive that the *forte* stream *dominates* the subordinated *piano* stream.

In this case, the theory proved to be suitable to indicate the behavioral relationship between streams.

3.3 Counterpoint

Another approach on the relationship between streams is Wishart's counterpoint idea. In the *Gesture and Counterpoint* chapter of his book [5], Wishart is concerned with a contrapuntal structure in *continuum-based music.* According to the author's perspective, music has historically been developed in a tridimensional *lattice* with discrete values for pitch, duration, and timbre. In such lattice-based music, pitches are organized in semitones; durations, in divisible values; and timbre, in different fixed-timbre instruments. Nevertheless, there is a continuum between a semitone and a tone (evident in the glissando, for example); there are infinite duration values between an eight and a quarter note; and, as demonstrated before, there is a gradation in timbre. The continuum is even more apparent when we consider sounds that are not made with conventional instruments, very common in electroacoustic music. If the lattice provided a structure to evolve a tonal counterpoint, thence the question is how to establish a contrapuntal structure to work in the continuum.

Wishart distinguishes two independent principles to achieve a contrapuntal structure: (1) An *architectural principle*, which offers reference points in the global progression of music material. In tonal music, it corresponds to the key structure (the return to the tonic tonality is often an important point, for example). In the continuum, this architectural principle will be 'the concept of transformation from one timbral and sound-morphologic area to another' [5, p.117]. And (2) a *dynamic principle*, which determines the nature of motion. In tonal counterpoint, it is related to the ebb and flow of rhythm coordination and harmonic consonance-dissonance, that is, it refers to the manner that the notes of one voice are placed in relation to the notes of other voices. In the continuum, instead of this, the dynamic principle will be the "gestural evolution and the interaction between the separate streams" [5, p.117].

The *gestural evolution* is related to the horizontal features of the dynamic principle: the type of gesture[6] used, the sequence of individual gestures and the average rate of gestural activity in a stream. In this context, it is important to consider especially the features of the gestures.

[5]Extracted from [4, p.119].

[6]The term *gesture* is used here according to Wishart's concept, as an "articulation of the continuum" [5, p.17]. On the basis of this concept relies the idea that the intellectual-physiological gesture can be translated in sound-morphology and, inversely, the sound-morphology evidences the intellectual-physiological gesture [5].

The *interaction between the streams* is related to the vertical features of the dynamic principle. Wishart offers a solid theoretical development in this ambit. It is possible to consider: (a) the number of gestures that occur in all streams in an observed period; (b) the homogeneity of the gestures among the different streams (homogeneous or heterogeneous); and (c) if the gestures of one stream seem to interact with that of the other streams or if they seem to behave independently. From these relations the author establishes six archetypes of gesture organization (see Figure 2): (1) *parallel*: akin to *tutti,* the gestures have the same characteristics in all streams (it is not about spectral characteristics but about gesture structure); (2) *semi-parallel*: the parts follow the same logic but not synchronously; (3) *homogeneous independence*: the parts seem to behave independently; (4) *heterogeneous independence*: the gestures are independent and different; (5) *interactive*: especially related to causal and imitative connexions between events in different parts; and (6) *triggering*: the gesture of one part is the onset of a gesture in another part or it causes a modification in another part quite clearly, it is the *causality,* to use the Smalley's term.

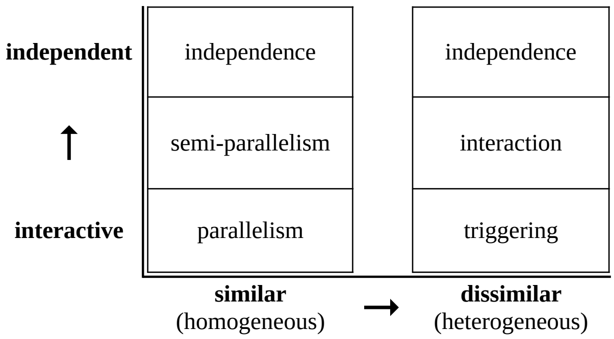


Figure 3. Six archetypes of vertical gesture organization (dynamic principle)[7]

Wishart describes a compositional strategy from these ideas:

> I found this approach to be a powerful heuristic tool for composing with this kind of material. It was possible to lay out the structure of the overall density of events on the score, then compose the gestural structure of a section using elementary symbols [...] and then, working from the overall plan of timbral and articulatory development, score in the details of the individual sound-events in each voice. [5, p.123]

Tritan Murail's *Desintegrations* (1982), for computer-synthesized tape and 17 instruments, can be an example to elucidate Wishart's theory[8]. The piece is based on spectral analysis of instrumental sounds (piano, brass and cello sounds). Therefore, tape and instruments share the same origin, according to Murail's previous notes, "their relationship being one of complementarity" [12]. The tape part frequently "exaggerates the character of the instruments, diffracts or disintegrates their timbre, or amplifies the orchestral effects" [12]. That the relationship is one of complementarity it does not mean that we will perceive just one shared stream. In fact, like in Stockhausen's *Kontakte*, Murail works with the disintegration of sound in different streams that are shared among instruments and tape.

Murail explains that there are eleven moments, eleven distinct "stages" in the piece, and the change of stage happens by "transition-transformation or by the unleashing of a 'threshold effect'" [12]. This transition evidences the *architectural principle*, the transformation of one sound-morphologic area to another. The composer explains the transformation process:

> Each moment emphasizes a different kind of spectrum treatment, each stage makes it evolve from the harmonic to the inharmonic or vice versa. This creates movements of shade and light, accompanied by movements of increasing or decreasing agitation, of rhythmic ordering or disordering. [12]

Therefore, in architectural principle, it is possible to observe the transformation of a harmonic sound-morphologic area to an inharmonic one or vice versa.

During the third stage of the piece, the *architectural* principle is characterized by a transition from harmonic to inharmonic spectra and from a high range to wide one. Figure 4 shows part of the beginning of this stage.

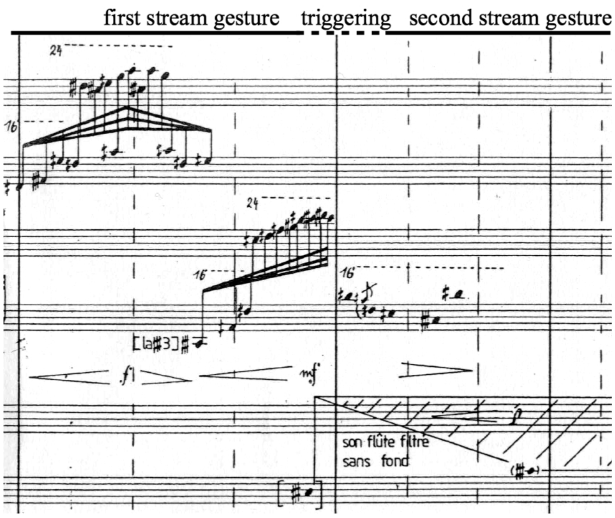


Figure 4. Triggering relation between gestures in different streams in Murail's *Desintegrations*.[9]

In this stage, piano, crotales, glockenspiel, and tape share a stream. The gestures in that stream are composed of short and resonant high notes that are rhythmically active at the beginning of the gesture and present just the resonance in its termination. In a first moment, the *gestural evolution* happens by a temporal expansion of the rhythmically active part of the gesture. At the same time, the resonant part of the gesture engenders gradually a second stream, first only in tape, after in the instruments. In relation to the *interaction between the streams*, we listen as if the high notes gesture in the first stream is initiating

[7]Extracted from [5, p.122].

[8]In his book, the author demonstrates the theory in his own compositional experience in *Vox-I* (1982), for 4 amplified voices and tape.

[9]Extracted from [12, p.30] and analyzed.

the resonant gesture in the second stream (*triggering*) that become each time more *independent*. Figure 4 shows the representation of the two gestures in the tape part analyzed.

Finally, it is important to underline Wishart's attention and imperativeness to the importance of the listening experience in compositional practice. 'No notational logos can in itself justify a musical procedure' [5, pp.123-125] According to the author the listening experience is unmediated, unprejudiced, and must to be the responsible for the ultimate validation of any musical procedure.

4. DISCUSSION

The objective of this study was to demonstrate the analytical and creative potential in approach the interaction between instrument and electroacoustic from a perspective of interaction between streams.

Both approached theories are pretty similar. The largest difference is that specifically Smalley's *behavior* concept does not consider the aspect of development in time of the behavioral relationships as does Wishart through the *architectural principle.* This could be easily changed if we considered other aspects of the *Spectromorphology* such as motion and growth process, structural functions, and others. It is possible to perceive that Smalley's *behavior* intends to describe a state, a character rather than a process, while Wishart's approach creates a space in which the streams can change from one state of interaction to another, going through the concepts of Figure 2.

In addition, the two theories exposed and applied in concise analyses demonstrate pertinence and achieve categories of analysis that others do not considerate.

Menezes's article [1] discusses the issue of fusion and contrast between instrument and electroacoustic, however, do not relate this morphology with the disposition of streams. Nevertheless, as demonstrated, the Menezes's study is a useful tool to observe the splitting and converging of one stream. The fusion-contrast continuum present direct relation to the similar-dissimilar, homogeneous-heterogeneous axis in Wishart's table (see Figure 3).

Bachratá's [2] study presents the interaction by a multi-perspective on musical gesture. The study presents a catalog of several and very specific gestural interactions between instrument and electroacoustic. It concentrates on what we can consider specificities of the general terms of both Smalley's and Wishart's theories. And, because of that, maintain a local rather than a global perspective, with exceptions. Her research can be used in order to specify how are the *relationship modes* [4], or how the *six archetypes of gesture organization* [5] are in punctual descriptions.

Souza [7] approaches the interaction between instrumental and electronic sounds by several perspectives. By a semiotic perspective, he distinguishes the *marked sounds* which we are habituated to attribute to the instruments (e.g. piano sounds) from the *not-marked sounds*, sounds to which the relationship with the source is not recognized (e.g. some granular sounds). This perspective call our attention to the fact that the stream can be articulated through this two types of sounds but it does not mean that marked-sounds were produced on the instrument and the not-marked sounds in electronics. This fact also directs us to think an alternative category to that of simple sound production distinction between instrument and electroacoustic. By the viewpoint of this article, the stream concept is an interesting one, as demonstrated.

Although this three mentioned studies configure a limited referential to outline trends, it is possible to perceive that the authors tend to introduce new concepts or a multi-perspective approach in order to demonstrate different aspects, in different levels of the same interaction. It is important to point out the need to concatenate these approaches in order to build a shared and complementary knowledge about the issue. In this sense, it is especially necessary to consider the previous approaches to mixed and electroacoustic music such as the Smalley's and Wishart's ones. We could go beyond and question if there are no other theoretical approaches to instrumental music that think the interaction, likewise the mentioned *Klangfarbenmelodie*.

In addition to this centripetal force, we could think in a centrifugal one that lead us to investigate aspects that are not covered in this studies. For example, based on the extrinsic behavioral references [4], was possible to explain how we perceive this relationships. Nevertheless, we need to ask if there is no other references that explain better certain behavior in question. Or, in the case of composition, we could ask what other extrinsic references, present in other lived experiences, can be projected in the relationship between streams.

Further studies can also focus on the influence of technology on the configuration of streams and their interactions. It could be investigated how does the relationship between two types of interaction: on the one hand, interaction between streams and, on the other hand, the interaction between performer and machine (HCI) in performance.

5. CONCLUSION

The interaction between instrument and electroacoustic can be approached by a perspective of interaction between streams perceived in listening. This perspective proved pertinence in analysis and possibly in composition of mixed music. Future studies are needed to, on the one hand, combine the already done studies in this area and, on the other hand, to expand the issue with questions related to the influence of technology and the use of other extrinsic references to interpret interaction.

Acknowledgments

The research for this paper was financially supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível

Superior (Capes). I also thank to Felipe de Almeida Ribeiro, my supervisor, who contribute with helpful input in developing the ideas presented here.

6. REFERENCES

[1] F. Menezes, "For a morphology of interaction," *Organised Sound*, vol. 7, no. 3, pp. 305–311, 2002.

[2] P. Bachratá, "Gesture Interaction in Music for Instrument and Electroacoustic Sounds," Doctoral thesis, Universidade de Aveiro, Portugal, 2010.

[3] R. C. Souza, "Da interação entre sons instrumentais e eletrônicos," in *Criação Musical e Tecnologias: teoria e prática interdisciplinar*, Pesquisa em Música no Brasil serie, vol. 2, Goiânia: ANPPOM, 2010, pp. 149–179.

[4] D. Smalley, "Spectromorphology: explaining sound-shapes," *Organised Sound*, vol. 2, no. 2, pp. 107–126, 1997.

[5] T. Wishart, and S. Emmerson. *On Sonic Art* (Contemporary Music Studies). New York: Routledge, 1996.

[6] R. Weale, "Stream", *ElectroAcoustic Resource Site*, 05-Jan-2005. [Online]. Available: http://ears.pierrecouprie.fr/spip.php?article185 [Accessed: 26-Apr-2018].

[7] J. Rushton. "Klangfarbenmelodie." *Grove Music Online*. 20-Jan-2001. [Online]. Available: http://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000015094 [Accessed: 26-Apr-2018]

[8] C. Lippe. Program Notes of Music For Flute and Computer. [Online] Available: https://www.cortlippe.com/uploads/1/0/7/0/107065311/lippe-flutenotes.pdf [Accessed: 26-Apr-2018]

[9] K. Stockhausen, "Four criteria of electronic music," in *Stockhausen on Music: Lectures and Interviews*, New York: Marion Boyars, 1989, pp. 88–111.

[10] J.-J. Nattiez, *Music and discourse: toward a semiology of music*. Princeton: Princeton University Press, 1990.

[11] E. Simon, and D. Smalley. "Electro-acoustic music." *Grove Music Online*. 20-Jan-2001. [Online]. Available: http://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000008695 [Accessed: 26-Apr-2018]

[12] T. Murail. (Composer). *Desintegrations* [Sheet music]. Paris: Henry Lemoine, 2004.

MOVING TOWARDS SYNTHESIS: THE INCORPORATION OF TECHNOLOGY IN MUSIC THROUGH A CASE STUDY OF THE FLUTE

Gerardo Lopez

University of Redlands
Redlands, CA, USA
gerardo_lopez@redlands.edu

ABSTRACT

The following paper explores the integration of technology in music, specifically through the scope of flute composition and performance. By surveying pieces from the flute repertoire, the existence of three emergent stylistic categories that have distinct techniques and aesthetics will become apparent. The idea of a fourth category will also be discussed, but this particular category remains speculative as it has not emerged empirically, except for a few precursory examples. The importance of understanding this classification and categorization rests with being able to examine the evolving nature of the human-technology relationship within an artistic space, specifically, live performance.

1. INTRODUCTION

This story begins with a car ride on the infamous 10 Freeway in Southern California. After an afternoon in Los Angeles, a few friends and I found ourselves discussing technology and music, having spent most of the afternoon visiting the Griffith Observatory. During the course of the conversation, I jokingly brought up the question: “So when do you think we will be replaced?” This was in response to the idea that the current rate of automation and advancement in AI indicates total integration within certain sections of our society and complete takeover in others will happen in our lifetime [1]. My question toyed with the possibility that the advancing technology would replace performers, once it becomes sophisticated enough. I received scoffs as an answer and then: “That’s not going to happen. What are people going to do? They wouldn’t let it happen, they would actively resist it.” Shortening the discourse that ensued after that statement, we came to two conclusive predictions: technology would either dominate major aspects of music performance or, there would be a synthesis between human and machine.

2. CLARIFICATION OF TERMINOLOGY

A recurring issue in the intersection between music and technology is terminology. Before continuing, it would be best to define the parameters for some of the following terms.

2.1. Electro-Acoustic Music

The term *electro-acoustic music* is defined as “[m]usic in which electronic technology, now primarily computer-based, is used to access, generate, explore and configure sound materials, and in which loudspeakers are the prime medium of transmission [2].” An elaboration on “electronic technology” would be useful to ensure readers understand that this term encompasses a diverse array of hardware and software, such as distortion pedals and MAX/MSP patches, not just computers as might be suggested. Also, the term “amplification system” may be better suited than “loudspeakers,” because one may be incorrectly led towards imagining monitor speakers on a stage as the only manner for electronic amplification. This blocks the inclusion of other systems, such as portable mini-speakers and those that have not been created yet. The importance of this clarification also rests with understanding a critical split in sound production: *acoustic music* is created and transmitted by “naturally resonating bodies,” as compared to *electro-acoustic music*, which has components of its composition that use electronic technology for its production and transmission [3]. With this definition, the realization can be made that there are aspects of our current musical society classifiable under *electro-acoustic music*, even though it may not have been formally thought of as such before [3].

2.2. Subgenres

The discussion within this paper mainly focuses on works of *live electronic music*, which involves the use of electronic technology within a live performance space [2]. Best thought of as a collaboration between itself and other agents, this technological incorporation can include the voice and acoustical instruments, electronic instruments, and other devices and controls in a variety of ensemble combinations [2]. This is in contrast to *acousmatic music*, which has been created in a studio and exists in recorded tape form [2]. Under these two categories exist more defined categories, as illustrated in the figure below.

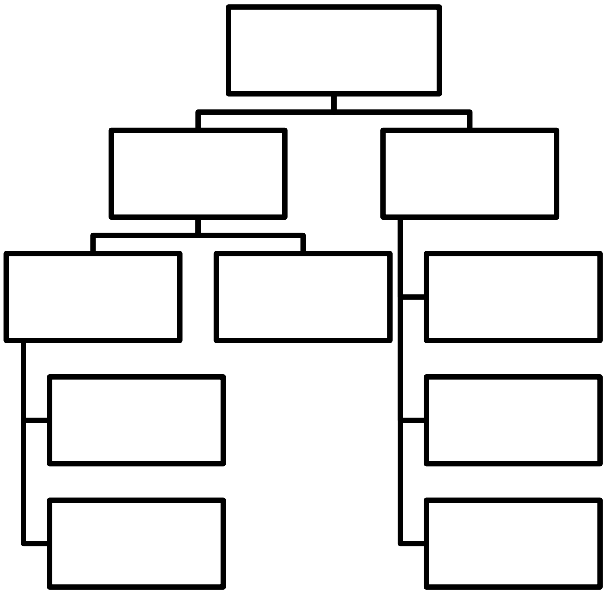


Figure 1. This is borrowed from Schrader (see [3]), with the edit of "acousmatic" instead of "tape." He mentions it as by no means a complete categorization of the field but uses it a framing tool and the reason for including it here is the same.

3. RELATIONSHIPS

At the core of technological integration is the evolving relationship between the human and machine. Traditionally, listeners have been able to associate the production of sound to the performer(s) on stage as a direct result of an action enacted through their respective instruments. With the introduction of certain technologies, such as "tape," this is no longer the case. The split is captured well by the "local/field" distinction [4]. These terms are defined as two spaces created in a live performance; the "local" level deals with the performer's actions and the resulting sounds, whereas the "field" level encompasses the whole of the environment in which the "local" is a part of [4]. Expanding on the "tape" example, in a performance the "local" would be the instrumentalist playing their instrument as per the composed score and the "field" would include the elements presented in the electronic track. The electronic track exists in the "field" as another "agent," but is not another "local" since it is neither human and it breaks the "action-sound" relationship of the human performer's actions [4]. The manner in which a composer, and in some cases the performer, goes about managing these relationships is what ends up creating the work's unique aesthetics and definable elements.

This negotiation between human and technological elements can be understood through levels of interaction. The definition of interaction within the context of this discussion can be understood as the crossing of elements between the "local/field" split. Termed another way, the manner in which either the performer encounters the electronics or how the technology adapts to the performer. With the inclusion of technology, the perception of these relationships can be manipulated through the use of "real" and "imaginary" causal interactions [4]. On a "local" level, such examples can include the transmutation of the acoustical sound of an instrument into a completely different one through the use of effect pedals or software, or the triggering of specific electronic track events with the aid of motion sensors. This can also occur in the "field," such as with the creation of artificial "call-and-responses" in which the instrumentalist or electronic track "responds" to the other in a seemingly quasi-improvised manner.

3.1. Static

Drawing upon the "tape" example again, within this relationship, both the human and technological components are static in their interaction. The performer mainly functions in the "local" and the electronic track mainly in the "field." This is a common critique of some compositions that involve tape and live acoustic instruments, the tape serves merely as an accompaniment or that the instrumental component is subservient to the electronic track [4]. However, this would miss the idea of a composer playing with perceptual relationships. Though a perceivable "real" interaction may be created through creative scoring, it remains "imaginary" since no actual relationship exists, since neither of the components cross the "local/field" spaces.

3.2 Active

A more active relationship would require that one of the elements make a "real" relationship between the "local/field." For example, a work that uses pressure or movement sensors to trigger certain events has a "real" relationship between the motions occurring in the "local" and what ends up manifested in the "field." This is different in the case of a work that uses effects pedals to distort the sound of an instrument; in this scenario, the "field" has been reduced into "local," since the actions are not used to trigger, but have become enhanced in order to shift the focus solely onto itself.

3.3 Interactive

In short, it can be understood as a return to the traditional relationship that exists between musicians in a chamber group; the creation of another "local" within the "field," or the inclusion of another "agent" in the "local." This new "local/agent" however, is no longer a human, but an autonomous or highly-interactive system. Within the context of these relationships, this new element is easily absorbed into the paradigm as another "local/agent," and the question of how to understand or deal with this relationship is answered: constructive and collaborative.

4. CATEGORIES AND SUPPORTING REPERTORIE

While the examples of supporting repertoire will be pulled from the flute literature, the terms and relationships are by no means solely applicable to the instrument; a future exploration of the universality of these categories into other domains may provide interest points of contrast or similarity.

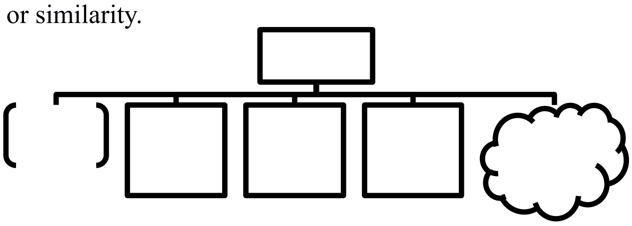


Figure 2. These are the proposed emergent categories within the flute literature.

The bracketed *Music for Live Electronics* will not be developed because many of the other categories already use live electronics (pedals, software, etc.). However, it is not being discarded from the diagram since it captures the idea of "electric flutes."[1] Still, the integration of technology concerns the traditional acoustical flute, and the "electric flute" would extend beyond the scope of this paper, though this would be an interesting area for future research. Also, an argument can be made for the existence of subgenres underneath these proposed categories. Again, for the purposes of this paper, that discussion would be more detailed than necessary since we are mainly looking at broader style characteristics.

4.1. Fixed Media and Flute

This category has the longest history among the other categories, appearing as one of the first methods for incorporating electronic technology into a performance setting. One of the earliest pieces that featured both live acoustical flute and electronic technology, in this case a tape recording, was *Musica su Due Dimensioni (Music in Two Dimensions)* by Bruno Maderna in 1952, written for flute, percussion, and electronic sounds on tape [5]. A quick note on the term "tape," which has traditionally been used in describing a fixed element that accompanies a live performance. Within this discussion the term "fixed media" is offered as an update since the technological advancements have far surpassed the original parameters of "tape." The update takes into consideration other forms of recorded mediums, such as CDs and digital files.

4.1.1. Pray (2010)

"Pray" by Allison Loggins-Hull was written for solo acoustic flute and an electronic track, with its conception coming "from a need of serious reflection and patience [6]." As a consequence, many of the sounds on the electronic track signify religious elements, such as with the inclusion of "organs, whispers, flute harmonics and chants… the sounds of various places of worship [6]."
Using some of the language suggested by the "local/field" discussion, the relationship between the performer and the electronic track in "Pray" is supportive and responsorial [4]. For example, the fragments of melody are created within the context of call-and-response. In the track, there are snippets of chant calls used that the flute melodic material frequently seems to respond to. If not specifically chants, there are acoustical "events" that take occur, such as an electronic swell, that puts the flute part into context. The rhythmic profiles of each section are also solidified: in the freer sections, there are layers of disjointed rhythms and the pulse is occasionally lost, as compared to other sections that have a constant pulse and stable rhythms.

The performer-technology relationship is also static. The creation of the responsorial environment suggested above is an "imaginary" relation, one that the composer has used in order to create an interesting and engaging "field." This would not be possible without the use of timestamps, which help guide the performer through the notated score and ensure synchronization with the electronic track. The ultimate success of performing the work rests with being able to unite both components. In this respect, it can be observed that neither the "local" nor the "field" cross or interact, rather they function separately in creating the whole of the work.

This also speaks to changes in practice/rehearsal pedagogy. Since the performer would need a solid "feel" of how the electronic track fits with the notated score to have a successful performance, it can be assumed that they have practiced the whole of the work multiple times. While in a traditional setting this may be dismissed, since both performers can adapt in the moment, within this context that is harder to do since one component is nonresponsive.
The use of electronics also achieves something that would have been harder to put together acoustically; the creation of a "digital" ensemble. For example, something as simple as the inclusion of chants, traditionally, would have involved not just an actual performer on stage but would require someone with the specific skill set to sing the chant material in style. Also, rather than having this envisioned performer, or others who would be needed to replicate the track such as a bass player or synthesizer, these people and their equipment would take up room on the stage. The use of an electronic track allows for mobility since it only now requires a flutist and the electronics. This allows for greater flexibility in terms of the performance venue, which is apparent in the reference video [7].

4.2. Enhanced Flute

The emergence of this category comes more from the influence of the popular music sphere rather than the classical.[2] The term of *enhanced flute* includes technology that works specifically with altering/modifying the

[1] An example of this can be seen in the work of Bryan Jacobs at http://bryanjacobsmusic.com/mechanicalflute.html

[2] Specifically referencing the rise of the electric guitar and the use of microphones in jazz.

sound produced by the flute, such as microphones for amplification, distortion and looping pedals, and software. While there exist a variety of pieces that use microphones to amplify the flute, there are fewer that use pedals or software.[3]

4.2.1.*Eruption* (2012/2014/2016)

"Eruption" by composer/performer Melissa Keeling is a transcription of rock guitarist Eddie van Halen's solo [8]. Aside from the electronic components, which will be mentioned shortly, the piece includes the use of improvisation, "multiphonics, singing and playing, harmonics, residual tones, tremolos, and circular breathing [9]." In short, this piece should be "performed with a sense of power, ease, and virtuosity [9]." Also, just a small aside on the use of the "glissando headjoint," which is a relatively new acoustical hardware introduced to the flute by Robert Dick, which allows for noticeable bending of the pitch as if it were a guitar "whammy [10]."

This work can be performed without electronics, just solo flute with a glissando headjoint. However, the element that makes this work really stand out is the use of the pedals. Right from the beginning and throughout the work, the distortion pedal is set and the flute sound projected from the speakers is no longer the traditional flute but has taken on the sound of a distorted electric guitar. The other elements mentioned above in conjunction with the other pedal effects combine to accomplish the composer's goal, an almost exact replication of the original rock guitar solo on a foreign instrument.

While it may seem that the use of distortion pedals is something novel that has appeared within the last decade, there is a history from which this emerges: some of the earliest *enhanced flute* works used software for distorting or manipulating the sound, such as Kaija Saariaho's *Laconisme de L'aile* (1982) and Judith Shatin's *Kairos* (1991) [11,12]. However, the direct use of distortion pedals in a formalized setting such as this is relatively new.

Also, a noticeable difference from this category to the previous one is in regard to the involvement of the performer with the technology, since there are now more parts. With this work, for example, a BOSS VE-20 Vocal Effects Processor is used for a "flanger effect," a BOSS DD-7 Digital Delay pedal for delay, and BOSS DS-1 for distortion [8]. While there are more components involved than an electronic track, the piece is still quite accessible when a performer has the will to practice with the accompanying electronics. This in contrast to similar works previously mentioned which produce some of the same results but through software, involving complicated lines of code that the average performer may not be familiar with.

In this framework, the relationship becomes more active. The idea of the "field" being reduced into the "local" is created when they seem to become one-in-the-same. There are no other "agents" to consider, the singular "local" has been enhanced to extend beyond the normal parameters, in this case distorting the sound to that of an electric guitar. Rather than having this work be a duet between a live electric guitar and solo flute, the composer has opted to bypass the physical and acoustical limitations through the use of technology. This example showcases one of the benefits of technological integration, the expansion of the sonic palate. Now, composers and performers are not just limited to the traditional acoustic properties achievable on the flute but can reach beyond to include an endless possibility.

4.3. Augmented Flute

As compared to the previous categories, this one is still quite new in its development. The term *augmented flute* encompasses technology that functions independently of the tone and involves adding a device(s) to the flute and/ or flutist. This would be in contrast to the *enhanced flute*, which deals specifically with manipulating the tone; in this context, the sounds produced by the technology may be produced by other aspects of the performer, such as physical gestures and biometric information. This would include technology such as motion sensors that detect the position of the flute in space, pressure sensors that measure the force of the fingers on the keys and can even be extended to include physiological measurements through galvanic skin response sensors and EEG headsets. This technology then uses the raw information passes it through software that turns it into sonic material that can be used in a musical context.[4]

4.3.1 Le cercle des catharsis (2010)

Work in this area is prevalent in the compositions of Cléo Palacio-Quintin. As flutist and composer, she has created and advanced what has become known as the *hyper-flute* [13]. What can be better described as a system, she retrofitted a standard flute with sensors in various locations each capturing unique information [13]. It is of note that developments in with this system have been rather new, and Cléo Palacio-Quintin has been a champion, the only one in actuality, composing works for this instrument. The composition being focused on here is "Le cercle des catharsis," which is played on hyper-bass flute.[5]

One of the issues she dealt with was deciding where to place the sensors. Some of the sensors she attached to the flute included: an ultrasound sensor to measure the distance of the flute from the computer; pressure sensors to detect pressure levels in the left hand and thumbs; and mercury tilt switches to measure the tilt and rotation of the flute [14]. She mentions how limiting the space can be when considering the element of mobility and placement in a manner that does not adversely affect the performance [14]. This is a reason for her liking of the hyper-bass, which is a bigger instrument and provides more room for more sensors [14].

[3] For reference, there is an extensive catalogue compiled by Bassingthwaighte (see [5]).

[4] For more information on the process of sonification, reference T. Hermann, "Taxonomy and Definitions for Sonification and Auditory Display," *Proceedings of the 14th International Conference on Auditory Display (ICAD 2008),* P. Susini and O. Warusfel, eds., Paris, France: IRCAM 2008.

[5] Similar to the hyper-flute system except on bass flute.

The continued discussion of electronics in this work becomes a bit more tangled than the previous ones. Mainly, the manner in which the sensors are used can be argued as an extension of the *enhanced* category, since in some instances it uses a triggering mechanism similar to a foot pedal. For example, in the figure below the down arrows indicate to the performer that they should lower their foot joint to allow the pre-recorded words notated above to play. However, the key distinction between this use of motion and the pedal is the intention; previously it seemed that having to press a foot pedal was extraneous motion to the music. In this instance, it has become a part of the music. This becomes more apparent with the use of lopsided "U" with "RevFX," which indicates that the performer should tilt their embouchure in that direction which activates the reverb.

While I was not able to provide a solid example of a work that truly uses sensors in manner described under the *augmented* label, the composer does have works that fit within these parameters. This is particularly true of some of her improvisations, which are hardly notated [15]. In corresponding with her, she expressed the difficulty with not only notating the nuances of this system into notation (such as how one would go about notating pressure levels), but also the fact that she has composed in a manner that is not intended for general sale since she is the only *hyper-flute* player [16]. However, the existence of this system and the fact that work is currently being done within this domain allows for this discussion to have some evidence on which to make the categorical argument and classification.

Similar to the previous example, this category provides another active relationship. In this situation, there is a "real" relationship being created between the "local" and the "field." The sensors are using the physical motions of the performer as the action by which the sounds are produced. While the sounds maybe exist as imaginary, since the flute is not capable of electronic sounds, the movements of the performer have been elevated into being part of the music rather than being a byproduct; it has been enlarged into the "field."

4.4. The Fourth Category

As had been stressed early in the paper, this area is the most speculative. A better title for this section might be "Directions in Music Technology," but with the previous discussion, one might be able to see this as a very probably "next step." Firstly, AEE stands for "autonomous electronic entities", which includes AI and highly interactive systems. While advancements have been made in this area, it has not reached a stage where the technology is actively being used in the performance sphere. The software components of this system may in fact be ready, but the presentation of it takes the shape of previously presented mediums, such as computer interfaces or software. These are the interactive systems, which can be better thought of as pre-autonomous systems, which have already been used in some works, as presented below.

4.4.1 Out, Out (2013)

Composed by JP Merz, this work serves not only as a representative for the fourth category but as well as a synthesis for some of the tension surrounding human-machine relationships.

"Out, Out-" captures the idea that was previously mentioned with the categories not being mutually exclusive to each other. Here, a foot pedal is used to activate electronic tracks as well as to trigger effects (reverb). The interface was also created in a manner that allows relatively easy access to any performer interested in presenting this work, with the software packaged as a stand-alone MAX patch.

The distinction comes at the end of the work when the software accompanying the performer begins to "improvise." This example is an example of what can be better described as pre-autonomous systems. They can take the shape of highly-improvisatory software or technology and give the illusion of "improvisation" despite adhering to strict coding parameters. While it may be reaching towards the interactive relationship described earlier, it is still really more within the active category. In this instance, the "improvisation" is taking place with the material being feed into the software by the performer, which is then manipulated and presented through the speakers. Here, the "local" is seen as directly affecting the "field," though the resulting "field" may be substantially different every time (not just in terms of nuance: notes, dynamics, articulation, etc.). However, the use of an "imaginary" relationship here is not just a tool to create an illusion but inadvertently serves as the propping up of an ideal not yet here.

What is this ideal? It is being presented within this discussion as a system that is able to mimic the expressive qualities of a human performer. This whole branch of technological incorporation breaks from the constraints put upon musical interpretation by composers, best captured by the historical emergence and evolution of musical notation [17]. Beginning with the earliest notation and later constrained/refined in the early Romantic era and into the 20th Century, the score was developed with finer and finer instructions to be realized by a performer resulting in less and less creative input into the composition by the performer [17]. Technology is now being granted the ability to creatively realize a composition rather than being the ultimate performer or "Platonic dream," which stems from the idea that technology allows composers to eliminate the human element of physical bodies and subsequent "flaws" from the music [18]. Now, technology is allowed to be as "flawed" as humans.

5. CONCLUSION

It is without note that the idea of having a fully functioning humanoid robot performing in a Brahms quintet alongside other humans is still ways away. However, dismissing the idea or refraining from discussing it may not be in our best interest.

This is particularly true within the educational/ pedagogical spheres: technology is not leaving music but instead will continue to expand. Understanding this realization, it becomes pertinent that music curriculums should include classes or instruction familiarizing their students with these emerging technologies. As may have been noted in the discussion of the pieces above, they require a set of skills that traditional education does not focus on, leaving many to figure things out on their own. While it has worked so far, this must be overcome in order to more performers accessibility, which in turn would allow for a reception of these works by a larger audience. It also has certain implication within the practice space: imagine having a robot next to a piano that you can program to play the piano part to a flute and piano sonata, while it may not replace a human in a live performance setting, it would have great applicability in terms of rehearsals. This is just one idea of many that can be drawn from envisioning the future role AEEs in the musical domain. Taking one step into a really speculative area, this integration can be an example of how the future synthesis of technology into our society can continue. Music may be an area where humans one day might be surpassed by technology, a future where, for example, people would be more willing to purchase a robotic string quartet for their homes rather than paying to see human performers in a concert hall. While a hypothetical, it is important to understand that technology in music cannot be ignored; we must focus on fostering a collaborative relationship these technologies and create spaces for incorporating them into our existing paradigms.

Acknowledgments

Many thanks to Dr. Mark Dancigers for helping overseeing this small project and for his guidance with moving towards exploring the different relationship dynamics that exist as a core idea in human-technology integration.

6. REFERENCES

1. R. Zinn, “The Future Of AI And Automation In The Workforce," *Forbes*, May 5, 2017. [Online]. Available: https://www.forbes.com/sites/forbestechcouncil/2017/05/05/the-future-of-ai-and-automation-in-the-workforce/#4dcf44004e36. [Accessed Sept. 21, 2017].

2. S. Emmerson and D. Smalley, "Electro-acoustic music," *Grove Music Online*, *Oxford Music Online*, Oxford University Press.

3. B. Schrader, *Introduction to Electro-Acoustic Music*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1982, pp. 1-3.

4. S. Emmerson, *Living Electronic Music.* Aldershot, UK: Ashgate Publishing, 2007, pp. 92-95.

5. S.L. Bassingthwaighte, "Electroacoustic music for flute," Doctoral dissertation, University of Washington School of Music, Seattle, 2002, Chapter II. [Online]. Available: http://www.subliminal.org/flute/dissertation/TOC.html.

6. A. Loggins-Hull, *Pray*, Brooklyn, NY: Flutronix Music, 2010. [Music Score}.

7. flutronix, “Flutronix - Pray - Live at The Cell,” *YouTube*, Nov. 20, 2010 [Video file]. Available: https://www.youtube.com/watch?v=k6TW6F0zJlU. [Accessed: Sept. 16, 2017].

8. sonyqTV, “‘Eruption’ (Van Halen/Keeling) - Melissa Keeling, electric flute,” *YouTube,* Aug. 16, 2016 [Video file]. Available: https://www.youtube.com/watch?v=WJrrZRNkVjs. [Access: Sept. 16, 2017].

9. M. Keeling, *Eruption*, 2012/2014/2016. [Music Score].

10. R. Dick, "Glissando Headjoint from Robert Dick," Robert Dick. [Online]. Available: http://robertdick.net/the-glissando-headjoint/. [Accessed: Sept. 16, 2017].

11. K. Saariaho, *Laconisme de L'aile*, Paris, FR: Edition Wilhelm Hansen, 1982. [Music Score].

12. P. Spencer, "The musical shape of technology," *The Flutist Quarterly*, vol. 19, no. 3, pp. 47-51, Spring 1994. [Online]. Available: http://judithshatin.com/kairos/. [Accessed Sept. 23th, 2017].

13. C. Palacio-Quintin and M. Zadel, "Interactive Composition and Improvisation on the Hyper-Flute," in *ICMC 2008, Belfast, North Ireland* [Online]. Available: http://hdl.handle.net/2027/spo.bbp2372.2008.088. [Accessed: Sept. 17, 2017].

14. C. Palacio-Quintin, "Eight Years of Practice on the Hyper-Flute: Technological and Musical Perspectives," in *NIME 2008*, *4-8 June 2008*, *Genova, Italy*.

15. Département de musique de l'Université de Moncton, “Improvisation avec la Hyper-flûte – Cléo Palacio-Quintin,” *YouTube*, October 10, 2014 [Video file]. Available: https://www.youtube.com/watch?v=X1zJca_VaBw. [Accessed: Sept. 17, 2017].

16. C. Palacio-Quintin, [Email], Dec. 10th, 2017.

17. I. Pace, "Notation, time and the performer’s relationship to the score in contemporary music," in *Unfolding Time,* D. Crispin, Ed. Leuven University Press, 2009, 151-192. Available: http://openaccess.city.ac.uk/5418/

18. S. McClary, “This is Not a Story My People Tell: Musical Time and Space According to Laurie Anderson,” in *Feminine endings: Music, gender, and sexuality*. Minneapolis, MN: U of Minnesota Press, 1991, p.136.

MUTATIONEN III FROM CLAUDIO SANTORO: A CASE STUDY OF ANALYSIS AND REINTERPRETATION OF A MIXED PIECE

Fabio Wanderley Janhan Sousa
Music School – Federal University of Minas Gerais
fabiojanhan@gmail.com

Margarida Borghoff
Music School - Federal University of Minas Gerais
guidaborghoff@gmail.com

ABSTRACT

Attempting to break the coldness of sitting in a room listening to fixed music, composers of electroacoustic music have been trying to add some vividness to the concert by adding live electronics, images, lighting or a human performer. Mixed music is the terminology we use to define the practice of composing for acoustic instruments accompanied by a pre-recorded electronic part. The present article describes a practical experience of reconstructing the pre-recorded part of the piece “Mutationen III” [1], written for pre-recorded sounds and piano by the Brazilian composer Claudio Santoro. With this experience and the research of new sonorities based on the documentation presented in the score, a new version of the work was developed using some of the most recent tools for audio manipulation. This exercise proved to be an excellent opportunity to develop both analytical and compositional skills, as well as the ability to the usage of digital audio manipulation tools.

1. INTRODUCTION

What we are used to call “acousmatic music” has a format that is fixed by the composer in a physical media, be that analogic or digital, in which is contained and realized the whole artistic planning of pitches, timbres, rhythms, gestures and textures parameters and developments. The listener, who simple hear the reproduction of it, is already relating himself with the plenitude of the piece. However, resources such as live spatialization, for example, historically come into play to solve the big problem of the acceptance of this kind of concert.

According to Amorim [2], “electroacoustic music, in its most primitive form, eliminates the interpreter and attributes to the composer a new function, that is to be responsible not only for the construction of the sound material as a formal structure, but also as a physical object which contains the work performance in a latent state”.

Throughout time this purely acousmatic music, most times considered cold and sterile from the listener point of view, searched for ways to relate with other elements, mostly manipulated in real time, to be added to the performing practice. These manipulations normally follow three different approaches. The first one would be the spatialization in real time through multichannel systems difusion, making the hearing of the piece a performance through the space of listening, most commonly executed by the composer himself, as explored by Vande Gorne [3] for example.

The second one would be what we usually call “live electronics”, which presents some kind of interface, be that an acoustic instrument, a computer, or any other device, to interfere with electronic sound, or to be interfered by them in real time during the performance.

A third one would be what we call “mixed music”, which can be put in between the other two by introducing the vividness of an interpreter to the cold performance of the tape music but avoiding the unpredictability and the limitations of the “live electronics”. Usually this type of music presents a pre-recorded part and a part that is assigned to an acoustic instrument for the interpreter to perform live.

According to Zattra [4], the first experiments on mixed music were made by Edgar Varèse (“Désert” – 1954), and Karlheinz Stockhausen (“Mantra” - 1969). The first one being an example of what we call “mixed music”, presenting the interpolation of electronic timbres with acoustic ones, and the second being an example of “live electronics” with sounds being generated and manipulated in real time, interacting with the acoustic sounds.

Claudio Santoro, a Brazilian composer born in 1919 in Manaus – Amazonas, wrote twelve pieces that could be defined as “mixed music”, presenting a pre-recorded part and a score for the instrument to be played on the stage. These pieces were written during his stay in Berlin and Heidelberg, in Germany. Among these twelve pieces, denominated “Mutationen”, the third one was written for piano and tape and has been the object of a few works in both composition and performance fields such as those from Guerra [5], Ventura [6] e Martins [7].

Since the tape part of this piece is relatively easy to make, a few of the above-mentioned authors propose that the performer could create it by himself. Although it is perfectly possible, our proposal was to create something beyond this simplicity, taking the most out of the piece through the collaboration of the performer with a sound engineer, which is also a composer.

2. RECREATING THE ELECTRONIC PART

The electronic part, describe in the score of “Mutationen III”, for piano and tape, is made with sounds recorded from the piano itself that are manipulated by processes which are relatively simple such as accumulations, acceleration, or slowdowns (with the change of the pitch), reverse, attack cuts and eventually literal reproductions of what was previously recorded.

As does Zattra [4], “we think that the knowledge of the historical period and instruments typical of the musical repertory is fundamental, because electronic music equipment, with their potentials and limits, influence the typology of sound, the compositional process, the performance and the listening”.

It is important to point out here that, although we intend to make a technologically informed recreation of the composition, we did not use tape manipulation itself. The main reasons would be the unavailability of these resources at the recording labs and the intention of developing an exercise that could be performed by students in music and technology classes, using relatively common materials and equipment. Therefore, the technological informed recreation is limited to the processes and not to the media or equipment.

We understand that those processes we called “simple” were the most technologically advanced tools the composer had available at the time of the composition, but still, to begin our experiments, we perform the creation of the tape part in the most faithful way it was possible, obtaining a “literal” version of what was suggested by the original score.

Although we used protools software to make the electronic part, it could be accomplished at any DAW (digital audio workstation) such as Reaper, Sonar, Cubase, Digital Performer, etc, since the manipulations the piece requires are common to all of them. Cutting, accelerating, slowing down, reversing or accumulating sound tracks are “easy to make” processes at all the above-mentioned softwares.

Although the process of creating the electronic part of the song following the composer’s orientations present itself as a very interesting task, during the process, we found out excerpts that would benefit enormously from contemporary tools of digital audio manipulation to better translate the suggested graphics the composer present in the score. This led us to the development of a second version, we may say, “up to date” technologically.

3. CREATING A REVIEWED VERSION

As a research tool of the compositional process, we choose to generate a version as described by the composer and reflecting the limits we believe he had at the time of the composition. As an extension of the research process, we also propose ourselves to review the tape part discussing and proposing new solutions, which could be more interesting through the usage of a few contemporary tools. Following are the descriptions of three excerpts in which the recreation presented itself most improved and evident.

In the first system of the piece, the composer presents an electronic gesture that is generated from successive attacks of a chord at the piano keyboard by the performer that goes slowing down through time.

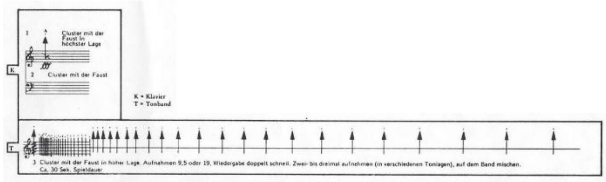


Figure 1. First system of the piece “Mutationen III”.

Link 1. www.youtube.com/watch?v=X_wVXguVRpc

According to the composer’s orientations, this gesture should be recorded in different piano regions, accumulated and played back simultaneously. While generating this electronic gesture, we faced two basic difficulties. The first difficulty is that the recording of multiple versions to be accumulated requires a lot of the musician’s musculature who has to repeat the same gesture multiple times in different regions of the piano. The second difficulty is the physical limitation and training of the musician that leads to very similar results regarding the initial time of repeated chords and the time between the chords while slowing down.

In the reviewed version we realized the electronic part as suggested by the composer but we also feed delay lines with feedback times that slows down trough time, as does the time parameter between the repetitions, accomplished by manipulating an automation line. This effect presented itself as richer in color and more linear rhythmically (fig 2).

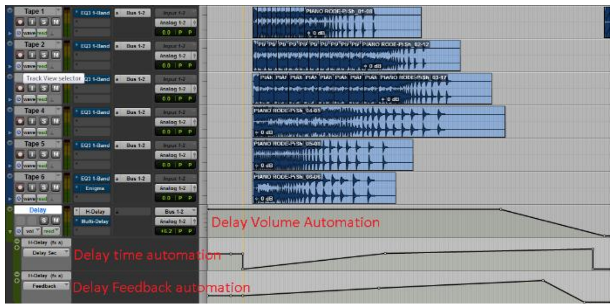


Figure 2. Accumulation of the recordings added to the delay line with Time, Feedback and Volume automation.

Link 2. www.youtube.com/watch?v=xMUcYUwRYNQ

In the half of the third system of the piece, we observe an electronic gesture, which, if performed using only recording methods as suggested by the resources described at the original score, can be extremely difficult to be executed by the pianist. At the score the composer wrote: “Staccato accelerando and diminuendo (Staccato stroke faster and faster and softer and softer), recording 3 3/4 or 7 1/2 ips, playback at double speed. The diminuendo can be produced during the recording with the aid of the volume control”.

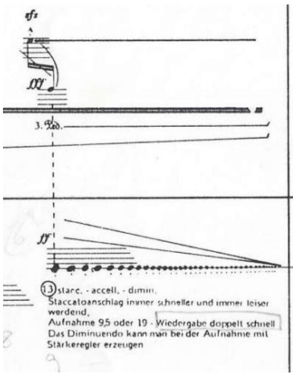


Figure 3. Gesture of the third system of the piece “Mutationen III”.

Link 3. www.youtube.com/watch?v=ll9MJo_5kiM

At the reviewed version of this excerpt, we made the recording of a single strike and feed a delay plugin with it. By controlling both “time” and “volume” parameters through automation, the sound of the stroke repeats slowing down and getting quieter. This process is much simpler and controllable when performed by the computer than when performed by the musician.

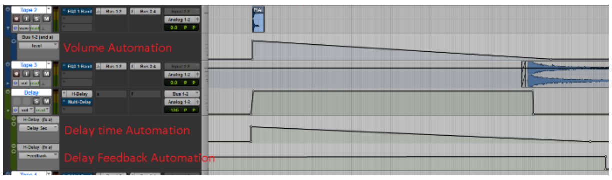


Figure 4. Implementation of the gesture in fig. 3

Link 4. www.youtube.com/watch?v=tCZ_6MHfFVY

Another observation we can make about this same excerpt is that, according to the composer’s orientations, the tape must be played back in the double speed. Such process, in the digital domain, is extremely simple to perform if a time compression / expansion tool is used, followed by a pitch shift in an octave higher than the real sound. The pitch shift manipulation is necessary since we must consider that, while manipulating tape recordings, this is exactly the effect that we would perceived. The sound is not just the double speed but also the double frequency, therefore, an octave higher than it sounds originally.

At the fourth system of the piece the composer suggests that we record several times an excerpt of the piano player touching the strings with his fingers and that this recording be repeated and accumulated. This results in a granular texture, which is also suggested by the graphics of the score itself.

Granulation tools have been developed mainly after the 1990’s and they perform the task of accumulation of sounds much more effectively and interestingly than just repeating and overlapping recordings, as suggested by the composer of the piece in several excerpts. In our reviewed version, we recorded the excerpt as suggested but instead of accumulating it with other copies, we feed a granulator plugin with it - in the occasion the Crystalizer from Sound Toys (fig. 6). This plugin cuts small samples between 15 ms to 25 ms long, alternates, repeats and toggles them in the time domain as well as in pitch, creating a much more interesting texture than the tools from the period of the composition could make possible.

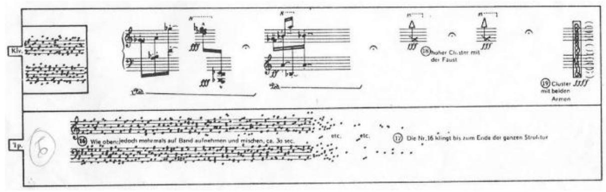


Figure 5. Fourth system of the piece “Mutationen III”.

Link 5. www.youtube.com/watch?v=I4JJEU9vs0Q

Figure 6. Crystalizer, plugin for audio granulation.

Link 6. www.youtube.com/watch?v=8VIYeKBkZTY

4. MAKING THE BEST OUT OF THE MUSICAL GESTURES

By exploring new tools to recreate musical gestures suggested by the graphical score of the piece and the notes made by the composer, we decided to add gestures in the electronic part and to elaborate the already existing ones, having in mind not to deprive the piece or corrupt the original material of the composition.

The composer himself, Claudio Santoro, is very open to such reinterpretation of his pieces and suggests that we explore such possibilities as reported by Lívero [8]: “I have always been like this, I have never been a composer who wrote something to be played exactly as I think it should. I always let the performer to recreate the piece and give something from himself. I always thought this way and I believe that is why I have been both interpreter and composer”

The experience of recreating the piece enriched itself through the collaboration between a sound engineer who is also a composer and the performer. We created gestures more interconnected, related to each other and elaborated, making the piece as a whole something much closer to the contemporary aesthetics.

As an example of this contribution, we can point out the gesture presented at the end of the third system (fig. 7) and beginning of the fourth (fig. 5). According to the score, the recorded part present a second minor glissando with a coin performed directly at the strings of the piano, which should grow up in intensity and connect to a granular texture. The crescendo of a tone, which naturally decays such as those created by this specific gesture at the piano, can easily be implemented with the usage of a volume automation line.

In most times this raw manipulation sounds very artificial and obvious.

Therefore, in the reviewed version (fig. 8), we added to the original sound of the second minor glissando, the same second minor reversed, in such a way that it generates a crescendo that reaches the attack and then continues with its resonance. This resonance becomes granular through some random volume automation, which introduces the granular material that follows, which in turn, has its granular characteristics reinforced by a granular plugin, already mentioned.

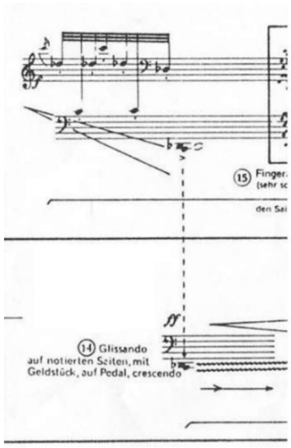


Figure 7. End of the third system of the piece “Mutationen III”.

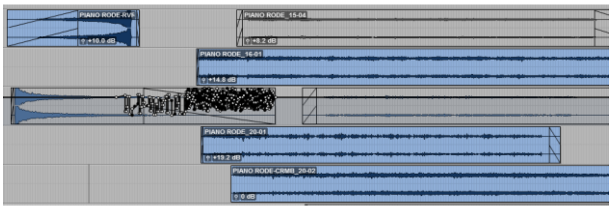


Figure 8. Implementation of fig. 7 and fig. 5 at Protools.

Link 7. www.youtube.com/watch?v=4T52c41HObY

A second example of these musical gestures improvements can be observed at the end of the fifth system (fig. 9) and the beginning of the sixtieth (fig. 10). In this excerpt, we used the pre-recorded harmonic sounds added to some extreme long resonances and followed by the attacks of these same resonances in reversed to culminate in the low pitch notes attacks (fig. 11). The creation of this texture, only suggested by the score, was developed by listening and analyzing the piece during the recording sessions. It enriches the pre-recorded part by blending the acoustic and tape parts, and expands the usage of the materials created by the composer.

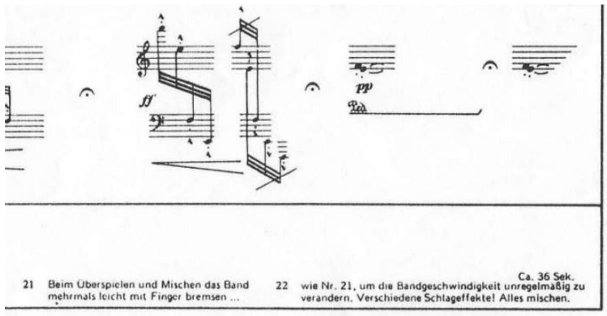


Figure 9. End of the fifth system of the piece “Mutationen III”

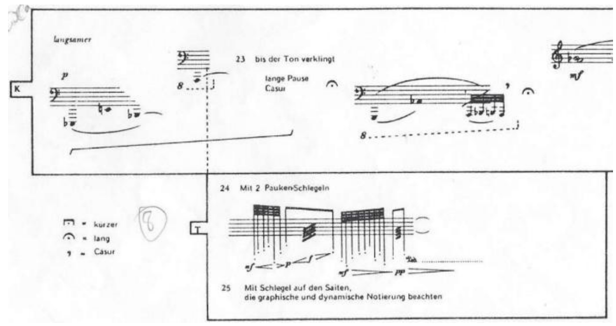


Figure 10. Beginning of the sixtieth system of the piece “Mutationen III”.

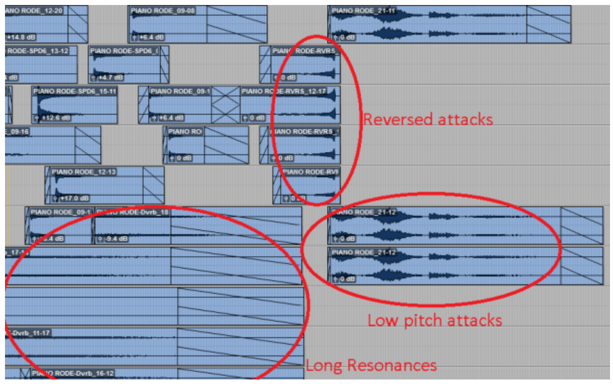


Figure 11. Implementation of fig. 9 and 10 at Protools.

Link 8. www.youtube.com/watch?v=tRTbYa5yNUo

5. EDUCATIONAL EXPERIENCE WITH THE GENERATED MATERIAL

The rhizomatic characteristics of the described experience is worth been noticed since it starts as an analysis exercise, goes through the usage of technological tools of digital audio manipulation, demands the development of a musical performance, inspire some composition skills and goes back to musical education in the form of an exercise that was proposed inside the classroom.

The main intention of making this experience also a music technology exercise for composition students is based on the perceived need for them to manipulate digital audio in the composition studies as well as in their music production activities, imposed by recent media communication.

The recorded material generated by the reported experience was used as an exercise during an undergraduate class at the Music School of the Federal University of Minas Gerais called “Digital audio manipulation”. The present author gave these classes during his teaching internship as a PhD student.

In this exercise, the students were asked to make the tape part using the original recordings, made by the author, and following the score. To master the proper digital audio manipulation tools to perform the cutting, accelerating, slowing down, reversing and accumulating sound tracks processes was the main goal of the classes. To discuss historically informed pieces of the XX century, the transformations of the notated references to sound manipulation and to assembly it in a digital audio workstation, as well as the formation of an audience aware to the aesthetics present, were secondary goals that was explored in the classes.

The acceptance of the exercise by the students was excellent since it brought them closer to practice and generates a production worth to be presented to the public. Topics such as the use of delays and granulation, which were also goals of the above-mentioned discipline, were received with much more interest since they are presented as practical solutions for compositional issues.

6. CONCLUSIONS

The delay is a phenomenon that is usually observed acoustically as an “eco” when its time is too long. Its usage in audio suffered a revolution and become more versatile after the process of digitalization that made possible the manipulation of very short and very long times. The usage of the delay as well as the automation of its parameters show themselves as important tools now a days that would certainly be used by the composer.

The granulator, as a recent developed tool, present itself as an alternative to generate sound textures which are very common in contemporary compositions and which, if it was available at the time of the piece, surely it would be very exploited by the composer, as the score itself suggested.

The process of recreating the electronic part of a piece after reflecting about its proposals and to seek new solutions using recent tools, showed itself to be an extreme thought-provoking and self-accomplisher task. Possibly if adopted in the training of contemporary composers, this experience could add as much in the development of their abilities of manipulation digital audio as in their analytical and creative abilities.

Developing activities that can cross different areas from analysis to composition, performance, research and musical education presented itself a very interesting and important method of improving teacher and students’ skills.

Acknowledgments

Acknowledgments to the pianists Guida Borghoff, for being available for the experience, to the Music Conservatory of Minas Gerais and to the School of Music of the Federal University of Minas Gerais, where the recording sessions took place.

7. REFERENCES

[1] Santoro, Claudio. “Mutationen III für klavier und tonband”. *Studienpartitur*. Tonos International, Darmstad, 1971.

[2] Amorim, Felipe. “A Interpretação Fixada”, in *Revista Modus*, 8 / 13 Belo Horizonte – MG, 2013, pp. 51-58.

[3] Vande Gorne, A. V. “L’interprétation spatiale”, in *Revue DEMéter- Revue Électronique*, Université de Lille-3, 2002.

[4] Zattra, Laura. “Analysis and analyses of electroacustic music”, in *Proceedings of the Sound and Music Computing International Conference*, Italy, 2005.

[5] Guerra, Anselmo. “MUTATIONEN III de Cláudio Santoro: uma Releitura Eletroacústica”, available at <www.academia.edu/831932/MUTATIONEN_III_de_Cláudio_Santoro_uma_Releitura_Eletroacústica> Access in 09-12-2017.

[6] Ventura, Fabio. “Mutationen III de Cláudio Santoro”, in *Anais do congresso da Anppom*, São Paulo - SP, 2007.

[7] Martins, Denise A. F. “Análise da obra eletrônica Mutationen III de Cláudio Santoro”, in *Revista Eletrônica OPUS* 8, Brazil, 2002.

[8] Lívero de Souza, I. “Santoro: uma história em miniaturas”. *Master Dissertation*, Campinas University – São Paulo, 2003.

ON TREVOR WISHART'S MUSICAL TRADE-OFFS: COMPOSITIONAL GOALS, MATERIAL'S SELECTION AND SUITABLE SOUND PROCESSES SUPPORTING THE LISTENER PERCEPTION.

Fellipe Miranda Martins
f13martins@gmail.com

ABSTRACT

Based on Wishart's literature, this paper investigates the most frequent and significant recorded voice transformation processes, and its correlation with the composer's compositional goals, presented on four Trevor Wishart's acousmatic pieces: Vox-5, Tongues of Fire, Globalalia and Encounters in the Republic of Heaven. Although Wishart developed new sound transformations using the phase-vocoder, the selection of the materials takes a key role in determining which sound metamorphosis would fit his aesthetic goal: retain the source credibility from the listener perspective. This articles shows how his compositional goals in each piece demands both specific recorded materials and precise transformation processes in order to achieve his intention with respect to the listener. Hence, one of the main relevant aspects of Wishart's work for electroacoustic composers is to reveal – from the audience point of hearing – subtleties between natural and artificial transformations of voice samples.

1. INTRODUCTION

1.1 Background

Trevor Wishart is a freelance composer/performer based in York, UK, who made an extensive theoretical and compositional work on human voice and sound metamorphosis. He is the main designer of Sound Loom and Composers Desktop Project: softwares focused on transformations of sound recordings.

The Composers Desktop Project (CDP) is an open source and collaborative software that dates back from 1986, when Wishart started an artist residency at Institut de Recherche et Coordination Acoustique/Musique (IRCAM). During this period he devised his firsts computer tools based on the phase vocoder: spectral stretching and sound morphing. This collaboration gave birth to CDP and to the piece Vox-5.

1.2 Literature Review

Wishart main theoretical works are On Sonic Art, Audible Design and Sound Composition [1, 2, 3]. Though he claims that the contents are rather speculative and non-scientific, his deep and vast discussions of electroacoustic music are prestigious and largely cited in the academic world.

Topics like psychoacoustics; voice production and extended vocal techniques; sound morphology; lattice music and sound continuum; real time and non real time composition; relations between sound and image, and its theoretical concepts, occupy a central position in his work.

The deep exploration and usage of human voice is a backbone of the composer's whole work. From the exploration of extended vocal techniques for live improvisation to the computational transformations of vocal sounds into sounds from the nature, Wishart unveiled the limitless creative possibilities offered by this multi-purpose human apparatus.

Wishart is strongly interested in the sounds of the natural world, in opposition of synthesized sounds. Therefore the CDP is strongly oriented towards processes that transform samples of audio recordings – instead of being a platform for audio synthesis, algorithmic composition or live coding.

As a natural consequence of the composer interest in vocal sounds, a large set of the CDP functions is somehow devised for working with voice. Pitch-trackers, formant extractors and imposer, FOF[1] extractor and imposer, spectra combination, vowel imposition, vocoder, independent time-stretcher and pitch-shifter (phase-vocoder) and many others assist the composer's aim: transform the human voice.

[1] Fonction d'Onde Formantique (FOF) or Formant Wave Function Synthesis is a spectral synthesis method developed in 1984 at IRCAM by Xavier Rodet as part of CHANT project.

2. COMPOSER'S PERSPECTIVE: SOUND METAMORPHOSIS AND SOUND CONTINUUM

2.1 Criticism on notation and super-conceptual thinking.

Throughout his literature, Wishart makes strong criticisms on the oppression that notation and philosophical thinking had imposed to music. Firstly, he claims that notation imposes and creates certain musical priorities, and not the way around. Secondly, it imposes that many musical parameters – pitch, rhythm and loudness – a strictly obeyer to lattice logic. As a consequence, numerous possibilities of musical expression have been mowed by this system.

The author states that this notation tradition withdrawn music for its basic purpose: being heard. He attacks late Twentieth Century composers arguing that their music suffers of "obsession with complicatedness" because they lay out so many complex mathematical and relational ideias for the listener without careful testing if the results are truly audible [1]. In his own words:

> For me, on the other hand, a musical experience which appears aleatoric is aleatoric. The experience that the listener has is the music and the composer's methodology, no matter how rational it may be, is a heuristic device realising the source of that experience in sound. (Wishart, 1986, pg. 43)

Wishart also claims that this notation priorities and lattice logic in tradicional music also imposed a secondary role to timbre. One of the consequences is that music instrument design and their relative performance techniques have been imprisoned in the dogma of being sound sources of stable timbre and variable pitch. Thus, the piano is the ultimate outcome of this lattice conception due to its severe submission of timbre to discrete pitches and to its absence of possibilities regarding continuous sounds.

Summarising, this lattice structure attempt to negate the importance of recognition of the source and focus the listener attention upon the discretized parameters [1]. Consequently, this leds late Twentieth Century Western Music to become super-conceptual, abstract and distant from the appreciation of the audience.

2.2 Sound Continuum

On Sonic Art (1986) proclaims the full liberty of usage and exploration of sound parameters, allowing them to express the continuity that is typical of natural sounds in the natural world. It urges composers to free themselves from tradicional lattice based music, which restricts the infinite possibilities of music to series of discrete steps events.

Hence, the composer proposes a characterization of the sonic continuum as well as a methodology for working with it. Wishart does not intend to make a exhaustive description of sound objects – in a way that the Groupe de Recherches Musicales (GRM) did – nor he wants to postulate meticulously a morphology of sounds in the continuum. Despite of that he discusses non-standard metrics for the continuum, points some conceptual approaches to the material and list some of the perceptible and diferentiables archetypes [1].

He also adds that the importance of the sound continuum is related to our experience of the real world. He claims that the lattice concept and traditional notation are only rough approximations of what reality is:

> Music, however, cannot be divorced from the medium of sound and enters into our experience as part of an immediate concrete reality; it impinges on us and in so doing it affects our state. Furthermore, as Susanne Langer remarks in Feeling and Form, in its articulation of the time-continuum of concrete experience, it corresponds directly with the continuum of our experiencing, the continuous flux of our response-state. (Wishart, 1986, pg. 16)

2.3 Sound Metamorphosis

Wishart has debated extensively in his theoretical work about sound metamorphosis and its difficult to point out a piece of him in which this idea it is not presented. As often mentioned, his main goal in doing so is to retain source credibility, keeping the listener aware of what is being transformed and metamorphosed.

> We can proceed seamlessly from one point to another (a continuous transformation) or by discrete steps (a sequence of discrete metamorphoses, more akin to traditional motivic variation, perhaps). I use the term metamorphosis to refer to the sonic manipulation of a sound to produce related sounds, while the term transformation refers to a process of sonic development through time; that is, the use of sonics relationships between events to build musical structures in time. This is also referred to as sonic modulation by analogy with motion between keys in the tonal system. (Wishart, 2000, pg. 22)

Wishart also makes criticisms on the attempts of tradicional lattice based music to represent natural sounds. He claims that figurative music, especially by some canonical orchestral composers like Tchaikovsky, Respighi and Beethoven, never presents truly audible results, even when it attempts to mimicry sounds related to orchestral instruments (e.g. birds). In the author point of view, this is once again an example of overvaluation of notation and theoretical thoughts in spite of audible phenomena [1].

As he has worked both on analog studio and with modern computer techniques, the ideia was pushed to its limits. Facing many problems in the analog studio and reaching its limits, Wishart proposed collaboration with IRCAM and developed a large set of computational procedures to aid or execute sound metamorphoses. As this topic transversely touches most matters concerning the composer's work, it will be discussed throughout the paper.

2.4 Red Bird: early experiments and results

Red Bird (1978) is an acousmatic piece made entirely on the analog studio. Here Wishart has achieved notorious results, both technically and aesthetically, concerning sound metamorphosis. In this piece, the composer begins his long journey to attempt to make sound metamorphoses using the human voice.

Despite of the well known problematic points of working with recorded sounds on a strictly analog studio, one important aural issue emerged from his technical difficulties: some sounds do not resemble a specific aural image, e.g. a book being slammed into a table can sound like an uncountable situations and objects. Thus, contextual cues often need to be presented in order to induce the audience (e.g. sound of pages being flipped).

Wishart also encountered many problems concerning the metamorphoses of vocal samples into both clearly defined and doubtable aural images, specialy due to the lack of possibility concerning the analog studio. For instance, two sounds that he declares that were problematic were flies and machines. The first one, when mimicked by the human voice only sounded natural when isolated, when in context with other sounds it did not sound natural. The second one, due to formant characteristics and continuity of vocal stream, when tried to be mimicked only sounded like human voice. It was difficult to achieve the mechanical and straightly repetitive behavior of machines.

3. LISTENING: THE MAIN GOAL AND PURPOSE

> The computer, having no ear or human musical judgement, can manipulate order sequences of any length and of any unit-size in an entirely equivalent manner. As composers, however, we must be clear about the relationships of such operations to our time experience. Equivalence in the numerical domain of the computer (of for that matter on the spatial surface of a musical score) is not the same as experiential equivalence. Audible design requires musical judgement and cannot be "justified" on the basis of the computer-logic of a generating process" (Wishart, 1994, pg. 64)

3.1 Aural Criterion

Considering the criticisms mentioned above, Wishart is naturally a composer strongly oriented towards aural and perceivable aspects of music, despite of the metaphoric and philosophical meanings. Thus, his main goal is to achieve processes that sound for the audience exactly as he wanted them to sound like.

He asserts that, technically, almost anything can be a metamorphosis and any technical metamorphosis can be proposed. However, perceptually, the listener perception is the only reliable criterion that can state what is a successfull metamorphosis [4].

A second point is how each metamorphosis is related to each other and how they are organized across a piece. He claims that he is not merely interested in creating soundscapes with dimensions of recognition or lack of it. Otherwise, he is interested in constructing compositions with interplay of recognisable aural images which is quite different from the use of degrees of recognisability in an otherwise formalised musical landscape [1]. This approach motivated some authors to classify Wishart's style as storytelling or narrative [5].

3.2 Source Credibility

Regarding the importance of listener perception for his viewpoint of music, the first issue that arises is importance of the source credibility in order to make a truthful metamorphose.

His point of departure can be any sound, since it establishes a connection with the audience. Usually this is done by sequenced repetition of the material – when it has a short duration – or by utilizing materials that have long continuations (e.g long vowels). This can be shown by the fact that Wishart hardly or never presents the transformed material and then goes step-by-step revealing the original source. His transformation route is mainly the way around: showing the raw material and then transforming it step-by-step.

The importance of the source credibility is also attested by the frequent usage of the human voice: surely the most easily recognisable sound for humans. This topiv will be more deeply analysed below. When the sound source does not resemble a strictly specific aural image, Wishart also gives contextual cues to help the listener, as mentioned in Red Bird example.

3.3 Intrinsic acousmatic problem

As mentioned before, Wishart basically works in two main areas: acousmatic music and live acoustical performance of extended vocal techniques. Concerning the listener importance in his viewpoint, one crucial intrinsic problem of acousmatic music for Wishart's goals is that it is acousmatic: the absence of visual, (realistically) spatial and tactile stimulus fools and confuses the audience recognition of the sound source and its physicality.

On Sonic Art (1986) debates what would be a physical parallel of the sound transformations that he makes, suggesting some mathematical, physical and biological theories that would support the idea. The composer also asserts that sound metamorphoses often exhibit a dreamy and supernatural characteristic due to the exclusive focusing on the sense of hearing.

3.4 What the audience can perceive?

Wishart has debated extensively in his literature what the audience can perceive. While on On Sonic Art (1986) he shows a more philosophical approach, on Audible Design (1994) and Sound Composition (2012) he discussed and exemplifies which CDP transformations is more suitable for attending both compositional goals and fit audience perspective.

He often quotes and discusses achievements of psychoacoustics in his literature, however in his writings his is more concerned with testing and exploring the boundaries of these achievements, always providing sonic examples.

Once again, he claims that he is not interested in presenting ideas, sctrictly philosophical thoughts, notation procedures, etc for the listener. He is interested in presenting sounds, organized materials and structures:

> In composing, I differentiate between the procedures I use to generate compositional materials and the structures I lay out for the listener. It's not important for the listener to know anything about my compositional procedures, but I hope they inform the formal and dramatic structure that I do hope the listener will perceive and appreciate, though not necessarily consciously (Wishart, 2000, pg. 22).

> (...) if the organization of our music is to be based on the audible reality of the listening experience, the music must be organised according to *perceived* relationships of materials or *perceived* processes of derivation of materials.(Wishart, 1986, pg. 66)

4. SELECTION OF MATERIALS

> (...) Achieving a convincing transformation between two sounds is a practical problem of sonic art. The transformations between the sounds of different instruments playing the same note (e.g. oboe-flute) are very convincing as sonic transformations but unfortunately totally uninteresting as metaphors. The transformation voice -> bird-sounds is metaphorically quite interesting but much more difficult to generate.(Wishart, 1986, pg. 166)

4.1 Voice

As investigated above, On Sonic Art (1986) makes strong criticisms on the tradicional instruments and their absence of timbral possibilities. Wishart opposes the keyboard – which represents the ultimate point of this lattice conception – against the human voice, which is can be viewed essentially as a continuous timbral articulator:

> "The human larynx and vocal tract as well as the birds' syrinx, however, are altogether more malleable systems, allowing control through the pitch continuum (as with unfretted string), the continuum adjustment of formant spectrum and transformations in the direction of noise, grain, grit or of harmonic and multiplexed textures. Subtle muscular control of the physical apparatuses which govern the emission of sounds ensures that intentional gestural input may be exceedingly subtle and multi-dimensional. Hence the voice lies at the farthest extreme from the typical musical instrument developed in the West" (Wishart, 1986, pg. 103)

Wishart elects the voice as his prime material because of malleability and instrinsically natural human recognition. Clearly, it is the best choice to fits his viewpoints of sound continuum and sound metamorphoses and also his necessity regarding the listener perception.

All through his work he explored the numerous possibilities offered by the voice: vocal extended techniques; mimicry; sound of the text and relations with its meaning; phonetics and exploration of syllables across the world languages; age, gender, nationality and other intrinsic parameter of individual's voices, to list a few. In each of these compositional goals, he searched for specific transformation processes to best fit the audience perspective.

4.2 Departure and arrival materials

Wishart's metamorphoses do not strictly goes from one strongly specific aural image to another strongly specific one. His aim though is to make the departure from a sound as clearly audible as possible and to make the transition to the arrival sound as convincing as possible.

He firmly states that he is interested in *source-focused transformation* (where the nature of the resulting sound is strongly related to the input sound, (e.g time-stretching of a signal with a stable spectrum, retaining the onset unstretched) in spite of *process-focused transformations* (where the nature of the resulting sound is more strongly determined by the transformation process itself, e.g. using very short time digital delay of a signal, superimposed on the non-delayed signal to produce delay-time related pitched ringing) [2].

He claims that transformations focused in the source retain the same infinite potential that the infinity of natural sound sources offers. On the contrary, process-focused transformations soon easily become overused because they tend to be used inadvertently with any source [2].

This interest is also show in many of Wishart compositions where he allows himself to "run with the result" [4]. For him, the main point is listening to the whole process and perceptually evaluate if the source and goal sound are related, despite of which process was done [2]. Moreover, the composer also claims that the sounds in the middle of the process are extremely important, but the process itself might be transparent to the audience.

5. TECHNICAL AND AESTHETIC GOALS

Wishart declares that each of his pieces proposes both a new poetic/aesthetic goal and a new technical challenge [6].

> I usually have a poetic idea in the background when I make a piece but I also need a technical challenge because I get bored easily. (Wishart, 2013, pg. 513).

Regarding this, his compositional method follow three basics steps in order to develops these goals. His main compositional stages are:

1. Idea itself. Why start on a piece at all?
2. Transform the sounds
3. Determine the structure of the piece. (choosing the piece's form) [7]

Each piece technical and aesthetic goal will be analysed in the next section, relating the to the suitables sources and processes to achieve this goals.

6. PIECES AND PROCESSES

Wishart has made a extensive description of his processes and sources throughout his literature [1, 2, 3]. Thus, here we will point out some key processes and sources on the chosen pieces as well as his technical and aesthetic goals. He carefully selects each computational process regarding a sound source and a compositional goal:

> In particular the sound examples in this book are the result of applying particular musical tools *to particular sound sources*. Both must be selected with care to achieve some desired musical goal. One cannot simply apply a process, "turn the handle", and expect to get a perceptually similar transformation with whatever sound source one puts into the process. Sonic art is not like arranging (Wishart, 1994, pg. 9)

The four pieces analysed in this article were chosen both by their relevance in the author work and because of the idiosyncratic results and approaches that each of them present.

6.1 Problems of transforming voices

The effortless and natural identification of the human voice by human beings is a double edged-sword: the human hearing detects precisely when a voice is transformed. This means that extra caution must be taken when choosing the vocal samples, the transformation and the goal material or aural image.

One classic problem is the preservation of formants in the transformation processes. Wishart has debated it extensively in his literature [1, 2, 3] and proposed some solutions and techniques. As he is constantly using time-stretching methods, and they tend to distort formant structure, this is a constant concern.

Issues concerning vocal attack are also a central topic in his literature and pieces. In his literature, he states that the attack is directly linked with the physicality and causality of the sound, thus it is intrinsic related to the source recognition. Wishart frequently merges vocal attacks with the tail of other sounds in order to achieve a vocal-like sonority; in order to blend materials that might be recapitulated or to reach unique transformations but retaining some source credibility. Alternatively, when he wants to destroy these instrinsic characteristics he often alters the attack of the sound.

The author is well known for his unique voice spectral tools devised using the phase-vocoder. This is a extremely extensive topic and can not be entirely covered in this article. Nevertheless, the analysis and resynthesis stages of the phase vocoder worth consideration. Wishart carefully selects the analysis parameters depending on the source and the spectral process to be applied. The same is true for the resynthesis stage. A negligent selection of these parameters can introduce many problematic artifacts.

Generally, all these problems appears together when working with recorded samples of voices and trying to transform them:

> Compositionally, we tend to demand different things of discontinuous sounds, than of continuous ones. In particular, if we *time-stretch* a continuous sounds, we may be disturbed by the onset distortion but the remainder of the sound may appear spectrally satisfactory. If we time-stretch a discontinuous sound, however, we will be disconcerted everywhere by onset distortion as the sound is a sequence of onsets. Often we want the sound (e.g. in the real environment, a drum roll, a speech-stream) to be delivered more slowly to us without the individual attacks (the drum strike, the timbral structure of consonants) being smeared out and hence transformed. We wish to be selective about what we time-stretch! (Wishart, 1994, pg.55)

Wishart also underlines a decisive point. Once a good transformation is done, and the listener is convinced, the composer can proceed more freely and alter the material more radically:

> (…) but nicest thing is, once you have reached this plausibility you can do implausible things. So once I've convinced you that this is happening for real I can, for example, make that rrrr sound go on forever or I can make it change in tempo, become very regular, i can make it more pitched (…) (Wishart, 2018, 19'00")

6.2 Vox-5

The Vox Cycle is a series of 6 compositions for amplified voices and 4 loudspeakers, except for Vox-5 which is an acousmatic piece based on voice recordings. The whole cycle is based on the possibilities of making sound metamorphosis between vocal sounds and non-human sounds through the use of vocal extended techniques.

Vox-5, commissioned by IRCAM in 1981 and released in 1986, is mainly focused on the computer realization of sound metamorphoses between human voices and commonly recognizable sonic images like crowds, bees, bells and horses. During this period Wishart devised his first computational functions for transforming vocal samples, which led to the development of the Composer Desktop Project (CDP).

Two main transformations where devised: spectral manipulation and spectral and spectral interpolation, both using the phase vocoder [9]. The spectral manipulation techniques consisted of a method of dividing the spectrum into two parts and shifting one of them called spectral shifting; it also consisted of a spectral stretching technique in which the the amount of shift was not constant for each spectral data, but dependent on the frequency position, thus changing the harmonicity of the spectrum. The spectral interpolation consisted of a function for interpolation spectral data from two diferent sound samples through a given time [10].

The piece tries to achieve consistent sound metamorphosis between vocal samples and sounds of bees, horses, crowds and bells. The bell transformation is a strongly important one because of the inharmonic content of bell sounds, which make them impossible to mimicry using only the natural voice. To achieve this goal, he selected a vocal of the syllable “ko-u”, which is somehow simliar to a bell atack and continuation. Then he spectrally stretched it, in order to achieve the inharmonic spectral content of a bell sound.

Wishart enumerates the factors which guarantees the source credibility:

> It is important to understand that this transformation "works" due to a number of factors apart from the harmonic\inharmonic transition. As the process proceeds, the tail of the sound is gradually *time-stretched* to give it the longer decay time we would expect from an acoustic bell. More importantly, the morphology (changing shape) of the spectrum is already bell-like. The syllable "ko->u" begins with a very short broad band spectrum with lots of high-frequency information ("k") corresponding to the initial clang of a bell. This leads immediately into a steady pitch, but the vowel formant is varied from "o" to "u", a process which gradually fades out the higher partials leaving the lower to continue. Bell sounds have this similar property, the lower partials, and hence the lower heard pitches, persisting longer than the higher components. A different initial morphology would have produced a less bell-like result.
>
> This example (used in the composition of Vox-5) illustrates the importance of the time-varying structure of the spectrum (not simply its loudness trajectory) (Wishart, 1994, pg. 35)

6.3 Tongues of Fire

Tongues of Fire was the first piece realized entirely with the CDP, released in 1994. At this period, the CDP has evolved to full suite of processes for temporal and spectral transformations. In this piece, new transformations techniques like waveset distortion, granular synthesis, brassage, spectral freezing, spectral averaging, and many others were used and develop during [3]. It worth noting that in this piece Wishart beggins a long exploration of time-stretching transformation and possibilities, which will play a central role throughout his whole work.

> The "theme" of *Tongues of Fire* is a rapid solo vocal utterance less than 2 sec in length. This theme was chosen both for expressive reasons (it is recognizably human, but slightly grotesque, slightly comical, and without any linguistic content in any existing human language) and for sonic structural reasons (it is a sequence of several spectrally complex and different sounds, thus making excellent raw material for many kinds of sonic metamorphoses). (Wishart, 2000, pg. 22)

The main aesthetic goal here is to transform a utterance in many complex sounds, which must be audible related to original material, but do not necessarily need to achieve a specific aural image. This is achieved through a form of theme and variations, similar to instrumental music tradition.

The main technical goal is develop the four nodes of variantions: rhythmic variations, *voismetal*, *gablcrowd* and *pichstak* and a climax sound of *fireworks*. All the names are composer’s references and cues to similar sounds. *Voismetal* has a percussive attack, it is gradually time-stretched and waveset distorted in order to generate a metallic sound; *gablcrowd* is a texture of gabbling voices, extending the solo voice theme material into a disgruntled-crowd-like event by superimposing several different variants through time; *fireworks* is a texture of stretched vocal glissandos and portamentos that sounds like fireworks.

The *pitchstak* metamorphosis is a series of synchronized percussive vocal attack with different octave transposition, which bring the feeling of a D chord. Here Wishart maintains the vocal attack of the sound and connects to its tails many different transformations in order to retain the source credibility of the processes, a key technique commonly used in his work. This pitch quality of the *pitchstak* sound is often used in Tongues of Fire to make ruptures and articulations between the other materials [4].

It worth noting that Wishart only achieved so many complex metamorphoses in this piece due to selection of a utterance “theme” that is resembles high uniqueness and complexity. He would not have been able to do so if he had chosen a vocal sample of a single syllable, a sung tune or typical spoken phrase.

6.4 Globalalia

Globalalia came out of collaboration with the composer’s friends, whom recorded radios from all over the world. The source material assembles 134 voices in 26 different languages. The aesthetic goal of Globalalia is to understand and reveal what we have in common as human speech communicators. Although the number of words in all human languages together can reach the order of millions, all of theses languages together are built from a much smaller set of sounds, the syllables [3, 11].

Basics technical challenges naturally emerges from this compositional goal: how to organize and catalog the material, divide the syllables, select and transform or group sets with large numbers of small subsamples, etc. Moreover, the composer claims that the main technical innovation was a set of programs to time-stretch vocal iteratives, like rolled "rr" sounds or vocal grit [11].

The problem is that those sounds cannot be time-stretched in a tradicional fashion because they will lose the quality of rapid and random attacks, hence losing the source credibility. The solution for this problem was:

> The new process first searches for the attacks in the sound (which repeat approximately every 50 milliseconds) with an appropriately sized envelope-tracking window (about three times smaller than the typical flap duration). This gives us an indication of where the tongue flaps are. The process then searches away from the peaks of the flaps to find the minimum energy troughs between the events and finally cuts the flaps apart at (identically oriented) zero crossings. (The zero cuts avoid introducing splices into the sounds, which might subtly alter the sonority at this time scale). We then reconstruct our source, using random permutations of this edited set of flaps [11].

6.5 Encounters in the Republic of Heaven

In his late work, Encounters in the Republic of Heaven (2011), the artist scrutinizes the musical aspects of speech and orality: melody and melodic contours, implicit harmonic fields, rhythm, tempo, bar and also the timbre of individuals. He also intended to represent the musical diversity of human speech across a entire community, thus he recorded diverse inhabitants in the region of Durham, north-east of England.

The aesthetic goal of this piece is to try extract the essence of someone's voice in some way. However, the composer came to the conclusion that this is impossible, due to huge complexity of human source recognition and differentiation [6].

The main technical challenge in this piece is the attempt to make a piece with the musical features of larger structures of speech, from spoken phrases, which will therefore have a pitch contour, a speed and a [6]. One piece acts deals with recordings of teenagers, so Wishart had a doubled and paradoxal challenge: hide or protect the individual characteristics of each teenager – due to ethical and law issues – and also reveal speech characteristics that resembles groups of teenagers, gossiping and some intrinsic personal qualities that differentiate them.

The solution came out by merging different approaches to different excerpt of the recorded conversation:

> I used an envelope follower with a large window set to recognize individual vocal syllables, then retained the centre of each syllable, discarding the onset and tail. The syllable cores were then rejoined in a rapid fixed-tempo stream. This reconstruction maintained the pitch contour, the vowel stream and the expressiveness of the speech line (e.g., laughter is perceptible) while completely disguising the semantic content. This material is then juxtaposed with clear text utterances (e.g., "ginger hair!!") with different processes applied. (Wishart, 2013)

7. CONCLUSIONS

Firstly, Wishart's strong criticism on traditional lattice based music demonstrates that this approach is incompatible with his main aesthetic and compositional foundations: sound continuum and free sound objects; hearing priorities despite of intellectual ones, sound transformations and metamorphosis. Wishart is strictly committed to his viewpoint that the audience hearing is what dictates the effectiveness of the composer's ideias and goals. Thus, the sound metamorphosis in the sound continuum comes to light as a system to retain the listener constantly making correlations between the materials and its respective transformations.

Secondly, as the listener perspective and sense of hearing is one of his main concerns, specific tools must be applied to specific sources in order to achieve perceptually successfull transformations. As he is also aspiring to achieve sound metamorphoses in the sound continuum, the voice emerges as the most suitable source for his goals. It is one of the most malleable natural sound sources and allows continuous timbral articulation and it also has highly intrinsic human recognition characteristics.

The composer's working method establishes at least two goals: one aesthetic and one technical. As shown throughout this paper these two goals constraints even more which material can be used and which transformation can be applied to them. As a result, many pieces of Wishart reveal what can or cannot be audibly done to some materials given the mentioned rules.

By the analysis of some of Wishart pieces, this system of trade-offs shows also that his acousmatic music tends to be strongly audibly organized, in a sense that the materials do not articulate freely or in a improvised fashion, but constantly correlating with each other and its transformations.

Acknowledgments

This study was possible thanks to the support and funding of SEMPRE society. I would like to expressly thank Caroline Ferreira Rosa for all her commentaries, suggestions, and reviews. Finally, I would like to thank the anonymous SysMus18 reviewers for their valuable commentaries and suggestions.

8. REFERENCES

[1] T. Wishart, *On Sonic Art*. Routledge, 1986

[2] T. Wishart, *Audible Design: A Plain and Easy Introduction to Sound Composition.* Orpheus the Pantomine, 1994.

[3] T. Wishart, *Sound Composition.* Orpheus the Pantomine, 2012.

[4] Wishart, T. (2000). Sonic composition in Tongues of Fire. Computer Music Journal, 24(2), 22-30.

[5] Landy, L. (1991). Sound transformations in electroacoustic music. *Composers' Desktop Project Manual. York*.

[6] Garrard, C., & Williams, D. (2013). Tools for Fashioning Voices: An Interview with Trevor Wishart. Contemporary Music Review, 32(5), 511-525.

[7] Vassilandonakis, Y. (2009). An interview with Trevor Wishart. Computer Music Journal, 33(2), 8-23.

[8] [CIRMMT]. (2018, February 1). Trevor Wishart – Composing the real. [Video File]. Retrieved from https://www.youtube.com/watch?v=SyixalrH7xo

[9] Fischman, R. (1997). The phase vocoder: theory and practice. *Organised Sound*, *2*(2), 127-145.

[10] Wishart, T. (1988). The Composition of "Vox-5". Computer Music Journal, 12(4), 21-27.

[11] Wishart, T. "Encounters in the Republic of Heaven" in *Toronto Electroacoustic Symposium* (TES 2012) Montréal 2012.

BRAZILIAN MUSICOLOG*IES*: RESOURCES, APPROACHES AND DESIGNATIONS

Felipe Novaes
Federal University of Minas Gerais
f.novaesr@gmail.com

ABSTRACT

Musicology as science, since its adlerian modern foundation to contemporary epistemological fields, presents itself in moving conceptual and paradigmatic frameworks. Nevertheless, the development of musicology in Brazil, in one hand, cannot be located in a parallel relation with German-English programmatic speeches and, on the other hand, it does not need to be placed in this north-south nexus with regards of comparisons. Therefore, as the first time that Brazil receives the SysMus conference, the issue of Brazilian musicolog*ies* presents itself as a good subject for discussion. Moreover, for understand these multiple musicological profiles this paper presents the analysis of the third volume of the series *Diálogos com o Som: Musicologia[s]* (2016), an edited collection with contributions of Brazilian scholars exposing a range of *solutions* and *agendas* towards Brazilian musicological praxis for the 21st century. Although outlined in a plural epistemological dimension, from its analysis it is possible to define the Brazilian panorama in three major fields of concern: *resources*, *approaches* and *designations*. However, the question for a contemporary musicology is not an exclusive issue of Brazilian academia and, within a postmodern condition, Brazilian musicolog*ies* can be perceived in a consonant movement with international discussions; the consonance within divergence as the incredulity of totalizing narratives.

1. INTRODUCTION

It can be affirmed that the adlerian *Musikwissenchaft* corresponded to a specific epistemological regime and the implications of its historical-systematic dyad outlined a field in which the two axes of this relation were faced in manners of a modern perception of science. Moreover, Adler's concept of *culture* mediated all branches and subsidiary disciplines [1] in his view of a comprehensive study of music: epistemological spaces of a *sound production* grasped in a comparative dimension concerned with the *non*-occidental – *Éthene* – or reduced to Central Europe chrono and geographical core.

Nevertheless, musicology as science, since its nineteenth-century binominal designation, crossed by twentieth-century concerns with the appropriate connective [2] – music *as* or *in* culture – and achieved in contemporary regimes presents itself in conceptual and paradigmatic moving frameworks. However, this mobility of values does not only underscores modern and postmodern musicology and, as well, do not bespeak a gradual and cumulative progress or development; in other words, from ancient Greece to present days countless perspectives and practices towards the understanding of music, each one related to a specific *episteme*.

Despite these movable frameworks, the development of musicology in Brazil, in one hand, cannot be located in a parallel relation with German-English programmatic speeches and, on the other hand, it does not need to be placed in this north-south nexus with regards of comparisons. Therefore, and being the first time that Brazil receives the *International Conference of Students of Systematic Musicology*, the issue of Brazilian musicolog*ies* present itself as a good subject for discussion at its 11th edition. Regardless the plural perception to designate Brazilian musicological praxis, the motivation for this paper is the questioning of which epistemological frameworks can be perceived in contemporaneity. Consequently, the objective of this paper is the analysis of the conceptual and paradigmatic frameworks that coordinate musicology in contemporary Brazil.

For understanding the issue, I assume as object of analysis the third volume of the series *Diálogos com o Som: Musicologia[s]* (2016); an edited collection with contributions of Brazilian scholars exposing a range of *solutions* and *agendas* towards Brazilian praxis for the twenty-first century.

Moreover, with a methodological cross-comparative analysis of each author's perspectives it was possible to delimited three major *fields of musicological concern* related to Brazilian contemporary praxis and realities. Nevertheless, these fields are perceived as fragmented and, paradoxically, in consonance with the international discussion.

- Resources: the perception of a theoretical, methodological and material gap and the need to overcome it.
- Approaches: the necessity to deal with the *object-music* with a large scope of perspectives, even though consolidated ones.

- Designations: the possibility of social impacts from an instrumental musicology circumscribed in pedagogical or civic roles.

2. PLURAL PERSPECTIVES

As described in its editorial note, the series[1] *Diálogos com o Som* is a field mapping publication with the objective of disseminate and hold the "avant-garde thinking around the knowledge that constitutes the universe of music and its relations"[3]. Moreover, its third volume, *Musicologia[s]* (2016), was conceived within the editors' perception of an Ibero-American context underscored by fragmented musicological academia profiles and the necessity to understand it. Consequently, the title with an *s* between square brackets, bespeaking an intentional addition: the plural perspective. This article sustains a consonant view.

It can be summarized with Rocha's question in the introductory letter *Musicologia e os seus afetos* [4].

[...] understand and questioning Musicology as a whole within an Ibero-American space: how to define Musicology, and why we should think Musicology?

Regardless the editorial note and Rocha's enunciation, to attend this paper questioning – which conceptual and paradigmatic framework outline Brazilian contemporary musicology? – the following section is related to an analysis of *agendas* and *solutions* presented by each one of the Brazilian[2] authors in *Musicologia[s]*. As the outcome of the analysis, it is possible to define three major categories.

The development of *fields of musicological concern* was delimitated by the conceptual dimension, perspectives and solutions intended to Brazilian musicology and suggested by each author. Therefore, the first procedure was the delimitation of guideline-statements: what is the author's concept of musicology? What is her or his comprehension of Brazilian musicological praxis in contemporaneity? What strategy is – or strategies are – suggested to improve musicology in Brazil?

The second procedure adopted was a cross-comparative analysis of the qualitative data cumulated – the author's statement(s) – and the conceptual definition of axes of disagreement or conformity among the authors' perspectives. As final procedure, the definition of categories capable of highlighting fields of proximity or distance.

Nevertheless, as a *Resources*, *Approaches* and *Designations* perspective towards a multiple field of suggestions, the converging aspects of these categories correspond to macro-structures; in other words, in each of these fields coexist similarities and contrasts (e.g the nature of a methodological procedure or a theoretical bias may differ, still, it can be underscore as a *resource* concern). Moreover, as a large articulation of fragmented profiles, each category is crossed by some conceptual feature in the author's perception (e.g. a theoretical concern within a *designation* agenda). However, due to a qualitative dimension of analysis, it is possible to underscore a central or peripheral – see figure 1 – enunciation.

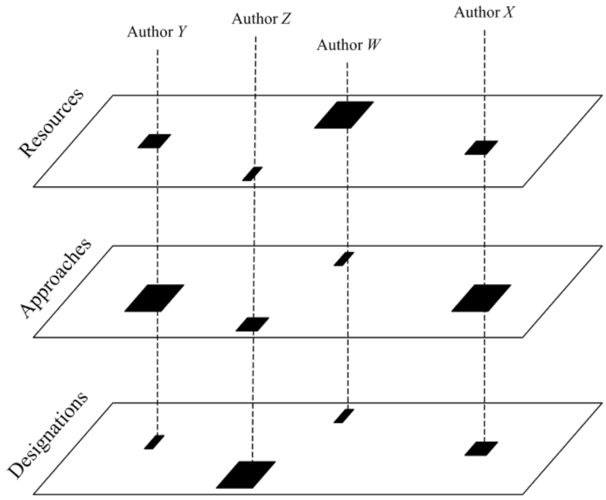


Figure 1. General representation: authors' central and peripheral concerns and its projections.

As final words before the analysis presentation and the description of each suggested field, it is important to say that *Musicologia[s]* represents one of many contributions of each author towards the development of musicology in Brazil. Moreover, is also important to underscore their academia activities as professors, researchers, lectures, scientific event organizers, book editors etc., dealing with and creating academia sociability networks. Therefore, it is necessary to circumscribe *Musicologia[s]* and its authors – as well the present analysis – in a historiographical dimension.

Another important observation is related to musicological field itself. Most of the Brazilian scholars listed are well known from their historical bias in musicological practice. Albeit the clear interdisciplinary perspective exhibit in their contributions, Brazilian praxis – in a large scope – can be defined as mostly historical. Nevertheless, Richard Parncutt definition of musicology [5] and G. Stanley, T. Christensen, V. Duckles *et al* [6] definitions of its disciplines underscores this paper analysis and its author's perspective.

2.1 Resources

It can be defined an academia profile outlined by a *resources* concern that gathers M. A. Volpe, P. Castagna

[1] At present, 2018, the series has four volumes: v1. A música dos séculos 20 e 21 [*The music of 20th and 21st centuries*] (2014); v2. Música e Educação [*Music and Education*] (2015); v3. Musicologia[s] [*Musicolog[ies]*] (2016); v4. Música, transversalidade [*Music, transversality*] (2017).

[2] Although born in Uruguay, P. Sotuyo Blanco has an expressive carrier in musicology with Brazilian subjects: researches about the musical heritage in Bahia state, member of the Brazilian RIdIM and RISM sections and, also, as professor at Federal University of Bahia. Therefore, Sotuyo Blanco integrates the group of authors in the present analysis.

and B.M. Castro[3]. This category highlights the authors' perception of a theoretical-methodological and material gap in Brazilian musicological praxis and the necessity to overcome it in consonance with Brazilian reality. However, each presented paths are unique.

Volpe suggest that this theoretical-methodological and material distance as symptoms of a contemporary epistemological regime in which *full-scale explanation* and *total discourses* are no longer capable of dealing with musicological practices of multi canonical realities, fragmented analytical systems and critical disposals towards musicology writing. Therefore, it must be fulfilled with the search and development of specific musicological tools that, in one hand, provide assistance to Human and Social sciences due to the *object-music* technical characteristics and, on the other hand, reinforce musicology institutionalization [7].

With another perspective about the nature of this lacuna, P. Castagna identifies its origins in an elitist culture devoted to academia instead of proximities with Brazilian reality, especially, with the fundamental bases for a musicological praxis in Brazil: musical archivology. Therefore, for musicology expands it scientific and social meanings – and become more efficient – it is required the development of theoretical tools and methodological procedures adapted to Brazilian specificities [8], such as "the development of the archivistical theory, creation of methods, systems, descriptive cards" [9].

Castro argues that the iconic one hundred years that separates Adler's modern definitions of musicology from Kerman's new-musicology criticism stands as boarders of two conceptual and paradigmatic agendas. Moreover, B.M Castro identify a void of institutional agenda in Brazilian musicology and, therefore, a lacuna dividing Brazil from the international musicological community; in other words, a symptom of an institutional distance. Diverging from Volpe and Castagna's perception, Castro do not define this void as an absence of intellectual exercise in academia [10], but as a yet not answered question of which agenda for musicological research in Brazil.

Nevertheless, B.M. Castro underscores some strategies towards the consolidation of this *agenda* for Brazilian musicology, specially, substantiated with the development of sufficiently systematized materials and instruments of reference capable of dealing with Brazilian concerns, as well, to sustain this suggested new agenda: the proximity with the *R-group* – RISM, RILM, RIdIM and RIPM – due to activities of a new musicological generation; the institutionalization of Brazilian sections such as the IAML-Brazil; and the foundation, in 2012, of the ABMUS – Brazilian Association of Musicology.

Regardless of Volpe's searching of specific musicological tools, Castagna's musical archivology as an efficient strategy towards the development of Brazilian musicology or Castro's concerns with an institutional agenda for Brazil, each one of these authors identifies a lacuna in Brazilian musicological practice. This void is perceived as a distance substantiated by a theoretical, methodological and material hiatus. Moreover, to surpass this lack it is decisive the development of tools, strategies or agenda in a consonant perspectives with the Brazilian reality, therefore, *resources* capable of dealing with musicological praxis beyond strictly relations with international values or conceptual or paradigmatic frameworks.

2.2 Approaches

It is possible to define a field of musicological concerns circumscribing Machado Neto, F. Barbeitas and A. N. N. Cardoso's central perspectives related to a necessity of re-evaluate, in a more rich or sensible *approach*, the handling of the object-music itself. Albeit these strategies towards a new comprehension of music are perceived as theoretical or methodological tools, they are not exclusively related to Brazilian reality – and its necessity of development or adaptation – and, as well, do not bespeak a local field adjustment.

Furthermore, Machado Neto, Barbeitas and Cardoso do not perceive this necessity as a distance or lacuna between Brazilian and foreign musicological practices, but constituted within the object's scientific treatment and perception.

Although one of Machado Neto's objective is "to discuss the statute of Brazilian musicological thought within a global musicological context" [11] with regards of the impacts of new-musicology, Machado Neto demonstrate a wide questioning towards the relation between musicology and musical analysis in post-critical global and local (Brazil) agenda. Therefore, this scope is read as a major enunciation in Neto's discourse. As a peripheral concern, Neto's historiographical analysis reveals that this relation impacted Brazilian musicological praxis and its academia institutionalization, at the end of 20th-century, due to a large conceptual and paradigmatic framework in musical theory and analysis: the Modernist agenda and its 1960 and 1970's criticism.

Nevertheless, Neto's conclusion is emphasized by the necessity to perceive music as an understandable discourse based on symbolic relations with socio-cultural and historical contexts and its constitutive meanings. The key-principle to achieve and understand this relation is changing the object's treatment itself, starting at the "musical linguistics structures, in other words, an analysis of the sound phenomena organized as discourse" [12]. Therefore, despite Neto's perception of a distance between musical analysis and musicology in Brazilian con-

[3] All the following references are exclusively related to the authors' contribution in *Musicologia[s]* (2016). Therefore, when an author's statement is presented and analyzed it concern only with her or his enunciations in its respective paper contribution. The adoption of this guide-line was based on the objective of avoiding multiple *year* references of the same bibliography; e.g Author X (2016), Author Y (2016), Author X (2016) etc. However, every quotation presents itself in the proper reference format and is listed at the *References* section.

text, the nature of a new sensible and rich *approach* towards the comprehension of music relies on a treatment reviewing: music as language, music as discourse.

With another perspective, F. Barbeitas indentifies within an academia crisis in post-modernity his "question for music" [13]. Barbeitas underscores a tension between the presence and operation of a modern institution at a contemporary context circumscribed by different precepts, values and social ordinations. Furthermore, music as an object capable of scientific treatment, presents itself, among other vectors of this question-tension, within the perception of a fragmented sovereignty of musicology in dealing with music at a contemporary comprehension of science. Moreover, a context in which full-scale and definitive explanations are no longer part of academia guidelines; "if everything which science affirms about music *can* be true, it will never be, however, *all* the truth" [14].

Therefore, the re-evaluation of music's scientific treatment has to be made within the perception that "the real strength of music relies on a moment" [15], an instant capable of "transport us", "moving us" and "dissolve us"; a poetic dimension of music free from scientific light. Nevertheless, this metaphysical dealing do not bespeak a de-legitimation of musicology, on the contrary, it outlines the necessity of a critical posture – this poetic moment can be understood by its poetic strategies, its social meanings, its physiological and psychological operations, its historical and cultural relations, its physical organization; this poetic moment can be measured, understood, classified, systematizes etc. However, as science cannot illuminate all branches of music as an inquired object there is a metaphysical level of the object as well and, consequently, other *approach*.

Ângelo N. N. Cardosos evaluate the movable definitions of ethnomusicology as a scientific field responding to adoptions of new methodological procedures, interdisciplinary contacts and theoretical convergences. However, Cardoso's view of the development of ethnomusicology and its relation with traditional methodological procedure – as the graphical representation of sounds with traditional notation understood in ethnomusicological sphere as symbols of a cultural domination – outlines a distance with a more rigorous approach.

Although Cardoso's perspective be related to a methodological concern, his perspective towards the re-evaluation of ethnomusicolgical praxis is a major statement in his discourse: reconsidering the European traditional notation of sound as a rich tool and ethnographic procedure; enlarging the scope of objects, not only non-occidental cultures; the critical analysis of the autochthonous narrative; and a more careful dealing with interpretation – in Geertzian terms – substantiated with statics analysis. Therefore, the re-evaluation of methodological precepts in Cardoso's discourse can be underscores as an *approach* concern circumscribed in a procedure debate in ethnomusicology. Changing the approach to enhance the scientific rigor [16]

Regardless Machado Neto, Barbeitas and Cardoso's diverging perspectives towards musicology it can be define a field circumscribing their concerns: the necessary changes in the object's scientific treatment and perception. This re-evaluation can be achieved with the comprehension of music as an organized discourses and its symbolic agency; within its metaphysical instance of musical experience; or, still, with a rigorous handling capable of enhance its scientific treatment.

2.3 Designations

A last *field of musicological concern* can be defined in M. Páscoa and P. S. Blanco's instrumental perspective of musicology: a context in which musicology can act as a scientific apparatus towards social changes within a pedagogical or civic role. Therefore, musicology is perceived circumscribed in a *designation* status.

Páscoa states that the conservatorial model at present academia cannot attend to contemporary issues, therefore, a tension between reality and institutionalization can be outline. Although with a similar perspective to Barbeitas, M. Páscoa does not locate his discussions at the object-music handling in a more sensible approach in contemporaneity. The author advocates musicology as an agent of social change. These social transformations can be achieved when musicology is perceived as based on cultural diversities. Therefore, the reformulation or re-evaluation of musicology has to be placed in its own scientific constitution as well in its praxis.

Nevertheless, Páscoa suggest that the rethinking of musicology has to be made in consonance with another way of understanding and seeing [17] cultural diversity: first, a disruption with a so called *enunciator*-center and, consequently, the reformulation of a scientific field within a heterotopical dimension – a perception of the canonical occidental discourse as an elimination of diversity for the *unit* conformity; and, secondly, the necessity of a musicologically oriented Musical Education capable of dealing with new sensibilities, perception and creative psychological-pedagogical processes [18]. Therefore, for M. Páscoa musicology can be an agent of social transformation insofar as discourses are reformulated in consonance with cultural diversity.

P. S. Blanco demonstrates, similarly, an instrumental perspective towards musicology and its social impacts. Comparing the legal and constitutional Brazilian regulation of heritage rights with the concept of musical patrimony, Blanco indentifies in Brazilian musicological contemporary and historical praxis a friction between academia research and patrimony's rightful heirs.

Blanco emphasizes that this friction can be surpassed with a re-evaluation of former publications related to the *history of music* in Brazil: outdated pedagogical, formative and informative materials and the necessity of correction [19] in its methodological and theoretical bias; the necessity to complete musical inventories capable of providing to its heirs the rescue of their historical-musical memory; a decentralized musicological activity fulfilled by a new generation of musicologist graduated and post-graduates at Brazilian universities; as well another professional ethics in musicology.

The key-concept of *cultural citizenships* underscores Blanco's perspective: as the heirs of this musical memory are neglected in Brazilian musical historiography their cultural heritage and rightful patrimony is muted. Therefore, a short, medium and long-time strategy is crucial to correct this social reality; in other words, an instrumental perspective of musicological praxis in Brazil with regards of a civic duty.

Albeit M. Páscoa and P. S. Blanco's perspective diverge at the nature of this new action for musicology, their understanding of the social impacts related to a new kind of musicological practice undertake both authors in a *designation* category. Therefore, this last *field of musicological concern* identified in Brazilian musicolog*ies* has the principle of a scientific area – even though a reformulated one – capable of social changes; in a pedagogical formula or as a civic role.

3. CONCLUSIONS

The first statement of this paper is an affirmative one. Adler's academia institutionalization of musicology corresponds to an understandable epistemological regime. Moreover, in its historical-systematic definition Adler's methodological and theoretical discourse can be perceived as consonant with the mind and desires of an Austrian scholar at the end of the 19^{th} century. However, Brazilian musicology cannot be placed in a parallel relation with German-English programmatic speeches insofar as Brazilian musicological development diverges in its own writing and thinking processes[4].

However, the question for a postmodern musicology, its methodological and theoretical precept and praxis, is not an exclusive issue of Brazilian academia reality or contemporary thinking – and, also, not a new subject for international community. As defined by David Beard and Kenneth Gloag [20], musicology is also a process and therefore it dependents on how music is perceived within distinct contexts and concepts throughout time and spaces. When approached at contemporary perception of a pluralism of the present, musicology discourse itself is tossed at multilevel epistemological dimension: gender and sexual studies, oraganology and iconography, archival, historical, performing practice etc. The driving force that tensions this scientific area – and, consequently, underscores its praxis and conventions – is the incredulity with totalizing narratives.

Moreover, if it is possible to enunciate that the development of Brazilian musicology presented itself in a disagreement with the international community at the past, it can be denied that this friction persists at present. Furthermore, the lack of methodological, theoretical and material conditions or a common agenda capable of answering a uncertain reality; the necessity of changing its own parameters and scientific treatment towards a better understanding of its object; or the strength underneath its non realized social potential, outlines Brazilian musicology in a fragmented epistemological movement that can be perceived, paradoxically, in consonance with the international discussion since[5] J. Kerman's impact [21].

Additionally, it is also irrelevant measuring its resemblances or distances insofar as cultural relativism sustains these differences at non linear, gradual or formative pertinences as well with an insignificance of scientific value. Moreover, the unit among scholars cannot be reduced to geopolitical frontiers or civic status – native, foreign, immigrant etc – considering that the perception of plural identities is shaped in a globalized world.

Nevertheless, this second statement – the consonance within divergence – does not bespeak an inverted musicological Babel Tower in which the common language is an indecipherable linguistic composite. The comprehension of a consonant dimension based on plurality can be, on the contrary, a bridge between local and international musicological discussions; a broad cognitive world underscored by the necessity of musicological debate and delineated by the perception of science in postmodernity.

Therefore, Brazilian musicolog*ies* and its three major fields of concern –named as *resources, approaches and designations* – are perceived as symptoms of a contemporary condition: the perception of scientific praxis and precepts as no longer circumscribed within totalizing narratives.

Acknowledgments

I am grateful to Professor Dr. Edite Rocha for the enlightening suggestions and tutoring lectures at the Post-graduate Program in Music of the Federal University of Minas Gerais, to Dr. Loque Arcanjo for advising on the

[4] Cf. Carla Blomberg, "Histórias da Música no Brasil e musicologia: uma leitura preliminar", in *Projeto História: Revista do Programa de Estudos Pós-Graduados de História*, 2011, pp.415-444. Also see: Diósnio Machado Neto, "Curt Lange e Régis Duprat: os modelos críticos sobre a música no período colonial", in *Revista Brasileira de Música*, 2010, pp.73-94; "Em vão vigiam as sentinelas: cânones e rupturas na historiografia musical brasileira sobre o período colonial", in *Tese de Livre Docência Universidade de São Paulo*, 2011; "Discursos críticos da musicologia contemporânea: modelos analíticos sobre a música colonial", in M.A. Volpe (Ed) *Teoria crítica e música na atualidade*, 2012, pp.145-156; "O 'mulatismo cultural': processos de canonização na historiografia musical brasileira", in M. Santos. E. Lessa (Ed.), *Musica Discurso Poder*, 2012, pp.287-308; "Em torno da musicologia brasileira: ensaio sobre a relação com a análise musical", in E. Rocha. J. Zille (Ed.), *Musicologia[s]*, 2016, pp.99-112

[5] Joseph Kerman's *Contemplating Music* (1985) was translated into Portuguese and published by Martins Fontes in 1987, therefore, a short lapse of time between the English and Portuguese versions.

topic of this communication, to my colleagues at the Postgraduate program, and to *Coordernação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES* for the scholarship grant.

4. REFERENCES

[1] P. Bohlman. “Music and Culture: historiographies of disjuncture”, in *The cultural study of music: a critical introduction*. Routledge, 2003, pp.44-55.

[2] Op. Cit., p.54.

[3] E. Rocha, J. A. B. Zile, *Musicologia[s]*. EdUEMG, 2016.

[4] E. Rocha, “A musicologia e os seus afetos: carta introdutória”, in *Musicologi[as]*, 2016, p.22.

[5] R. Parncutt, “Systematic Musicology and the History and Future of Western Musical Scholarship”, in *Journal of the Interdisciplinary Music Studies*, 2007, pp.1-32.

[6] G. Stanley *et al.*, “Musicology, §II: Disicplines of musicology”, in *Grove Music Online*, Oxford University Press, accessed March 23, 2017.

[7] M. A. Volpe, “Musicologia em contextos institucionais interdisciplinares: desafios, integrações, um testemunho brasileiro”, in *Musicologia[s]*, 2016, pp.159-172.

[8] P. Castagna, “Desenvolver a arquivologia musical para aumentar a eficiência da Musicologia”, in *Musicologia[s]*, 2016, pp.191-243

[9] Op. Cit, p.231.

[10] B. M. Castro, “Musicologia brasileira: corpo, problemas, perspectivas”, in *Musicologia[s]*, 2016, pp.113-146.

[11] D. Machado Neto, “Em torno da Musicologia brasileira: ensaio sobre a relação com a análise musical”, in *Musicologia[s]*, 2016, p.103

[12] Op. cit, p.111.

[13] F. Barbeitas, “A universidade e a pergunta pela Música”, in *Musicologia[s]*, 2016, pp.41-51.

[14] Op. cit, p.48

[15] Op. cit, p.50

[16] Â. Cardoso, “Apontamentos críticos sobre tendências atuais na Etnomusicologia Brasileira”, in *Musicologia[s]*, 2016, p.70.

[17] M. Páscoa, “Música e Musicologia: nota crítica em direção a uma heterotopia”, in *Musicologia[s]*, 2016, p.154.

[18] Op. cit, p.156.

[19] P.S. Blanco, “O desafio da herança na história da música no Brasil”, in *Musicologia[s]*, 2016, p.184.

[20] D. Beard, K. Gloag, *Musicology: the Key Concepts*, Routledge Press, 2005..

[21] J. Kerman, *Musicologia*, Martins Fontes, 1987.

A MUSICAL ANALYSIS OF MUTANTES' BALADA DO LOUCO

Juliana Altoé de Oliveira
Universidade de São Paulo
j.altoe@uol.com.br

ABSTRACT

The aim of this study is to propose a formal, harmonic and voice-leading analysis of the song *Balada do Louco*, written by the group *Os Mutantes*. The formal analysis was made according to concepts presented by William Caplin of inter and intrathematic functions and adapted for rock music by Walter Everett. Harmonic analysis was done according to the syntactic criteria, that is, it is the approach in which the chords design a certain function in line with its context within the form of the song (NOBILE, 2016, passim)[1]. Harmony studies and Schenkerian rock analysis initiated by Christopher Doll are also considered in this study.

1. INTRODUCTION

Musical analysis of the Rock genre is one of the fastest growing areas in US music research. In Brazil, Rock is studied mainly through sociocultural approaches, and it is rare that the genre is approached by musicology, music theory, analysis and harmony. In addition, popular harmony taught in conservatories and music colleges are focused on Jazz and Brazilian music that resemble harmonically and formally Jazz music, but it is not able to satisfactorily explain Rock. (TEMPERLEY, 2011)[1] [2] Thus, the aim of this study is to propose a formal, harmonic and voice-leading analysis of *Balada do Louco* by the group *Os Mutantes* following the concepts of William Caplin, Walter Everett, Drew Nobile and Christopher Doll. The band *Os Mutantes* was chosen as the representative band of rock made in Brazil due to its historical path: in the 1950s, rock in Brazil was based mostly on versions of English and American music and remains with this predominant characteristic until the "Jovem Guarda". The group was chosen considering their presence at the beginning of rock written by Brazilians (original compositions). In addition, the importance of its members and their contribution to tropicalism and rock in Brazil are factors that strongly contributed to the choice of the group for this study. To clarify the terms used in this article, "Rock" will be used to refer to the genre developed in England in the 50's and 60's derived from Blues, while "rock in Brazil" is refering to the genre that became popular in the country during the 70's and 80's. Roman numerals represent chord degrees.

[1] The survey looked at the "500 best songs of all time" according to Rolling Stones magazine and it was notable that the IV-I movement is more common than V-I in rock. How is it possible to traditionally analyze a genre that doesn't follow the most used cadence in tonalism (V-I) as a general rule?

2. ANALYSIS

Even today there is no agreement between an appropriate methodology for analyzing songs from the rock repertoire. Therefore, I opted for analyzing the form of the song according to concepts of inter and intra-thematic functions according to William Caplin and adapted to Rock by Walter Everett.

The harmonic analysis will be done by degrees to elucidate along the song the voice leading movements. It will also be done according to syntactic criteria. In syntax harmony, Christopher Doll considers chords as function projectors, but the chord will be classified as tonic or pre-tonic according to the concepts of preparation and stability. Chords that are stable enough to not require a resolution project the tonic function, whereas pre-tonic chords are those that through voice-leading goes to stability. (NOBILE, 2014) [3]

The Schenkerian analysis was chosen because it reveals important harmonic and melodic relations in this song. Despite being useful in this song particularly, Rock is a genre that encompasses many composers and sub-genres, and therefore, not all songs analyzed can be "embedded" in this type of analysis. As we know, the number of composers Schenker has established for analysis is very small and it is not possible to establish a Urlinie (fundamental line) in Rock as Schenker proposed in his theory. (NOBILE, 2014)

Like Nobile (2014, passim), I also believe that these relations that the Schenker analysis shows are as fundamental to Rock as to the repertoire analyzed by Schenker and despite the lack of an unifying Urlinie of the genre, the relations evidenced are sufficient for a deeper analysis of harmony and melody than analysis by degrees alone. Finally, I believe that Schenker's analysis can bring an important harmonic overview of a piece and clearly shows the parts of tonic prolongation and emphasis on the dominant in the different sections of the music analyzed.

Finally, I will make some considerations relating form, harmony and instrumentation.

2.1 Formal Analysis

The 32-bar AABA ternary form in popular music originates from Tin Pan Alley, a form used in the 30s and 40s, showing variations in the 50s and 60s with the beginning of rock. Thus, there is also the AABA

(BA) form of 48 bars, as is the case of *Balada do Louco*, in which the sections, instead of 8 bars, have 12 (resulting in 3/2 of original Tin Pan Alley). An example of this form in Rock repertoire is the song *I Want to Hold Your Hand* by *The Beatles.* It is also present in several songs from the beginning of the band's career and, due to their great influence, defined the pre-psychedelic era (NOBILE, 2014). This song has great similarities to *Balada do Louco*: both begin with an introduction, and the A sections of both songs are made of 12 bars.

In *Balada do Louco*'s A section, it is possible to hear in the melody the formal type sentence being used in what Everett calls the Statement-Restatement Departure-Cadence (SRDC) model. "SRDC is the basis of a verse with refrain." (EVERETT, 2009, p. 140)[4] In fact, what we hear is a basic idea or Statement (c. 3-6), the repetition of this idea or Restatement (c 7-10), the phrase that distances itself from the idea or Departure (c.10-11) and Cadence (c.11-12), which is repeated in the form of the refrain. Nobile also states that it is common in SRDCs to "reinterpret the cadential", just as it occurs at the end of the A section of *Balada do Louco*. See Fig. 1. (NOBILE, 2014, p.137)

Figure 1: Section A's formal analysis.

This section is divided into main theme (c.3-10) and expanded cadential (c.11-12). Thus, the first ten bars can be compared to a main theme. Caplin states that only the main theme is tight-knit in a song, because it follows the ensuing characteristics: tonal-harmonic stability, cadential confirmation, melodic-motivic unit, sentence symmetry and efficiency on functional expression. (CAPLIN, 1998, p.17) [5]

In fact, this theme of the A section may resemble a main theme, because in both the phrases of presentation and continuation, as well as in the cadential, the emphasis is reinforced, constructing a harrmonic tonal stability. Harmony is confirmed at the end of the theme, which would only result in a more stabilized harmony if a perfect cadence was used instead of the plagal one. The melodic-motivic unit is observed when we note that all motives derive from the basic idea and we see the formal efficiency of this theme when we try to remove some of its parts and we realize that the meaning of the section is impaired.

The post cadential of the section may also resemble a codetta, although with a difference. A codetta may contain melodic-motivic material from the cadential and harmonically, the codetta prolongs the root. These features are compatible to what occurs in the post cadential of the section, however, the codetta is usually preceded by a perfect cadence, and not a plagal one as we have heard in the song. (CAPLIN, 1998, p.16)

Section B is marked by its division between the theme and the interlude. This instrumental part that separates the A and B sections has a framing function, that is, separate the A section (quieter and tight-knit) from the B section (looser and chaotic). Just like the expanded cadential of the A section, the interlude also has a post-cadential function.

Harmonically, the B section focuses on dominant harmony, since its theme, rather than focusing on the root, ends in V, to start the interlude. This emphasis on the dominant is characteristic of contrasting middles from small ternaries. (CAPLIN, p.13, 1998)

This typical sound difference of the contrasting middle is usually associated with changes in textures, melodies and instrumentation (CAPLIN, 1998). An example of this is in the song *Every Breath You Take* by *The Police*.

In the A section of *The Police'*s song, we can hear guitar, bass, drums and voice. On entering the B section, we hear the addition of a keyboard and the guitar playing open chords instead of the fingering style. Despite the meaningful contrast, the biggest difference we feel between one section and another in this song is the loose form how section B poses in relation to section A, since B is characterized by groups of asymmetric phrases, some harmonic instability because the theme ends in V rather than completing the cadence and a certain inefficiency of functional expression, that is, the parts are not highly interrelated, different from the A section. (CAPLIN, 1998, p.17)

2.2 Harmonic analysis

Harmonic analysis will be divided in parts for better understanding: introduction, A section (divided into theme and expanded cadence) and B section (theme starting in anacruse and interlude). In each section, each chord and its functions will be analyzed, as well as the cadences. Harmonic analysis will have two strands: separate analysis of each chord by degrees and analysis of sections based on tonic prolongation (section A), and emphasis on the dominant (section B).

This last type of analysis is a syntactical one. In syntax harmony, Christopher Doll considers chords as function projectors, but the chord will be classified as tonic or pre-tonic according to the concepts of preparation and stability. Chords that are stable enough to not require a resolution project the tonic function, whereas pre-tonic chords are those that through leading voices conduct to stability. (NOBILE, 2016, p. 32) Doll also asserts that predictive functions can have different levels of proximity to tonic, especially considering temporal position and stepwise resolutions that each degree escale anticipate (DOLL, 2007, p.17) [6]. For example, in a song in the key of C major, a movement from a D minor chord or a triad B diminished do predict a "stronger" pre-tonic than a F major. But the terms dominant and subdominant might not be appropriate because eventually a IV might represent a pre-tonic function stronger than a V considering its position and importance within the song, as we will see in this song.

2.2.1 Introduction and A section

Figure 2: Harmonic analysis of the A section.

According to Caplin (1998), a thematic introduction contains the following characteristics: short (between two and four bars), weak or nonexistent melodic-motivic component, and usually emphasizes the harmony of the tonic. In this case, the introduction of the beginning of the song resembles a thematic introduction as described by Caplin. The two bars played on the piano begin with a diminished chord (VII) followed by the I, without building any motif or melody. Thus it is possible to note its role in the song: to start the music without starting genuine ideas, so that the basic idea starts in the A section. This introduction, however, cannot be considered a thematic introduction, because it does not repeat itself when the A section is repeated throughout the song.

Section A begins with the A major chord with ninth (IV), sounding like a cluster, because in the piano it's played the notes A, B and C# in the same octave. This chord is then followed by VII in the same measure. The IV chord acts as a pre-dominant that evidences the voice leading of the B note as a contrapontistic suspension that resolves on the diminished chord with dominant function. Already in the third and seventh beat of the theme, the diminished chord starts to have the pre-dominant (or pre-pre-tonic) function to the V that follows it. Considering the syntactic criteria, both the diminished chords are in a pre-tonic position, as well as the V, since both go to stability.

The first section is formed by one theme only, similar to the classic sentence defined by Caplin (1998, p.10). Its presentation phrase extends from c. 3 to the beginning of c. 10, having a basic idea of four bars that repeats itself. The first motif of the theme (c 3-4) begins with the note F # and ends at the second measure of the section on a V chord. The second part of the theme retains the rhythmic pattern of the first, but instead of leading to the fifth of the tonic (B), it leads to the third (G#), creating a sense of question and answer between the two sentences.

From the fifth measure of the A section, the basic idea is repeated, but the last note does not extend to the end of the measure, because instead of using I, the progression was extended through III, I6/4, # IVo until reaching IV and finishing on the plagal cadence. The cadence is then repeated in the last two measures of the section, as well as the lyrics. These last two bars are considered as an expanded cadential as defined by Caplin (1998, p. 19-20). The cadence is repeated because the first time the melody ends, it ends on the third of the chord, and when it is repeated, the melody reaches the tonic. In our analysis, we consider c. 13-14 as an expanded cadential. The lyrics "and (He/She) is not happy, is not happy" of the last four measures of the section is always repeated, creating what is defined by Walter Everett (2004) [7] of refrain or hook. The only time we have a change of these lyrics is on the last repetition of the A section, at the end of the song: "and (He/She) is not happy, I am happy."

One aspect that reiterates this song as part of the rock repertoire is its harmony, especially the plagal cadence (IV-I) at the end of the section. The repetition of the cadence serves to reinforce its intention of ending the section, since the melody in the first time ended on the third of the tonic and thus creates an impression of a question that demands a response made by the expanded cadential. Although it seems a weak movement, it is curious to notice that the IV-I movement is much more common than the V-I in rock. A research analyzed the "500 best songs of all the times "according to Rolling Stones magazine and have arrived at the following result, see Table 1: (TEMPERLEY, 2011):

Ant \ Cons	I	bII	II	bIII	III	IV	#IV/bV	V	bVI	VI	bVII	VII
I	0	25	132	94	44	1052	2	710	104	302	470	16
bII	31	0	0	0	2	0	0	0	0	0	0	12
II	120	1	0	2	20	58	0	97	0	24	10	0
bIII	50	6	6	0	0	64	2	2	67	0	41	0
III	16	0	39	0	0	46	0	6	0	60	3	4
IV	1162	14	30	98	45	0	4	514	57	72	90	4
#IV/bV	7	0	0	6	0	10	0	0	0	0	0	0
V	788	0	36	6	17	392	4	0	6	191	48	0
bVI	208	0	1	20	0	22	6	22	0	10	78	0
VI	144	0	87	0	32	260	0	124	21	0	3	0
bVII	386	0	0	11	2	188	2	26	114	6	0	0
VII	18	0	0	0	12	0	4	0	0	3	0	0

Table 1: Chord transitions occurences in the "500 best songs of all times" (TEMPERLEY, 2011)

Alike *Balada do Louco*, the *Eagles*' 1975 song *Lyin 'Eyes* presents a plagal cadence in the continuation phrase of the first theme (NOBILE, 2016). Thus, we can understand the aspects of syntactical analysis: the IV in these two cases act as dominants even though, theoretically, they normally assume the function of subdominant.

The function of the diminished chords is also a prominent subject in this song, since they are always followed by I or III, suggesting a dominant or pre-dominant function when they appear before V. Another point that can be raised is the few occurences of VII chords (usually diminished ones) found in rock songs in the same research. These chords represent less than 1% of the chords used in the songs of the analyzed list. They are also the least recurrent (along with III and #IV) to be followed by the tonic, as occurs in *Balada do Louco* countless times.

Root	Number of occurrences: Proportion of total	Proportion of total, excluding tonic	Number of occurrences in pre-tonic position (proportion of total, excluding tonic)
I	0.328	—	—
bII	0.005	0.007	0.010
II	0.036	0.053	0.041
bIII	0.026	0.038	0.017
III	0.019	0.028	0.005
IV	0.226	0.336	0.396
#IV/bV	0.003	0.004	0.002
V	0.163	0.241	0.269
bVI	0.040	0.059	0.071
VI	0.072	0.107	0.050
bVII	0.081	0.119	0.132
VII	0.004	0.006	0.005

Table 2: Cromatic distribution of tonics in the "500 best songs of all time" (TEMPERLEY, 2011)

In summary, the analyzed music contains in its first section a harmony based on tonic prolongation, similar to that of a presentation phrase, as defined by Caplin (1998). That is, a harmony that may contain passing chords or neighbouring ones, but that its base and final direction is the tonic. (CAPLIN, 1998, p.10)

2.2.2 B section and Interlude

The B section is also composed of 12 bars. However, it is divided as follows: the theme and the interlude. According to Everett (2009, pp. 147-8), B sections, in relation to As, are less varied and has the function of creating a contrast (be it harmonic, melodic, textural, lyrical, etc.). One of the most common ways to achieve this effect is through harmony with an emphasis on the dominant.

Figure 2: Harmonic analysis of the B section.

In the first measure occurs the transformation of I into the secondary dominant (V7 / IV) to start the theme. Then it goes to IV and alternates between I and IV. Finally, we get to the F# major chord (V / V) followed by B major (V). Instead of ending on I, the theme ends on V, suggesting an emphasis on the dominant, main characteristic of contrasting middles of small ternaries, according to Caplin (1998, p.13). This emphasis on the dominant is also a feature of rock, and this particular type of harmonic progression (IV-I-IV-I-V/V-V) is a mixture of the "classical" types of bridges defined by Trevor de Clerq (2012, p.77) [8].

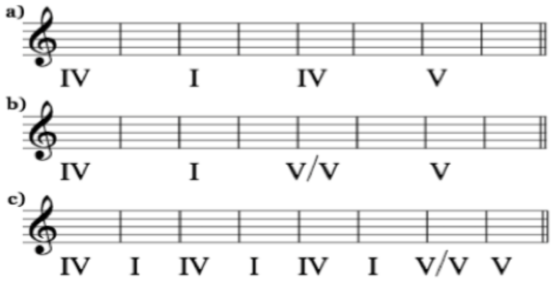


Figure 3: "Classical" bridge models by Trevor de Clerq (2012, p. 77)

In comparison to the theme of the A section (tight-knit), the theme of section B is relatively looser. The five-bar interlude keeps the bass pedal on E and toggles between the IV and V / IV chords until it returns to section A again. The function of this interlude is to separate the two sections and to create different atmospheres, reaching an apex of euphoria and madness, as the name of the music itself suggests, to return to the verse.

2.3 Schenkerian Analysis

From the piano transcription and voice melody, a Schenker analysis was adapted to rock following the model proposed by Nobile. At each level of analysis, it is possible to see important details of voice leading movements that reaffirm the issues of tonic prolongation and emphasis on the dominant. It is important to remember that in popular music the rules for voice-leading are not always followed. As in the harmonic analysis, the sections were also divided for better visualization. The images contain the analysis and some comments will be made on each level of observation.

2.3.1 Introduction and A section

Figure 4: A level analysis of the introduction and A section.

A Level: At this level, it is possible to notice the passing chords being used to prolongate tonic, as they return to I, as we see in the bass leading voice keeping the tonic (E) and contralto the fifth (B). The first chord of the A section is a melodic emphasis chord, just holding the B note from the previous chord and solving it in the chord of VII that follows it. In relation to the melodic contour, it is possible to perceive that the ideas always return to the triad of E major. The pre-tonic chord V is found only at the end of the basic idea leading to I. Ascending fourth leaps in the bass are prominent in this section, as well as the chromatic descending bass at the end of the section before the expanded cadential.

Levels B and C: At the intermediate level it is possible to observe the tonic prolongation in most of the section and at the end to go to IV there is a descending leading voice movement. At the structural level, in the end of the A section, the piano makes a passage from the tonic (E) to the fifth (B) with prominence in the song and that is almost repeated when going to section B, but, instead of the fifth, it is directed to the seventh. At this level it is also possible to clearly notice the I-IV-I cadence as the section structure. See fig. 5 and 6:

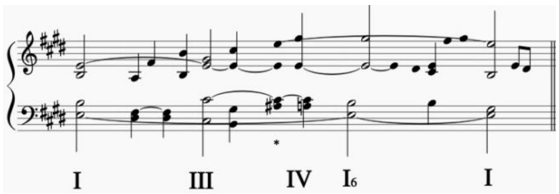


Figure 5: B level of the introduction and A section.

Figure 6: C level of the introduction and A section.

2.3.2 B section and Interlude

A Level: In section B, it is clear the melody is at a higher register. The use of secondary dominants draws much attention in this section, suggesting an emphasis on the dominant, however, returning to the prolongation of tonic in the interlude. Also in the interlude it is possible to observe an interesting voice leading between the V/IV and the inversion of the IV. See fig. 7:

Figure 7: A level analysis of the B section and the interlude.

B and C Levels: Voice leading movements by joint degrees on the contralto and the bass leaping in fourths. The emphasis on the dominant in this section is observed both by the melody always returning to the note B, and in the presence of the secondary dominants. The interlude, due to the presence of the low pedal, makes us think of a prolongation of tonic, despite the use of the chord of first degree with seventh minor. This quality of chord may suggest a blues influence in this passage, a genre in which it is common to have a major chord with a minor seventh as tonic. See fig. 8 and 9:

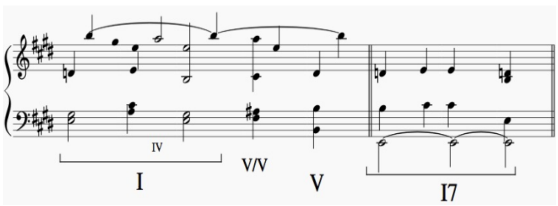


Figure 8: B level of the B section and the interlude.

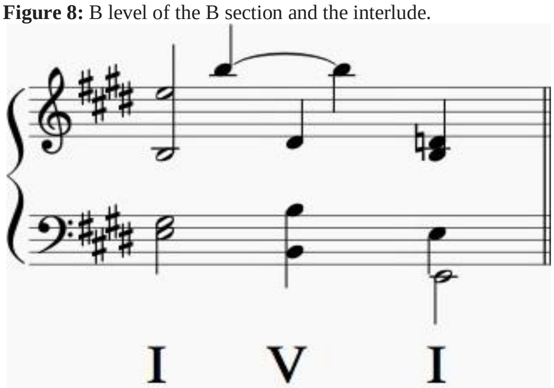


Figure 9: C level of the B section and interlude, as well as the song's structural chart.

Considering the form of the song (AABABA), we can observe that the structure of this song can be drawn in a graph I-V-I as proposed by Schenker, since section A and interlude are tonic prolongations and between the first A section and the last one, we have section B of emphasis on the dominant, very similar to the structural chart of section B and interlude (Fig. 9).

2.4 Considerations on form and instrumentation

The instrumentation on this song is highly related to its form, since the first separates the sections as instruments are added. Thus, we have in the introduction only the piano. The first section contains piano and voice. In their repetition, the drums and the bass begin their participation. As we have seen previously, the B section can be considered a contrasting middle according to Caplin (1998), since it is common that this section contains changes in textures, instrumentation and accompaniment patterns. In the case of this song, the chorus and the synthesizer join the layers already built throughout the music, causing great impact, since they are "above" pre-existing layers. The interlude mixes sounds of the synthesizer with piano, drums, bass and some parts of not spoken voice (screams, onomatopoeias, etc.). Upon returning to the verse, the chorus and synthesizer give way to the piano, drum, bass and voice lull, just like the second verse. The verse is not repeated as in the first part of the song and we return to section B with the same characteristics as before: shocking, sudden changes of layers and interlude with the function of separating these two atmospheres.

Finally, the song ends with the verse in a different way: the sitar is added. The use of sitar suggests an interesting topic present in psychedelic rock and possibly lysergia. This instrument is used in emblematic *The Beatles*' songs like *Norwegian Wood, Love You To and Within You, Without You*. The band *Os Mutantes* is known for its influence of *Beatles*, its irreverence, originality and, since the album *Mutantes E Seus Cometas no País do Baurets* began the period of influences of progressive rock in their compositions. It is curious to notice that the lyrics, in two of the four repetitions of section A, the verses begin with the structure: If they are X, I am Y. However, in the last A section, the singer starts the verse admitting "Yes, I am very crazy". At the end of each A there is also the refrain that is repeated "and (He/She) is not happy, (He/She) is not happy". In the last repetition of A, however, the lull of the piano and voice is accompanied by the sitar and the phrase becomes "and (He/She) is not happy, I am happy", creating a sense of completion and transcendence of the singer, accepting his madness and reaching a state of happiness.

3. CONCLUSIONS

Based on the study done by Drew Nobile and using his method (analysis of harmony through the syntactic criterion, application of notions on form by William Caplin and Schenkerian analysis of songs), it was possible to establish similarities between *Balada do Louco* and songs from the English, American and Canadian rock repertoire studied by Nobile.

Concerning harmony, it was possible to identify the plagal cadence in section A, the most common cadence seen in rock according to Temperley's study (2011). The issue of tonic prolongation observed in section A and the dominant emphasis in section B are also recurrent in the genre. In addition to these similarities, the use of cadential chords common to classic tonal music such as I6/4 and the expanded cadential are notable.

The form of this song is similar to several rock songs of the 60s, with its origin in the Tin Pan Alley. In addition, the melody analysis showed the use of the sentence form defined by Caplin and renamed for rock as SRDC by Walter Everett. This form is widely used in the genre and is usually associated with a refrain (hook) that repeats at the end of each section A, just as it occurs in this song. Section B, with typical aspects of contrasting middles also suggests the approach of rock songs to music of the classical period analyzed by Caplin.

The instrumentation is another important subject to be mentioned. Although it has several aspects similar to a rock song and even sound as a song of the genre, it is not possible to hear guitars in it. This fact suggests that rock is considered an autonomous genre, but not only by its standard instrumentation composed of guitars, bass and drums, there are other elements that unify the genre such as its form and its harmony and that are not mentioned as part of its identity and uniformity.

Finally, it is possible to notice that the layers used by *Os Mutantes* are typical of a rock band when passing through the sections, however, some harmonic and formal aspects of the music that could not be explained by rock studies suggest a certain connection to genres of Brazilian music. The use of diminished chords are extremely rare in rock, but very common in Brazilian music. In addition, the AABA form chosen by the band to structure the song was used a lot in the beginning of rock, but in the 70s the genre already used other forms, like the ones that today we know by verse-chorus and verse-pre-chorus-chorus. Thus, it can be said that *Os Mutantes*, despite its explicit influence of British rock, opts for some uses of chords and typical forms of bossa nova, style that preceded the group in Brazilian music history.

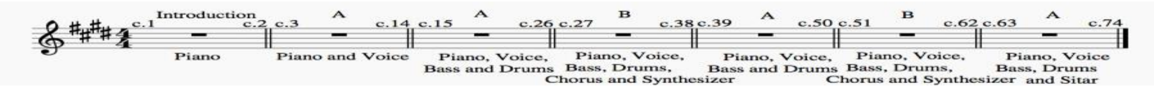


Figure 10: Instrumentation x Form in the song *Balada do Louco.*

Acknowledgments

Thanks to Ciro Visconti for helpful feedback on an earlier draft of this paper.

4. REFERENCES

[1] D. Nobile. “Harmonic Function in Rock Music: A Syntactical Approach.” in Journal of Music Theory 60:2. Nova York, October 2016.

[2] D. Temperley. “The IV Cadential in Rock.” Music Theory Online, Volume 17, Number 1, April 2011.

[3] D. Nobile. “A Structural Approach To The Analysis of Rock Music.” in CUNY Academic Works. New York, 2014.

[4] W. Everett. The Foundations of Rock: From Blue Suede Shoes to Suite: Judy Blue Eyes. Nova York: Oxford University Press, 2009.

[5] W. Caplin. Classical form: a Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven. Nova York: Oxford University Press, 1998.

[6] C. Doll. Listening to Rock Harmony. Ph.D. dissertation, Columbia University, 2007.

[7] W. Everett. “Making Sense of Rock’s Tonal System.” Music Theory Online, Volume 10, Number 4, 2004.

[8] T. de Clercq. Sections and Successions in Successful Songs: A Prototype Approach to Form in Rock Music. Ph.D. dissertation, Eastman, 2012.

COMPARATIVE ANALYSIS AS A TOOL TO IDENTIFY ASPECTS OF ELECTROACOUSTIC MUSIC COMPOSERS LANGUAGE

Dilson Colman Cassaro
Federal University of Parana
dilsoncolman@gmail.com

ABSTRACT

The present article provides a description of how the comparative analysis of Pierre Couprie can be an useful tool in order to help identify compositional strategies by electroacoustic music composers. In the essay entitled *Analyse comparée des Trois rêves d'oiseau de François Bayle*, Pierre Couprie [1] analyses in terms of material, morphology, shapes and graphic representation three pieces by François Bayle constituting the work *Trois rêves d'oiseau*. Besides providing a listening score showing structural aspects of the work, Couprie creates charts and tables reporting correlations among the three pieces by means of, for example: materials articulation table, sound units classification table, morphologies chart, etc. This article has its basis in the following tables and charts elaborated by Pierre Couprie: "The articulation of materials", "Classification of different sound units" and "The distribution of durations". The three pieces featured in this article are *L'oiseau Moquer, L'oiseau triste* and *L'oiseau zen*, composed between 1963 and 1971. Such pieces are discussed through complementary charts created in order to point out possible compositional strategies adopted by François Bayle.

1. INTRODUCTION

The wide range of compositional possibilities provided by electroacoustic music led graphic notation to be replaced by notations aiming to translate composer's intentions by means of other tools.

When analyzing electroacoustic music, we use aural transcription, in which the analyst focus on the detailed hearing of the piece, constructing an analytical thought.

The methodology of Pierre Couprie analyzes the work in three stages: from ideas to form, through morphologies. Based on the *Traité des objets musicaux* [2], Couprie creates graphs and tables in order to compare different pieces of the same composer, looking for similarities between them. In this case, the pieces present features quite different from one another. Therefore, we will use the comparative methodology to understand how these three pieces are gathered in a single work by identifying compositional processes.

The tables created by Couprie show the following aspects: articulations of the materials; typomorphological classification of sound units (addressing spectral complexity and duration); study of different morphologies; structure of the work (splitting it into sections and commenting it). The graphs created by Pierre Couprie address: division of durations, classification by types of sound, sound envelopes features, average duration and variation of the three pieces.

Couprie, in the article [3] entitled "Trois modèles d'analyse de *'L'Oiseau moqueur'*, un des *Trois rêves d'oiseau* de François Bayle", clarifies the way the three pieces were integrated: *L'oiseau moqueur, L'oiseau zen* (composed in 1971 with the title Uirapuru) *and L'oiseau triste* were brought together in 1971 to compose the *Trois réves d'oiseaux* created on 6 February at the Guimet museum in Paris (Bayle 1993).

In the same article, Couprie addresses one of the pieces through three different analysis. In the article entitled "Trois modèles d'analyse de *'L'Oiseau moqueur'*, un des *Trois rêves d'oiseau* de François Bayle", the following methods of analysis were used: identification and classification of sound objects using a Schaefferian typology, comparison of objects with one another following a paradigmatic approach, and presentation of the results obtained through enumeration of a certain number of values (duration, density of objects, etc.).

Through Schaefferian Analysis, Couprie made a record of the most pertinent sound objects, dividing them into four categories: a) manufactured sounds (for example, thick held note, scratch, glissandi, etc.); b) instruments; c) bird; d) man's laugh. Pierre Couprie identified thirty-two sound objects, of which twenty-four were retained for the final classification. Through the use of Acousmographe, it was inferred the use of montage as a compositional feature of François Bayle's *L'oiseau moqueur.*

The paradigmatic point of view presents the variations of an object or similarities between different objects. By looking at sound objects from such a point of view, Couprie separated them into three categories: (a) simple or very simple variations (pitch, duration, etc.); b) objects rarely present (between 1 and 3 times) and c) objects present a number of times with significant variations.

The last part of the article presents an approach based on the detailed and precise measurement of several values, such as importance, relationship with other objects, structural role, etc.). The graphs created in this last approach provided information about the behavior of objects in relation to each other (from highlighting to certain structural relations between objects).

Using three different approaches, Couprie provided different information. With this, we could notice the importance of different approaches in the analysis of electroacoustic music: comparing different methodologies or comparing different pieces of the same composer.

2. DISTRIBUTION OF ARTICULATIONS WITHIN THE THREE PIECES

In the graph representing the different elements of articulation present in the pieces, Couprie identifies and classifies five articulations: a) silence: short silence between two textures or different morphologies; b) punctuation: short sound between two different textures; c) morphological rupture: abrupt passage between two very contrasting morphologies (spectrum rupture or intensity rupture); d) transition: progressive passage between two quite different sounds by means of texture; e) roofing: gradual passage from one texture to another. Couprie points out the set of articulations and their arrangement throughout the three pieces in the following two graphs:

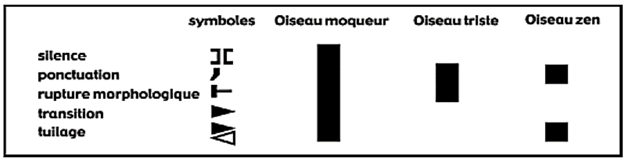


Figure 1. Articulations distribution over the three pieces.

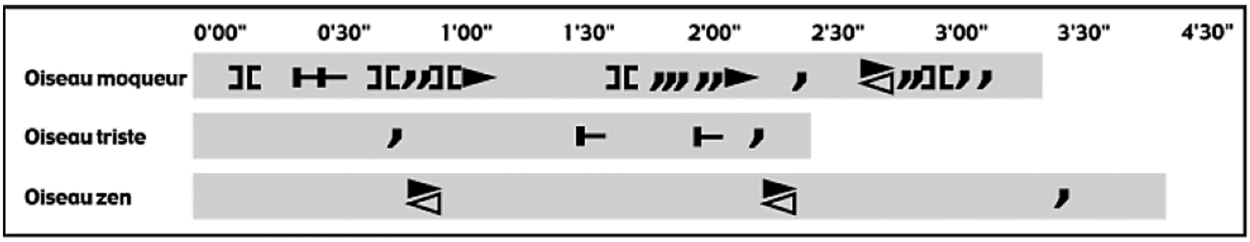


Figure 2. Timeline showing the occurrence of articulations.

Focusing on articulations, we created the chart below emphasizing the way Bayle addresses the focus on moving from one sound object to another. We then organized the articulations into two groups: articulations whose transition is made without intersection between textures and articulations presenting intersection. In Group I we included: Silence, Punctuation and Morphological Rupture. Group II includes: Transition and Roofing.

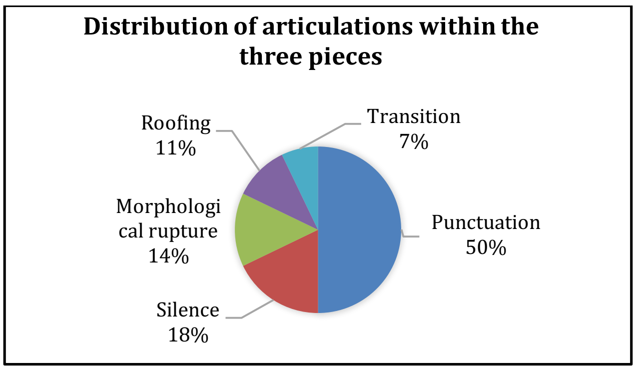


Figure 3. Chart representing the articulations used over the three pieces.

On creating the chart above, we notice that punctuation is the most explored articulation in the work and since it is the only articulation present throughout the three pieces, we can identify it as the unifying feature of the pieces. In addition, we identified the majority of Group I (50% + 18% + 14%) in relation to Group II (11% + 7%), suggesting the duality between articulation groups as a feature of Bayle's compositional process, besides the use of punctuation as an unifying element.

3. SPECTRAL COMPLEXITY

When we draw our attention to the spectral classification regarding the duration of sound units, we find the following table developed by Couprie according to the morphological classification proposed by Pierre Schaeffer in the work Traité des objets musicaux: according to its spectrum and its duration classification, the light gray color corresponds to the balanced sounds, that is, moderately long duration and moderately complex spectrum. The eccentric sounds have a long duration and/or a very complex spectrum. In the first two pieces, is predominant the balanced sounds and, in L'oiseau zen, the sound units are distributed over different degrees of complexity:

Figure 4. Table representing the typomorphological classification of sound units

Based on table above, we created a graph (X-axis: level of spectral complexity, Y-axis: number of sound units) revealing important information about the compositional process of François Bayle:

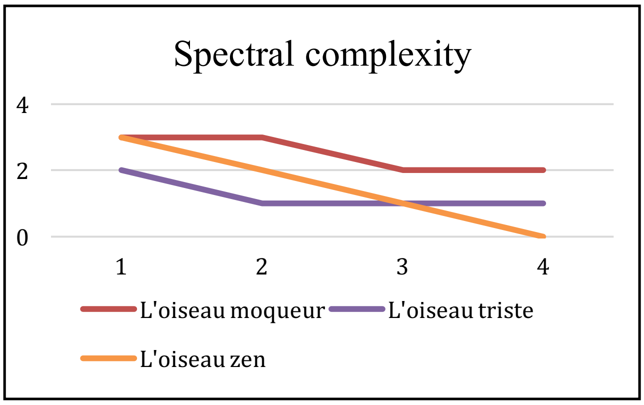


Figure 5. Graph representing the spectral evolution over the three pieces

Every piece starts with a certain complexity and all the three pieces increase their degree of complexity whereas the number of sonic units explored in the pieces decrease. Therefore, we notice again the aspect of progressive decline. This time, in relation to spectral complexity.

4. CONCLUSIONS

With the help of the representations created (Figure 3 and Figure 5) it is possible to see how François Bayle organized some compositional processes used in the three pieces, creating a link between them, uniting them in a single work: *Trois rêves d'oiseau*.

We divided the articulations present in the work into two groups: Group I (great differentiation in the passage from one sound object to another) and Group II (little differentiation in the passage from one sound object to another). From this, we notice how Bayle explores the group in a gradual way: the three articulations present in Group I are the most recurrent in the work and then, the articulations of Group II are explored. We can infer that the progressive passage from Group I to Group II is a compositional process of Bayle.

When we observe the development of spectral complexity, the graph (see Figure 5) indicates the following behavior during the three parts: the complexity of the sound units increases, however, the number of sound units decreases. That inversely proportional behavior can be considered another compositional process explored by Bayle in *Trois Rêves d'oiseau.*

The present article briefly pointed out how the comparative methodology developed by Pierre Couprie can be better explored in order to identify, by combining different pieces belonging to the same work, characteristics belonging to composers creative processes. With this, we intend to expand the understanding of procedural approaches developed by electroacoustic music composers.

5. REFERENCES

[1] P. Couprie, "Analyse comparée des *Trois rêves d'oiseau* François Bayle", in *Revue Démétér*, December, 2002. [Online]. Available: http://demeter.revue.univ-lille3.fr/lodel9/index.php?id=500 [Accessed Feb. 14, 2018]

[2] P. Schaeffer, *Traité des objets musicaux*. Paris: Le Seuil, 1979.

[3] P. Couprie, "Trois modèles d'analyse de *'L'Oiseau moqueur'*, un des *Trois rêves d'oiseau* de François Bayle", in *Les cahiers de l'O.M.F.*, 1998, pp. 50-70.

DMITRI SHOSTACKOVICH AND GYÖRGY LIGETI STRINGS MUSIC TEXTURAL SIMILARITIES

Carlos Alberto Bastos Oliva
UNESP
pollacooliva@gmail.com

ABSTRACT

This article compares two music excerpts which denote apparent sounding similarities. The first is in Dmitri Shostakovich's Symphony nº 2 (1927) and the second is in György Ligeti's *Atmosphères* (1961). A deep analysis was made in both excerpts using parameters such as rhythmic evolvement; chromatic voicing; instrumentation; harmony and texture in order to develop this technical study that might explain why sonorous results are so much alike. The analytical approach is based on the specific sounding texture present in both excerpts and suggests that music cognition through auditory process may reveal subjective musical issues that can be inferred as a result of an analytical process which takes human perception into account. The holistic sounding phenomena require an analytical concept capable of demonstrating how sonority turns into a functional dimension and may assume a structural role in music analysis (Guigue, 2004). After all it is important to realize that there is a link between the past traditions and new technics in arts once the later has its roots in the former (Schoenberg, 1984).

1. INTRODUCTION

About 34 years apart Dmitri Shostakovich's Symphony nº 2 written in 1927 and György Ligeti's *Atmosphères* finished in 1961 present a string music passage in which certain parameters such as instrumentation, pitch combination, rhythmic disposal, articulation and dynamics seem to support a similar sonority effect.

The musical aspects of the resulting sound were analyzed using some of those musical parameters with a special attention to the chromatic harmonic approach within a continuous rhythmic evolvement in a very slow tempo present in both excerpts. Even though the musical material was crafted in different ways by the two composers, the overall sounding texture represents a very similar result for about 10 to 15 seconds. That can be perceived through an attentive hearing and traced with an analysis using traditional music parameters.

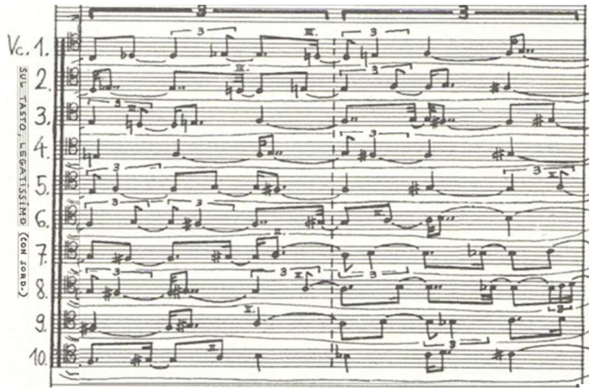
In addition to that, the matter of interest here is to throw some light to the emergent hidden components of the sound field in order to bring new elements to the music analysis, hearing perception, sound cognition and other aspects of general music study and research.

2. THE MUSIC

Dmitri Shostakovich's 2nd Symphony was written as homage for the 10th anniversary of the Bolshevik Revolution and was first performed in Moscow in November 1927. In the opening section of this one movement symphony between measures 16 and 22, a dense and compact superposition of interdependent melodic lines within the string section of the orchestra performs a type of musical texture in which each one of the voices evolves a particular variety of a rhythm figure with diatonic melodic fragments. "The symphony opens mysteriously, on the threshold of hearing, as a web of sounds on muted strings begins to grow" (R. Blokker, 1979). Different amount of notes per beat allied to a legato articulation *con sordina* in all string instruments creates a dense and compressed type of continuous sound exposure (W. Berry, 1987), see Figure 1.

Figure 1. Measure 20 from Shostakovich's 2nd Symphony introduction.

Commissioned by the Southwest German Radio György Ligeti's full orchestra *Atmosphères* was first performed in October 1961. Resulting in similar sounding qualities, an excerpt of the piece starting on letter 'H' between measures 44 and 48, presents the string section of the orchestra performing a canonic structure of chromatic lines fully *legato* and *sul tasto* where each voice appears with a different rhythm value, with the note attacks occurring in different beats of the measure and also on different parts of the beat. That creates likewise the same type of a continuous sounding carpet (G. Ligeti, 1965). "By layering his structures on top of each other Ligeti creates a sound colossus that sometimes appears static, sometimes shimmering …" (V. GroBberger; J. Voit, 2013), see Figure 2.



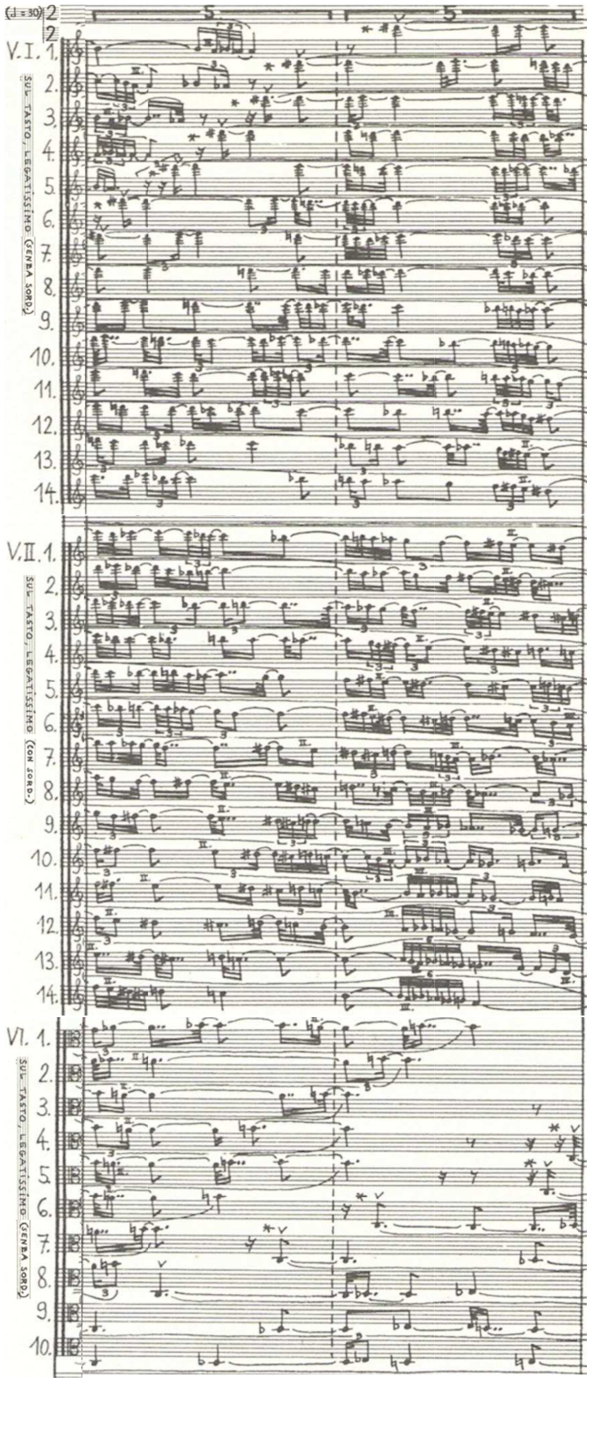


Figure 2. Ligeti's *Atmosphères* measure 47

3. METHODS

Once identified the sounding similarities between the two excerpts, measure 20 of Shostakovich's 2nd Symphony and measure 47 from Ligeti's *Atmosphères* were chosen to be analyzed according to parameters such as: the chromatic path that is formed from the evolvement of each one of the voices individually; the vertical axis clustered harmony instead of a tonal embedded one implying in a dense textural sounding fabric (W. Berry, 1987); the opposition between the rhythmic aspects of the interdependent melodic lines. While in Shostakovich's excerpt the rhythm figures respect the beat order but vary in groups of quarter, half, eight notes and different groups of semiquavers, Ligeti's excerpt has a canonic approach that otherwise has its note attacks spread out within each measure on different parts of the beats, along with various durations amidst the voices in order to create a sort of chromatic field with internal rhythmic-intervallic movement, see Figures 1 and 2.

3.1 Rhythm

Figure 1 shows an example of the rhythm values performed per beat by the instruments of the string section on a 4/4 signature. From the bottom to the top: Contrabass plays quarter notes; Cellos half notes; Violas 8th note triplets; Violin II, second group, plays semiquavers; Violin II, first group, and Violin I, second group, alternate one 16th triplet and two semiquavers; Violin I, first group, plays only 16th triplets. That evolves on 46 bpm and, along with the *legato* articulation, assures a linear type of musical texture as if it were one single layer of sound perceivable as being static rhythmically.

On Figure 2 is possible to look at all different note attacks performed by each instrument of the string section on 2/2 time signature, 60 bpm and *legato* articulation. At a first sight the voices seem to be disposed on a rather random chaotic lay out, showing all the possible subdivisions of one quarter note. Nonetheless it has been organically organized by the composer in such a manner that creates a great sounding path formed by an intense rhythm movement opposed to a notes sounding field apparently static caused by the excessive accrual of superposed pitches.

3.2 Melodic Approach

The excerpt from Shostakovich's Symphony nº2 presents each instrument of the string section performing a melodic line within a specific chromatic range. The following example shows Violin I at a range between A5 and E4; Violin I (2) C5 and A3; Violin II from Gb4 to D3; Violin II (2) Ab3 to G2; Viola from D3 to C#2; Cello B2 to G1and Contrabass Ab1 to E1, see Figure 3.

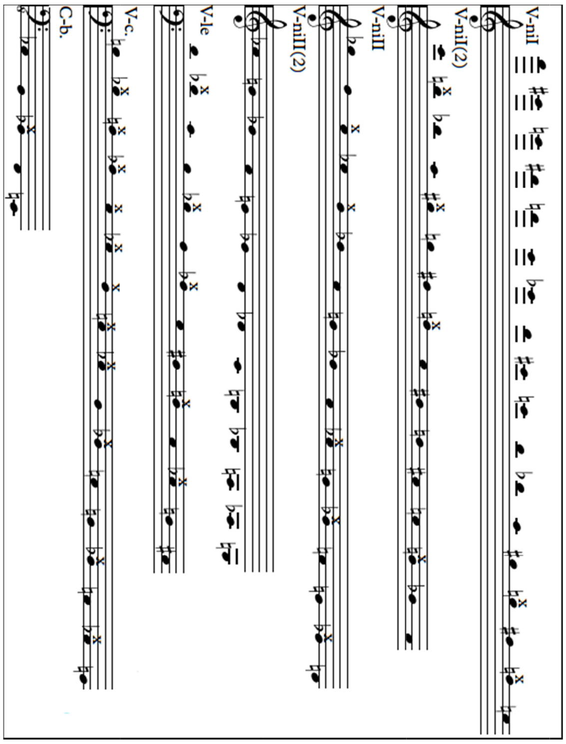


Figure 3. Chromatic range of all voices on Symphony nº2 excerpt.

In Shostakovich's excerpt it is possible to observe each string instrument performing a melodic line within its range going up and down in a wave type of movement ascending and descending partially over diatonic note scales, see Figure 4. The result is a fully chromatic environment in which becomes impossible to identify the individual lines. Besides, there are some pitches marked with an 'x' that are the ones missing in one specific range. Later on it will be demonstrated how those missing pitches will appear played by another instrument, this way covering a large chromatic range from A5 down to F#(Gb)2 and downer to E1 with missing pitches F2, Eb2, Bb1 and Ab1.

If we look at the Figure 2 it is possible to perceive the canonic path in a chromatic evolvement that each one of the polyphonic voices performs. On Figure 5 we may observe those voices in one large range from A5 to C2 with a distinct contrary movement between 1st and 2nd Violins going downward while Violas and Cellos move upward.

Figure 4. Groups of key related notes as indicated.

Figure 5 shows the fully chromatic range occupied by all forty eight string instruments that perform canonic lines starting on letter 'H' on the 47th measure of György Ligeti's *Atmosphères.* It totals fourteen 1st Violins performing within a range between G5 down to D# 4; the same amount of 2nd Violins occupying a range starting on E5 down to D3; Violas 1- 8 within a range starting on Ab 3 going up to B; Violas 4 - 10 from C2 up to E2 and the Cellos going up from C2 to D3, see Figure 5.

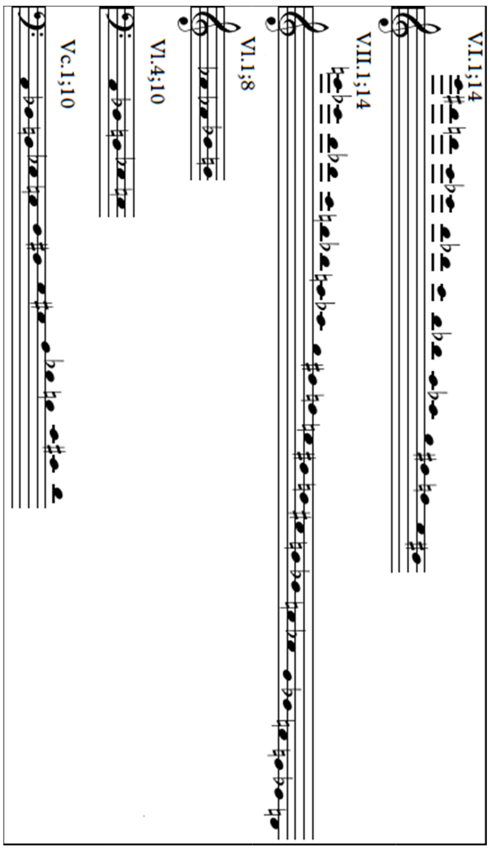


Figure 5. Chromatic range of all voices on *Atmosphères* excerpt.

3.3 Harmonic Approach

The accrual of voices in both examples along with their linear polyrhythmic performance creates a dense sounding environment which makes impossible to the human perception to distinct any kind or a minimal reminiscence of a tonal harmonic field. Instead what we hear is a sound mass emerging from the tight superposition of pitches and their multiple rhythm dispositions what ends up enhancing the amount of density and compression (W. Berry, 1987); (D. Guigue, 2004).

Besides, a fully chromatic sonorous ambience is created once in both cases there is clear evidence that the large vertical range existing in each of the excerpts is occupied in its whole amount of possible pitches. As it may be seen in the following example there are some missing pitches in the individual range of each string instrument in Shostakovich's excerpt. Those are marked with an 'x' and their absences are compensated in one of the other voices. These creates an entire chromatic field from an A5 at the top through a F# (Gb)2 on the bottom, that goes downer to an E1 but missing the pitches F2, Eb2, Bb1 and Ab1, see Figure 6.

In Ligeti's excerpt a likewise large chromatic range from G5 down to C2 is formed as a result of the composer's micropolyphonic technic of superposing canonic lines half step apart using a succession of minor and major second intervals, see Figures 7 and 8.

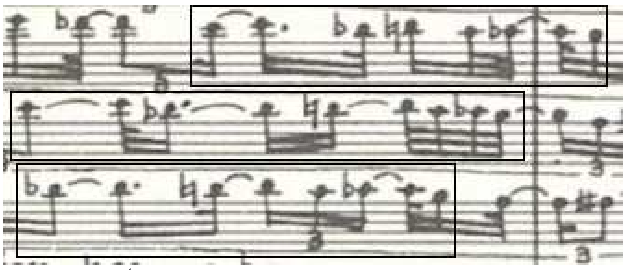


Figure 7. 1st Violins voices denoting horizontal and vertical minor and major 2nd intervals.

4. RESULTS

The chosen excerpt from Shostakovich's Symphony nº 2 amounts three measures (20 -22), has a 4/4 time signature and the metronome mark is at 46 bpm. Thus it is possible to conclude that the sounding texture generated then lasts for about 15,6 seconds. In the other hand Ligeti's excerpt endures for five measures (44-48) under a 2/2 signature and 30 bpm metronome mark what would make it last for at least 10 seconds. There would be time enough to hear and perceive in both examples a polyrhythmic and fully chromatic combination of pitches and note attacks, resulting in a type of sonorous texture which human perception might only attempt to it as being a single layer of a sound mass. Although using much different compositional technics, this two excerpts present sonorous similarities at the perception level due to the way that the musical material was crafted.

5. CONCLUSIONS

Dmitri Shostakovich's Symphony nº 2 was written more than three decades before György Ligeti's *Atmosphères*. The music excerpts presented in this article point out to the similarities between them which could suggest that:

5.1 Music Education

A profound research on the musical auditory experience might reveal aspects perceived through the sounding qualities that are not apparent in the written musical notation and might be demonstrated by specific analytical methods including those of music perception and cognition. That might help develop new ideas regarding the study and practice of ear training; musical analysis; performance and general music education.

5.2 Musicology

György Ligeti made his early studies in music between 1942 and 1949, period in which Hungary - the composer's home country - was taken under the communist regime. There is a great possibility that the composer might have had the opportunity of attending to a Dmitri Shostakovich Symphony nº 2 performance, or at least, have heard it through a radio station. That might became a sound texture reference for the Hungarian composer, once it is possible to perceive the same textural sounding effect in his works during the 1960's.

"Only a thorough knowledge of the styles makes one conscious of the difference between 'mine and thine'. And accordingly one cannot really understand the style of one's time if one has not found out how it is distinguished from the style of one's predecessors" (Arnold Schoenberg, 1938).

Acknowledgments

Marcos José Cruz Mesquita, Phd

All members of research group COGMUS

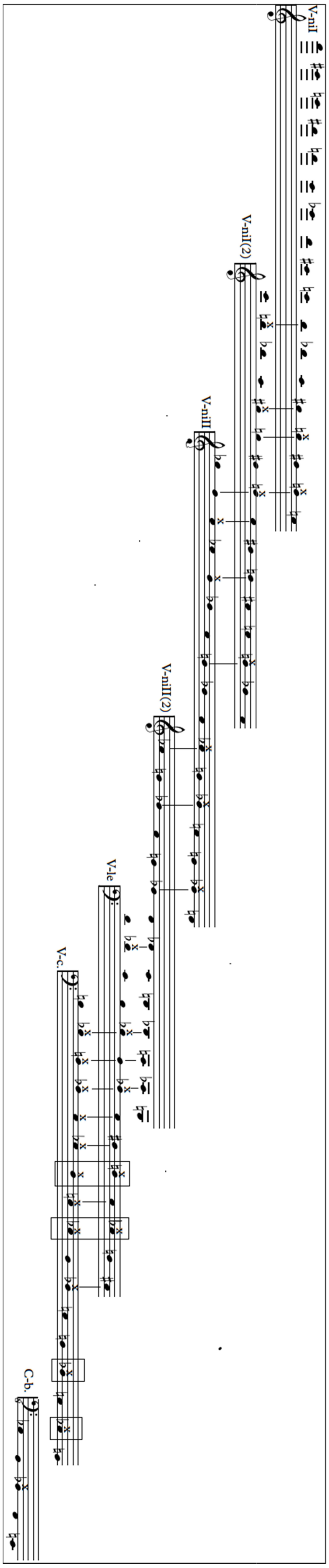


Figure 6. Chromatic range of 2nd Symphony excerpt.

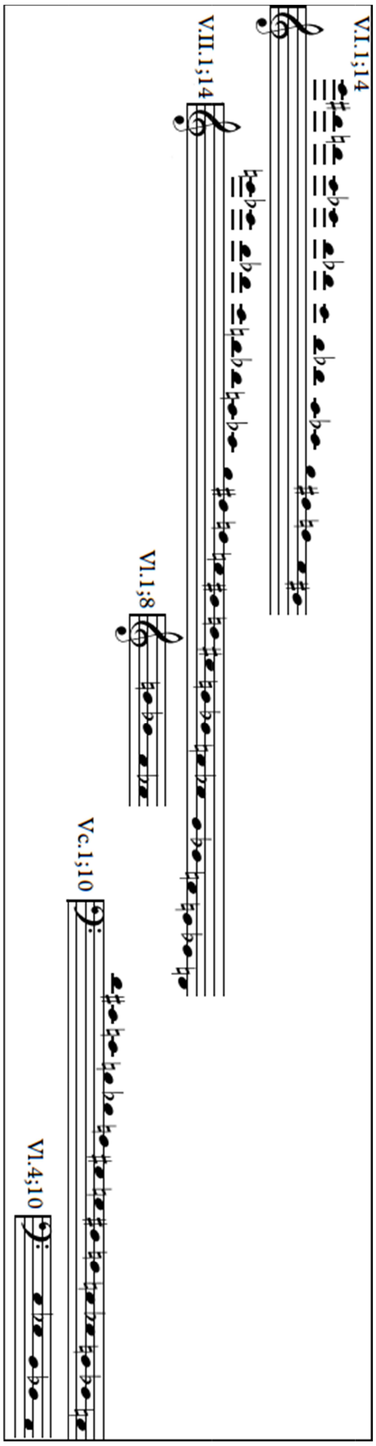


Figure 8. Chromatic range of *Atmosphères* excerpt.

6. REFERENCES

[1,8] D. Guigue, *Esthétique de la sonorité: l'héritage de Debussy dans la musique pour piano du xxe siècle.* Editions L'Harmattan, 2004, pp. 31-41.

[2,9] A. Schoenberg, in Stein, Leonard (ed.). *Style and Idea. Selected Writings of Arnold Schoenberg*, University of California Press, 1984, p. 76-79.

[3] R. Blokker, *The Music of Dmitri Shostakovich*, Associated University Presses, Inc., 1979, p. 48-50.

[4,7,8] W. Berry, *Structural forms in music*. Dove publications, 1987, pp. 184-299.

[5] G**.** Ligeti, *Metamorfhose of Musical Form,* in *Die Reihe*, Vol. 7 Bryn Maw Presser, 1965, pp. 5-19.

[6] V. GroBberger; J. Voit, *Listening Lab, vol. 2,* Universal Edition A.G., Wien, 2013.

AWÊ AND SPETACULARIZATION: THE DIFFERENTIAL LINE

Daniele Fischer
UFMG
danielefischer15@gmail.com

ABSTRACT

The present discussion is a dismemberment of our master's research (Fischer, 2017). Such an endeavour consisted in the realization of an ethnography focused on the contact between an indigenous group from the Pataxó ethnicity and a school community. Our fieldwork was developed in ten schools from Belo Horizonte. The goal of this study is to propose reflections starting from the concept of spectacularization proposed by Carvalho (2010) in contrast to the performance practice of Awê such as this Pataxó musical practice was experienced and shared in several schools of the research field. Carvalho (2010) shows us different questions that appear when musical and cultural expressions of traditional groups are parted from their specifics contexts. Among the author's questionings, it highlights reflections about temporal compaction and ritualistic dissociation when such expressions are distanced from their sacred contexts. Even if Awê is experienced in the villages in a more expanded way related to its duration (Vieira, 2016), its use is more compacted and its detachment from other ritualistic moments (when sharing it in schools), must not be evaluated towards the analysis keys proposed by Carvalho (2010), given the importance of sharing these expressions in multiple contexts as a way to disclose the Pataxó's culture and struggle.

1. INTRODUCTION

This work is the unfolding of the research made for my Master's Degree (Fischer, 2017). Actually, I must confess that its motivation appeared in an academic context and it was caused by a discomfort I felt in relation to it.

In 2017, three months after the presentation of my final dissertation, I found myself with some colleagues and professors from the university (UFMG) watching work presentations during Gret[i].

After my presentation, where I talked about my research (giving initial information on the field development and some conclusions), I was questioned by a doctorate student when we opened the time for questions (as it is usual in these kind of presentations). She wanted to know if the Awê (which will be explained later) – such as it is in schools – could be considered a kind of spectacularization.

I was not expecting this kind of question and immediately started to think of a way of answering it trying to justify the formatting of this expression, since when reading the article by Carvalho (2010) we see the perversity involved in the context of spectacularization. Following the author's considerations, a whole panorama is uncovered where we begin to understand the interest that so many promoters and spectators have in the field of popular culture, and how the concept of spectacularization becomes inadequate when we think about how one is objectified this way.

Facing this questioning I felt motivated to think about it, and I share in this article some of my thoughts with the intention of broadening such a necessary discussion that was started by Carvalho.

2. CONTEXTUALIZING OUR RESEARCH FIELD

This research is consisted in realizing an ethnography focused on the established contact that happened between an indigenous group, the Pataxó, and the school community. We focused on ten regular schools from Belo Horizonte and surroundings, doing what Prass (2009) calls a multi situated work.

Through a project called Kijetxawê[ii], a group of indigenous from the Pataxó has been developing a rich work of cultural interchange with school community, at the same time they broaden the visibility of the fight of the indigenous people in general, mainly in relation to the fight of the Pataxó for their rights.

In return for the work, the school offers a place where the Pataxó can expose and sell their handicraft, which is nowadays their main means of supporting (according to their key interlocutors, as well as it is possible to find references in Grunewald, 2001 – Melatti, 1995 and Porto, 2014).

3. SHORT DESCRIPTION OF THE FIELDWORK

In my research I had the opportunity of following two different indigenous groups. Even with a different approach to each one, the Awê was a common element between them and it was present in all of our visitations.

One of the groups worked more the general aspects and curiosities about the indigenous culture and the Pataxó culture, and the other had a more political approach, conceptualizing the PEC 215[iii] and making an appeal for the students to speak up against it in their social networks. However, the Awê was always the element that made possible the first contact between the indigenous and the school community, and it was the only element that, in spite of the different kinds of approach, kept some homogeneity in the different schools.

In order to elucidate this moment, I share here an excerpt of my dissertation[iv], where the Awê is described with more details:

When the Pataxó arrived in the schools, they first organized the exposition of their handicraft, and only after it was carefully arranged – making sure everything was very well organized in a harmonic way – they went to the place the school had pointed that they would wait or be expected by the students. In that place they officially started their work (I always had the impression that the organization of their exposition was a previous moment where they were preparing themselves, only then to show themselves officially present, since the products would only be sold after the first contact with the students), presenting themselves as indigenous of the Pataxó ethnicity (using their indigenous names) and soon saying they would share a little of their culture, starting to use the Awê.

Most of the time, they started chants with a prayer called Kanã Pataxi Petoi[v],which, according to my interlocutors, they have a moment not only to thank, but also ask for the protection of Niamisu (motivation also cited by Vieira, 2016*).*

Bending down during the prayer, one of them sang the chant as a solo followed by the others with the same intention, a corporal movement of contemplation and resignation. After singing their prayer, they stood up and said it was the Awê moment.

When they stood up to do the Awê, the indigenous put themselves in the position of a circle and a line. The corporal movements were made walking through the place, but with a typical stomping of the feet that was more emphatic in each second beat of a binary metric, which was also executed concomitantly by the maracás.

The majority of the chants were executed in Patxohã.

The average of the chants had about four parts after the initial prayer. Each one was repeated three to six times.

4. THE AWÊ AND ITS IMPORTANCE IN THE CONTEXT OF THE PATAXÓ CULTURE

The Awê achieved a special highlight in our research in relation to the cultural elements of ethnical demarcation. During the bibliographic research, we could also verify the strength of this musical practice in the indigenous Pataxó communities as a unifying element (after a long dispersion period caused by the 51 Fire[vi]), as well as a part of the cultural updating and revitalization of the Patxohã (see more in Santana, 2016).

From our observations it was possible to determine the importance of musical practice to the Pataxó (practices that receive the name of Awê) in their imbricated relation with the processes of identity reaffirmation and ethnic demarcation.

When taking the Awê to schools – even if with a reduced formation in relation to the bigger ritual that is realized in their villages – the Pataxó show the aspects of their culture and elicit in the students the debate about the indigenous matters at the same time they gain more visibility to their fights and their own survival.

5. SOME REFLECTIONS ABOUT THE AWÊ IN OPPOSITION TO THE CONCEPT OF SPECTACULARIZATION

Anthropologist and researcher José Jorge de Carvalho (2010) brings an important contribution to the field of thoughts concerning the commercial relations surrounding the groups of popular culture[vii], elucidating the asymmetric relations that appear in this field, in what concerns the valuing of such expressions to the detriment of the real acknowledgement that they reach in scenarios that involve power relations.

Carvalho (2010) signals some elements that constitute what would be the spectacularization of certain cultural manifestations and points some components like temporal summarization and the split of the ritual aspects from their original context, besides the fact that spectators that watch the presentations, however, do not share a bigger involvement[viii].

Even if in many cases the spectacularization of traditional expressions happens like this, we want to amplify the discussion to the fact that not everything that is compact and not everything that is separate from a bigger ritualistic context should be said so.

The keys for interpretation proposed by Carvalho (2010) cannot be automatically imported so that we interpret the *modus operandi* of the Awê out of the villages. However, we do not erase the argument and the validity of its contributions, since there are inequalities in our country in what concerns the access to property and services, fees paid to many cultural groups, as well as the asymmetries of power that underlie the relations involving determined groups and the scenarios they transit.

In our analysis, however, we do not want to disqualify the great contribution given by Carvalho (2010) in his reflections. We wish only to amplify the discussion concerning the field of popular and traditional cultures, bringing more elements that can redirect our look to the Pataxó context and their cultural elements related to negotiations and interchange.

6. CONCLUSIONS

When we proposed this discussion, we did not deny the asymmetries concerning the commercial relationships involving the traditional and popular groups in relation to the power structures that hire such groups or enjoy their work.

However, we want to personalize such propositions by calling attention to the fact that in each group there are specific characteristics that involve their musical and cultural expression, and only the keys of temporal compaction and the dissociation of a bigger ritualistic context are not enough to give readings concerning the field of spectacularization of culture.

It is pertinent to point that the presence of the Pataxó in schools does not solve by itself the problem they face. They keep suffering with discrimination in many places they attend, they are sometimes invisibilized in talks about indigenous representation and so many other problems like that, but we cannot deny that their presence in schools is essential, and many times they are the only representation of the Law 11.645/08[ix]. Because of them the students are aware of the indigenous in the country, in an urban context, in the reality as a whole, and "find out" that "indigenous" are not only figures of a remote past.

The Awê, as a mechanism that is largely shared by the Pataxó, works as an icon of indigenous representation, and even happening outside its ritualistic place, in a more compacted way, we cannot deny the huge importance of its diffusion as an element that is notably related to the existence and resistance of the Pataxó.

7. REFERENCES

[1] Carvalho, José Jorge. Espetacularização e Canibalização das culturas populares na América Latina. Revista Anthropológicas, Pernanbuco, ano 14, vol. 21: 39-76. 2010.

[2] Fischer, Daniele Damasceno. Práticas musicais dos Pataxó nas escolas regulares de Belo Horizonte: contatos interétnicos e afirmação identitária. Belo Horizonte, 2017. [166f.] Dissertação (Mestrado em Música). Universidade Federal de Minas Gerais – UFMG, Belo Horizonte, 2017.

[3] Grunewald, Rodrigo de Azeredo. ***Os Índios do Descobrimento: Tradição e Turismo.*** Rio de Janeiro: 1999. [362f] Tese (Doutorado em Antropologia Social). Universidade Federal do Rio de Janeiro – UFRJ, Rio de Janeiro, 2001.

[4] MelattiI, Júlio César. Áreas etnográficas da América Indígena. In: **Conferencia presentada al Seminario "Arqueologia-Tendências Atuais.** Universidad de Brasilia. 1995. Edição de 2016.

[5] Pataxó, Povo. **Inventário Cultural Pataxó: Tradições do povo Pataxó do Extremo Sul da Bahia.** Bahia: Atxohã/Instituto Tribos Jovens (ITJ), 2011.

[6] Porto, Helânia Thomazine. T. Linguagens e significação: análise semiótica do artesanato dos Pataxós do extremo sul da Bahia. In: BARZOTTO, Valdir Heitor et al. **Linguística e Ensino de Língua Portuguesa.** Adriana Santos Batista, Aline Maria dos Santos Pereira, Celso Kallarrari [Organizadores]. São Paulo: Opção Editoras, 2014.

[7] Prass, Luciana. Maçambiques, quicumbis e ensaios de promessa: musicalidades quilombolas do sul do Brasil. Porto Alegre: 2009. [313f.] Tese (Doutorado em Música). Universidade Federal do Rio Grande do Sul – UFRGS, Porto Alegre, 2009.

[8] Santana, Cleidiane Ponçada. **Cantos Tradicionais Pataxó na Língua Patxôhã.** Belo Horizonte, 2016. [68f]. Trabalho de conclusão de curso (Formação intercultural para educadores indígenas). Universidade Federal de Minas Gerais – UFMG, Belo Horizonte, 2016.

[9] Vieira, Vislandes Bonfim. A importância do canto dentro do ritual Awê Pataxó. Belo Horzionte, 2016. [49f.] Trabalho de Conclusão de Curso (Formação Intercultural para Educadores Indígenas – FIEI). Universidade Federal de Minas Gerais – UFMG, Belo Horizonte, 2016.

[i]UFMG's Ethnomusicology Group, guided by professor Glaura Lucas and Professor Eduardo Pires Rosse.

[ii]Kijetxawê in Patxohã (language that has been revitalized by the Pataxó) means *school.* More in PATAXÓ POVO (2011).

[iii]PEC – Constitutional Amendment Project that threatens indigenous lands through new rules of demarcation. See more in: http://www.cimi.org.br/pec2015/cartilha.pdf. Accessed in January 18th, 2018.

[iv]More in Fischer (2017).

[v] According to Vieira (2016:19) this prayer can be "Kanã Pataxi Petoĩ" or "Goá Miãga".

[vi]More in Fischer (2017).

[vii] Carvalho conceptualizes popular culture as "a whimsical set of cultural forms – music, dance, dramatic autos, poetry, handicraft, science about health, ritual forms, traditions of spirituality – which were created, developed and maintained by thousands of communities of the country in distinct historical moments". (2010, p. 44)

[viii] Check antagonism between "livingness" and "experience" (cf. Carvalho, 2010 p. 48).

[ix] Law that deliberates about the obligation of teaching matters related to indigenous and afrobrazilian culture in all regular schools within national territory. Available in: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/lei/l11645.htm. Accessed in May 26th, 2018.

APROPRIATION AND INAPPROPRIATE TRADITION IN THE PRACTICES OF SAMPLING AT RAP MUSIC IN THE CITY OF BELO HORIZONTE

Michel Brasil
State University of Minas Gerais
michelbrasil83@gmail.com

ABSTRACT

Appropriation is a recurrent element in the context of Hip Hop. It can be enacted in a number of ways, but the tradition of sound sampling is a very significant creative resource driving music in the Hip-Hop culture. The use of sampling started in the 1980's and helped to configure the traditional sonority of rap. It helped the consolidation of the Hip Hop culture and Rap as a world music genre, but also promoted the development of local styles of Rap. How the original action of borrowing music samples evolved in such a diversity of local cultures of Rap? How do local agents of the cultures of rap produce idiosyncrasies from the idea of sampling? In this study, we investigate musical practices and sampling carried out by beatmakers in Belo Horizonte, based on an ethnography among artists from the city of Belo Horizonte.

1. INTRODUCTION

This article is the result of ethnography among rap music agents from the city of Belo Horizonte, Brazil. The main purpose is to investigate the use of samples of other musical works in the musical production of rap beats. We are interested in knowing the opinion of the local artists about this technique, what symbolic constructions they inculcate in this procedure, in what ways this procedure (and the sonority that derives from it) inhabits the processes of beatmaking.

Born in the 1970s, in New York, in the Bronx region, Hip Hop is currently a global phenomenon. Over the last few decades, rap music, as well as break and graffitti, has spread through various locations, becoming a transnational artistic language. Although rap has many characteristics of Afro-American culture and musicality, the genre currently embraces a multitude of local characteristics. According to Toni Mitchell:

"Hip-hop and rap cannot be viewed simply as an expression of African American culture; it has become a vehicle for global youth affiliations and a tool for reworking local identity all over the world. Even as a universally recognized popular musical idiom, rap continues to provoke attention to local specificities. Rap and hiphop outside the USA reveal the workings of popular music as a culture industry driven as much by local artists and their fans as by the demands of global capitalism and U.S. cultural domination". [1]

This ethnography sought to observe the postures and creative options that comprise the use of samples in the beats creation process (or the option of not using them). During the research I talked to several producers, beatmakers, rappers and individuals from Belo Horizonte connected to rap music. I also participated in several events and visited studios, where I observed processes of musical production of raps. In addition Thus, I collaborate with the production of a radio show[1] specialized in rap, which provides me contact with artists and agents linked to the local rap music scene. Throughout the research I recorded notes, took pictures, and recorded nine interviews with rap agents, including eight music producers and one rapper.

The consultants who collaborated with this research are male, most of them over the age of 30, residents of Belo Horizonte and neighboring cities. Although the rappers were also able to collaborate with the discussion, I prioritized the contact with beatmakers and producers. This focus is due to the fact that they experience daily musical production, and provide a more immersed view on this issue. Another element that contributed to this choice is the fact that few researches about rap music use the beatmakers as interlocutors, or investigate aspects of the musical production of the genre. Rappers and the lyrical element of raps are often privileged as sources of information and analysis.

It was possible to observe that the sampling, besides being an aesthetic option that results in a type of sonority, also constitutes in one of the places of interaction of the artists of the region with the Hip Hop culture. Local producers' views and positions on sampling reveal the ways in which local actors deal with the history and the geographic shift of this culture and its musical genre. The value they attribute to Hip Hop demonstrates the desire to follow a tradition, even if it is imperative to adapt it for making it possible to be followed.

[1] The radio show Hora RAP is produced by Clebin Quirino and the author. Since 2014, the program has been broadcasted live by Rádio UFMG Educativa, an FM radio station linked to the Federal University of Minas Gerais. www.ufmg.br/radio

2. RAP MUSIC AND ITS MUSICAL PRODUCTION

Rap is an outgrowth of the combination between DJs and MCs, and it evolved from the improvisation of the MCs, who used to perform over the beats played by the DJs in the Bronx block-parties. After 1979, with the commercial release of the first raps, what was a live performance also became a genre of popular music, a studio-produced musical record, to be sold and consumed by a large audience.

As it moved toward the recording studio, rap music faced a technical and creative challenge: how to record and produce the musical foundations under which rappers rhyme? During the parties, the MCs improvised during instrumental stretches of the songs selected by the DJs. These parts were known as breaks, and were chosen because they contained a strong rhythmic foundation [2] [3]. The DJs and dancers from the Bronx enjoyed the breaks because these were moments in which the harmonic and melodic elements stopped playing, so drums and percussion elements gained prominence. The DJs started to repeat the breaks in loop, through the use of two turntables with the same record. As soon as the break ended in one record, the DJ would start the same break in the other turntable, and so on, switching between the turntables. This technique created a repetitive cell and expanded the break for a long time. As in the Jamaican street parties in the 1960s [4], MCs began to improvise verses on the breaks, and this practice gave rise to rap music.

From the earliest recordings to the present, different techniques, tools and procedures were used to make the beats, which are the musical foundations that support rapper/MC performance. Originally, they preferred to record the performance of a band playing some grooves inspired by the breaks played by the DJs. Some time later, the electronic drums gained space as a tool in the beatmaking process.

In the mid-1980s, rap music producers started to use keyboards and other hardware devices known as samplers. Samplers were devices that had the ability to digitally record, store, edit, and manipulate any sound. The excerpts recorded and digitally manipulated in these devices are called samples. The process of digitalization and manipulation of sounds for musical production came to be called sampling. Mark Kats defines sampling as:

"[…] a type of computer synthesis in which sound is rendered into data, data that in turn comprise instructions for reconstructing that sound. Sampling is typically regarded as a type of musical quotation, usually of one pop song by another, but it encompasses the digital incorporation of any prerecorded sound into a new recorded work". [5]

Initially, the samplers were designed as keyboards, in which it was possible to insert digital samples of sounds and to play them using the keys. The intention was to increase the timbral range available to the musicians and producers, without the necessity of a high investment on equipment or musicians. Later on, the samplers gained other designs, and came into existence with other shapes and functions. What was initially thought to save time and money, meant a technical and procedural innovation in the context of popular music. Tricia Rose states that:

"Prior to rap music's redefinition of the role samplers play in musical creativity, samplers were used almost exclusively as time- and moneysaving devices for producers, engineers, and composers. Samplers were used as short cuts; sometimes a horn section, a bass drum, or background vocals would be lifted from a recording easily and quickly, limiting the expense and effort to locate and compensate studio musicians. […] In fact, prior to rap, the most desirable use of a sample was to mask the sample and its origin; to bury its identity. Rap producers have inverted this logic, using samples as a point of reference, as a means by which the process of repetition and recontextualization can be highlighted and privileged". [6]

Rap musicians recognized in the samplers the opportunity to extend the logic of the Bronx DJs. The samplers enabled the producers to expand the foundation of the breaks. Nelson George expounds that:

Sampling's flexibility gave hip hop–bred music makers the tools to create tracks that not only were in the hip hop tradition but allowed them to extend that tradition. For them the depth and complexity of sounds achievable on a creatively sampled record have made live instrumentation seem, at best, an adjunct to record making. Records were no longer recordings of instruments being played—they had become a collection of previously performed and found sounds. [7]

The transition from the stages to the studios also meant a change in the DJs' activities. The advent of rap as a musical product brought the necessity of beatmaking. This is declared by Joseph Schloss:

"When hip-hop expanded to recorded contexts, both of these roles [MC and DJ] became somewhat more complex. MCs began to create increasingly involved narratives using complex rhythms and cadences. And although deejays continued to make music with turntables when performing live, most also developed other strategies for use in the studio, and these eventually came to include the use of digital sampling. As these studio methodologies gained popularity, the deejays who used them became known as producers". [3]

Gradually, the DJs took over the beatmaking tasks, since conventional music producers did not have the aesthetic sensibility, neither the expertise needed for the job. At the beginning, nearly all beatmakers were DJs. Today this connection is not so strong anymore. Many beatmakers have never acted as DJs, and although a large number of DJs are dedicated to making beats, many of them are only committed to live performance.

Between the second half of the 1980s and the first half of the 1990s, most of the rap beats were composed by the appropriation of sound elements from other musical works [8] [9]. The sample-based beatmaking became a sort of standard procedure, and coincided with a period of creative expansion and consolidation of the genre in the popular music market.

However, throughout the 1990s, the use of samples in rap's musical production started to decline. One of the reasons that contributed to this decline was the legal proceedings against rappers and producers, due to the use of phonograms and musical works belonging to third parties. The "Grand Upright Music, Ltda. Warner Bros. Records Inc., " case, which occurred in 1991, was a landmark with regard to legal interpretation of the use of samples of other works in musical creation. This is a lawsuit filed by singer-songwriter Gilbert O'Sullivan against rapper Biz Markie, who sampled an excerpt from O'Sullivan's song "Alone Again (Naturally)" on a track from the album "I Need a Haircut "(1991). After this case, it was determined that any use of musical samples should be pre-approved by the copyright holders of the original work. Such positioning interfered directly in the processes of rap music production, since it presented a considerable impediment to the viability of the records. Rights holders of sampled songs (usually record labels and copyright management offices) began to demand large amounts for samples clearance, which began to undermine the producers' borrowing enthusiasm. This is what Williams reports:

“Because of the tightening of copyright legislation for sampling in the late 1980s and early 1990s, collage-style albums like those from Public Enemy and De La Soul would be too expensive to make commercially in the mid-1990s and after”. [10]

Thus, producers and rappers sought other ways to produce their music tracks, which contributed to the consolidation of other processes and tools that were already part of rap music, but which did not have as much evidence as sampling. “With the imposition of so many limitations, not to mention the fine line between creativity and a lawsuit, sample-based hip-hop artists have adapted their production approaches in many ways”, states Amanda Sewell (2014. p. 300). This adaptation made room for an increase in the use of synthesizers and the return of the use of studio musicians. Anyhow, the obstacles imposed by copyright law cannot be considered the only reasons for reducing the use of samples in rap production. The socio-technical context - the genesis of tools and production equipment - and personal artistic choices are also significant factors for grasping the issue. Albeit sampling is not as decisive for rap's sonority today, it still operates as a timbral and procedural reference in different ways.

3. SAMPLING, PURISM AND BEATMAKING

Joseph Schloss [3] is responsible for a relevant research on musical production of raps based on the use of samples. Based on ethnography among American beatmakers that prioritize the use of samples in the production of beats, the author sought to understand the beatmakers on their own terms, focusing on the sampling process rather than on the final result (the beat, the music).

The author seeks to highlight the creative agency of musicians and producers, and states that the choice of using samples is not a pragmatic option, but aesthetic: […] sampling, rather than being the result of musical deprivation, is an aesthetic choice coherent with the history and values of the hip-hop community” (p. 21) [3]. Beatmakers do not sample because it is convenient, but because they consider it musically beautiful. In addition, he believes that the distinction between raps produced from samples and raps that do not use samples, rather than a technical distinction, is a distinction of genres (or subgenres) within rap music.

Schloss identified a sort of purist behavior among the community of beatmakers surveyed, whose sampling role is central in this perspective. His research claims that beatmakers regard sample-based production as an evolution of hip hop DJ practice. In this sense, the ways of manipulating the samples and the aesthetic assumptions that guide their production processes are the same as the DJs did. The samplers allowed the possibilities envisioned by the DJs to expand and become more complex, and also made the work easier to execute. Another extreme of this purist behavior is the little (or no) importance attributed to the use of live instrumentation in beats production. According to Schloss, the use of such artifice does not sound authentic for the producers consulted during his research.

In an excerpt from the conversation I had with rapper and producer Easy CDA (Hertz Bento), he remembered a kind of purist behavior existing in rap music a few years ago. In this passage, he explained the importance of breaks for the construction of rap beats and, in a sort of historical link, he recalled a behavior similar to what Schloss explains:

“Because the break is born from this [funk] groove. And the rap music comes from the break… because the break was for dancing.. and later, for singing. And then it turns to a colossal thing. Today, the sky is the limit. Because today you can use anything to do it. Back in the days we had that stuff: “Hey, it ain’t rap if it’s not made in the MPC, if it’s not sample-based”.

M: “Where did they use to say that?”

E: “All over the world”.

M: “Here in Brasil, too?”

E: “Everywhere, specially here”. [11]

4. BEATMAKING AND SAMPLING IN BELO HORIZONTE

The city of Belo Horizonte has one of the most vibrant and diverse scenes of Brazilian rap. Alongside São Paulo, the city was the first in Brazil where the presence of b-boys and rappers could be noticed, in the mid-80s. Although the artists of Belo Horizonte do not have as much visibility and national circulation as the artists of São Paulo and Rio de Janeiro, the local scene shows a constant, diversified and vibrant production. It is possible to suppose some explanations for the fact that the local artists do not reach so much repercussion, and they range from economic, social and historical questions. However, further research and further study is needed, and this is not our focus. In any case, there is an increasing number of local artists who overcome the difficulties and achieve national repercussion and acceptance. In the early days, it was the turn of the group União Rap Funk and MC Pelé.

More recently, artists like Flávio Renegado, Das Quebradas, DV Tribo, Gustavo Djonga, Clara Lima and Duelo de MC's, a bloc-party that includes all the Hip Hop elements as well as rap battles.

The first recordings of local rap groups took place in the early 1990s. The first beatmakers were DJs who spinned records at the black-music parties, like DJ Joseph and DJ A Coisa. There was restricted access to equipment and recording studios so, until the end of this decade, few musical producers appeared in the local scene. At the beginning of the 2000s, there was an expansion of the producers' work. Despite the low purchasing power, this generation had access to basic computers and software, which enabled them to experience the beatmaking process more easily. DJ Spider, Easy CDA and Enecê are the first producers that emerged this time, followed later by Clebin Quirino, DJ Giffoni, among others. There was an even greater expansion of beatmaking activity in the last decade. The number of producers who started their activities recently is much higher than in the previous decade. This number grows regularly, due to full access to information, equipment and musical production resources. Currently, the beatmakers local scene is sufficiently structured, and offers rappers a rich variation of styles, sonorities, and technical resources.

The field research allows to affirm that the local beatmakers do not present a purely purist behavior (in the sense attributed by Schloss) in relation to the use of samples, but the use of samples is still very present in the beats production / composition processes. The study did not identify any beatmaker that produced solely based on the use of samples. No artist with whom I spoke to stated such a proposal, nor was it possible to determine a beatmaker with such a nature. On the other hand, the producers rarely say they don't use samples at all. To a greater or lesser extent, everyone makes use of this procedure, either due to aesthetic choices, an external demand (for instance, by the rapper), or inability/insecurity to deal with other music production tools.

There is a local tendency to merge different production and composition techniques and tools. It was possible to observe that sampling, both in the pragmatic and musical scopes, continues to be a reference for producers, but is coalesced with other production and composition resources, such as the use of virtual synthesizers and the recording of live performance on instruments such as guitar, electric bass and piano.

In addition, the data collected in the ethnographic process suggest that the use of samples is seen as a timbristic and symbolic link to the traditional sonority of hip-hop music. Several of the consultants made this connection explicit. Some producers, like DJ Giffoni (Sérgio Giffoni), relate the sampling to the essence of the genre: "I think that's part of rap music. This sampling stuff is part of rap" [12]. DJ A Coisa (Paulo da Silva Soares), one of the oldest DJs and local producers, has a similar opinion: "[…] sampling has everything to do with hip hop. Hip hop is shaped this way, you know? When you take something nobody remembers or, often, nobody knows, and you give it a new look" [13]. The MC and journalist PDR (Pedro Valentim) also develops a similar reasoning:

"I think [sampling] is a major principle, dude. I'm passionate about this sample stuff, indeed. I've been a beatmaker too, I've had my beatmaker season, I'm very much interested in the practice of making beats. I think it's foundational, like a lot of practices that come from the background of Hip Hop culture, that makes you… it makes you give new meanings to the things, from what you have in your hands, from the possibilities that are presented to you. And it's creating your history too". [14]

The producer and DJ Preto C (Carlos Henrique da Silva) was even more emphatic in expounding this link between sampling and Hip Hop:

Michel: "You said you like to use samples. Why do you like this tool? [Why do] You like to work with..."

Preto C: "Because I think this is part of the Hip Hop history. I think this is being conservative with ideology. I think this is part of the history… since the beginning of Hip Hop, since the b-boy, since when they used to make loops with soul music [records]. When the DJ used to make a kind of back-to-back and the loop came back to the beginning of that part… I think the sample is part of the history. I've been researching a lot, this is part of the beatmaking art. […] For me, the sample is actually a prime factor in my beat. I may use the other instruments but I won't stop using even a voice [from a sample]. Sometimes I may not use a melodic sample, I just use a voice. I may borrow that melody, maybe recalling it in a different way… but the sample, for me, is primordial in a beat. I think this is part of the culture, Hip Hop culture". [15]

The producer Enecê explains his understanding of the issue in a testimony about his initiation as a beatmaker. In the following passage, he goes from a narrative about his search for information until the discovery of the use of samples, and the importance of this procedure for the Hip Hop culture:

"I started to understand [the beatmaking process] by paying attention to the music, to the sample, [and] how the guys used to sample, the gringo guys… both the gringos and the Brazilian [beatmakers]. And I used to search a lot for samples, you know? I liked the sampling technics a lot since the beginning. I think the essence of the rap beat is the sample, and it will always be like that. I think that's something that's never gonna die. Maybe, today, I do not use as much samples [as I used to], but I think it still is the substance of rap music". [16]

The rapper and producer Clebin Quirino (Jefferson Cleber dos Santos Costa) brings up two different arguments to justify the for choice of samples. He does not necessarily talk about "essence" or "foundation", but refers indirectly to Hip Hop culture, when it values the possibility of appropriation itself, and the logic of creation that lies on the reconfiguration of a preexisting musical material. He then argues that sampling enables a person who does not master any musical instrument to express himself musically, and this is a prime factor in his choice. He reports as follows:

"Michel: Why do you use samples? Is there any special reason for that"?

Clebin: "First of all, it's because I think that the possibility of borrowing music from the past, or music from the current time, music that… brings you some memories, or some section of the song reminds you something pleasant… I think this is very cool. I think it's nice when you try to reshape a song that already exists, by turning a small sample of it into something else, which sometimes has to do with it and sometimes has nothing to do [with it]. So, that's the first reason. Another feature that makes me enjoy sampling is because I don't play any musical instrument. The only musical instrument I play is the turntable. Sometimes I try to make some melody, at the studio, but I'm not a musician that picks up an instrument and play along with someone who's singing. […] the sampler gives me the possibility of not knowing how to play anything but to be able to use a sample with features that please me, and to make a song that I like". [17]

This section of the conversation with Clebin, regarding the use of samples, is unusual among the consultants who collaborated with the research. Although he mentions the question of appropriation, his justification does not clearly refer to a purported connection to the essence of Hip Hop. In addition, his second argument introduces another layer in the discussion. Although most beatmakers consulted demonstrate knowledge on the foundations of Hip Hop culture and rap music, what other elements influence their personal choices on tools and methods?

As previously stated, the beatmakers demonstrated a thorough understanding of the history and aesthetics of Hip Hop. However, this knowledge does not necessarily engender purist behavior (using the Schloss term) between them. The use (or not) of samples involves a relationship with the Hip Hop tradition, but other elements come into play when it comes to this choice. In this sense, local producers move between purist behavior and less traditional technical and aesthetic options in the context of Hip Hop.

Producer Easy CDA (Hertz Bento), for instance, recognizes the centrality of sampling in Hip Hop but, at the same time, he claims that he prefers to create the melodies by playing it live at the keyboard. At the same time, he says he likes to work with samples and does not impose a restriction on himself in that sense:

"Michel: Do you still use samples? Did you use to use it more?

Easy CDA: Yes, I use it. But I use it less, because I come from this… I really like to create it. I like the sound of something created, I like… making my own arrangements, [I like] searching for my own sound at the track. I use virtual instruments, I use a conventional keyboard, synthetizers. Sometimes, when it's needed, I make something with samples. […] something remarkable about my work is that you'll always listen to remarkable keyboard lines, something clean. I'll hear the keyboard and the bass together. It's quite this way. It's a feature from my beats. I'll listen to the synths… It's the keyboard stuff, it'll always have a keyboard or a synth, either a real or a digital one. Sometimes I also use live instrumentation. Sometimes I play the bass, sometimes I play the *cavaquinho* guitar […]" [11]

Although Easy is characterized by the blatant use of timbres and arrangements of keyboards, it has several sample-based beats, both in older works and more recent productions. During our conversation, he argues that he does not make frequent use of samples because he does not have time to devote himself to musical research of sample sources, which is a prime factor for this type of production. According to Easy:

"[…] because I like to use samples too, but there is something about it, that doesn't allow me using it frequently. As I work a lot, I work with a lot of people… so I don't have time to make it with samples, I don't have time to work as a traditional beatmaker, like… entering into the studio, locking the door, start digging in the crates, tracks… and sequencing, making a collection of samples and beats. I can't afford such a luxury. Because, for me, the traditional beatmaker is a real researcher. He, or she, is a researcher. You must have time, and it's a hard-working thing. Sometimes you make it and it doesn't work. That sample doesn't work and you must delete everything. You gotta have time to slice the samples. You gotta research for sounding, for creating your own beatmaker identity". [11]

What this passage of the conversation with Easy points out is the specific working method that production from samples imposes. Producing beats from samples involves: a constant search for sample sources; familiarity or interest in digital editing and manipulation tools; as well as a specific creation framework that is based on borrowing. Beatmakers that work with samples have a rationale, a workflow, which is based on the transformation of pre-existing material. Their creativity evolves within a supposed limitation, which involves dealing with sound and musical elements, or features present in the original track that they sample.

In this sense, the option to work with samples involves a personal positioning, on the part of the producer, as to which method of creation they feel most willing to express themselves. Some, like Easy CDA and DJ Spider, believe that they express themselves better, or they can be more creative, more productive, working with traditional tools, like live instrumentation. It is interesting to notice that working with more traditional musical concepts and tools, such as musical notes and live performance recording, departs somehow from the Hip Hop tradition, which from the beginning valued the appropriation and use of reproduction and programming equipment, to the detriment of instrumental performance.

5. CONCLUSION

What is observed is that, although the use of samples is not as blatant as in the classical period of style, the musical identity produced by sampling continues to be accessed by local producers. In this sense, beatmakers continue to use samples in their productions, but they are not limited to producing exclusively from samples (like some North American producers). There is a local trend to mixing samples with virtual synthesizers and live performance recording. While beatmakers recognize that sampling is linked to the essence of the genre, they feel the need to use other tools available in order to develop an artistic uniqueness that positions them in an authentic way in the local community of beatmakers.

The aesthetics of rap music is in constant transition, which is observable both in its sonority and in the processes of production. Hip Hop culture offers a context of philosophical and aesthetic creation and reference for this musical genre, but rap is no longer limited to the limits of this culture. And while Hip Hop embraces change and adaptation at its core, perhaps some of the developments that rap music presents are not confined to this tradition. Sampling is one of the ways in which rap agents in Belo Horizonte dialogue with this tradition. We observe that this procedure links them to this culture, but other options and personal and local elements are mixed in the processes of composition and musical production of rap. Such elements approximate traditional procedures in the production of popular music but distance themselves from the traditional production procedures of rap music.

Acknowledgments

I would like to thank Paula Manzali, all the consultants that colaborated with the research, and Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG) for the support.

6. REFERENCES

[1] T. MITCHELL (org.). *Global Noise: rap and hip-hop outside USA*. Wesleyan University Press, 2011.

[2] B. BREWSTER and F. BROUGHTON. *Last Night a DJ Saved my Life: the history of the disc jockey.* Groove Press, 2006.

[3] J. SCHLOSS. *Making Beats: The Art of Sample Based Hip Hop*. Wesleyan University Press, 2014.

[4] D. HEBDIGE. *Cut 'n Mix: culture, identity and caribbean music.* Comedia, 1987.

[5] M. KATZ. *Capturing Sound: how technology has changed music*. University of California Press, 2004.

[6] T. ROSE. *Black Noise: rap music and black culture in contemporary America.* Wesleyan University Press, 1994.

[7] N. GEORGE. *Sample This*. In M. FORMAN and M. A. NEAL. “That’s the joint: the hip hop studies reader book”. New York: Routledge. 2004.

[8] G. KISTNER. *Hip Hop Sampling and African American Music: an analysis of Nas’ ‘Get Down’*. Dissertation. Université Laval. Québec: Manuscript. 2006.

[9] A. SEWELL. *A Tipology of Sampling in Hip Hop*. PhD Thesis. Musicology Program Jacobs School of Music, Indiana University. Indianapolis: Manuscript. 2013.

[10] J. WILLIAMS. *Musical borrowing in Hip Hop: theoretical frameworks and case studies*. PhD Thesis. Nothingham University. Nothingham: Manuscript. 2010.

[11] EASY CDA. Interview by author. Digital recording. Belo Horizonte, Brasil. 13 October, 2017.

[12] DJ GIFFONI. Interview by author. Digital recording. Belo Horizonte, Brasil. 6 October, 2017.

[13] DJ A COISA. Interview by author. Digital recording. Belo Horizonte, Brasil. 2 November, 2017.

[14] PDR. Interview by author. Digital recording. Belo Horizonte, Brasil. 4 April, 2018.

[15] DJ PRETO C. Interview by author. Digital recording. Belo Horizonte, Brasil. 29 March, 2018.

[16] ENECÊ. Interview by author. Digital recording. Belo Horizonte, Brasil. 10 October, 2017.

[17] CLEBIN QUIRINO. Interview by author. Digital recording. Belo Horizonte, Brasil.11 October, 2017.

Aesthetic judgments and musical awe: linking appraisal models and the BRECVEMA framework

Landon Peck

University of Oxford
landon.peck@music.ox.ac.uk

ABSTRACT

There are striking commonalities between the hypothesized criteria for aesthetic judgments in music (beauty, expression, novelty, emotion, skill, message, style, and the sublime) and the proposed hedonic appraisals (threat, beauty, exceptional ability, virtue and the supernatural causation) theorized for the experience awe [2]. Beauty and skill/ability are listed in both sets, suggesting both models may be related. Furthermore, the sublime is listed as a type of musical aesthetic judgment, while psychological and philosophical literature relating to awe suggest that the sublime is intrinsically related to sates of awe. These similarities may suggest that both models are targeting similar psychological appraisals, and these may best be understood as part of one another. This investigation explores the relationship between the two models advocates for an expansion of the BRECVEMA [1] framework to more accurately represent the aesthetic emotion of awe.

REFERENCES

[1] Juslin, P.N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews,10*(3), 235-266. doi:10.1016/j.plrev.2013.05.008

[2] Keltner, D., & Haidt, J. (2003). Approaching awe, a moral, spiritual, and aesthetic emotion. *Cognition & Emotion,17*(2),297-314. doi: 10.1080/02699930302297