Musical Explorations Through Spaces

Guzmán Calzada Llorente

Master's thesis

Institute of Sonology

The Hague

May 2020

Abstract

The artificial application of the spatial qualities of a space to an audio signal constitutes a usual procedure within the audio production paradigm, which is also applied to model acoustic situations, distinguishing between sound source and spatial qualities. Within the project, I explore the boundaries of the concepts of spatial qualities and sound source, and such exploration guides my artistic practice. The process of adding the sound characteristics of a space to an audio signal is usually made by means of convolution. Using convolution for this purpose requires an impulse response, which provides information about how the space reacts to every frequency. However, an impulse response can relate to any kind of system—which means any defined situation where we can know, or wish to know, the impulse response. This research explores audio files as speculative 'spaces' and investigates on how 'impulse responses' can be extracted from them. Throughout the project, I also developed a compositional strategy which consists of mechanism-like behaviours encompassed within a modular structure, which emerges from musical automata examples. Finally, concepts like convolution, deconvolution, impulse response, filter, musical automata and subvocalization are elaborated through my artistic practice and they dialogue within the pieces of the project.

Acknowledgements

Firstly, I am deeply grateful to Richard Barrett, Raviv Ganchrow and Gabriel Paiuk for their permanent guidance and generous support. I would like to thank Peter Ablinger and Peter Pabon for their advice and strong guidance in very important parts of this project. A special thank to Kees Tazelaar for his support and dedication to the institute.

Many thanks to my teachers Justin Bennett, Graham Flett, Bjarni Gunnarsson, Paul Jeukendrup, Joel Ryan and Johan Van Kreij.

Thanks to Marcelo Rilla for his advice and support.

I would also like to extend my sincere thanks to Lucie Nezri and Alexandru Pantazica for being part of this research project.

I would like to express my appreciation to my fellow students Tony Guarino, Michael Kraus, Eunji Kim, Toby Kruit, Simone Sacchi, Jad Saliba and Ernest Vilsons.

Thanks a lot to my close friends and family for their continuous support.
Lastly, thanks to my parents Inés and Daniel for their unconditional support.

In memoriam Inés Llorente.

CONTENTS

INTRODUCTION

The current thesis focuses on a series of pieces I have been developing during the period of my master studies. As the process has been continuous I am going to present the pieces in chronological order. The compositional ideas and their dialogue with a number of procedures constitute the core of the research.

The chapters work as ramifications of the main compositional ideas and they are ordered according to the order of the pieces that they include. This is intended to underline the connections and continuity between the pieces. To better emphasize the research, I start the chapters by focusing on a number of fundamental concepts that give a panorama of the subject, and then narrow the discussion towards the pieces I have developed. The title of the chapters refers to these fundamental concepts I will elaborate through my artistic practice. Additionally, the pieces are not only a consequence of the discussion within the chapter that they belong to, but they are gathering points for the perspectives expressed in every chapter.

One of the concepts under scrutiny is digital reverberation. In particular, I will discuss the role that the spatial qualities of a space play within it. Digital reverberation within the audio processing realm can be thought of as an after-effect, a way of adding a sense of spatiality to any audio file. In a typical digital audio workstation situation, audio files pass through various different effects like reverberation, delay, filters, etc. However, that does not mean the audio file—particularly a recording—does not already contain the spatial qualities of a room or an object depending on how it was recorded. A modular way of thinking—consisting of an *audio file+spatial qualities*—is frequently applied in order to model acoustic situations and to remove undesired physical reverberation by means of deconvolution. For this purpose, the paradigm *audio file+spatial qualities* is understood more generally as *sound source+spatial conditions*. I will, in contrast, discuss the notion of sound source and spatial qualities as non-divisible parts in a sound file or acoustic situation, proposing that the sound source could be approached as a set of spatial qualities as well.

Convolution is one of the main tools for the experimentations I have carried out in this research. It is usually implemented to add the spatial characteristics of a given room to an original signal, as in the case of convolution reverberation. However, my main interest is to expand the concept of space within this context, proposing a compositional strategy based on working with sound files instead of spaces for convolution-reverberation purposes. I speculate with the possibility of extracting the 'spatial qualities' of an audio file instead of the spatial qualities of a room or an object. Therefore, I will investigate strategies to extract the 'spatial qualities' of an audio file.

The other compositional strategy I will explore is applying a modular structure to the pieces. The pieces start always from a given audio file which is to undergo a process. These audio files do not appear in their original form within the pieces. Instead, there are live electronics which attempt to imitate the audio files. These synthetic sounds arise through a long chain of modules, which are either filters that shape synthetic sound in order to imitate the audio file, or convolution reverberations that add the spatial qualities of a certain space.

Pieces are structured following a mechanism-like behaviour comprising a modular structure whereby each module carries out a single task within the whole. My interest is in relating the mechanism-like behaviour of the composition with musical automata. Namely, I explore the modular typology in terms of how the mechanism becomes—or can become—audible in the outcoming sound. Automata in the form of musical instruments express an aesthetic fingerprint, encompassing the imperfections of its operating mechanism, which I will explore in my own pieces. Furthermore, some pieces feature live performers, who are also approached as being modules of the mechanism of the composition.

Audio file - Audio source

The pieces discussed in this thesis are always based on audio files. Such audio files will be used as an input for the modular chain I have described. Moreover, the current project involves audio files which might be recordings of pieces with disparate performance styles, historical relevance, recording techniques and audio quality, as well as voice recordings and synthetic sound materials created by myself.

My compositional interest resides in tradeoffs between a new sonic entity—the pieces contained in this research project—and the initial audio files. The main goal for the pieces is the possibility of audibly recognizing elements from the initial audio files still being very different. I believe this research is a creative way of composing pieces and an exciting way of discovering hidden aspects of those initial audio files. There are underlying structuring strategies and classifications of materials that may be used to organize and condition the way a piece of music is heard and analyzed within its context. Consequently, once a piece is taken out of its context and put together with another material in this electroacoustic music context, it can be heard under new listening paradigms. This is to say that pieces from the last 600 years from disparate cultures here are brought together in the pieces involved in this research project, which focuses on sound and discusses the boundaries between sound source and spatial qualities. Furthermore, they pass through a modular mechanism as if they are played by a musical automata.

Although most pieces start from an initial audio file, two pieces' starting point is actually silent speech, which in fact consists of 'imagined' sound. These pieces also have audio files—voice recordings of the performers—which are the starting point for the modular process. However, these voice recordings allude to the inner speech of the performers. Therefore, as the term audio file does not itself encompass inner speech, I will be also using the expression *audio source* as a more general concept.

Finally, some pieces in this project deal with two sources rather than one. Firstly, there is an audio source as I have just explained. Additionally, some pieces involve a process of live synthesis that imitates the sound source. Consequently, there is a secondary input, providing a synthetic sound that will be shaped according to the sound source. Such a secondary input might or might not be a rich audio source with an evenly distributed energy within the spectrum. It might not be possible for the spectral characteristics of the first audio source to be adequately applied, because of the characteristics of the second one. Therefore, a very recognizable spectral profile in the second input might remain, and the output would involve a mix between both.

Finally, the second audio source could also be changing over time, which introduces time as a new determinant parameter.

CONVOLUTION

> Convolution is a mathematical way of combining two signals to form a third signal. It is the single most important technique in Digital Signal Processing. Using the strategy of impulse decomposition, systems are described by a signal called the impulse response. Convolution is important because it relates the three signals of interest: the input signal, the output signal, and the impulse response. [1]

Convolution is central to this thesis, as applied both in the domain of frequency and in that of time. My research into convolution is mainly related to questions concerning the role and perception of the spatial in sound as a core aspect of my compositional strategies. Convolution can be used to implement digital reverberation: its standardized use in the domain of reverberation 'effects' allows the application of spatial qualities of a specific location or an object to an audio file. However, my main interest is to expand on the concept of space or object that could take part in this process. Furthermore, I will discuss in this chapter the indivisibility of sound source and its spatial qualities (considering a recording or a physical situation).

According to the Digital Signal Processing definition of convolution, it is a process that combines two signals. There is an input signal—any audio source—which is convolved by an impulse response. This last one features the spatial qualities of the room or object that is wanted to be added to an audio file. The impulse response is the outcoming signal of a system when the input is an impulse.[2] It is a guide to any

---

[1] Steven W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing* (San Diego, California: California Technical Publishing, 1999), p. 107.
[2] *ibid.,* p. 108.

system, describing how it reacts to every frequency. Therefore, in convolution, the impulse response is applied to every sample of the incoming signal in the same way.

Although in the context in which we consider reverberation, we usually refer to rooms, venues or even objects as spatial qualities that are added to our sound files, within audio engineering, the term *system* refers to any defined situation where we can know, or wish to know, the impulse response. This terminology allows the speculation about other systems besides rooms or objects that can be projecting their spatial qualities to an audio file. This led me to experiment with audio files as systems, as it will be discussed in the chapter called *Timelessness.*

Impulse response measurements usually take place in concert halls, which allow the recording of the venue's reverberation qualities. Such reverberation qualities involve reverberation time—defined as the time it takes for sound to decay by 60dB—and how every frequency is boosted or attenuated within it. The length of an impulse response used for convolution purposes is usually between 1 and 2 seconds for concert halls.

Apart from working with this kind of impulse response, I deal also with much longer ones. The purpose is to investigate their time dimension as a crucial aspect for the compositions. In particular, their spectrum may change critically over long periods of time, which does not usually happen in the impulse responses used for artificial reverberation.

Convolution is the basis of many signal processing techniques. "For example: Digital filters are created by designing an appropriate impulse response. Enemy aircraft are detected with radar by analyzing a measured impulse response"[3]. Many audio processes can be carried out by means of convolution, and these usually focus on the sound characteristics of the combined signals resulting from the process. The present research is concerned with the referential or poetic dimension of the signals being combined as much as with their sonic qualities.

In order to introduce the investigations of my own which I will be discussing throughout the thesis, I will refer to the earliest investigations on deconvolution. Deconvolution is the inverse process of convolution. In convolution reverberation, an input signal is convolved by an impulse response, and the outcoming sound consists of the input together with the spatial characteristics of the impulse response. Instead, deconvolution works by removing the impulse response from the outcoming sound, 'getting the input signal back'. One of the first experiments on deconvolution, carried out by Thomas G. Stockham, Thomas M. Cannon, and Robert B. Ingebretsen[4], focuses on restoration of old recordings by the famous opera singer Enrico Caruso (1873-1921). The singer had recorded *Vesti la Giubba*—a tenor aria from *Pagliacci* (1892), by Ruggiero Leoncavallo (1857-1919)—with a long tubular horn, according to the recording techniques then available. This tubular horn was consequently filtering the voice, enhancing the frequencies corresponding to its resonance modes.

---

[3] Steven W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing* (San Diego, California: California Technical Publishing, 1999), p. 123.
[4] T. Stockham, T. Cannon, and R. Ingebretsen, "Blind Deconvolution Through Digital Signal Processing," proc. *IEEE* vol. 63, no. 4 (April, 1975): 678-692.

One of the premises of this experiment on deconvolution is the modeling of physical situations under audio processing paradigms. In other words, Caruso's recordings are approached as consisting of Caruso's voice plus the tubular horn spatiality. This system resembles any audio processing software or hardware, where reverberation can be added to an incoming sound, whatever it is. The spatial qualities of the long tubular horn into which Caruso was singing—with its strong resonant frequencies—can be added to a sound file in the same way as the spatial qualities of a concert hall can be. For this purpose, an impulse response measurement should be done in the long tubular horn. Then, any audio file can be convolved by the impulse response of the long tubular horn. Since the strong resonances of the tube upon Caruso's voice can be modeled as a result of convolution, deconvolution was applied in order to remove this undesirable "convolution".[5]

[5] Steven W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing* (San Diego, California: California Technical Publishing, 1999), p. 300.

"In terms of functions, the singing waveform s(t) was convolved with the impulse response of the recording mechanism h(t) to produce the recorded waveform v(t). Except for surface noise, it is this latter waveform which is available to us today by playing an old recording.". [6]

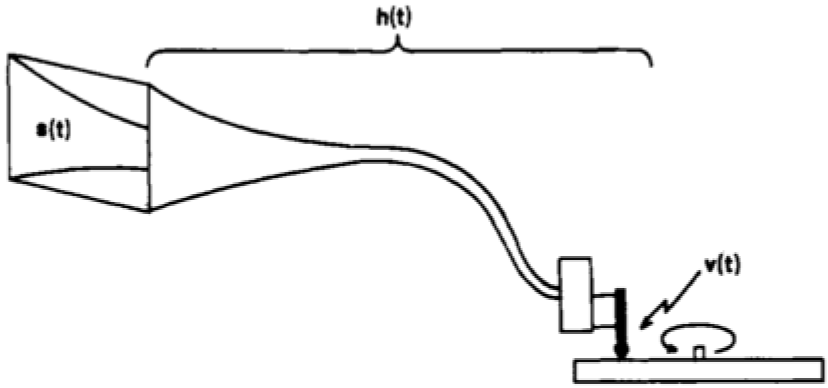


Fig. 1. A recording setup typical of those before the mid 1920's.

Figure 1 - Convolution diagram extracted from *Blind Deconvolution Through Digital Signal Processing.*[7]

The former experiment on deconvolution reveals what the researchers regarded as Caruso's actual voice. Restoration—the act of returning something to its former or original condition, site, owner, etc—means to recover Caruso's voice excluding the sound qualities added by the tubular horn. However, I would take issue with this static notion of Caruso's voice. Certainly, it seems that what we regard as the actual voice is the air vibrations happening as close as possible to the singer, but, for instance, not those inside his mouth or vibrations in another medium. Moreover, the aforementioned static notion does not account for the embodied spatial qualities within the voice itself. In fact, the human voice can be modelled as the outcoming sound of the vocal cords coming through the cavities of the head. Therefore, just as

[6] T. Stockham, T. Cannon, and R. Ingebretsen, "Blind Deconvolution Through Digital Signal Processing," proc. *IEEE* vol. 63, no. 4 (April, 1975): 678-692.
[7] *ibid.*

the whole recording was divided up into Caruso's voice and a long tubular horn, Caruso's voice itself could also potentially be divided up into several spatial qualities.

Thus, the recovered voice of Caruso corresponds to how the researchers would have recorded him—air vibrations as close as possible to the singer, involving a certain microphone position and microphone response. Such conditions seem to be related to the current paradigm of recording as far as the researchers were concerned. Therefore, a particular recording technique is understood as constituting the original voice, which can be achieved by reconstructing the desired type of recording. Deconvolution is presented as a tool capable of bringing back Caruso's voice—more recent technologies seem to be able to get closer to past events.[8]

Therefore, if desired, another tube, or any other moldable space or object, could be added to or removed from the recording through convolution or deconvolution. This opens up many important questions and compositional possibilities for my research. In fact, any kind of audio file could be manipulated by adding and removing the spatial qualities of objects and spaces of any kind, even if this is hard to imagine physically. Since the only requirement is to measure the impulse response of the chosen space or object, the pieces of this research speculates with the spatial qualities of sound files thought as 'spaces'. This is because I propose creative ways of designing 'impulse responses' of sound files, trying to achieve ones that can resemble the main characteristics of the audio in a few seconds. This framework also allows us to think further about other non-physical sites, and how these could be recorded/measured

[8] The researchers propose this solution, but they can not develop it because they do not have the long tubular horn into which Caruso had recorded. Therefore, they can not measure an impulse response. Finally, they use blind deconvolution—a procedure that can model how the tubular horn would have been by comparing Caruso's recording with later recordings by Jussi Bjoerling of *Vesti la Giubba*.

and involved in the aforementioned processes. To sum up, I mentioned this experiment on Caruso's voice because it works with convolution with specific reference to its application for removing an object's spatial qualities. Furthermore, it opens up the question of what is a sound source and what is reverberation—or the spatial qualities of a defined system—, which is a starting point for the current research.

With regard to resonances and convolution, two matters should be clarified. Firstly, the fact that any physical structure can be approached in terms of its frequency response permits the design of filters based on any object or space. However, convolution with impulse responses can be used in many other ways just for filtering a source without any relation with a physical structure, which allows the developments of experiments like the 'impulse response' of a sound file, which does not come from a physical structure, but its sound qualities are added to another sound files as the ones of a venue.

The compositional ideas I will be developing throughout the thesis emerge from presenting in the form of pieces my own 'impulse responses' of speculative 'spaces' as well as convolving them with other audio files. Moreover, some pieces involve a physical space where the pieces or installations take place. The audio files and the physical spaces which I will be using for every composition are approached in two ways. On the one hand, they interact with each other through their sound qualities by means of convolution, filtering, impulse response measurement, etc. On the other hand, they always relate to each other in a poetic dimension of the composition, sketched out by the history of the venue and the context of the audio files—the composer, the instruments, the text, etc. This is why I have previously emphasised the poetic or referential qualities of my approach, and it will be properly explained for each piece of this project.

FILTER

The purpose of this chapter is to explore the notion of a filter in a different context from the usual one. As stated in the previous chapter, any object or space can project its spatial qualities over an audio file by means of convolution. For this purpose, objects or spaces are modelled as custom filters conforming to their spatial characteristics. Therefore, convolution can also be approached in terms of its filtering nature. Filtering has a broader meaning than the ones which can be implemented with convolution or deconvolution: namely, the performers themselves are conceived as filters, as will be discussed in the chapter *Subvocalization*.

The pieces involved in this research feature modular processes, and the majority of these modules are filters. The pieces always involve a sound source coming through a series of successive modules. My interest resides in simultaneously showing different links in the chain of filters, rather than contrasting the source and the signal that results from the entire process. Additionally, while the source is being played, the whole modular process is ongoing. Therefore, the music is structured by interactions between different stages of the filtering process in real time, rather than unfolding a linear sequence of iterations.

Gandini and the pre-existent material

In that regard, *Eusebius* (1984) by Gerardo Gandini (1936-2013) provides an interesting example, where the composer works with one of Robert Schumann's (1810-1856) piano pieces. *Eusebius* consists of 4 Nocturnes for piano, with the same duration and metric structure as the *Davidsbündlertänze* op.6 no. 14 (1837) by Schumann. In the whole piece, Gandini used each note from Schumann's piece only once (the note's attack is exactly in the same place of the metric structure as in the *Davidsbündlertänze*). Consequently, notes are distributed between the 4 Nocturnes such that if these were played at the same time Schumann's piece could be heard. To sum up, *Eusebius* consists of four pieces which seems like *Davidsbündlertänze* op.6 no. 14 but with filtered-out notes.

Regarding the interest of the composer in using the kind of compositional strategy that works with other music as pre-existent material, Gandini wrote the following in the programme notes of the premiere: "Schumann's figure, his music and his myths, his different identities became an obsessive presence to me in 1984. And I tried to exorcize his spirit through several pieces written in this year."[9]

Moreover, Graciela Paraskevaidis—a musicologist who has conducted research into Gandini's work—reveals a layer of interaction between what Gandini composed and the original source:

> To escape from the stereotypes of new music, Gandini tries mainly to produce a sound structure in which the presence, confrontation, inclusion, processing and finally the relationship with the great European tradition become its core and crucial point. Whether this tradition is called Dufay, Lasso, Frescobaldi, Rameau, Bach, Scarlatti, Mozart. Schubert, Schumann or Schoenberg, it is always interwoven in Gandini's music as a looked-for relationship, as a conscious elaboration, as an obstinate incorporation, as a wished for confrontation, as a presence not to be driven out. It is the strange effect of musical processes of a kept-alive past that act deeply in the subconscious and that, through its irresistible

[9] Graciela Paraskevaidis, "The Own and the Other. The Argentinian Composer Gerardo Gandini," *World New Music Magazine*, no. 3 (September, 1993), accessed May 29, 2020, http://www.gp-magma.net/pdf/txt_i/Gandini-WNMM.pdf

> power of attraction and fascination invoke Arnold, Eusebius and Florestan, Franz and Wolfgang as fatherly models.[10]

Eusebius and Florestan were alter-egos created by Robert Schumann. In fact, he essentially divided his artistic persona into these two different characters. While Florestan—a character from Ludwig van Beethoven's (1770-1827) opera *Fidelio* (1805)—represents a masculine and fearless person, Eusebius refers to Saint Eusebius, a Christian martyr.[11]

Eusebius is an interesting example for this research due to the way the composer deals with the pre-existent material. Such pre-existent material is the same as what I have called audio sources for my own research. Both for Gandini and for my own compositional purposes, the audio source—which usually belongs to composers from a different context from our own—is the starting material for the piece, and this is remarked in the title and the score of the piece. In Gandini's *Eusebius*, the first part of the score explains how the piece has been composed upon Schumann's *Davidsbündlertänze*. Moreover, the title *Eusebius* refers back to Schumann, and even intends to reveal something beyond *Davidsbündlertänze*—a Schumann alter-ego.

The piece itself is concerned with the pre-existent material, its composer and its context; this pre-existent material is the core of the piece, rather than constituting a simple anecdote about the provenance of the material. Filtering out notes is intended by Gandini much more as a technique for enhancing aspects of the piece and the

---

[10] Graciela Paraskevaidis, "The Own and the Other. The Argentinian Composer Gerardo Gandini," *World New Music Magazine*, no. 3 (September, 1993), accessed May 29, 2020, http://www.gp-magma.net/pdf/txt_i/Gandini-WNMM.pdf
[11] *ibid.*

composer, rather than for making it unrecognizable. This is exactly the way I approach the pre-existent material within my own research. Additionally, the obsessive relationship between Gandini and Schumann's work also reflects my connection with the material I used for the pieces of this research, a connection which has been present in my artistic practice since long before this project.

Spatial speculations on *Eusebius*

The first time I listened to Gandini's piece I had an intuition about the way notes were split up between the 4 nocturnes. The notes for each nocturne seemed to have been chosen according to the resonance frequencies of four different venues. It resembled Schumann's piece being performed within four different rooms, where the resonance qualities of every room would act to enhance or attenuate different frequencies. Therefore, such notes as remain in each nocturne would be based on the boosted frequencies and the rest would be removed.

This speculation is based on the *piano* dynamics of Gandini's piece, and the amount of silence between sounds, which allows us to hear the soft attack and release of each note. In order to delve into my intuition, I myself made some try-outs which consisted of convolving *Davidsbündlertänze* by the impulse response of a room several times.[12] The sound file obtained bears similarities to Gandini's nocturne. As a result, the recording becomes much softer, while occasionally there are loud sounds with soft attack and release. These sounds feature an extreme narrow band spectrum and their

[12] Schumann's piece recording has been convolved with an impulse response of a room which has been previously convolved with itself 4 times. This is to say that the impulse response was convolved with itself and the outcoming signal was convolved with itself and so on. As it is being done, a few frequencies are extremely boosted and the rest is almost removed. This audio file can be found in the attached audio files of the thesis.

frequencies correspond to the main resonances of the room. Gandini's piece might be considered as unfolding Schumann's piece in four different and complementary spaces—each one of which would reinforce some notes and remove others.

According to the convolution model as in the example of Caruso's voice, I proposed that what initially seemed to be the sound source—Caruso's voice, separated from the sound qualities added by the long tubular horn—could be thought as a set of spatial qualities as well. Certainly, one of the main points of this research is to propose the embeddedness of spatial qualities within audio sources. My spatial speculation on *Eusebius* is an expansion of the previous discussion, and just depicts a creative proposal of how this matter would constitute a compositional strategy or even a tool for musical analysis.

Moreover—following what Friedrich Kittler wrote on "Goethe speaks into the phonograph" by Salomo Friedlaender (1916)[13]—Kittler has speculative thoughts about recording indoor spaces and amplifying those loud enough to listen to what was said in the past. In particular, Kittler writes about Goethe's voice, and how it could be recovered by means of extremely loud amplification and a tuned receiver—the latter for distinguishing Goethe's voice from the rest of the sound that ever occurred there. As a matter of fact—even if what he proposes would be achievable—the sound would have already been reflected back and forth for many years within the enclosed space, and we would probably hear much more of the resonances of the location than the original sound. The result would not be very different from the final iteration of *I am sitting in a room* (1969) by Alvin Lucier (born 1931). This piece consists of a

[13] Friedrich Kittler, *Gramophone, Film, Typewriter* (Stanford, California: Stanford University Press, 1999), p. 59.

recording of the composer reading a text played back into a room repeatedly. The first time consists only of his voice reading the text within the room, and the piece is structured by showing the same recording being recorded and re-introduced into the room again and again, until the resonances of the room get increasingly boosted and the voice of the composer can not be distinguished anymore.

According to Kittler, Goethe's voice would be graspable by means of extremely loud amplification and a tuned receiver. As in the Caruso example, there is an implied static notion of Goethe's voice. It seems to be Goethe's voice vibrating in the air, from a close distance, as if he would have spoken into the phonograph. Nevertheless, the outcoming sound from this extremely loud amplifier would be much more about the indoor space than about Goethe's voice, as can be inferred by listening to the last iteration of *I am sitting in a room*. In short, the attempt to separate the supposed original sound source—Goethe's voice in this case—from its spatial characteristics seems to be unachievable. Instead, I would like to propose the creative interest of conceiving any source as a set of spatialities that could be removed and added for compositional purposes. In that manner, *Eusebius* becomes a tool for modelling Schumann's piece, and implies a compositional strategy arrived at by borrowing notions of convolution from the composition practice of western music.

Nadie encendía las lámparas[14] *(2017)*

The following is a description of an early work of mine which is a site specific installation and involves the main resonances of two rooms—the living and dining

[14] The name of the piece (*Nadie encendía las lámparas*) is extracted from the title of a short story by the Uruguayan writer Felisberto Hernández.

room of an old house in Montevideo, Uruguay. The piece incorporates an audio programming patch which analyzes a musical recording—the audio source in this case—and plays sine-wave oscillators under certain circumstances. The patch analyzes in real time the fundamental frequency of a certain audio source—a recording of a piece of music in this case—and uses this signal to trigger several sine-wave oscillators when the fundamental frequency of the audio source coincides with the resonant frequencies of the rooms[15]. The sine-wave oscillators' frequencies are extracted from the main resonances of the room. Each oscillator uses an envelope with a very soft attack and release, and sounds for around three seconds. The analyzed audio file is not played in the room, but only works as a signal for triggering sine-wave oscillators within the patch.

The compositional algorithm I developed was first implemented in a piece named *Nadie encendía las lámparas (Nobody lit the lamps)* (2017). It took place in *Espacio Cultural Alberto Soriano*, an old house in Montevideo (Uruguay) which was home to the composer Alberto Soriano Thebas (1915-1981). The piece uses a four-channel sound system, with two speakers placed in the living room and the other two in the dining room of the house. The audience was in the interior courtyard of the house, which is directly connected to both living and dining rooms (Figure 2). Therefore, the electroacoustic sound coming from both rooms was audible in the courtyard.

[15] The algorithm described works independently either in the living room and the dinning room.

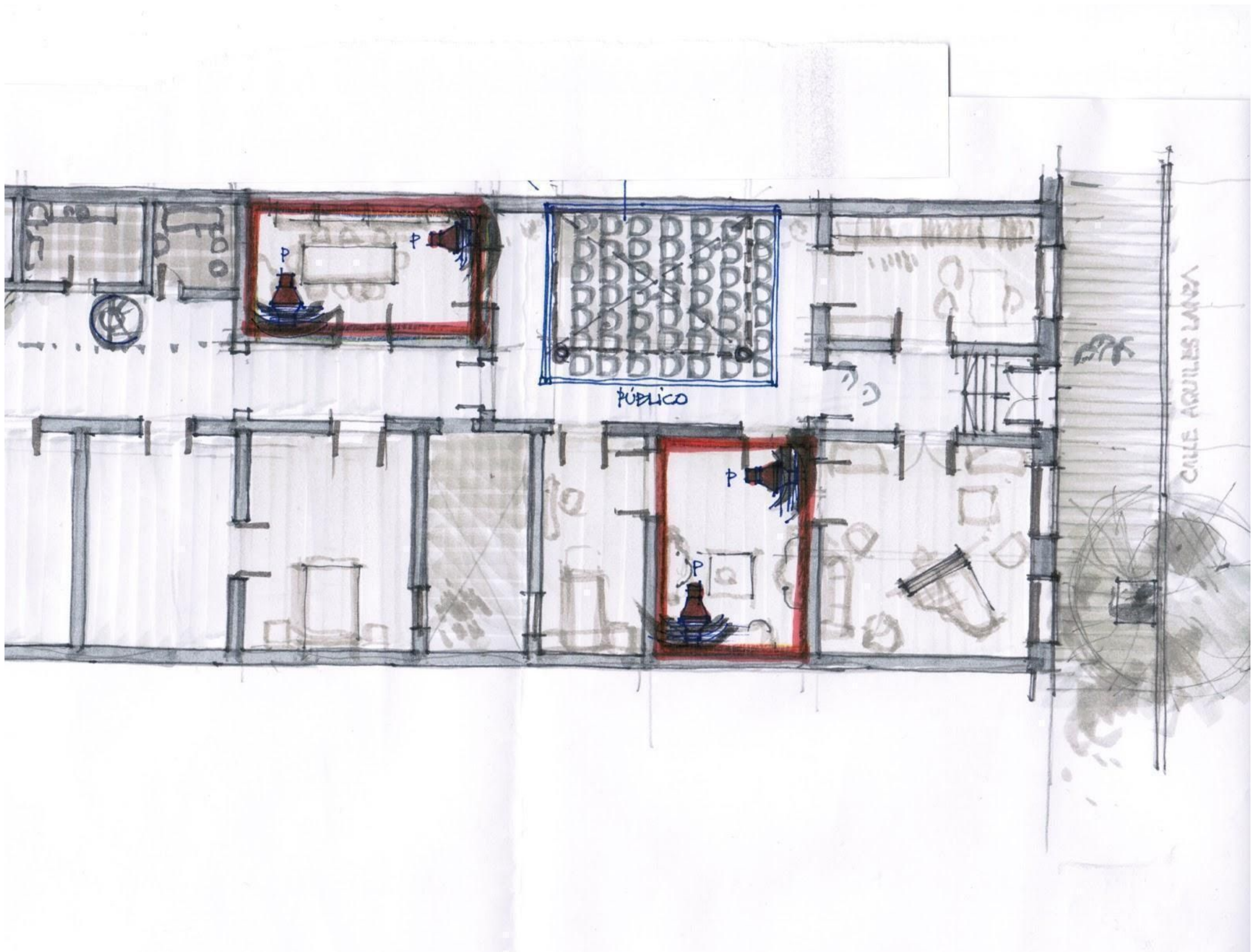


Figure 2 - Floor plan for *Nadie encendía las lámparas* outlining the dining, living room and audience space.

The audience is not inside the rooms where the sound is being played, but is seated in the courtyard, which is between the rooms, while many soft sounds emerge from them. However, there are no visible loudspeakers and no performers in the rooms. The music invites a sense of mystery which I hope causes the audience to speculate about the sounds resonating throughout the living and dining rooms which they are not able to look inside. Many small objects like crafts and pottery as well as some glass windows vibrated sympathetically with the sine-waves.

The audio files used for triggering the sine-wave oscillators, as previously explained, were two guitar pieces by Alberto Soriano Thebas, the composer who lived in the house for two decades (the 1960s and 1970s). The material chosen for triggering the acoustic aspects of the rooms is connected with the history of the house and the people who lived in it, by linking the material to a composer who conceived and played his music there. The resonances of the rooms of the house were excited by his music when he lived there, and my piece imitates how these resonances had sounded, but this time without Soriano's pieces being played in the house. The inner connection between Soriano's music and the acoustic aspects of the house is therefore, to some respect, materialized in my work, which shadows and echoes that music. Thereby, the piece put together audio files and a space and they interact through their acoustic properties as well as through their contexts—the audio files are recordings of pieces which were composed by a composer who lived in the house.

Nadie encendía las lámparas #2 **(2019)**

Nadie encendía las lámparas #2 (2019) is an installation composed for the Symposium *Transformations of the Audible* at WEST, an art space in the former U.S. Embassy in The Hague (The Netherlands). Regarding the electronics, I adjusted the same system I described before to the corresponding resonances of the venue. The installation took place in a room where I set up two loudspeakers, a storage room of the art space which I used because I was interested in working with the acoustic properties of some objects I found there as well as the room itself. This expanded upon the role of the objects in *Nadie encendía las lámparas*, as I decided to play back the oscillators through the objects themselves, thereby also involving the resonances of each object in the installation.

My focus has turned here towards the interaction between these objects' resonances and the playback, exactly in the same manner that the sound file within the patch interacts with the room's resonances. In other words, a sound file (playback within the patch) is filtered by the room, and the outgoing signal of this process is filtered by the objects. The sound file—which could be seen as the first stage in the modular chain—is never played; but the sine-waves and the objects attached to exciters can be heard simultaneously.

The objects inhabiting the room vibrate sympathetically with specific sine-waves, depending primarily on their frequency. On the one hand, the melody resulting from the sequence of sine-waves was split up between the objects, as if the hocket technique had been applied to them. On the other hand, such objects provide a wide variety of timbral qualities, which work as an orchestration of the sine-wave melody.

Nadie encendía las lámparas #2 involved a process of experimentation with all the available objects in order to find interactions between them and the room's resonant frequencies. A visual or sculptural dimension appears here in my work for the first time. Not only did I have to make decisions from a visual point of view, but I had also to decide on the location of the objects according to how the playback through them interacts with the loudspeakers. In fact, the 3 objects—attached to audio exciters—work as loudspeakers in the piece, which, added to the other two standard loudspeakers add to 5 loudspeakers in the room unfolding a spatialization of the sine-wave playback. Each sine-wave has a random amplitude assigned for each one of the 5 channels, and the objects themselves critically change the sound quality because of their strong resonance modes.

One last element of the piece that should be mentioned is that the audio source—the audio file ongoing within the patch—I worked with consisted of several Uruguayan folk dances played on the bandoneon, recorded by René Marino Rivero. Apart from being used to trigger the system, as previously explained, these recordings were processed and also played in the installation. While two loudspeakers and three exciters played sine-waves according to the coincidences between Rivero's recordings and the room's resonances, one exciter attached to the smallest object (a small pot) played the bandoneon recording. This recording, however, had previously been convolved with a boosted impulse response of the room[16], resulting in a version of the recording which was highly filtered according to the room's resonant frequencies. Looking back at the modular organization of the piece—a sound file (playback within the patch) is filtered by the room, and the outgoing signal of this process is filtered by the objects—the installation comprises a simultaneous playback of its sound source—highly filtered—as well as the room's resonance filtration and the objects' filtration.

Rivero's solo bandoneon recordings involve a very rich timbre in terms of spectrum—although a homogeneous one—and an unstable tempo. The rich timbral aspect of the recording makes the highly filtered version very different from the original one, since much of the sound has been filtered out. Moreover, the narrow

[16] This is achieved by convolving the impulse response of the room with itself.

band spectrum around certain frequencies—a consequence of the high degree of filtration—made this playback quite similar in sound to the sine-waves. The unstable tempo helped in avoiding regularity, making it easier to hear the recording and the sine-waves—which are also irregular in rhythm—as integrated with one another. The aforementioned characteristics of the recording allowed me to push it away from its folk music nature and bring it closer to the electronic music realm just by using a filter. I intended to create a balance between the recording and the sine-waves in which the audience could only recognize the recording by getting very close to the object where it is being played.

To sum up, these pieces rely on approaching a musical recording as a sequence of possible spaces. Both the room's resonances present in the frequencies of the fixed sine-wave oscillators and the small objects inhabiting the room stand as plausible constitutive spaces within the bandoneon recordings, in the same manner as was explained in the first chapter.

TIMELESSNESS

'Timeless' is the name I have given to the outcome of a process for decoupling any audio file from its time dimension, and is one of the results of my extensive experimentation with convolution and impulse responses. The process has an input which is an audio file and an output which is its 'timeless' version. In the current chapter I will describe this process and explain how it has been applied in two pieces. Furthermore, I will develop one of the main ideas of this research, which is to expand the concept of space in the context of convolution in a creative way. In particular, I propose that audio files can work as 'spaces'. For this reason, I developed their own 'impulse responses' so they can add their 'spatial qualities' to any other audio file by applying convolution.

In the previous chapter, the piece *Nadie encendía las lámparas #2* was described in terms of how it deals with the filtering of an audio source (Rivero's bandoneon recording). In short, the piece can be described as the interaction of three modules. Firstly, there is an audio file—Rivero's bandoneon recording—running within an audio programming patch. Secondly, it is being analyzed and as a consequence the patch triggers sine-waves, which have fixed frequencies corresponding to the resonances of the room where the installation took place. Finally, the sine-waves are played through two loudspeakers and through exciters, attached to some objects situated in the room.

Nadie encendía las lámparas #2 can also be analyzed by reference to two dimensions—time and spectrum. The time dimension is controlled by Rivero's recording while the spectrum dimension is based on the room's acoustic qualities. As previously mentioned, the frequencies of the sine-waves—which constitutes the playback that can be heard in the installation—are defined by the main resonances of the room. Therefore, the spectral dimension of the piece is based on the main resonances of the room.

An audio file is running through the patch—not being played audibly—and it is analyzed in real time in order to detect its fundamental frequencies. The patch uses the audio file to trigger several sine-wave oscillators when the fundamental frequency of the audio file coincides with any of the fixed frequencies of the sine-waves. Therefore, the time dimension of the piece is controlled by the audio file, which works as the trigger for the playback. Although not every note from the audio file triggers sine-waves, all of them were triggered by a note in the audio file. As a result, the audio file controls the time dimension of the piece—the duration and the distribution of the sound events along the course of the piece (sine-waves) are based on some of the notes in the audio file.

I propose now a compositional strategy that emerges from swapping the role of the audio file and the room as described for *Nadie encendía las lámparas #2*. The main proposal embodied in the next piece discussed in this chapter is the following: the time dimension is based on the acoustic qualities of a room while the spectrum dimension relies on the audio file.

The time dimension of the piece will be based on a room. For these artistic purposes, I propose that the time dimension of a room can be 'represented' by its impulse response. I decided that the piece's duration will be the same as the duration of the impulse response. Moreover, the impulse response provides the dynamic envelope for the piece, with a peak of energy at the beginning and a decay until the end.

The audio file should provide the spectral dimension of the piece, taking the role of the space in *Nadie encendía las lámparas #2*. One of the premises of the project is that the time dimension of a space relies on its impulse response. Generally, the spectral qualities of a space in the impulse response appear simultaneously in the first instant of the impulse and do not extend very much in time (the reverberation time of concert halls is usually between 1 and 3 seconds). This means that the impulse response does not critically change over time but is more a matter of the dissipation of energy of every frequency. Therefore, the audio file should be condensed into a few seconds, preserving its main spectral qualities. In other words, the main frequencies of the whole audio file should appear in the first instant of the piece and afterwards die out gradually, as if they were dissipating into an enclosed space.

Impulse response

The impulse response constitutes a central concept for various of the pieces I composed in the context of my research. However, the inspiration for each piece comes from different creative insights about the impulse response. The reason for developing these approaches has to do with the entity of the impulse response—which could be thought of as a representation of the amount of every for every frequency in a room, or as a sonic object itself.

Although the impulse response is involved in convolution processes, it also has its own materiality and unity, and it can be heard independently of any convolution process. In fact, every impulse response of a room would sound different from every other. For instance, on any digital audio workstation, when a reverb is applied, the impulse response is usually heard convolved with an incoming signal. It is quite easy to try out many different reverbs in terms of sound quality, duration and other parameters related to the physical properties of a venue, if appropriate. However, in this situation, the impulse response remains hidden from the users—it is an audio file which is involved in typical music production situations and is not commonly listened to.

With this in mind, I will discuss three artistic insights into the concept of impulse response, which together explain how I have developed my compositional ideas. Such approaches are not intended to be engineering approaches in any way, but instead are elaborations through my artistic practice. The impulse response can then be thought of as the result of a procedure, as one part of the convolution process or as a sonic object in itself.

Procedural insight

This section focuses on the procedure of obtaining the impulse response of a room, and describes some experiments in which I have creatively tweaked that impulse response measurement procedure. The starting point for this section of the thesis resides in the methodology which was applied to get the impulse response of the room for the installation *Nadie encendía las lámparas #2*. A 10-second sine-wave sweep between 10 and 12.800Hz was played in the room and recorded. Subsequently, the initial sine-wave sweep and the one recorded in the room were convolved, but first the initial one was reversed. As a result, the outcome of this convolution is the impulse response of the room.

Firstly, I was surprised by the fact that convolution is capable of transforming the time dimension of audio material in that way. The outcome of convolving a synthetic sine-wave sweep of 10 seconds and a recorded one is an impulse of 1.3 seconds. This can be explained in the following way: the sine-wave sweep can be thought of as an impulse which passes through a system that provides a delay for each frequency—the delay holds the frequency a bit longer if it is higher. As a result, the outcoming signal will be one with a constant increasing frequency over time (sine-wave sweep). Once it is convolved by a reversed one (which can be seen in the same way but the delay holding longer the lower frequencies), the overall time delay will be the same for all frequency components, and all components will focus (peak) again at the same moment in time, which constitutes an impulse.[17] Aside from featuring disparate durations (the impulse response and the sine-wave), an impulse has a large concentration of energy and a decay until the end, while a sine-wave sweep has constant amplitude, and in terms of its spectrum is always a sine-wave.

The recorded sine-wave sweep—which covers a large part of the audible frequency spectrum—works by surveying how the room reacts to each frequency that is played. In fact, it works as a guide to the room—frequencies are ordered from the lowest to the highest and each has an amount of energy based on the response of the room. In

[17] Personal correspondence with Peter Pabon, May 18, 2020.

short, the recorded sine-wave sweep already has the same spectral information of an impulse response of the room, but the response of the room for each frequency appears unfolded in time (from the lowest to the highest frequency), instead of simultaneously.

Once the recorded sine-wave sweep is convolved with the reversed synthetic sine-wave sweep, the impulse response emerges. It works as if the recorded sine-wave sweep would be folded back and all the frequencies with their corresponding amount of energy are brought to the same instant in the beginning of the audio. Therefore, the impulse response corresponds to the recorded sine-wave sweep in terms of its spectral qualities, even if the outcome has crucial differences in respect to the duration and dynamic envelope.

The transformation from the sine-wave sweeps to the impulse response by means of convolution—which transforms the time dimension of the audio files but keeps their spectral qualities—has opened up a field of creative ideas in which this procedure can be applied to other audio files besides the sine-wave sweep. Therefore, the outcoming audio would condense the spectral characteristics of an audio file—as it is usually for a room. The proposal is to replace the sine-wave sweep by another audio file. Therefore, it would be convolved with a reversed version of itself. The outcoming audio of this experiment has a static texture, as if the time dimension of the audio had been erased. This static texture is made out of the main frequencies of the audio file, as a sort of average of its spectrum—this is what I named the ‘timeless’ version of an audio file for my artistic purposes.

After the convolution process described above—an audio file convolved with a reversed version of itself—a piece of audio twice as long is obtained. In any convolution process, the duration of the outcoming file equals the sum of the durations of the two incoming audio files. The process can be carried out as many times as desired by convolving the result of the first convolution by a reversed version of itself, and so on[18]. As a result, on every iteration, the audio file doubles in size, and while the main frequencies get increasingly boosted, the others are almost removed.

My goal is to get 'impulse responses' from these speculative 'spaces' which are audio files. The 'timeless' versions of audio files emerge from the impulse response measurement procedure whereby the venue has been replaced by the audio file so to speak. However, typical impulse responses of concert halls are relatively short, while these 'timeless' audios are always twice as long as the initial audio file. In other words, an audio file of two minutes will have a 'timeless' version of four minutes, which would be extremely long in comparison with the duration of the impulse response of a room. Furthermore, the 'timeless' version of an audio file does not feature a dynamic envelope as does the impulse response of a room, as can be seen in Figure 3.

[18] This process of recursive convolution is explained further in section 5 *Timelessness Studies*—detailed procedure.

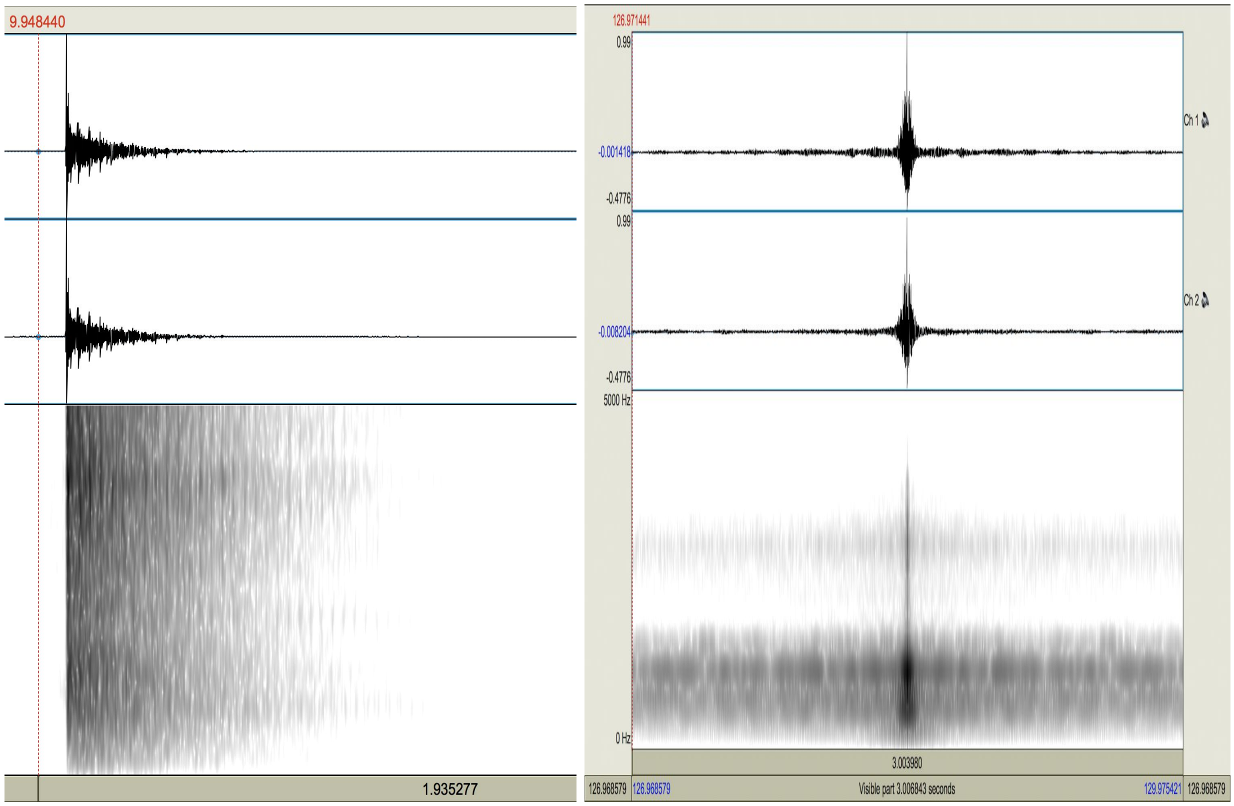
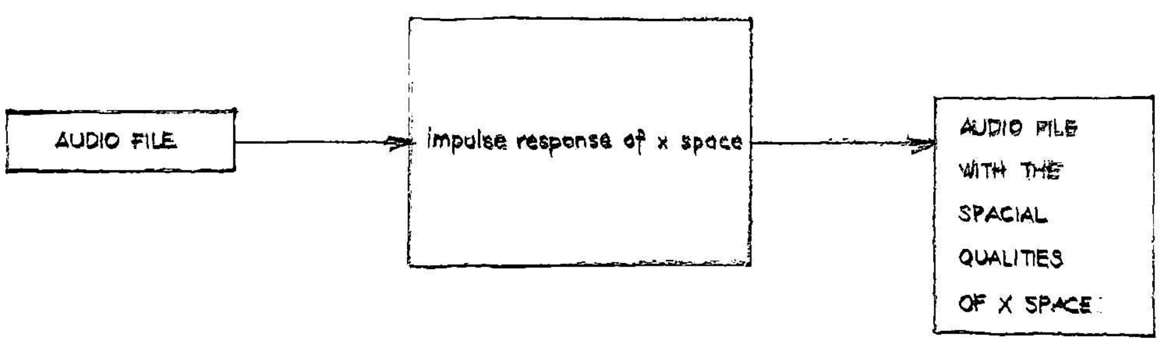


Figure 3 - 'Timeless' version of an audio file (left) and impulse response of the room where *Nadie encendía las lámapras #2* took place (right). Oscillogram (above) and spectrogram (below).

Furthermore, I wanted to convolve this 'impulse response' of an audio file with another incoming audio file in order to add their 'spatial qualities', as an impulse response of a room does. However, I wanted to do so without critically affecting the duration of the incoming audio file. In other words, if an incoming audio file is convolved with this 'timeless version' of four minutes, the outcome would be 4 minutes longer.

As previously mentioned, the 'timeless' version of an audio file has the desired static sound qualities to be convolved as if it were an 'impulse response', but it is too long and does not feature the desired dynamic envelope. For these reasons, I propose two methods to obtain a sound with the desired duration and dynamic envelope typology of an impulse response from these 'timeless' versions. The first solution would be to cut an excerpt of the 'timeless' version. This excerpt will begin from the zero point—within the oscillogram—right before the main energy peak[19]. In this manner, the extracted fragment will feature a peak of energy at its beginning. Afterwards, the rest of the audio excerpt will be shaped in terms of dynamic envelope and duration in the image and likeness of the impulse response of a room[20]. Another procedure to extract an impulse response from the 'timeless' version is to apply the output of an envelope follower to it. This envelope follower should be tracked by an impulse response recorded in a room. For this option, I suggest that it is better to apply the envelope follower to a fragment of the 'timeless' version which features homogeneous amplitude. A homogeneous amplitude will make sure that the envelope will be applied properly and the outcoming sound will have the desired dynamic envelope. The second option is the one I finally used for *5 Timelessness Studies*.

Convolution insight



[19] This is explained in detail in the section of this chapter called *5 Timelessness Studies*—detailed procedure.
[20] This is explained further in the section of this chapter called *5 Timelessness Studies*—detailed procedure.

The second insight I found interesting for my research consists in focusing on the relationship between the three signals which are involved in convolution. This approach comprises experiments on replacing the impulse response for any other audio which is not an impulse response—not even this 'impulse responses' of the audio files that I developed. I was investigating longer audio files than typical impulse responses of rooms (in the order of minutes), and how their duration and sound qualities affect the outcoming signal. This opens up an underlying question: Is there a physical correspondence with any existent or imaginable space?

Convolution between an audio file and a reversed version of itself, which I used for obtaining the aforementioned 'timeless' versions, can be seen as a convolution with long audio files. The experiments I have done involved audio files longer than 4 minutes. For reasons of clarity, I will apply the term *initial audio file* to what I previously simply termed the "audio file", *second audio file* to the one which replaces the impulse response, and *outcoming audio file* to the outcome of the convolution process. In general, since the second audio file is longer than a few seconds, the sense of hearing the initial audio file with the sound qualities of an acoustic space is partially lost. To the extent that the duration of the second audio file becomes similar to that of the initial audio file, both acquire a similar presence in the outcoming audio file.

The dynamic envelope of an impulse response is a consequence of sound behaviour within an enclosed space. Generally, sound is reflected back and forth by the walls or other surfaces while dissipating energy. Energy is dissipated because of the absorption qualities of the surfaces and air friction. In spite of the fact that some frequencies will remain more than others, sound energy can not increase once the sound source has stopped. For instance, another technique for impulse response measurement consists of recording an impulse within a room—after the instant when the impulse is played, the sound energy in the room can only decay until it disappears.

While experimenting with using audio files instead of impulse responses, I wondered if these can resemble the acoustic qualities of acoustic spaces. Definitely, the sense of hearing the initial audio file with the sound qualities of an acoustic space—as in convolution reverberation—is partially lost. I found out it has to do with the duration of the second audio file, but also with the dynamic envelope. This is because a sound file with a dynamic envelope that features any noticeable increase of energy other than at the first instant, behaves like an impossible acoustic space. Additionally, a duration in the order of minutes (which is much longer than the reverberation time for common rooms) also makes it lose verisimilitude with the behaviour of an acoustic space.

However, I became interested in convolutions with long audio durations because through them I developed the 'timeless' versions. Apart from using excerpts of them for the *5 Timelessness Studies* as explained, they will be presented in their long form in one of the pieces discussed in the upcoming chapter. Furthermore, even if convolution with long audio durations can lose verisimilitude with how acoustic spaces behave, I find the results of my investigations as interesting starting points for speculation concerning abstract spaces, namely by approaching polyphony as a kind of abstract space. One melody within a polyphony—which has been composed to be played at the same time as the others—can be thought of as the initial audio file of a convolution process. Therefore, I approached the rest of the melodies as 'spaces' inhabited by the first of them. I made an experiment with an homophonic four-part piece, where a singer performs only one melodic line. The other three parts become present by convolution. This is to say that the performer is singing one part and this is being convolved with these three other parts. The outcome of each convolution is played in different loudspeakers. Additionally, I made another experiment, which included 'impulse responses' of recordings of the other three parts, in order to convolve them with the melody performed by the singer in real-time. The outcome of each convolution is played in different loudspeakers as well.

Impulse response without convolution

I am also interested in treating the impulse response as a sonic entity outside any convolution process. My intention is to bring the impulse response to the core of my compositional thinking as well. This means not only thinking of it as a tool capable of adding spatial qualities to something else, but using it as a composition in itself. This is mainly how *5 Timelessness Studies* were conceived.

Since we assume that impulse responses constitute pieces in themselves, there are two more things to consider. Firstly, they can be composed like any other piece. Previously I explained how the idea of impulse response can be creatively expanded turning audio files into 'impulse responses' of themselves. Moreover, once we obtain the impulse response of a room or an audio file, it can pass through any other digital process—e.g. filtering, compression, etc—in order to achieve the required sound quality.

Secondly, the dynamic envelope of an impulse response becomes yet another material that could structure a composition. This is to say that keeping the dynamic envelope fixed, sounds from any provenance can be mixed together under the proposed dynamic envelope typology. Furthermore, such typology can be stretched as desired, and turned into a structural strategy for a longer piece.

5 Timelessness Studies

The first piece I will talk about is *5 Timelessness Studies* (2019). This is a set of 5 studies which work with recordings of 5 preexistent recordings of pieces and one room. The impulse response of the room provides the time dimension—the duration and dynamic envelope of the pieces is identical to an impulse response recorded in the room. Regarding the spectral dimension, I used excerpts of 'timeless' versions of the five recordings, as described. Therefore, the material for the pieces consists of audio files which basically comprises the principal frequencies of the whole recordings of the pieces. These principal frequencies emerge within a subtle texture containing rapid interactions between them. Such a textural fabric depicts rhythmic patterns that seem to relate to textural characteristics of the pieces.

5 Timelessness Studies can be conceived as very condensed versions of musical pieces. Each piece has been condensed to an approximately 2 second duration, with the dynamic envelope of an impulse response, with a main peak of energy at the beginning and a release of energy until the end. I would like to compare this Studies to *Weiss/Weisslich 22: Haydn, Mozart, Beethoven, Schubert, Bruckner, Mahler* (1996) by Peter

Ablinger (born 1959). This is a 4-minute fixed media piece made out of 45 hours of symphonic music. In particular, all the symphonies by Beethoven, Bruckner, Schubert and Mahler were included, as well as a selection of Haydn and Mozart's symphonies.

> Ablinger had been looking for a method of transforming the linear experience of time into a momentary or perhaps aimless/formless, static one. He and the technicians finally found a way of condensing the stored information in a way that made the linear time-line tip over into a vertical column of condensed information. They looped and folded the music's horizontal time line into a vertical sound column exactly forty seconds wide. Instead of a few hours of Beethoven in a straight line, Ablinger turns it around on top of itself at 90 degrees into a 40 second sound column.[21]

As the quote suggests, *Weiss/Weisslich 22* is also an exploration of forms of condensation of long audio recordings in short periods of time. This necessarily involves thinking of a transformation of the linear experience of time. For *5 Timelessness Studies*, I decided to use the time dimension of a room—which I represented with the impulse response—as an alternative proposal for structuring the experience of time. In fact, the time dimension of a room appears as the most condensed way to experience its spectral material—the impulse response. The impulse, recorded in the room, is the shortest signal that can be conceived in digital audio, and it activates every frequency in the room simultaneously.

Ablinger's starting points are the symphonic pieces, which are overlapped until an extreme accumulation of them. As a result, the sound qualities of Ablinger's piece become similar to noise. Noise is a central concept in Ablinger's work, and in this case

[21] Christian Scheib, "Static's Music - Noise Inquiries," *CalArts*, (1997/1998), accesed May 29, 2020, https://ablinger.mur.at/noise.html

noise is a signal full of information[22]—an important part of the western music heritage—rather than a random audio signal. *5 Timelessness Studies* starts from the impulse, another basic audio signal, which theoretically has an infinite spectrum—it encompasses every frequency. This impulse is metaphorically recorded into a 'space' which is the sound recording turned into a 'space' itself. White noise, as well as an infinite impulse, is a signal that theoretically contains every frequency. However, noise is the result of the process in Ablinger's piece while in my own work the impulse is the starting point.

The 5 pieces of which I used recordings are as follows:

Timelessness study #1: Deo gratias
Deo gratias à 36 (36-part canon) - Johannes Ockeghem (1410-1497)
Huelgas Ensemble - Paul Van Nevel

Timelessness study #2: Vltava
Má vlast - Vltava (The Moldau) (1879) - Bedřich Smetana (1824-1884)
Berlin Philharmonic Orchestra – Herbert von Karajan

Timelessness study #3: La piojosa
La piojosa (1999) - La Sonora cienaguera
Banda Conmoción

Timelessness study #4: Sonata K 333 in Bb
Sonata K 333 in Bb (1783) - Wolfgang Amadeus Mozart (1756-1791)
Javier Toledo - fortepiano

[22] Christian Scheib, "Static's Music - Noise Inquiries," *CalArts*, (1997/1998), accesed May 29, 2020, https://ablinger.mur.at/noise.html

Timelessness study #5: La negrina

"Florida estaba la rosa" (tenor part excerpt from *La negrina*) - Mateo Flecha el viejo (1481–1553)

André Pérez Muiño - tenor

Timelessness study #6: Retirada

Retirada del 1967 (1967) - *Asaltantes Con Patente*

El tablado del tiempo (2009)

These five musical pieces do indeed come from disparate contexts. Audio recordings open up the possibility to bring these disparate examples together and put them through the same digital processes. My main interest is in investigating how they work together under the same electroacoustic music framework, which focuses on sound quality in a different way than in their original contexts.

The *Deo gratias* canon is a piece for 36 voices from the tradition of Franco-Flemish polyphony. This piece in particular involves a large number of parts, and pushes the possibilities of polyphony to the extreme. In fact, when many voices are present, it turns out to be quite difficult to recognize the simultaneous melodic lines. From my point of view, therefore, the piece can be heard in two layers, a thick stratum of vocalic sound and short interventions of consonantic sound. The large number of parts make the melodies unintelligible and they become a dense layer of vocalic sound. Upon this sound layer, the consonantic sounds, which are a few sounds in each melody, come up to the top of the musical texture. It resembles a fixed-media piece conceived from sound quality difference between vowels and consonants instead of the result of a dense contrapuntal fabric. Apart from that, this piece was conceived for venues with long reverberation times, which enables a high degree of integration of the voices. These aspects, which relate the piece with space and sound qualities—within a fixed media context—have drawn me to use the piece as a source material for *5 Timelessness Studies*.

The Moldau fits with these studies because I wanted to include a big orchestral piece in order to test how its dense timbral qualities would work on its 'timeless version'. Furthermore, *The Moldau* includes instruments playing within a wide register, resulting in a much richer spectrum, and thus contrasting with the other selected pieces.

La piojosa is a cumbia song. Cumbia is a folk music genre spread throughout South America, and more recently also in Europe. From the sound quality perspective, this version by Banda Conmoción includes many wind instrument performers playing almost the same melodic line, with some variations in ornamentation and tiny differences within the line. The variety of instruments introduces an interesting timbral quality that I wanted to include in these studies. From a personal perspective, I have been playing cumbia, and particularly this song, from the time of my adolescence.

I wanted to include a Mozart Sonata for fortepiano because I wanted to include a piece for a solo instrument, but also a tonal piece. Of the two solo pieces I selected, this one encompasses a wider register. Furthermore, I decided to include a fortepiano version because I am interested in the noisy quality of its sound production. In particular, I am interested in the noisy sound attack of the instrument and the disparate timbral qualities through its range. Therefore, I wanted to include them in them to try out how much of these qualities remain in the 'timeless' version.

La negrina is the other selected solo music, and contrasts with the Mozart sonata because it involves a narrower register. In this case, only an excerpt of the tenor part is included in the piece. The piece by Mateo Flecha el viejo belongs to a genre called Ensalada. These kinds of compositions involve music from disparate contexts, for different purposes and in different languages, put together under the same contrapuntal style. This feature relates to the *5 Timelessness Studies* as a whole, where music from different contexts, for different purposes and in different languages are brought together under convolution processes.

Murga is the name of a music genre from Uruguay and belongs to the carnival tradition of that country. *Retirada* is the last part of their shows. My interest in including this piece resides in its peculiar timbral quality. Continuing with questions of contrast, this is another piece for a large choir but it does not involve more than two or three parts, and the timbral quality of the voices is completely different. Additionally, besides the choir there are cymbals, snare drum and bass drum. This instrumentation seems to be divided into two layers—choir and drums–and I wanted to test out how much of this layering remains in the 'timeless' version.

Strings, clicks and convolution (in progress)

The second piece (in progress) briefly described in this chapter is for two violins, one violoncello and electronics. The main exploration of this piece is how the *5 Timelessness Studies* are convolved with other audio files, in order to apply the 'spatial qualities' of themselves as 'spaces'.

The string players are only playing *clicks,* achieved by pressing the bow hard on the strings and carefully releasing the pressure. These tiny clicks feature pitched sound, and their pitch relates to the point on the string where the bow is placed. The clicks appear with a background of harsh noise because of the residual bow pressure while preparing to play the clicks. I include clicks because of their shortness, so that the qualities of the impulse responses within convolution can be heard more clearly. They resemble impulses, such as are used for impulse response measurements.

For this piece, I used *Timelessness study #3: Sonata K 333 in Bb* and *Timelessness study #2: Vltava* to convolve with the clicks. The string instruments—violin and violoncello—belong to the tradition of classical music as well as the chosen pieces, and are even present in one of them (*Timelessness study #2: Vltava).* The discussion of the paradigm of convolution reverberation, which separates sound source and spatial qualities, is brought to this piece as well. The pieces—*Sonata K 333 in Bb* and *Má vlast - Vltava*—were turned into the 'spatial qualities' of the piece, while the performers—which usually would be playing the pieces—are just playing tiny clicks that works as triggers for the 'spatial qualities'.

5 Timelessness Studies—detailed procedure

Timelessness study #1: Deo gratias

Deo gratias à 36 (36-part canon) - Johannes Ockeghem (1410-1497)

Huelgas Ensemble - Paul Van Nevel

Procedure details:

- Deo gratias recording
- Deo gratias recording reversed
- Deo gratias recording *[23] Deo gratias recording reversed = 'Timeless' audio 1[24]

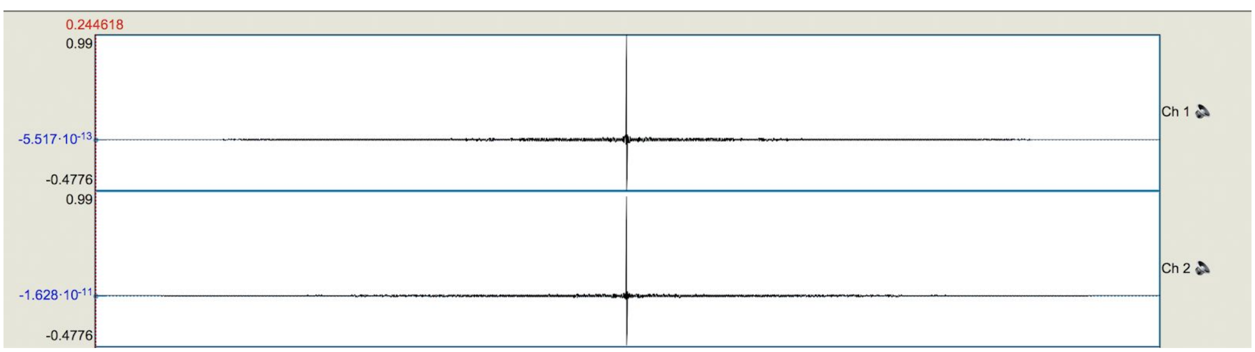


Figure 4 - 'Timeless' audio 1 (twice as long as Deo gratias recording)

[23] Convolved with (the convolutions of this section have been done using *Praat,* version 6.0.43).

[24] It can be found in the attached audio files—the parts without audio signal (at the beginning and at the end) were removed.

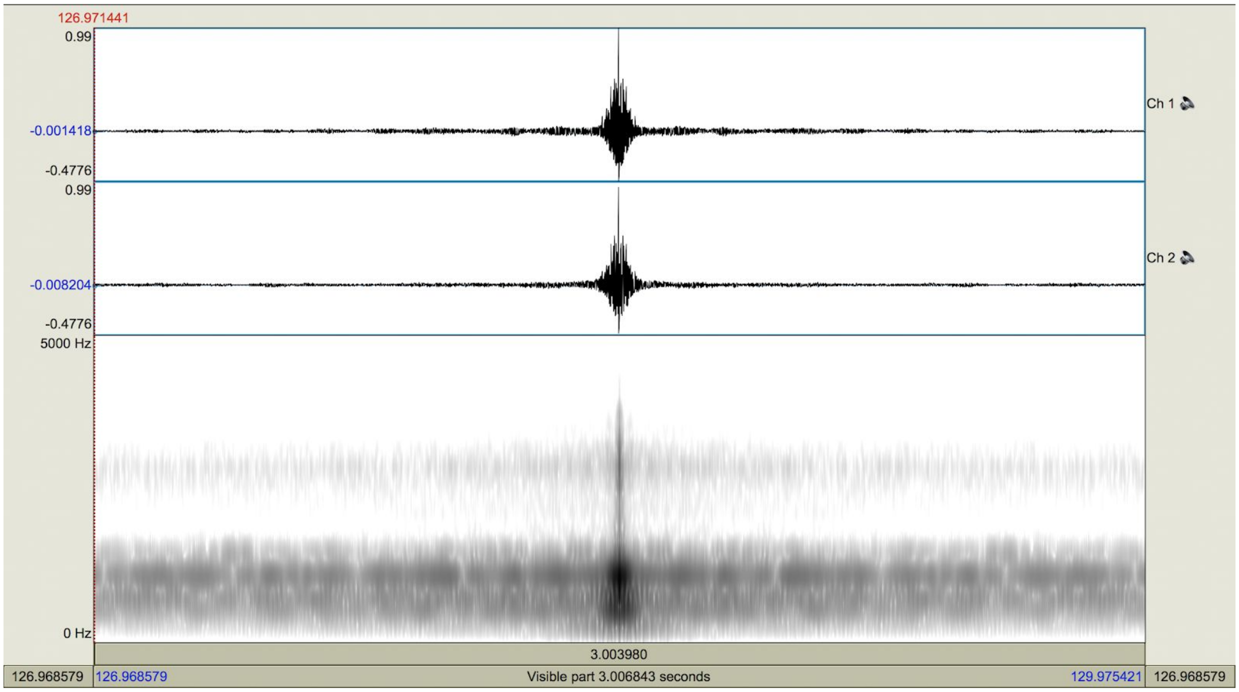


Figure 5 - Zoomed excerpt of its central portion, where the energy gets boosted

In cases of convolution between an audio and a reversed version of itself, the spectral characteristics are boosted and time references are destroyed. Moreover, the central peak is boosted as well. In order to obtain a more defined profile in terms of spectrum and higher central peak, I repeated the process as follows:

- 'Timeless' audio 1 (twice as long as Deo gratias recording) * 'Timeless' audio 1 reversed = 'Timeless' audio 2

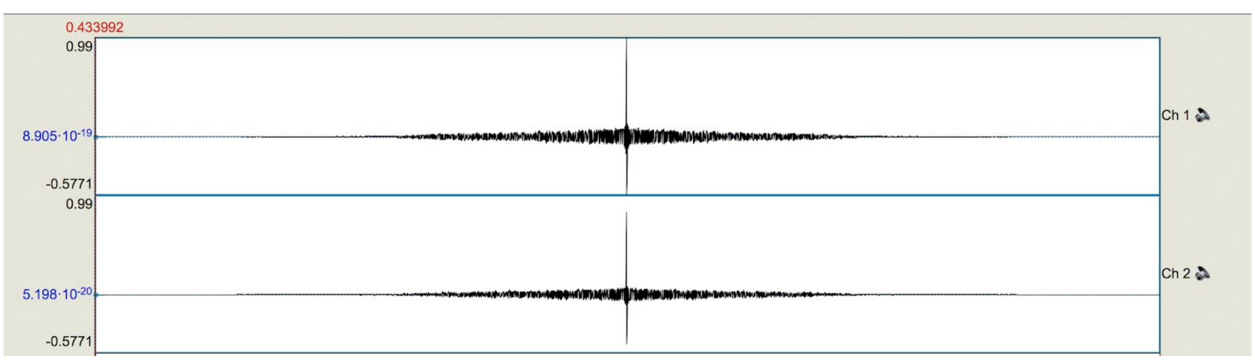


Figure 6 - 'Timeless' audio 2 (twice as long as 'Timeless' audio 1)

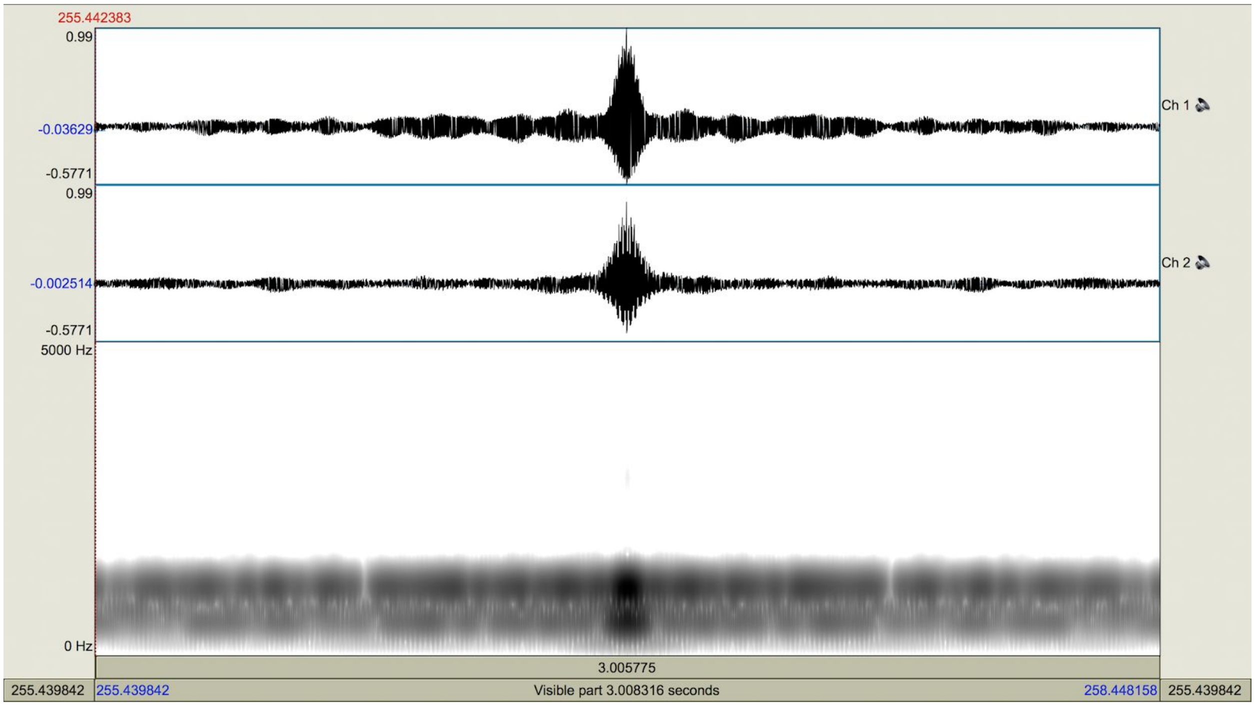


Figure 7 - Zoomed excerpt of its central portion

- 'Timeless' audio 2 (twice as long as 'Timeless' audio 1) * 'Timeless' audio 2 reversed = 'Timeless' audio 3

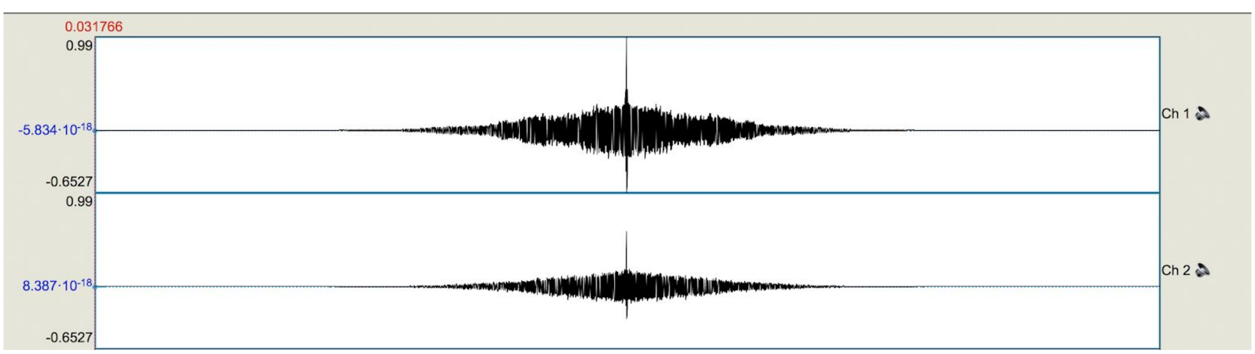


Figure 8 - 'Timeless' audio 3 (twice as long as 'Timeless' audio 2)

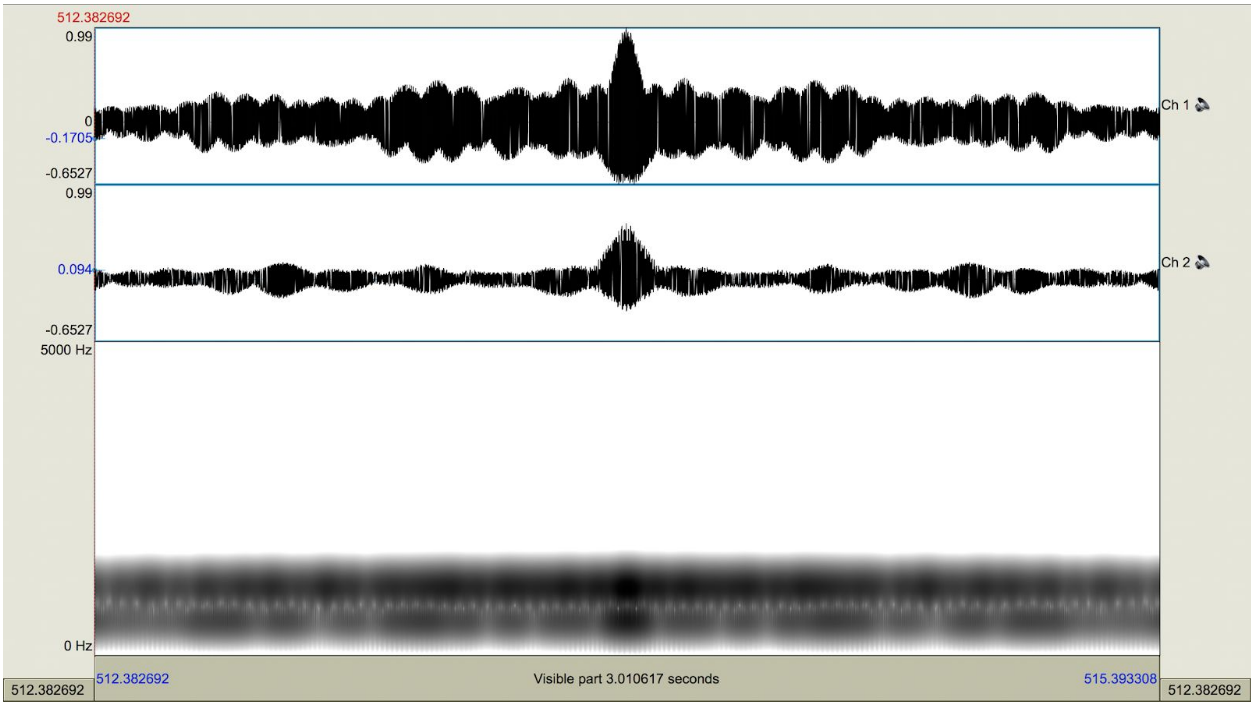


Figure 9 - Zoomed excerpt of its central portion

Finally, I cut off from the crossing 0 point before the main peak of the audio, right before the central part:

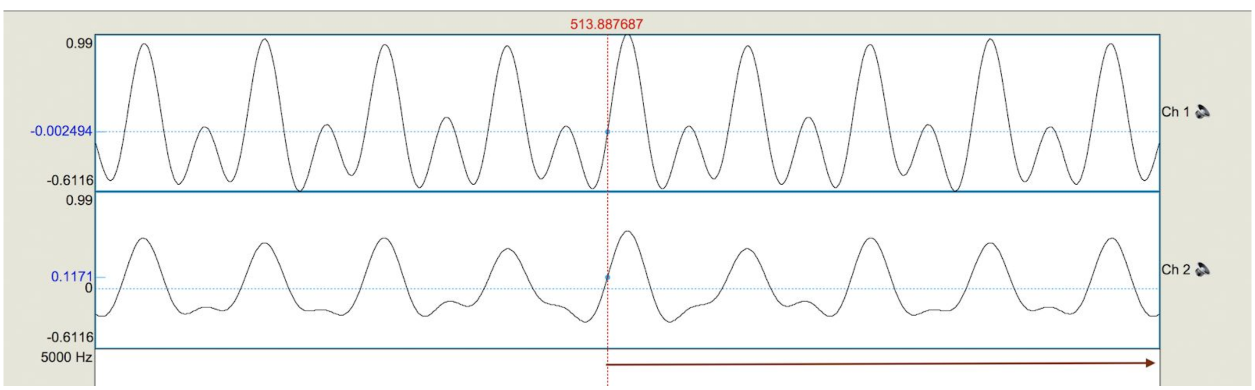


Figure 10 - Zoomed excerpt of the cutting point (red cursor)

Subsequently, I adjust the envelope and duration in order to achieve a similar one as an impulse response of a room.

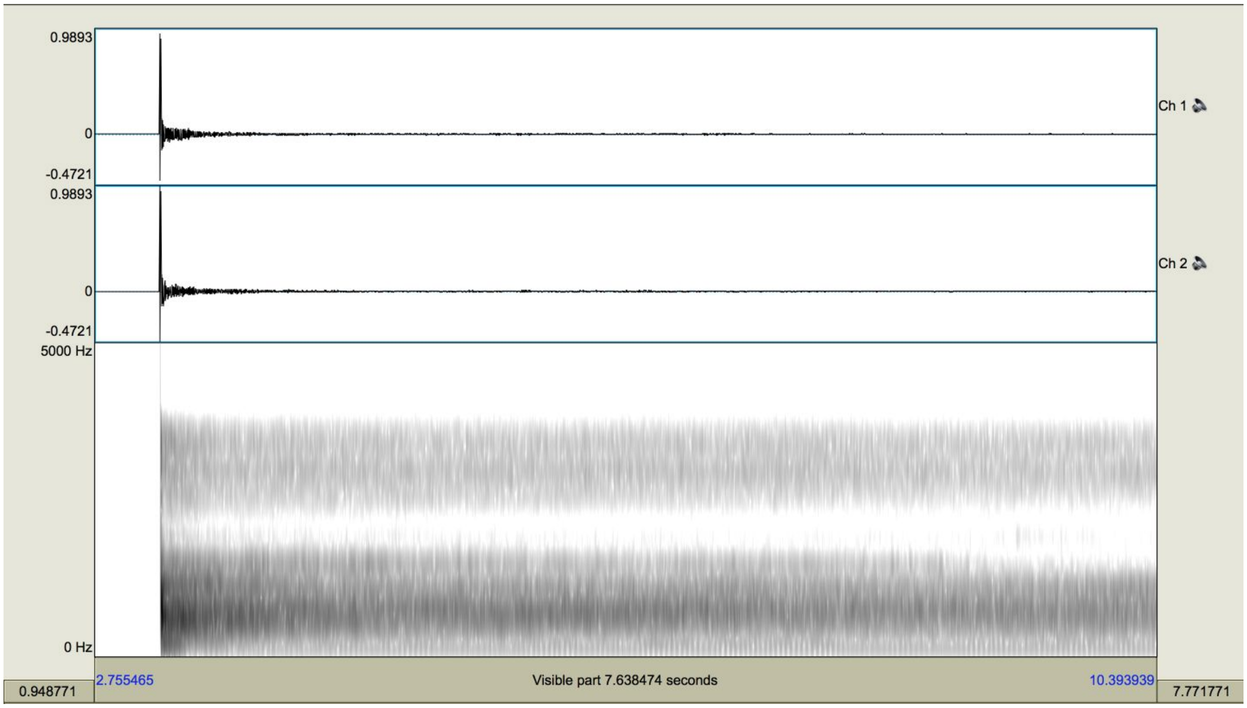


Figure 11 - Waveform and spectrogram of the *Timelessness study #1: Deo gratias*

FILTERING AND 'SYNTHESIS' ALGORITHMS

I will draw attention to conceptual and procedural similarities between two works, so as to emphasize how both pieces can be approached as chains of modules that involve a sound source and processes of filtering and 'synthesis'. This comparison, I hope, will introduce and frame some compositional strategies which were presented in the last pieces and will be specially present in the upcoming ones. The purpose of this text is to point out similarities between *Three Indigenous Songs* (1979) by James Tenney (1934-2006) and *LXVIII Canto Llamo de la Inmaculada Concepción de la Virgen María y Siguense Dies and Seis glosas sobre el Canto Llamo* by Francisco Correa de Arauxo (1584–1654) for the Fuente de la Fama in Seville (builded in the 17th century).

Modular structure is one of the main aspects of the compositions I discuss throughout this project. In particular, those structures following a mechanism-like behaviour, wherein each module carries out a single task within the whole, is what I am interested in for my own practice. I explore the modular typology in terms of how the mechanism becomes—or can become—audible in the outcoming sound. I will describe in this chapter how this mechanism-like behaviour emerges from two disparate examples.

Automata in the form of musical instruments always involve imperfections in their operating mechanisms, as well as the sound that the mechanism itself produces when in operation (mechanical sounds from cranks, gears, bellows or even motors). In this chapter I compare such musical automata with a chamber piece by James Tenney, in order to frame these inherent aspects of automata within a context closer to my practice—Tenney's piece deals with ideas like spectral analysis as a tool for his compositional strategy, as I do—. Finally, these examples also show a perspective on the likeness between the sound source and the piece. Even if the pieces are intended to imitate the sound source, the missing aspects of the sound source as well as the audibility of the mechanism and its imperfections result in a blurred presence of the sound sources. In that regard, Tenney includes the following comment in the score, which expresses clearly the kind of interaction between the sound sources—speech and sung voice—and his pieces: "The perceptual space introduced by Three Indigenous Songs is meant to be somewhere near the threshold between music and speech. Occasionally, perhaps, some semblance of the underlying texts may actually be heard."[25]

Three Indigenous Songs involves a particular approach to spectral composition. By using acoustic instruments, Tenney unfolded a strict algorithm to make his own spectral music, before digital analysis and synthesis tools had been widely developed. Therefore, the sound source could be followed throughout several compositional processes described by the composer in the score. Finally, it is possible to compare the sound source with Tenney's piece, because he includes a precise reference of the sound sources in the score. The compositional strategy can be seen as a series of processes of filtering and 'synthesis' which removes and adds information respectively, consequently they result in an interesting tension between the sound source and the piece, as I want to achieve for my own pieces.

[25] James Tenney, *Three Indigenous songs* (Toronto: Canadian Music Centre, 1979).

There are three sound sources used by Tenney in each part of *Three Indigenous Songs*. These are: *No More Good Water* (1927-1931) by Jaybird Coleman (1896-1950), the poem *Kosmos* (1860) by Walt Whitman (1819-1892)—read by Tenney himself— and *Hey when I sing these 4 songs Hey look what happens*[26] (1971) translated from the Iroquois by Jerome Rothenberg (born 1931). Since these sources mainly contain singing and speaking voices, the composer developed an algorithm in order to analyse and 'synthesize' vocal sounds with acoustic instruments instead of electronic ones. This algorithm can be seen as a chain of modules which basically involves the analysis of the sound source, the selection of certain sounds from it, a 'translation' of them to the instruments for which the piece is composed, and the performers. The process is described by the composer in the following way:

> The vocal sounds in each of these have been translated into a purely instrumental form by assigning the fundamental frequency of each vowel to the bassoon or tuba, and that harmonic (of the fundamental) nearest to each of the three major formant peaks for that vowel to the alto flute and piccolos. Consonant are represented by the two percussionists, using wood-blocks (for k, t, p) tom-toms (for g, d, b — and, with wire brushes, th, f, h), and suspended cymbals for s and sh. For the harmonica choruses in No More Good Water and for the antiphonal phrases in Hey when I sing... the instruments are used freely.[27]

Within this short analysis, I will explore how the aforementioned algorithm filters out some aspects of the sound source. Moreover, I discuss how this algorithm is used by Tenney to 'synthesize' the sound source, becoming audible in the piece by adding its own sound qualities and imperfections. On the one hand, Tenney's words suggest that he is considering the fundamental frequency and the three harmonics nearest to each of the three major formant peaks of the vowels, which, considering a typical speech

[26] The sound source is a composition by Tenney himself called *Hey when I sing these 4 songs Hey look what happens* which uses this translation of Rothenberg's text.
[27] James Tenney, *Three Indigenous songs* (Toronto: Canadian Music Centre, 1979).

synthesiser's algorithm, constitute the most important ranges within the spectrum in order to achieve vowel-like sounds. However, other harmonic as well as inharmonic components—e.g. a transient—are not considered in this interpretation. On the other hand, although the alto flute and the piccolo's fundamental frequency is used to play three harmonics of the vowel sounds, they have their own spectra, which does not totally match with the spectrum of the vowel sound. Therefore, spectral components from these instruments, such as some of their own harmonics and transients, are adding sound information within the 'synthesis' process.

The bassoon/tuba, the alto flute and the two piccolos are playing parts of the spectrum of notes from the sound source. Generally, the relative amount of energy between each harmonic and the fundamental frequency within an instrument remains more or less fixed. Therefore, each instrument has a fixed dynamic in order to maintain the same balance between the harmonics and the fundamental frequency within the notes they are 'synthesizing'. Specifically, the bassoon/tuba and alto flute play *mezzopiano*, piccolo II plays *piano* and piccolo I plays *pianissimo*. However, there is not another layer of dynamics in the piece which reflects the dynamic variation between the notes or parts within the original sound source. Apart from the intended dynamic variations in the original pieces, the spoken and sung voice always has its own natural dynamic fluctuations. As a result, the dynamic variations of the original source are removed. However, new dynamic variations added by the performers arise in Tenney's piece. This is mainly because of the performers' natural oscillations of volume related to the changes in the instrument's register, articulation, etc.

The time structure of the phonemes in the spoken and sung voice is extremely complex, which turns any transcription to the traditional notation system for music somehow imprecise (the fact that the transcription is approximate is explicitly stated by Tenney for the second piece in the score). Focusing on the imperfections of the algorithm which introduces deviations from the original pieces, it should also be said that there are always tiny rhythmic deviations from the score which performers naturally make while playing. All these aspects might go unnoticed in a different context, but they are highlighted in these pieces because the sound source is mainly spoken voice. Spoken voice is one of the most complex sounds we are used to hear and analyse in a very detailed way, which makes any tiny deviation from its natural behaviours noticeable.

Three indigenous songs, as a result of this particular procedure, finds its counterpart in an unlikely example from southern Spain. More specifically, in the Alcázar—a royal palace in Seville—where we can find Fuente de la Fama. Fuente de la Fama is a water organ that plays two pieces, belonging to the polyphonic renaissance music tradition which were composed by Francisco Correa de Arauxo, called *LXVIII Canto Llamo de la Inmaculada Concepción de la Virgen María y Siguense Dies* and *Seis glosas sobre el Canto Llamo*. These pieces were originally conceived for organ, and I call their versions on water organ *Water organ soundtrack*. The *Water organ soundtrack* can also be approached from a modular perspective, which includes the original piece for organ, the transcription of the score to the barrel, the pipes, the water organ mechanism and the water flow which drives it.

Francisco Correa de Arauxo's original pieces were probably conceived for church organ. I was told by Rodney Briscoe, the organ builder who recently recovered the Fuente de la Fama water organ, that the music was arranged to suit the pipes of the water organ and the length of playing time[28]. I was not told how it was arranged, but I speculate that it had to suit a least amount of pipes—because the water organ seems to be smaller than a church organ—and probably it had been shortened because the barrel has a limited space where the score can be recorded. The barrel activates a mechanism that lets air pass through organ pipes as is specified by the score (see Figure 12). The water powers a water wheel connected to the barrel and the amount of the water flow controls the speed of the barrel and thus the speed of the playback[29]. Therefore, the water flow controls the tempo, providing it with an important instability.

As a result, we can listen to a version of Correa de Arauxo's pieces where there are instabilities regarding tempo and volume—imperfections in the mechanism which let the air pass through the pipes provide volume differences. Although we as listeners probably do not know the original organ versions, we can appreciate that the water organ sounds different from what we might typically expect—the kind of instability in tempo due to the water flow is very different from tempo fluctuations that can be expected from a live performance, because it is not related to interpretative decisions. Moreover, the water passing through the organ contributes a very prominent sound (like a small waterfall), which becomes part of the sonic experience as well as the organ pipes' sound.

[28] Personal correspondence with Rodney Briscoe, April 30, 2019.
[29] *ibid.*

The original source or at least the reference to a certain genre—polyphonic renaissance music—remains recognizable. Furthermore, *Seis glosas sobre el Canto Llamo* is more active rhythmically, as could be expected, taking into account that the title refers to ornamentations or variations upon a plainsong. Such complexity, together with the imperfections in the playback system, results in a smudging and obscuring of the musical texture.

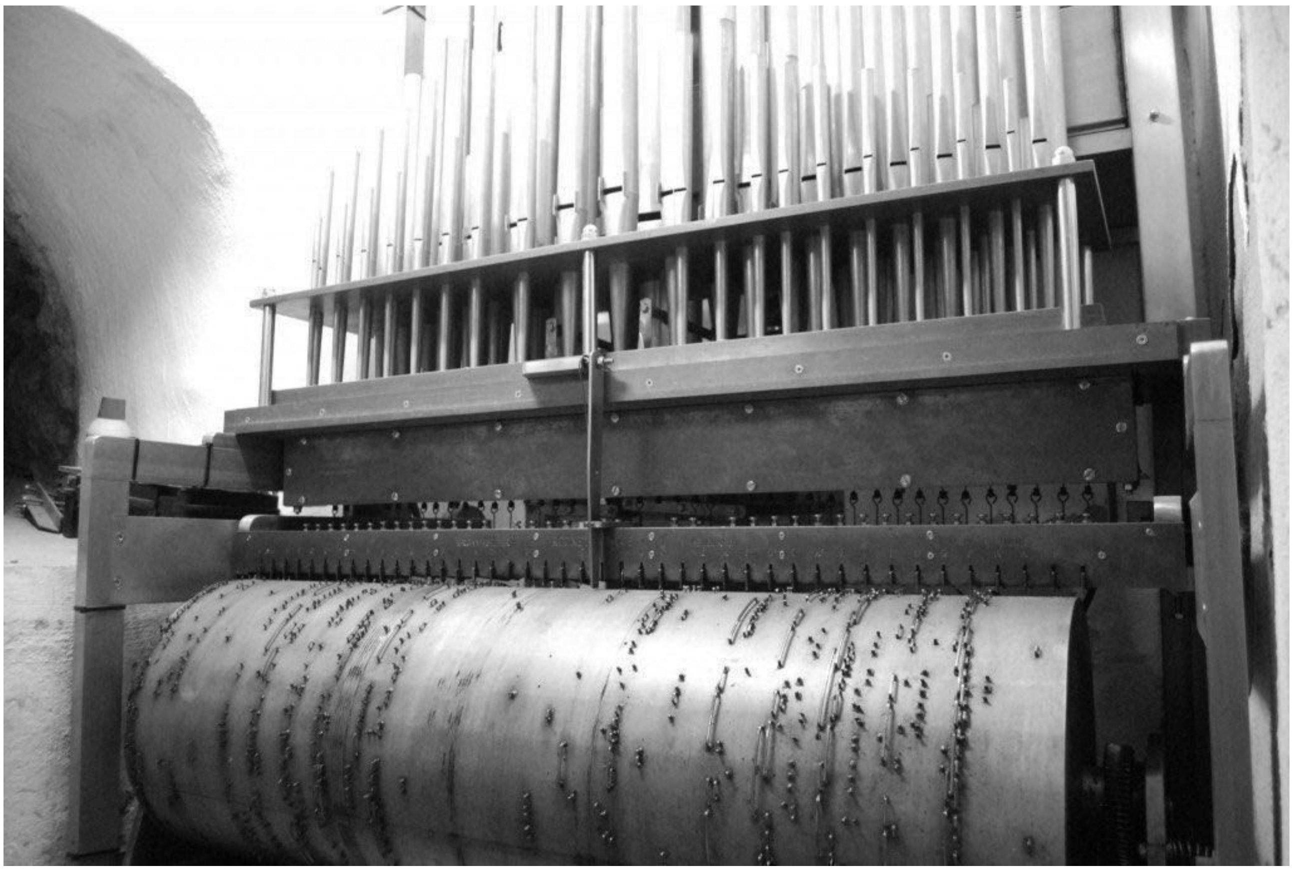


Figure 12 - Mechanism of the water organ *Fuente de la Fama.*[30]

[30] Real Alcázar website, "Puesta a punto del órgano de la Fuente de la Fama," last modified November 22, 2016, accessed May 29, 2020, https://www.alcazarsevilla.org/en/news/puesta-punto-del-organo-la-fuente-la-fama/

While at first these two pieces seem quite disparate in their origins, there are in fact many similarities in their algorithmic structure. I think *Water organ soundtrack* and *Three Indigenous Songs* are works that can be approached as sound sources passing through a chain of modules, which filters and 'synthesizes' them.

To sum up, I will point to three aspects of these examples which became starting points for the composition processes I will explain in the next chapter. Firstly, the way some modules of these pieces work as filters is relevant for my own practice. This mainly happens because there are aspects or parts of the sound sources which are not included in the 'synthesis' process, or because the new instruments are not capable of imitating them. For instance, in Tenney's piece all components of the vowel sound spectrum, other than the fundamental frequency and the three harmonics nearest to each of the three major formant peaks, were filtered out, along with the dynamic variations. Regarding the *Water Organ soundtrack*, we know that the sound sources were arranged to suit the pipes of the water organ and the length of playing time, which have involved removing notes or even sections of the original piece, as I speculate.

Secondly, the inaccuracy in some mechanisms of the 'synthesis' algorithm becomes audible in the pieces. This can be exemplified by the instability of the water flow in the *Water organ soundtrack,* which results in an unstable tempo. In Tenney's piece, the small inaccuracies in the transcription of the phonemes' rhythm, and the performers' natural oscillations of volume, are examples of imperfections in the mechanism.

Finally, I am interested in the sound qualities of the new instruments which are added to the piece and do not relate to the sound source. An example would be the spectral qualities of the instruments involved in Tenney's piece, which do not relate to the sound source (e.g. some harmonics and transients in the woodwind instruments). Another is the sound of the water passing through the water organ mechanism.

SUBVOCALIZATION

> *Subvocalization, or silent speech, is the internal speech typically made when reading; it provides the sound of the word as it is read. This is a natural process when reading, and it helps the mind to access meanings to comprehend and remember what is read, potentially reducing cognitive load.*[31] [32]

The two pieces I describe in this chapter involve performers which read by subvocalizing on stage—imagining their spoken voices—and enter part of the text they are reading into the computer. These thought words which come to the mind of the performers while subvocalizing are the starting point for both pieces. Besides the performers, the pieces include live electronics, which are inspired by the subvocalization of the performers as well. The live electronics are based on recordings of the performers speaking the same text they are subvocalizing, alluding to what they are hearing internally during the performance—as everybody, they hear their own voice when subvocalizing.

Both pieces were conceived with a modular structure (see Picture 12 and 13). Until now, audio files which are recordings of musical pieces have been used as sound sources (e.g. Rivero's bandoneon recordings, the recordings of *Deo Gratias* and *The Moldau*, etc.). The pieces start always from a given audio file which is to undergo a modular process. For the two pieces I will describe in this last chapter, the recorded voices of the performers play the role of the aforementioned sound sources. Additionally, the modular structure also includes the performers and the text which they are reading as modules. Throughout this chapter, I will explain in detail how

---

[31] Wikipedia, "Subvocalization," last modified March 28, 2020, accessed May 21, 2020, https://en.wikipedia.org/wiki/Subvocalization

[32] This is the beginning of the text which the performers read in *Reception Bells*.

modules in the compositional strategy work as filters and also add their own sound qualities to the pieces. Additionally, I will explain how some specific compositional decisions were made in order to enhance the audibility of the imperfections of the mechanisms.

Subvocalization is the internal speech that occurs when reading a word. The reader is able to imagine the spoken voice through the movement of the muscles involved in speaking. It is a natural process when reading which helps the reader to understand the meaning of the words. Subvocalization is sometimes related to the early stages of learning to read, which involves moving the lips and reading at a slow speed. However, even for fast readers who also associate meaning directly with the sight of words, micro-muscle tests have suggested that subvocalization is impossible to remove completely[33].

Bearing in mind that, the inner speech—which can only be heard internally by oneself—constitutes an attractive topic to work with because it cannot be heard by the audience. Involving subvocalization among my compositional materials pushes further the intention of using a sound source which does not appear in its original form. This is because the sound sources of these pieces (the recordings of the performers) alludes to another material—the inner speech of the performers—which is internally audible for them.

Within these pieces, there are performers reading a text by subvocalizing on stage. The programme notes explain that the performers are reading by subvocalizing and

[33] Psychology wiki, "Subvocalization," last modified August 19, 2013, accessed April 24, 2020, https://psychology.wikia.org/wiki/Subvocalization

how they are entering the text into the computer. Furthermore, the audience is invited to join the performers—the text which is read by the performers is attached to the programme notes. The sounds of the bells and the keyboard, as well as the playback, help the audience to follow the performers with their own inner speech. Also, the programme notes explain that the live electronics are based on recordings of the performers' voices, as a strategy to relate it to what the performers are experiencing internally (subvocalization).

The first module required for the pieces presented in this chapter is formed by the performers. They constitute modules of the mechanism of the composition, so they strictly fulfill the functions of subvocalizing the text, encoding and entering it into the computer. In *Keyboard Pieces* (2020), this is carried out by typing the words of the text that the performer is remembering, while in *Reception Bells* (2019) they are tapping bells when they come across vowel phonemes.

Live electronics, a real time synthesis shaped according to the sound characteristics of the sound source, are an important component of the pieces of this chapter. The sound source is a set of recordings of the voices of the performers. Within the recordings they are speaking all the phonemes which are present—and they have to react to—in the text that they are reading or remembering in the performance. Therefore, when they reference a phoneme by either pressing a key or playing a bell, real-time synthesis will be shaped according to the sound qualities of the corresponding recorded phoneme.

The live electronics for these pieces work mainly as a vocoder. This is to say that the sound source is analyzed and provides filters to be applied upon a synthetic signal. Thus, both pieces involve new instruments (synthetic ones) which somehow imitate the sound source, as acoustic instruments do in the examples of Tenney and the *Water organ soundtrack*. I label as the *second sound source* the synthetic signal which passes through the filters—as a reference to the terminology of a vocoder.

In order to introduce how the sound source and the second sound source work together, I will refer to the concept of the vocoder. This basically involves generating an outcoming sound signal from two input ones. Through vocoding, I will explain how I approach the sound source and the second sound source and what my interests regarding the relationship between them and the output are.

> It is possible to extract the (time-varying) spectral contour from one signal and impose it on another, a process originally developed in the analogue studios and known as vocoding (no connection with the phase vocoder).[34]
>
> The distribution of partials in a spectrum (spectral form) and the spectral contour (formants) are separable phenomena and limping differently on our perception. (…) If we therefore have now source with a clear articulation of the formants (e. g. speech), and another source which lacks significant formant variation (e.g. a flute, the sea), we can impose the formant variation of the first on the spectral form of the second to create a dual percept (vocoding).[35]

The concept of vocoding involves a synthesis process with two audio sources, one of which is used for designing the filters—which is the one I called the sound source within my research—, and a second sound source to fill them. My artistic research delves into the relationship between the two sources, and how both may appear at the output. Furthermore, the audio source is not considered as a static filter but as a

[34] Trevor Wishart, *Audible design* (Orpheus the Pantomime Ltd, 1994), p. 37.
[35] *ibid.*, p. 100.

source with changing frequency content over time. This means that both sources are considered not only in terms of their spectral characteristics, but also in their time dimension.

According to the common definition of vocoding, the second sound source is modified by imposing upon it the spectral characteristics of the sound source.

> As the spectral contour defining the formants must have something on which to "grip" on the second source, this process works best if the second source has a relatively flat spectral contour over the whole frequency spectrum.[36]

In order for the vocoding process to be as accurate as possible, the second source signal should be one with a rich spectrum, since energy needs to be spread over as much as possible of the spectrum to be filtered. Moreover, the aforementioned second input should have an evenly distributed energy within the spectrum. Consequently, the timbral characteristics of the second source tend to be less present in the actual synthesis because it works as a kind of pliable dough being shaped by the first input (sound source).

Success in vocoding seems to be related to a certain amount of likeness between the output and the sound source. However, the character of the outcoming sound has much to do with the characteristics of the second input (which might consist of square waves, pulse waves, triangle waves, white noise, pink noise, among others) and all the parameters which can be adjusted in these kinds of signals.

[36] Trevor Wishart, *Audible design* (Orpheus the Pantomime Ltd, 1994), p. 100.

Using live synthesis following the basic ideas of a vocoder is a strategy with two purposes. On the one hand, it is a way to keep some sound characteristics of the sound source in the playback of the piece, but blurred by the sound qualities of the second sound source. As previously mentioned, I intend to investigate a particular perspective concerning the likeness between the piece and the sound source, whereby some characteristics of the sound source could be traced back, but the sound source does not stand as an accurate synthetic version of it. This is to say that the second sound source should be present in the playback as much as the original source. On the other hand, the second sound source will play a role in the poetic dimension of the piece by being related to the read text, as I will explain later.

Figure 11 - *Reception Bells* performed at Sonology Discussion Concert (2019)

***Reception Bells* (2019)**

Reception Bells is a piece for five reception bells and live electronics. The set up consists of two performers reading the same text[37]—which is handed out among the audience—, five reception bells—differently tuned from each other—and live electronics.

[37] The text and further information about the piece can be found in the Appendix 2.

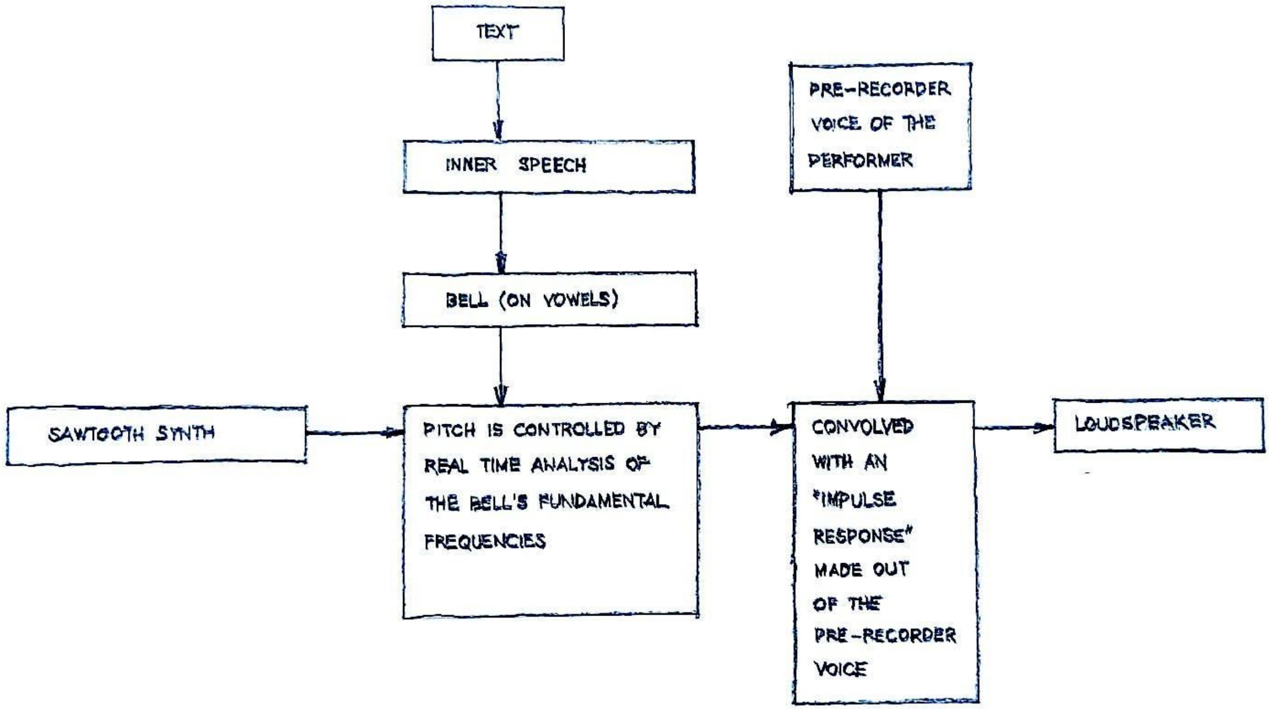


Figure 12 - *Reception Bells* modular diagram

The starting point for *Reception Bells* is a text which is read silently by two performers on stage behind a table. Each performer plays a certain bell every time they come across a specific vowel phoneme, the bells are recorded and their signal is analysed in a *SuperCollider* patch. As a result, only the vowel phonemes pass to the next link in the chain (the computer), while the other phonemes are filtered out—the performers work as a module which filters out parts of the sound source. There are five sets of phonemes of vowels proposed for the piece—the text is in English—which includes every possible vowel phoneme. As there are way more phonemes than bells, the performers made up these sets of phonemes according to similarities between the sound qualities of the phonemes which integrate each set. Every phoneme set corresponds to one of the five bells, and the performers have chosen the correlation between phoneme set and bells according to similarities in sound quality as well.

The next module of the piece is the *SuperCollider* patch, which synthesizes sound according to some characteristics of the pre-recorded voices of the performers. The bells are all recorded and each signal is analyzed in real time with a pitch tracker in order to obtain their fundamental frequencies. Subsequently, the outcoming information from the pitch tracker is used to control a sawtooth synthesizer. One of the reasons for using reception bells is their spectral qualities. It is quite inharmonic in the first milliseconds after the attack, and then becomes more harmonic. This feature brings the pitch tracker into error, basically because there is no clear pitched sound in the first milliseconds. The outcome of the algorithm of fundamental frequency recognition of the pitch tracker does not stop working when the sound characteristics get less harmonic, therefore, it still tries to find any harmonic relationship between the spectral components of the incoming signal and produces a number. Through experimentation with the bells, I discovered that when the fundamental frequency of the bells is unclear, the pitch tracker outcome is around 800hz. As a result, the pitch tracker outcome is a *glissando* between approximately 800hz to around 200hz, where the fundamental frequencies of the bells are. Therefore, the imperfections within the algorithm turn into one of the main audible aspects of the playback, which is the descending *glissando* of the playback after every attack.

The second audio source, the synthetic material of this piece, is a sawtooth synthesizer. Its outcoming signal is convolved with an 'impulse response' made out of the performer's voice pronouncing the corresponding vowel phonemes. Once a performer plays a specific bell, the fundamental frequency of this bell will be applied to a sawtooth synthesizer. Subsequently, this sound will be convolved with the 'impulse response' made out of the performer's pre-recorded phoneme[38] associated with this specific bell and finally sent to one loudspeaker (See Figure 12).

Finally, there are five loudspeakers extracted from 1990s' televisions, with the five different outcoming signals (one for each bell). These television loudspeakers feature a quite limited frequency range and a highly filtered quality. The low fidelity quality of the loudspeakers (which sometimes embrace distortion as well) enhance the audibility of themselves as a module in the chain which is adding its own fingerprint. The loudspeakers are similar to the bells in terms of shape and size, which is important because there is also a balance between them in volume level. At first sight, because of the set up, the loudspeakers can be thought of as intended for amplifying the bells. However, the playback is an imperfect imitation of the speech, and it interferes with the bells' acoustic sound because these and the playback are simultaneous and at similar volumes.

One of the main interests of this piece resides in exploring how the playback and the bells' sounds interferes with the subvocalization process of the performers. When I conceived the piece, I wondered if the performers would get used to the bells' sound and the playback enough to expect a specific bell's sound when they come across the correspondent phoneme. Performers are listening to their own inner speech simultaneous to the playback, which is reacting to the bell sounds with live synthesis based on recordings of the phonemes. The performers' own pre-recorded voices

[38] I recorded each performer saying the set of phonemes corresponding to each bell, mixed these sounds and applied the previously explained method for extracting 'impulse responses' from audio files.

while speaking these phonemes are included in the process as what might be called 'spatial qualities' convolved with the sawtooth synthesizer signal. The tiny movements of their throats and mouth—provided by subvocalization—and the action of playing a certain bell constitute a kind of motor circuit, which is added to hearing the outcoming sound from a specific loudspeaker with their voices' sound qualities. Certainly, the experience of performing the piece confirms they get used to this correlation and develop an expectation for certain sounds from the bells and the playback. Finally, they have to read following each other, because one is playing three bells corresponding to three groups of phonemes and the other one is playing the remaining two. For this purpose, the association which the performers make between vowel groups and subvocalized phonemes is very important.

Another important aspect of the piece is the role of the audience. The text read by the performers is available for the audience as an attachment to the programme notes. In this manner, the audience is able to follow the performers, and encouraged to do so. Once the audience decides to follow the performers, the experience becomes extremely different because the attached text becomes a score for the audience, which implies a precise text to subvocalize with a determined speed (following the performers by means of listening to the bells). As a result, subvocalization 'completes' internally the sonic experience of the piece for the audience, by adding their internal speech to the playback and the bells' sound. Reading (at a very slow speed) while identifying every phoneme—and collating this identification with the bell sounds in order to follow the performers—requires an important focus on an immediate level of hearing because of how demanding the activity is. Additionally, subvocalizing at an unusually very slow speed, paying attention to every phoneme allows us to create awareness of timbral and motor aspects of the speaking habits and how they are blended together. As soon as a listener ceases to follow the text, he/she can listen from another perspective not without their inner speech as an internal component of the piece, which allows focusing on the sound qualities of the interaction between the playback and the rhythmic fabric determined by the distribution of the vowels within the text.

This is the first piece that includes performers in the project. The performers are a reference for the audience, in order to make it possible to follow the text, but they are also involved in an experimental process while reading and figuring out when they have to play the bells. Moreover, they—as well as the audience—develop an awareness of their subvocalization through reading the text. Having two performers on stage makes it more achievable to play bells on every vowel phoneme. Moreover, I included more than one performer because it allows the possibility of reading the text following each other. This makes them slow down the reading flow—otherwise it becomes extremely difficult to follow each other—and together with the focus on their inner speech, the aforementioned movements at their mouth and throat become more noticeable. Additionally, it makes the reading traceable by the audience, which would not be achievable if they were reading quickly or in independent streams.

The text is encompassed in the narrative of the piece and therefore relates to subvocalization as well as the live electronics. It is a mix between an excerpt of the *Wikipedia* definition article about subvocalization—intended to introduce the subject—and an excerpt of a poem by Mario Santiago Papasquiaro (1953-1998) called *Advice from 1 Disciple of Marx to 1 Heidegger Fanatic (2013)* translated from Spanish

(*Consejos de un discípulo de Marx a un fanático de Heidegger)* by Cole Heinowitz and Alexis Graman.

One more thing about the performers relates to their kind of presence on stage. Their performance embraces a highly demanding cognitive activity as well as tiny concise actions—playing two or three reception bells placed in front of them. In that manner, their role on stage resembles the algorithmic processes carried out by the computer in the piece because they are analyzing their inner speech and playing bells (triggering sound) if certain conditions are met (vowel phonemes). Therefore, they become another module in the structure of the piece. I decided to use reception bells because they just have a single button and belong to a non-musical context, much more related to being a signal. Furthermore, this constrained performativity expressed by tiny movements (playing bells) is inspired by subvocalization, which also embraces tiny movements but in the throat and mouth.

***Keyboard Pieces* (2020) (in progress)**

Keyboard Pieces is a project which involves several pieces for one performer typing on a computer keyboard (the keyboard sound is amplified) and live electronics.

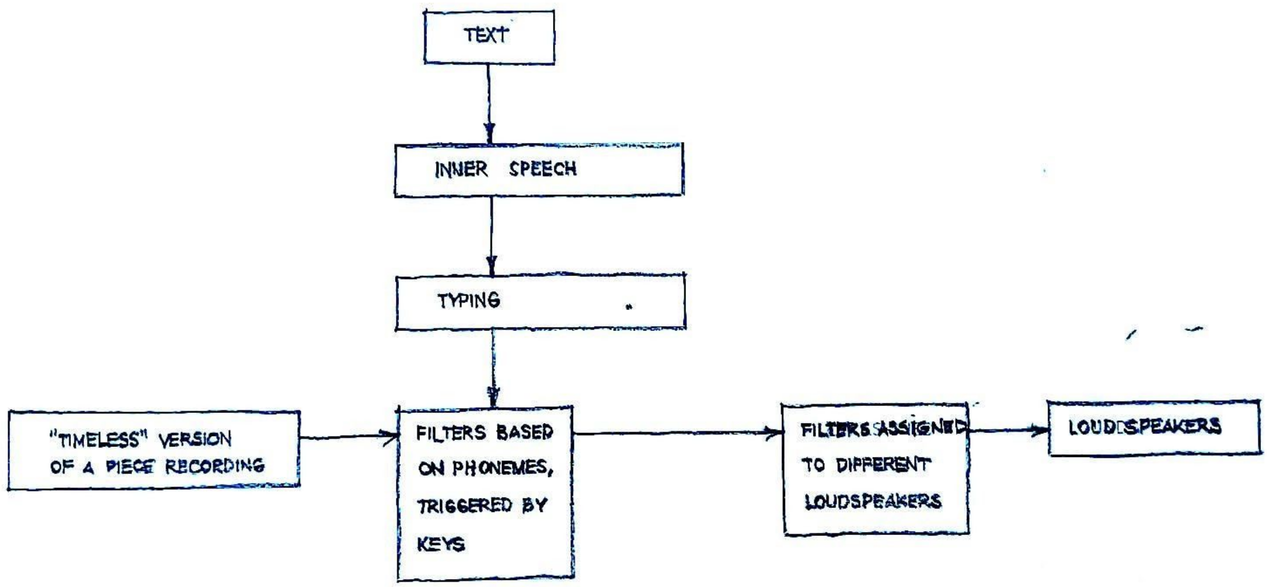


Figure 13 - *Keyboard pieces* modular diagram

These pieces were also conceived with a modular structure, similar to *Reception Bells*. The first module is the performer remembering a text by heart, subvocalizing on stage and typing it. In this case, not just some phonemes but every letter is entered into the computer. As well as *Reception Bells*, the sound source in these pieces are recordings of the performer speaking phonemes, which alludes to his/her inner speech during the performance. According to the vocoder paradigm, I designed a filter based on each phoneme of the performer, which is activated when the correspondent key is pressed and turned off when the key is released. Moreover, an audio file—which I generally called the second sound source—is always played in the piece and once a certain key is pressed, it passes through the correspondent filter. Finally, depending on which key is pressed, the filtered sound comes through a specific loudspeaker.

The piece was conceived in Spanish, which has an almost entire phonetic alphabet so that almost every letter has only one corresponding phoneme. There are just a few cases where a letter relates to more than one phoneme depending on the following letter. For these cases I just assigned the most common phoneme for the letter. Consequently, I recorded the performer speaking almost all the existing phonemes in the language in order to design the filters. These few missing phonemes bring the algorithm to some 'mistakes'—when typing a letter which should sound with the other phoneme corresponding to the letter—which are part of the imperfections of the modular structure I'm interested in for these pieces.

Every stage of the modular structure behaves by filtering out aspects of the sound source as well as adding its own sound characteristics. Firstly, the memory of the performer plays its own role in the piece, because sometimes there are small hesitations before remembering the following word, which adds an irregular behavior to the first module of the process. Secondly, typing involves an extra layer of interaction involving finger movements over the keyboard. It determines the speed with which letters are entered into the computer. In particular, this speed is not constant but it relates to the distribution of the letters on the keyboard or even to the usual letter combinations that we are more skilled in making.

In order to make these inherent aspects of typing more audible, some decisions were made. Firstly, the sound of the keyboard is amplified. Secondly, pressing and releasing a key becomes more noticeable because of the decision of avoiding crossfades between the filters which they are activating. This also filters out an important aspect of sound source—the transitions between phonemes are not incorporated in the synthesis. Finally, the keys are assigned only to one of the 8 loudspeakers, which makes the distribution of the keys on the keyboard more noticeable. This is another decision that points towards the physical aspects of the computer keyboard rather than the sound source.

The next module in the piece is the *MaxMSP* patch, which works very much like a vocoder. Within the patch is a set of filters based on the characteristics of the sound source. These filters are triggered by the keys, and have been designed according to the phonetic sound characteristics of the performer's voice speaking every phoneme which corresponds to a key. Regarding the second audio source which fills out the filters, I used the 'timeless' version of audio files, as previously described: a sound file that has passed through a process of blurring in its time dimension. However, instead of cutting out an excerpt from it as in 5 *Timelessness Studies*, I use the entire audio file. The 'timeless' audio is always played within the piece. Once a key is pressed, the audio comes through the corresponding filter instead of going directly to the loudspeakers. If more than one key is pressed, the audio source is filtered in parallel and assigned to the correspondent loudspeakers.

I composed two *Keyboard Pieces* so far. They use the lyrics of *Gaucho de florida* (1991) for bandoneon by Marino Rivero, and *Pues que jamás olvidaros* (choir) by Juan de la Encina (1468-1529) respectively. While the 'timeless' sound file which is used for filling out the filters are recordings of these pieces in its 'timeless' version, the text which is typed by the performer is the sung text of the corresponding piece. As a result, typing attempts to 'reconstruct' the sung text by activating filters which shapes the 'timeless' version of the piece. The 'timeless' versions as second audio sources do not have a rich spectrum and energy evenly spreaded over the spectrum, which would be important in order to let the filters leave their spectral characteristics. However, an equal presence of the sound qualities of the second sound source and the filters (which are modelled according to the sound source) is intended.

Epilogue

Initially, the motivation for this project was my fascination with room resonances. Previously, I was obsessed with the idea of artificially creating a situation where a music piece would be performed in an enclosed space, but only the main resonances of the space which are activated by the music could be heard. I had also been interested in the connection between the main resonances of different rooms in a house, and how it could be involved in a piece of music or a performance. These couple of ideas brought me into impulse response measurement and convolution, which has been one of the main tools of my artistic practice in the last two years.

The majority of the pieces and experiments which I mentioned in the thesis are the result of many explorations I have done with convolution and deconvolution. The usual implementation of convolution reverberation within music production and the attempts to recover Caruso's voice by deconvolution (as discussed in Chapter 1) were the inspiration for my investigations with these tools. Convolution reverberation and the experiments on Caruso's voice took me to reflect about the notion of sound source and sound qualities and all the creative potential of convolution and deconvolution as compositional tools. I believe I succeed in exploring many creative possibilities with these tools, guided by the discussion about the distinguishment between sound source and artificial reverberation. Furthermore, apart from the pieces themselves, I provided a creative framework where different kinds of approaches to convolution, deconvolution, filter, impulse response, subvocalization and musical automata were developed and can be useful for other artists. Particularly, throughout the thesis I drew a line between the pieces in order to show the intermediate experiments between them, and through them suggesting other potential artistic paths which were undeveloped in this project.

When I started to experiment with imitations of room resonances, I developed simultaneously a modular way of structure the live electronics of my pieces. This compositional strategy was later expanded for imitating speech and finally I included performers as s modules in the compositions. The modular structure of the pieces and its effectiveness in achieving the main goals of the pieces has been discussed, and along this path I myself created awareness of the possibilities and limitations of this compositional strategy.

Subvocalization was an unexpected subject in the project which appeared while experimenting with 'impulse responses' extracted from phonemes. It brought me to composing pieces which work involving the audience by subvocalizing, following precise instructions. I believe this idea is in its infancy within my own practice as an artist and I am sure I will develop pieces which involve the inner speech of the audience in the future, which intends to reveal new types of listening by involving not only their inner speech but their internal hearing in general to the sonic experience of a piece. My intention, apart from describing the main goals of the pieces that work with the audience and subvocalization, was to frame the pieces within a broader context which involves musical automata and algorithmic compositional strategies, in order to suggest other potential developments of this project.

REFERENCES

Etkin, Mariano; Cancián, Germán; Mastropietro, Carlos; Villanueva, María Cecilia. "Cita y ornamentación en la música de Gerardo Gandini." *Latinoamérica Música*. Accessed April 9, 2019.
http://www.latinoamerica-musica.net/

Kittler, Friedrich. *Gramophone, Film, Typewriter*. Stanford, California: Stanford University Press, 1999.

Paraskevaídis, Graciela. "The Own and the Other. The Argentinian Composer Gerardo Gandini." *World New Music Magazine* no. 3 (September 1993). Accessed May 29, 2020.
http://www.gp-magma.net/pdf/txt_i/Gandini-WNMM.pdf

Psychology wiki. "Subvocalization." Last modified August 19, 2013. Accessed April 24, 2020.
https://psychology.wikia.org/wiki/Subvocalization

Real Alcázar website. "Puesta a punto del órgano de la Fuente de la Fama." Last modified November 22, 2016. Accessed May 29, 2020.
https://www.alcazarsevilla.org/en/news/puesta-punto-del-organo-la-fuente-la-fama/

Scheib, Christian. "Static's Music - Noise Inquiries." *CalArts*, (1997/1998). Accesed Friday 29, 2020.
https://ablinger.mur.at/noise.html

Smith, Steven W. *The Scientist and Engineer's Guide to Digital Signal Processing*. San Diego, California: California Technical Publishing, 1999.

Stockham, T., Cannon, T. and Ingebretsen, R. "Blind Deconvolution Through Digital Signal Processing." proc. *IEEE* vol. 63, no. 4 (April, 1975): 678-692.

Wikipedia. "Subvocalization." Last modified March 28, 2020. Accessed May 21, 2020.
https://en.wikipedia.org/wiki/Subvocalization

Wishart, Trevor. *Audible Design*. Orpheus the Pantomime Ltd, 1994.

Pieces

Ablinger, Peter. *Weiss/Weisslich 22: Haydn, Mozart, Beethoven, Schubert, Bruckner, Mahler*. Produced at IEM Graz, 1996.

Gandini, Gerardo. *Eusebius: Cuatro nocturnos para piano*. Published by the author, 1984.

Gandini, Gerardo. *Eusebius: Un nocturno para piano*. Published by the author, 1984.

Gandini, Gerardo. *Lunario sentimental.* Buenos Aires: Melos, 1989.

Tenney, James. *Three Indigenous songs*. Toronto: Canadian Music Centre, 1979.

Cd's and Youtube content

Lucier, Alvin. *I am sitting in a room*. Lovely Music, Ltd. – VR 1013, 1981.

Recording uploaded to Youtube by jive74. Water organ at Seville's Alcázar. Youtube video, 04:11. Posted August 21, 2015. Accessed April 9, 2019. https://www.youtube.com/watch?v=AEHn2O4LR8E

Appendix 1: Attached files

Davidsbündlertänze op.6 no. 14 recording convolved by an impulse response of a room which has been previously convolved with itself 4 times. This is to say that the impulse response was convolved with itself and the outcoming signal was convolved with itself and so on. As it is being done, a few frequencies are extremely boosted and the rest is almost removed. This audio file can be found in the attached audio files of the thesis. (Chapter 2)

5 Timelessness Studies—detailed procedure (Chapter 3)

- 'Timeless' audio 1
- 'Timeless' audio 2
- 'Timeless' audio 3

5 Timelessness Studies (Chapter 3)

- *Timelessness study #1: Deo gratias*
- *Timelessness study #2: Vltava*
- *Timelessness study #3: La piojosa*
- *Timelessness study #4: Sonata K 333 in Bb*
- *Timelessness study #5: La negrina*
- *Timelessness study #6: Retirada*

Reception Bells (Chapter 5)

Reception Bells performed at Sonology Discussion Concert (2019)

Appendix 2: *Reception Bells*
text to be read by the performers and handed out to the audience

Subvocalization, or silent speech, is the internal speech typically made when reading; it provides the sound of the word as it is read. This is a natural process when reading, and it helps the mind to access meanings to comprehend and remember what is read, potentially reducing cognitive load. This inner speech is characterized by minuscule movements in the larynx and other muscles involved in the articulation of speech. Most of these movements are undetectable (without the aid of machines) by the person who is reading. The exploration into the evolutionary background of subvocalization is currently very limited. The little known is predominantly about language acquisition and memory. Evolutionary psychologists suggest that the development of subvocalization is related to modular aspects of the brain. There has been a great amount of exploration on the evolutionary basis of universal grammar.

The world gives you itself in fragments / in splinters:
in 1 melancholy face you glimpse 1 brushtroke by Dürer
in someone happy the grimace of 1 amateur clown
in 1 tree: the trembling of birds sucking from its crook
in 1 flaming summer you catch bits of the universe licking its face
the moment 1 indescribable girl
              rips her Oaxacan blouse
just at the crescent of sweat from her armpits
& beyond the rind is the pulp / & like 1 strange gift of the eye
                                                            the lash

Maybe not even Carbon 14 will be able to reconstruct the true facts
The days are gone when 1 naturalist painter
could ruminate over the excesses of lunch
between movements of Swedish gymnastics

scores developed by the performers

A

I

i I aI ɔI Iə

E

eI e eə 3

Reception bells

Subvocalization, or silent speech, is the internal speech typically made when reading; it provides the sound of the word as it is read. This is a natural process when reading, and it helps the mind to access meanings to comprehend and remember what is read, potentially reducing cognitive load. This inner speech is characterized by minuscule movements in the larynx and other muscles involved in the articulation of speech. Most of these movements are undetectable (without the aid of machines) by the person who is reading. The exploration into the evolutionary background of subvocalization is currently very limited. The little known is predominantly about language acquisition and memory. Evolutionary psychologists suggest that the development of subvocalization is related to modular aspects of the brain. There has been a great amount of exploration on the evolutionary basis of universal grammar.

The world gives you itself in fragments / in splinters:
in 1 melancholy face you glimpse 1 brushtroke by Dürer
in someone happy the grimace of 1 amateur clown
in 1 tree: the trembling of birds sucking from its crook
in 1 flaming summer you catch bits of the universe licking its face
the moment 1 indescribable girl
            rips her Oaxacan blouse
just at the crescent of sweat from her armpits
& beyond the rind is the pulp / & like 1 strange gift of the eye
                                                                                the lash

Maybe not even Carbon 14 will be able to reconstruct the true facts
The days are gone when 1 naturalist painter
could ruminate over the excesses of lunch
between movements of Swedish gymnastics

Reception bells

A O
æ ə a: oʊ ʌ ɔɪ ɔ:

Subvocalization, or silent speech, is the internal speech typically made when reading; it provides the sound of the word as it is read. This is a natural process when reading, and it helps the mind to access meanings to comprehend and remember what is read, potentially reducing cognitive load. This inner speech is characterized by minuscule movements in the larynx and other muscles involved in the articulation of speech. Most of these movements are undetectable (without the aid of machines) by the person who is reading. The exploration into the evolutionary background of subvocalization is currently very limited. The little known is predominantly about language acquisition and memory. Evolutionary psychologists suggest that the development of subvocalization is related to modular aspects of the brain. There has been a great amount of exploration on the evolutionary basis of universal grammar.

The world gives you itself in fragments / in splinters:
in 1 melancholy face you glimpse 1 brushstroke by Dürer
in someone happy the grimace of 1 amateur clown
in 1 tree: the trembling of birds sucking from its crook
in 1 flaming summer you catch bits of the universe licking its face
the moment 1 indescribable girl
rips her Oaxacan blouse
just at the crescent of sweat from her armpits
& beyond the end is the pulp / & like 1 strange gift of the eye
the lash

Maybe not even Carbon 14 will be able to reconstruct the true facts
The days are gone when 1 naturalist painter
could ruminate over the excesses of lunch
between movements of Swedish gymnastics