Designing embodied human-computer interactions in music performance

Balandino Di Donato

A Thesis presented for the degree of
Doctor of Philosophy

Royal Birmingham Conservatoire
Faculty of Art Design and Media
Birmingham City University
United Kingdom

February, 2020

Dedicated to

My family

Abstract

Interfaces for musical expression are widely used for controlling and transforming sound in live performance. They aim to facilitate the interaction with a computer and empower the performer with a more expressive control over the sound. However, the actions made to control them have the potential to interfere with the musical performance, in relation to the instrumental technique, choreographic aspects or the physical characteristics of the played musical instrument.

To avoid this issue, modes of interaction and various devices have been designed and utilised in conjunction with interactive audio and visual software to control and transform audiovisual media. In particular, gesture sensing technologies have been successfully used in different musical applications. However, they, in turn, raise questions such as, how can musicians most effectively control and transform auditory, visual and lighting effects during a live performance through gesture? What interaction design considerations should be made that allow performers to interact simultaneously with an instrument and audio-visual-lighting processing? How can disruption during a live performance with embodied human-computer interactions be reduced?

The work presented in this thesis investigates modes of interaction with sound, visual projection and lighting effects during a musical performance that may result natural and embodied, and not dependent from a particular musical instrument, its sound or instrumental technique. For this purpose, using a User-Centred Design method, I realised 'MyoSpat' upon Music and Human-Computer Interaction principles. MyoSpat is an interactive system, which embeds Inertial Measurement Unit (IMU) and Electromyography (EMG) technology, for gesturally controlling audio and lighting processes during a musical performance. As part of this research, I also created Myo Mapper, a Thalmic Labs' Myo to Open Sound Control (OSC) messages mapper.

Outcomes of this research are presented in this thesis and through a portfolio of performances realised in collaboration with musicians.

Declaration

The work in this thesis is based on research carried out at the Royal Birmingham Conservatoire, Birmingham City University, United Kingdom. No part of this thesis has been submitted elsewhere for any other degree or qualification, and it is all my own work unless referenced to the contrary in the text.

Related publications

Chapter 4 may include fragments of text from:

- Di Donato, B. and Bullock, J. (2015) gSpat: Live sound spatialisation using gestural control. In *Proceedings of the International Conference on Auditory Display, Student ThinkThank*, ICAD 2015 STT, pages 7-8, Graz, Austria.
- Bullock, J., and Di Donato, B. (2016) Approaches to Visualising the Spatial Position of 'Sound-objects'. In *Proceedings of the Conference on Electronic Visualisation and the Arts*, EVA '16, pages 15-22, London, United Kingdom.
- Di Donato, B., Bullock, J., and Bledsoe, J. (2017) Myo Mapper, Myo armband to OSC mapper. In *Audio Developer Conference 2017*, ADC17, London, UK.
- Di Donato, Bullock, J., and Tanaka, A. (2018) Myo Mapper: a Myo armband to OSC mapper. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME' 18, pages 138–143. Blacksburg, Virginia, USA.
- Di Donato, B. and Michailidis, T., (2019) Accessible interactive digital signage for visually impaired. In *ACM CHI 2019 Workshop on Mid-Air Haptics Interfaces for Interactive Digital Signage and Kiosks.* Glasgow, UK.
- Di Donato, B., Arterbury, T. (2019) Embodied interaction with sound. In *Audio Developer Conference 2019*, ADC19, London.

Chapter 5 may include fragments of text from:

- Di Donato, B. and Bullock, J. (2015) gSpat: Live sound spatialisation using gestural control. In *Proceedings of the International Conference on Auditory Display, Student ThinkThank* ICAD 2015 STT, Graz, Austria, pages 7–8
- Di Donato, B. and Bullock, J. (2016) xDbox: A System for Mapping Beatboxers' Already-Learned Gestures to Object-Based Audio Processing Parameters. In *Abstracts of Porto International Conference on Musical Gestures As Creative Interface*, pages 49–50, Porto, Portugal.

Chapter 6 may include fragments of text from:

- Di Donato, B., Dooley, J., Hockman, J., Bullock, J., and Hall, S. (2017) MyoSpat: A hand-gesture controlled system for sound and light projections manipulation. In *Proceedings of the International Computer Music Conference*, ICMC 2017, pages 335–340, Shanghai, China.
- Di Donato, B., Dooley, J. (2017) MyoSpat: A system for manipulating sound and light through hand gestures. In *Proceedings of Workshop on Intelligent Music Production*, WIMP, Salford, United Kingdom.
- Di Donato, B., Dooley, J, Coccioli, L. (2020) HarpCI, Empowering Performers to Control and Transform Harp Sounds in Live Performance. In *Contemporary Music Review*, (38)6:667–686, DOI: 10.1080/07494467.2019.1706351

Related Musical Performances

Chapter 5 includes the following performances:

- De Amicis, V. and Di Donato, B. (2015) VoicErutseG. Frontiers Festival. Available from `https://youtu.be/xUo1TyymAQc` and `https://youtu.be/r7_NN8ppu3g`. Accessed 5 Jannuary 2021. Birmingham, United Kingdom, 17 March.
- De Amicis, V. and Di Donato, B. (2015) VoicErutseG. EmuFest. Available from `https://youtu.be/pQOY-YsKHPY` and `https://youtu.be/gdhnKbxUkUg`. Accessed 5 Jannuary 2021. Rome, Italy, 7 October.
- De Amicis, V. and Di Donato, B. (2015) VoicErutseG. ElectroAQustica 2015, L'Aquila, Italy, 14 June.
- Savage, G. and Di Donato, B. (2016). Music Gesture Beatbox. Studio Recording, Birmingham, United Kingdom, Available from: `https://youtu.be/DRFqCXpvfW0`. Accessed: 23 May 2016.
- Savage, G. and Di Donato, B. (2016). Music Gesture Beatbox. Music Tech Fest. Berlin, Germany, 29 May.

Chapter 6 includes the following performances:

- Turner, E. and Di Donato, B. (2017) The Wood and the Water. Studio Recording, Birmingham, United Kingdom, Available from: https://youtu.be/gYu4Za-1E48. Accessed: 17 February 2017.
- Turner, E. and Di Donato, B. (2017) The Wood and the Water. Birmingham Harp Day 2016. Birmingham, United Kingdom, 20 January.
- Turner, E. and Di Donato, B. (2017) The Wood and the Water. Cardiff Camac Harp Weekend. Cardiff, United Kingdom, 4 April.
- Turner, E. and Di Donato, B. (2017) The Wood and the Water. HarpCI Workshop. Northampton, United Kingdom, 15 May.
- Turner, E. and Di Donato, B. (2017) The Wood and the Water. Music Program of the Audio Mostly 2017 Conference. Available from: https://audiomostly.com/2017/program/music-program/ and https://youtu.be/3n_2y3OerVQ. Accessed 5 January 2021. London, United Kingdom, 25 August.
- Turner, E. and Di Donato, B. (2017) The Wood and the Water. Electronic Music Week 2017. Available from: https://youtu.be/p0_VR9N4x8w. Accessed 2 Jannuary 2021. Shanghai, China, 19 September.
- Devaney, K. Turner, E. and Di Donato, B. (2016). Star Cluster. Studio Recording, Birmingham, United Kingdom, Available from: https://youtu.be/9ToP33Ki2SE. Accessed: 30 September 2017.

Acknowledgements

I would like to thank my former supervisor, Dr Jamie Bullock, for giving me a life-changing opportunity and making me rethink the way I see music, sound and technology. I also thank my supervisors Prof Lamberto Coccioli, Dr James Dooley and Dr Jason Hockman for their support while writing this thesis. I am also grateful to Birmingham City University for the award of the funding that enabled me to undertake this work. I am profoundly indebted with Prof Atau Tanaka for his support and counselling during the preparation to my Viva Voce, and Prof Andrew Hugill for his mentorship help during the last editing of this thesis.

I owe a debt of gratitude to Vittoriana De Amicis, Grace Savage and Eleanor Turner for the great moments spent together performing around the world. Thanks to all participants who took part in interviews and user studies for their immense contribution to my research work. I want to express my gratitude to Berklee College of Music in Valencia, Camac Harp, Cardiff Metropolitan University, and Southampton University for hosting and supporting workshops that informed this work. My gratitude is also extended to the EAVI research group resident at Goldsmiths, University of London for giving me the opportunity to present and assess Myo Mapper during the EMG workshop funded by the European Union's Horizon 2020 programme, H2020-ICT-2014-1 Project ID 644862, *RapidMix*. I am deeply grateful to those who helped me in realising small parts of this work: Emanuela Mentuccia for the lighting design of *The Wood and the Water*'s performance at Shanghai Symphony Hall and Jefferson Bledsoe for his work on

Myo Mapper v3, Niccolo Garnieri, Matt Crowther, Matt Canty for helping me in realising the video recording of *Music Gesture Beatbox* and the Royal Birmingham Conservatoire's Music Technology Department for supporting my experiments, concerts and conferences. Thanks to Thalmic Lab for building the Myo armband, to Miller Puckette for writing Pure Data, Cycling 74 for creating Max, Roli for the JUCE framework, Rebecca Fiebrink for developing Wekinator and Jamie Bullock and Ali Momeni for creating ml.lib; I could not have realised this work without these tools.

I could not have survived the past three years without climbing and being outdoors. Thanks to the existence of Reiki, hang and trip-hop music, I sometimes remembered to slow down and enjoy doing this work. I think it is essential that I thank my long-term friends and companions for the nights out, travels and the tasty meals. I welcome this opportunity to thank my friends based in Italy, who supported me when I first moved abroad and made me feel like I never left when visiting.

Most importantly, I would not have been able to undertake this endeavour without the knowledge provided by previous professors and supervisors at Conservatorio Alfredo Casella and Centro Ricerche Musicali di Roma (CRM). A very special thank you goes out to Michelangelo Lupone for planting in me many dreams.

Thanks to my family, the love of my life. They have always been there to support me, reminding me that I could achieve anything and pushing me to realise my dreams.

Contents

List of Figures

List of Tables

List of Equations

Chapter 1

Introduction

1.1 Research motivations

During my undergraduate studies in Contemporary Electronic Music Composition, I worked on the composition and performance of various works for music with live electronics. At the end of that period of study, I presented a dissertation on the development and use in live music performance of an audio-visual engine for the Tangible User Interface (TUI): “Metis”.

Metis was created by the Nuova Musica per l’Educazione (NUME) group and presented at the Festival of Science 2010 in Genoa (Italy). The purpose of this interface was to offer a new interactive educational tool for musicians. The design and development of the GUI and the audio-visual engine to support the composition and performance of *Il Sogno nel Sogno* (2013), for tape and Metis by the author. The performance required a complex and fine control of the audio materials. Adopting a multimodal approach, I implemented complex mapping functions to control audio processing parameters through touch hand-gestures on the table and by moving plastic disks with fiducial markers attached to their bottom. The movement and orientation of the fiducials were tracked using a camera and the reacTIVision framework (Kaltenbrunner, 2013). Fiducials’ data were then transferred via the

Open Sound Control (OSC) protocol to two applications, one developed using Processing (Kaltenbrunner and Reas, 2020) for rendering the GUI, and another one for processing the sound developed using Max (Cycling '74, 2007).
With the work on the Metis, I learnt the importance of modalities of interaction with technology during a performance. Through the dedicated user interface, and a mapping strategy designed and implemented upon the performance needs, Metis facilitated the sound control and transformation of audio-visual materials through a more "natural" and "intuitive" interaction with technology. However, like many other TUIs, Metis does not always enable musicians to process the sound easily during a live performance. A musical performance often requires constant contact with a musical instrument through the hands. Thus, in the case of a live performance for a traditional musical instrument (i.e. brass, strings, woodwind, percussion, keyboard instruments) and electronics, the need to control and manipulate the sound through a touch interface may disrupt the musical performance or affect the instrumental technique.

After my graduation in Italy, I moved to Birmingham, United Kingdom, where I worked at the Royal Birmingham Conservatoire's Integra Lab as a debugger and programmer for the development of Integra Live (Integra Lab, 2013). Integra Live is a software that aims to facilitate the realisation of compositions and musical performances using interactive controlled audio processing. This software puts musicians at the centre of the design, by empowering them to take advantage of programming environments like Pure Data, Lua scripting and complex data mappings through an easy-to-use interface.
In this period, I learnt the importance of developing musician-centred, easy-to-use, intuitive and useful applications for users with no previous training in technology. I then decided to start on a new journey towards exploring natural and embodied interactions with sound, visual and lighting effects.

1.2 Context

This work proposes to musicians and researchers in the field of Music and Human-Computer Interaction (HCI) new ways of controlling sound, visual projections and lighting effects. This research brings the skills of musicians and researchers together to attain a greater understanding of musicians' interaction with their instrument and technology, to then find new ways to perform and express sound, visual and lighting transformations. This work explores a number of interaction design solutions and audio-visual-lighting processing possibilities.

To create a successful system and the relative modes of interactions with it, Tanaka and Knapp (2002) suggest that it should have certain characteristics, such as: satisfying both the performers and the listeners, empowering performers with a sense of *"articulative freedom and expressivity"*, offering an intuitive way of interacting with the system, allowing performers to express their musical ideas freely, and providing the required control over the musical instrument. According to Nielsen (1993), a system is usable if it is *"easy to learn"*, *"efficient"*, *"easy to remember"*, *"relatively error-free"* and *"pleasant to use'*. After analysing aims, requirements, challenges and solutions for creating *"sensing"* user-interfaces in the field of HCI, Bellotti et al. (2002) concluded that a successful system should address the questions: *"how do I address one (or more) of many possible devices? How do I know the system is ready and attending to my actions? How do I affect a meaningful action, control its extent and possibly specify a target or targets for my action? How do I know the system is doing (has done) the right thing? How do I avoid mistakes?"*. All these factors are important to design successful interactions with an interactive system. However, they do not consider a vital aspect of music performance: the audience.

Looking at systems from the audience perspective, Reeves et al. (2005) identify four design strategies built upon the question: *"how should a spectator experience a user's interaction with a computer?"*. In particular, they categorised the musician's use of a device (*manipulations*), and consequent sound transformations (*effects*). These can then be described as *hidden*, *partially hidden/transformed*, *revealed* or even *amplified*. From here, they identify four design strategies: *secretive* (*manipulations* and *effects* are *hidden*), *expressive* (*manipulations* and *effects* are *revealed*), *magical* (*effects* are *revealed* but not their *manipulations*) and *suspenseful* (*manipulations* are *revealed*, but effects only are *"revealed when the spectator gets to take their turn as a performer"*). These are grouped in a bi-dimensional plane represented in Figure 1.1.

Using the taxonomy of Reeves et al., in this work, I aimed to design interaction modalities with audio, visual and lighting effects that are *expressive*, so as to enable musicians to generate *amplified effects* and *manipulations*.

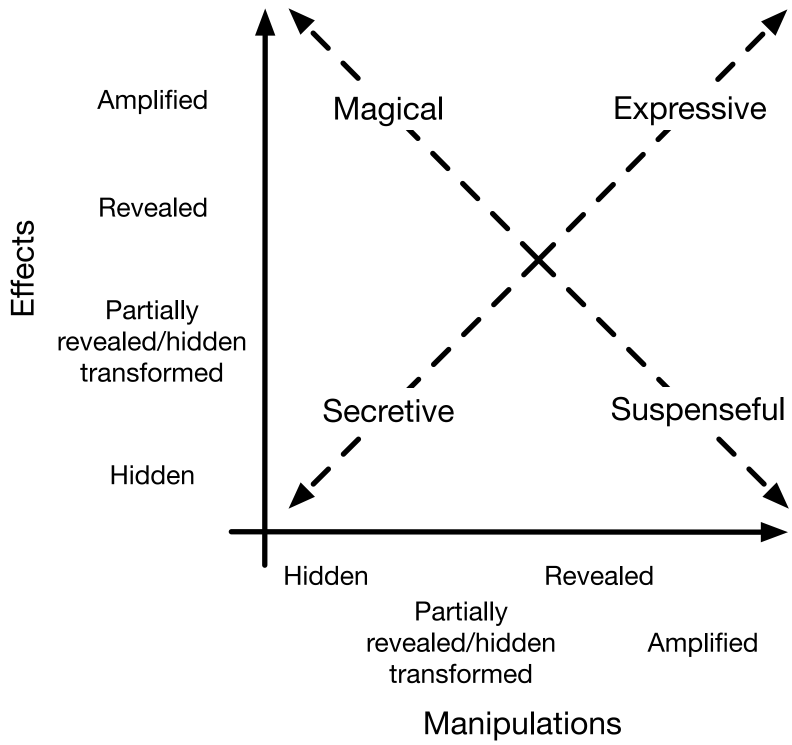


Figure 1.1: "Secretive, expressive, magical and suspenseful approaches to designing the spectator's view" (Reeves et al., 2005).

In the context of this research, it is important to consider the idea described by Leman (2007), that interfaces and modes of interactions should put the user in an ideal position where the interface itself disappears. It is necessary to specify that here the word "disappear" does not aim to promote the idea of an invisible system, but it refers to the interaction between the human and the machine (Dourish, 2001). Thus, the musician's and the audience's perception should not focus on the mechanics of interaction. Nevertheless, *"a musical communicative channel should be established that is catalysed by the modes of interaction, but not hindered by them"* (Tanaka and Knapp, 2002).

This research is characterised by musical works composed with computer-generated and transformed sounds, the which are commonly a family of sounds that are hard to associate with the way they have been created and the source that produced them. In this particular case, the interaction design plays a crucial role in establishing a relationship between the music being produced and the act of performing it. Interaction design is, therefore, a vital aspect of attaining musical satisfaction.

Body movements during musical performance enhance the experience of music, and the embodiment of the interaction with sound plays an important role to *"facilitate the encoding of expressive gestures into sounds, and the decoding of sounds into expressive gestures"* (Leman and Maes, 2015). Together with sound, gestures have *"the capacity to convey an emotion, a sentiment, a message, and many other things"*, in a context where *"expressiveness can be associated with physical gestures, choreographic aspects of the sounds resulting from physical gestures"* (Arfib et al., 2005).

Acknowledging the aforementioned research, I then welcomed the possibility to use body movements as means to design interactions with sound, visual and lighting process so that they result embodied.

1.3 Research aims

This research intends to address the following fundamental questions:

- How can musicians most effectively control and transform auditory, visual and lighting effects during a live performance through gesture?
- What interaction design considerations should be addressed that allow performers to interact simultaneously with an instrument and audio-visual-lighting processing?
- How can disruption during a live performance with embodied human-computer interactions be reduced?

This research will benefit musicians that aim to express their artistic ideas by controlling and transforming live sounds, visual and lighting effects through interactions with the computer that are "natural" and "embodied" in the context of musical performance; natural *"as being marked by spontaneity"* (Grandhi et al., 2011), and embodied as an extension and incorporation of human skills and abilities within the interaction design of a system (Dourish, 2001).
This work aspires to make the act of interactive performance natural by taking into account the performer's existing instrumental technique, their body movements and the physical and timbral characteristics of their instrument.

1.4 Contributions

This thesis aims to contribute to the multidisciplinary field of Music and Human-Computer Interaction. In this thesis, various modes of interaction with the computer, and their application in real-world scenarios such as music composition, rehearsal and performance are presented. This knowledge and the artefacts developed to answer the research questions contributes to the musical practice involving technology.

The experience gained in this work results in a portfolio of performances, and number of recommendations for the design of interactions with audio, visual and lighting processing parameters, taking into account composition and performance needs, as well as considering their movements in different situations, namely either being free to move or constrained by a musical instrument. Moreover, the MyoSpat system – a gesture-controlled interactive system for the live processing of audio, visual, and lighting effects during a performance – was created as enabling technology with the purpose to explore the research questions.

Performances and technological solutions are here the outcomes of a study focused on Music and HCI conducted through an iterative creative process, where the collaboration with musicians have informed the interaction design solutions and the development of MyoSpat.

All artefacts produced with this thesis aim to foster the adoption of embodied interaction through the use of body sensing devices in a wide range of performative creative computing applications.

1.5 Thesis structure

This research work is described in the next six chapters, the following: Literature Review, Methodology, Pilot Studies, Case study 1: Voice Gestures, Case Study 2: HarpCI and the Conclusion.

Chapter 2, Literature Review provides a landscape of the field of Music and Human-Computer Interaction, Interaction Design, Embodied Interaction and Machine Learning. Also included here is a review of technologies for gestural control of audio processing during musical performance. This technology has been analysed in the context of vocal performance, as well as instrumental performance, with a particular focus on applications using sound spatialisation.

Chapter 3 describes the methodology adopted to carry out the research described in this thesis. The User-Centred Design method, with the musician at the centre of the design, was utilised to find interaction design and technical solutions, to then support a number of performances.

Chapter 4 presents Myo Mapper, a software designed and developed by the author for using the Thalmic Lab's Myo as gestural controller device for the MyoSpat system. In this chapter, a series of pilot studies with the objective of exploring the potential of the Myo and different modalities of interaction with processing parameters are also reported.

In each of the next two chapters, two case studies are presented. The first case study, Voice Gestures (Chapter 5), describes two musical works, namely *VoicErutseG* (2015) (pronounced Voice Gesture) performed by Vittoriana De Amicis, and *Music Gesture Beatbox* (2016) composed and performed by Grace Savage. The former is focused on the control of sound parameters in a context in which the performer can enact free-space gestures without any constraint; in this case, singers using a wireless headset microphone system. In the latter musical work, the research was oriented towards supporting the interaction with audio, visual and lighting processing parameters, when the musician's movements are restricted by the use of a microphone to be held and the use of a loop-station device.

The sixth chapter describes the HarpCI case study. The objective of this case study is to apply and extend previous understandings, to a wide range of instrumental performances, with particular focus on harp performance. The HarpCI case study led to the creation of two music pieces: *The Wood And The Water* (2017) composed and performed by Eleanor Turner and *Star Cluster* (2017) composed by Kirsty Devaney and performed by Eleanor Turner.

In each case study, the iterative development of the MyoSpat system and an in-depth formal evaluation of the interaction design, MyoSpat and the user experience are presented in Chapter 6.

Finally, a summary of this work, the answer to the research questions, their limitations, and new questions emerging from the research are presented in the last chapter.

Chapter 2

Literature review

2.1 Introduction

This work builds upon findings from the field of Music and Human-Computer Interaction (HCI), defined as Music Interaction by Holland et al. (2013). This field looks at the design, use and evaluation of interactive systems in any musical activity and scientific aspects. There are certain elements of the field of Music and HCI that relate to my work, such as User-Centred Design (UCD), Interaction Design and the application of Embodied Interaction Design principles to music technology research. This work is also interconnected with the literature on design and use of devices and modes of interactions by researchers and practitioners for the control of interactive systems, with a focus on the implications of mapping strategies for the different interfaces.

Relevant literature is reviewed in the following sections.

2.2 Music and Human-Computer Interaction

Previous research has identified three phases in the history of HCI, called "waves", defined by paradigms built upon interaction metaphors (Harrison et al., 2007). The first wave was inspired by industrial engineering and ergonomics, and its paradigm focuses on the man-machine interaction, intending to optimise the interaction and solving related issues. The second wave looked at how our mind and the machine couple and interact, answering questions about how we perceive the information given by the computer, how we process them, how we react to them, and how we could improve this cycle. The third wave of HCI investigates questions such as: what existing situated activities in the world should we support? How do users appropriate technologies, and how can we support those appropriations? How can we support interaction without constraining it by what a computer can do or understand? What are the politics and values at the site of interaction, and how can we support those in design?

In the first and second HCI waves, researchers directed their work with a focus on engineering aspects of the interaction with a system and formal methods to evaluate it. In my work, I support contrasting works that show primary appreciation for the complexity of the real world and the activity at the centre of the design. In my work, as in Norman (2002), it is important defining the context in which culture interacts with HCI aspects and to systematically analyse them, and that an objective measure of system behaviour it is not sufficient to evaluate its acceptability. I also agree with Norman's deduction that a more holistic method that understands how people experience and judge information systems is needed. These concepts are relevant to my work to the extent that the evaluation of the human-computer interaction during a performance does not depend on the interface's behaviours, but also on creative possibilities that it promotes. For example, if and how it inspires and support the creative process of music creation and performance.

Sengers and Gaver (2006) argue that synergies with the arts, humanities, domestic and public environments field of studies promoted a shift from a design measured by the extent to which a single user interprets the system, to designing products for multiple interpretations, allowing users to appropriate and reinterpret the system.

Influenced by Sangers and Gaver's work, I design artefacts that can be reinterpreted by musicians according to the requirements of their artistic practice. In my work, technology and modes of interaction aim to enable complex and rich creative processes, rather than impose a specific workflow.

HCI principles have influenced substantially the work of music practitioners and researchers working with technologies. According to Hsu and Sosnick (2009), HCI methods should be used to address: (i) the usability in the context of performance, and (ii) whether such performance is musically interesting for an audience that is sympathetic to free improvisation. Although these methods provide a valid framework to evaluate the interface and its application during a performance, I find myself in agreement with Kiefer et al. (2008) who states that focusing on them only would exclude an important component in the evaluation of a device used during a performance: the musician's experience.

A multitude of different quantitative and qualitative methodologies have been adopted to evaluate New Interfaces for Musical Expression (NIMEs). These included musical tasks (Orio et al., 2001), questionnaires, interviews, observations, statistical analyses targeting the whole digital musical instrument (DMI), or single parts of it (input, output, mapping, feedback), from different perspectives (audience, performer, designer) and using different criteria (i.e., expressiveness, learnability, intuitiveness, fun, enjoyment, control) (Barbosa et al., 2015). In their review of the methodologies used in NIME works between 2012 and 2014, Barbosa et al. conclude that these are valuable methods, but there is no "one-size-fits-all" method that allows the evaluation of DMIs.

Different approaches to designing interfaces for musical expression have been taken over the three waves of HCI. In the case of the third wave, the user and the context in which she/he operates are of fundamental importance to determine design solutions. In the context of this thesis, the collaboration with musicians to realise musical performances is of relevant importance to create design solutions. In the following section, I review aspects of designing music interfaces with reference to the User-Centred Design methodology.

2.2.1 Music and User-Centred Design

The expression User-Centred Design (UCD) was formalised by Norman and Draper (1986) in their work entitled *User Centered System Design: New Perspectives on Human-computer Interaction*. They aim to shift the focus of designing systems towards people, rather than the technological artefact. They used a pluralistic approach that does not look only at solving singular tasks using a product, but considering the context in which it will be used, the social interactions that it might enhance or disrupt, the way we interact with these products taking into account bodily actions and emotions when perceiving their feedback or achieving the objectives. In a later work focused on the usability of a product, Norman et al. (1988) create guidelines including also the use of prior knowledge to simplify the structure of a task, to make the product and its feature visible and understandable, to foresee errors and to provide ways for fixing potential mistakes.

The UCD method has been extensively used by the HCI community and adopted by technologists in the field of music. In creating Musink, a tool that would ensure composers a smoother transition from a paper draft to the music software OpenMusic, Tsandilas et al. (2009) based their method on the work by Mackay et al. (1998), who observed workers on a daily basis gathering information about the way they operate and their work environment to inform the development of an interactive paper interface for controlling air traffic controls. Similarly, Tsandilas et al. worked in close collaboration with musicians to understand their processes and the role of working with paper in such a creative context. Through an exploratory evaluation conducted after the system development as a series of small workshops with individual composers, they learned that each participant had a unique approach and that they mix the use of paper and computers in different ways. They would

first create structures, which they would then fill in with details and subsequently go back to change them. So, in their later iterations, they added to their system the possibility to create structures and then play with them.
Another example is the work by Bau et al. (2008). They delivered a "participatory design workshop" to create a polyhedron-shaped, multi-channel audio input/output device. Sixteen non-technical users took part in two workshops. The first one was conducted through an interview-based exploration aimed at understanding participants' personal music player usage, followed by a structured brainstorming on communicating entertainment devices. The second workshop undertook a task-based evaluation of the perceptual characteristics of a multi-faceted multi-channel audio device and how it could be used as a technology probe and a musical instrument. Through this process, they designed an expansive platform for generating and exploring new types of interaction with sound.
The approach adopted in the cited works mentioned above allowed the authors to gather very useful information to the development of an interface that would satisfy the musician. In regards to my work, this approach offers the possibility to understand the musician, how they interact with their musical instrument and other technological artefacts. Thus, it can help answering the research questions of (i) how it is possible to design modes of interaction that allow an effective control of auditory, visual and lighting effects, (ii) to facilitate the simultaneous interaction with both the instrument and the computer, and finally (iii) to reduce the disruption of the musical performance as a result of the introduction of technology in musical performance.

Music is performed using different instruments, and it comes in different genres. Thus, in the UCD, it should be considered the possibility for any musician to adapt to the interface. According to Mackay (1991), "the use of software is a co-adaptive phenomenon, in which users not only adapt to new software but also appropriate it and adapt it to their needs". She writes that through studying this phenomenon, it is possible to find new ways of using a certain design solution, challenging basic assumptions about the purpose of the software and informing further design innovations.

Intending to reduce the degree to which musicians must adapt to the interaction design, her findings are relevant to this thesis. It is through observing musicians and how they adapt to the interface that I can iterate design solutions that reduce the degree of adaptability. One more consideration to make on this matter is that musicians' adaptation to the interaction with a system has the potential to develop a new instrumental technique or inspire musical idea around a specific feature of the system. Thus, it is necessary to consider if musicians' adaptation is an issue, for example, if it distresses the musical performance, or it is a feature of the system that should be highlighted so to stimulate their creativity.

2.3 Music and Interaction Design

2.3.1 Sound Affordances

In designing modes of interaction with an interface, we should consider various human factors such as the music genre we want to play, the instrument that the musician already plays, technological factors like available sensing technology, the computational power of our machines (Cook, 2001), and the affordances of an interface. This last aspect is particularly relevant to my work as affecting the interaction design. The term affordance was first used by the psychologist Gibson (1966), and its principles lie on the possible action that each object invites based on the characteristics of the objects and the capabilities of the subjects (Gibson, 1977, 1979).

Gaver (1991) expands the concept of affordances, separating them from the perceptual information available we have of them. This led him to distinguishing them between *correct rejections* (when there is no affordance and it is not perceived)

and *perceived* (the affordance is present and perceived), *hidden* (the affordance is present but it is not perceived) and *false affordances* (an affordance is perceived but it does not exist) (see Figure 2.1).
Adopting the taxonomy suggested by Gaver, sound has hidden affordances in the sense that we often cannot perceive them, or that they can rely on previous knowledge distant from our cultural background.

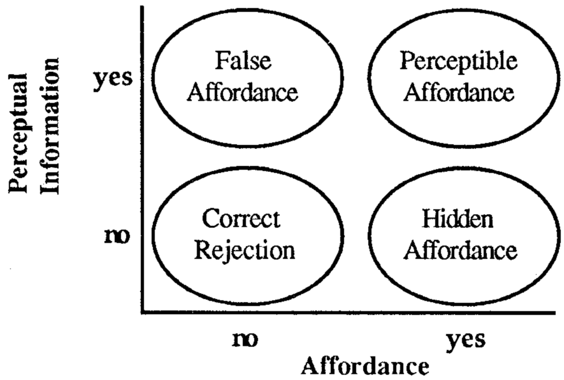


Figure 2.1: Distinction between different types of affordance. Adapted from Gaver (1991).

Godøy (2010) demonstrates that music can invite certain gestures that are often encouraged by timbral and dynamic qualities of the sound, by mimicking the action that might have produced them or gestures evoked by the music which not necessarily refer to the production or sound qualities. From here derives the concept of *affordance of musical sound.* Godøy et al. (2006) studied these aspects through a sound-tracing method in which participants were asked to trace gestures along with sounds. From this work, they conclude that the majority of gestures are evoked by sound morphology. For example, a sound with an ascending pitch would be described with a gesture tracing an ascending curve. Caramiaux et al. (2011) repeated the same experiment exploring differences between gestures evoked by causal and non-causal sounds, where the "causality" is the degree within which the listener can distinguish the sound's environmental cause. Results from their study showed that people's gestures aimed to mimic the cause of the sound but were inconsistent across participants. On the contrary, participant gestures were similar when tracing gestures along with non-causal sounds, which described the sound morphology.

Although these studies report important information about the relationship between gesture and sound, they do not consider the impact of gesture tracking device's affordance has on the execution of the sound-tracing exercise. In the case of my work, this aspect is particularly important, as the affordance of the interface can result in having a negative impact on musical performance.

This issue was taken up by Tanaka et al. (2012a) who explored gestural affordance with sound in relation to different devices. They realised a study in which participants were asked to use an Axivity Wax, iPhone and Wii Remote, to generate three different audio feedbacks: (i) to trigger a snare drum and a bowed violin sound, (ii) to control a granulation effect over violin sample and modify its pitch and loop speed, and (iii) to control the same granulation fed with a voice sample in a loop. From the results, they concluded that the combination of input devices and output feedback leads to the construction of highly complicated and sophisticated affordances, which require an equal complex construct of affordance.

Referring to the principle of *affordance of musical sound*, HCIs have the potential to foster of those processes at the origin of the sound, and lighting effects during a performance, so to imbue natural and direct interactions with the parameters that control them. But at the same time, it is important to consider the sensing modalities and affordances of the interfaces. However, differently than in Tanaka, Altavilla and Spowage's work, after a review of interactive systems in musical performance (Section 2.4), choosing an interface that minimises the implications related to its affordances and the disruption during live performance through the embodied human-computer interactions are an important consideration of this research.

2.3.2 Embodied Music Interaction

The concept of embodied interaction has its origins in the field of philosophy. In his *Meditations* (1641), Descartes worked on the assumption that everything can be doubted and that reason is the only thing to be trusted. He split the world into two models: a mental domain to which belongs the will and consciousness of a body, and a physical domain ruled by natural laws and that includes the human body. To escape from this Cartesian dualism of body and mind, the philosophers Martin Heidegger and Maurice Merleau-Ponty stated that we exist in the physical world and become aware of ourselves through interaction with the objects and subjects around us. In his major work entitled *The phenomenology of perception* (Merleau-Ponty, 1945), Merleau-Ponty introduced many of the principles today used in HCI. Svanæs (2000) summarises these as follows:

- Perception requires action. The experience of the interactivity of an object can happen through action. The object has the potential of being interactive, but it is through the interaction with the object that its interactivity appears to the user.
- Perception is governed by "pre-objective" intentionality. In perceiving the object, our actions are driven by "embodied" intentionality towards it. Perception hides this process from us going towards this "pre-objective" realm.
- Perception is embodied. Being perception activated through bodily action with an object, the perception here becomes embodied. *"The body is our general medium for having a world"* (Merleau-Ponty (1945) p. 146).
- Perception is an acquired skill. Being perception the results of the act of interacting with objects, to a large extent, it makes it a skill that can be acquired.
- The perceptual field. In Merleau-Ponty's work, the perceptual field, our first interpretation of what we perceive, is defined by our previous experience.
- Tool use. When we learnt to use a tool, it becomes part of our body "both as potential for action and as medium for perception" (Svanæs, 2000) p.89.
- Bodily space. We can see the world in two ways, as an object in the "external" world, or through our experiencing/living body. In this last case, we move within the constraints of our body.

- Abstract vs concrete movement. For Merleau-Ponty movement made on purpose are abstract movement, which they become concrete when they are made as part of a situation, it becomes concrete.

In reference to Svanæs' work, the interaction with an instrument for electronic music performance – intended as an open-ended system composed of different parts, the input device, the mapping algorithms, the sound synthesis engine, the compositional structure and the output system (Tanaka, 2012b) – raises a series of questions: (i) through the action with which part of the instrument do we experience the interactivity? (ii) Do we perceive the input device only? (iii) Does the interaction with the input device lead to the perception of all other parts of the instrument?
Verillon and Rabardel (1995) argues that the subject-object interaction is mediated through the interaction with the instrument. They call this Instrumental and instrumented Activity Situation model (IAS). In other words, during a musical performance, we engage with processes that generate those sounds through the interface.
I discard this model of interaction and aim to give performers the possibility to transform auditory and visual feedback through the embodiment of the audio-visual feedback, so through a corporeal experience of the interaction with the processes that originate it. In my work, technology is the medium through which the interaction is possible, and our body becomes the instrument that interacts with the object. My research will provide a possibility for musicians to transform themselves into the intermediate universe between the subject and the object while still maintaining subjective characteristics.
In line with the objectives of this work, Fishkin (2004) talks about the embodiment of the interaction as *"the extent to which input focus of a system is bound to the output focus"*. Upon this fundamental concept, Fishkin draws four levels of embodiment: (i) *distant*, where the input device is remotely controlling the output (i.e. organ keyboard and the pipes), (ii) *environmental* where the output surrounds the input but it cannot be touched (i.e. in acousmatic music, the mixing desk and

the loudspeakers), (iii) *nearby* where the input device is proximal the output (i.e. guitar and its amplifier), (iv) *full* where the input and output coincide (i.e. singing, violin). Building on this theory, my work aims to combine body gestures (the input focus) with audio, visual and lighting processes (the output focus), through a *full* embodiment of parameter control, here intended as embodied interactions.

2.4 Interaction design solutions for musical performance

Composers and performers have been exploring possibilities to facilitate interaction with audio processes and make them more theatrical since the mid-1930s. One of the first pieces written using a gesturally controlled electronic instrument was *Ecuatorial* by Varèse (1934) for solo voice (bass) chorus and ensemble, which included two Ondes Martenot (Erickson, 1975).

The Ondes Martenot, invented in 1928 by Maurice Martenot (Aimi, 2007), is one of the earliest gesturally controlled musical instruments. Its sound, generated by oscillators, can be controlled through actions on a keyboard, switches and a button situated on a drawer under the keyboard, and by sliding across the keyboard a ring, sustained by metal wires attached to the extremities of the keyboard, with a finger. One of the strengths of the Ondes Martenot is that it enables fine control of sounds through gestures already familiar to musicians. Pitches of the oscillators are defined by pressing the keys, and a vibrato effect can be triggered after hitting a note by wobbling the finger to change the pressure on the keyboard as if it would be executed on a string instrument. The ring allows continuous modulation of the pitch through a sliding hand gesture across the keyboard, as a metaphor of the glissando technique. A touch-sensitive glass on the drawer under the left corner of the keyboard allows control of dynamic of the sound, like the expression pedal of an organ.

The Ondes Martenot introduced new ways of performing and composing music through technology and, most importantly, new modalities of interaction with sound parameters, which are still in use today. An example is the keyboard of the ROLI Seaboard, with which (i) notes or chords are played by striking the keys, (ii) to modulate the dynamics alternating the pressure on the keys, (iii) and to continuously control pitch and other qualities of the sound by sliding the fingers horizontally and vertically across the keyboard and the single keys (Roland, 2014). Another example is *Touché* (Expressive E, 2018), built by Expressive E, built by Expressive E, a controller that extends control dynamics in MIDI instruments on the same principles as Ondes Martenot's touch-sensitive glass.

In the same year that the Ondes Martenot was invented, 1928, Lèon Theremin built an instrument called the "Theremin", after himself. It generates sound through a series of oscillators, whose amplitude and frequency can be controlled through free-space hand gestures around two perpendicular antennas. The vertical movement of the left hand over the horizontal antenna controls the amplitude, and the horizontal movement of the right hand from the vertical antenna controls the frequency (Theremin and Petrishev, 1996). This instrument today exists in both its original version and a more recent one, the Theremini (Moog, 2014). This last one maps the signal of the two antennas onto MIDI messages. With the Theremin, for the first time, gestural interaction is physically detached from the musical instrument itself.

The Theremin is one of the first instruments to give the possibility to interact with sound processes through free-space gesture. As for the Ondes Martenot, the Theremin's interaction design is still relevant to many systems today. Examples include the work of Porres and Manzolli (2007) who developed a gesturally controlled system which interaction design relies on the proximity to three antennas placed in the performing space. Geiger et al. (2008) developed the VRemini using Nintendo's Wii Remote to track movements of the dominant hand and the Nunchuck to track the non-dominant hand. A second version of the system was realised using a glove with attached a fiducial marker and a camera to track the hand movement in space. Johnson and Tzanetakis (2017) created the VRMin, a Mixed Reality (MR) system for enhancing theremin practice while minimising the reliance of feedback. While realising this work, they saw the potential of this interface as a pedagogical tool and for music expression.

The Ondes Martenot and the Theremin are the precursors of a new category of digitally musical instruments referred to as "gesturally controlled". Beyond historical facts, these two devices are testimony of the fact that innovation in the interaction design is of very high importance in music interfaces; at the point that after about one hundred years, the industry still builds electronic music instruments using the same modes of interaction.
Recent gestural controlled instruments have been built and played not only to generate but also to transform sound. In music performance, these are commonly generated electronically, with acoustic instruments or with our voice. In the next sub-section, I review the interaction with devices in the context of singing performance. This choice is made in order to review different design solutions on the interaction with the device without considering the potential simultaneous interaction with a musical instrument. In the following sub-sections, devices controlling sound spatialisation processing parameters regardless of their origin, this time taking into account the simultaneous interaction with a musical instrument during a performance, and multimodal approaches using biofeedback technologies in the last sub-section will be reviewed.

2.4.1 Interaction with sound processing parameters during vocal performance

An example of devices utilised in singing performance is the eMic. It was created by Hewitt (2013) as an extended microphone-stand interface controller used for live electronic vocal performances. It embeds a joystick microphone mount and pressure, distance, tilt and ribbon sensors. These allow the performers to control and transform the sound through gestures that are common among popular music singers. This device facilitates audio manipulation during a vocal performance but also sets a series of challenges. For instance, the user can manipulate the sound only through direct physical contact or by being close to the stand, which could disrupt the performance in case the musician plays an instrument. In collaboration with a choreographer, Hewitt (2011) realised *Idol* (2010), a performance that aimed to unify choreographic aspects with interaction with the eMic. Their work is an example of a common workflow adopted when working with new interfaces, that is the creation of an ad-hoc performance to explore the device. However, what if a composer or performer aims to produce new pieces of music? How can an electronic instrument support the performance of already composed pieces? How can this facilitate the composition of new pieces without the help of a choreographer? Another important observation from her performance is that unlike instrumentalists, whose movements are limited by their instrument's physical characteristics, singers have the freedom to move without any constraint, so why impose the limitations of a cabled device like the eMic? To make the performer free from a cable, Park et al. (2012) created VOICON, an interactive wireless microphone that includes one motion and two pressure sensors on the body of the microphone for controlling audio manipulations. With VOICON, singers can activate a vibrato effect by performing circular movements while holding the microphone, control a pitch shifter by tilting the microphone and by applying pressure on its body singers can activate a third customisable audio effect. Very similar to the VOICON is the Sennheiser's Concept Tahoe microphone (Schlessinger, 2012). This wireless microphone embeds a motion sensor and features three buttons on the body of the microphone to activate looping functions and different audio effects.

In the context of sonic interaction design, Rocchesso et al. (2016) created "miMic", an interactive microphone, that includes two buttons and an IMU device connected to a computer. miMic enables musicians to design electronic sounds, starting from a vocal sound. The designed sound can then be reproduced and explored through new vocal sounds and by moving the microphone. The combination of two different modalities of interaction, in this case, acoustic and gestural, opens a new realm of interaction possibilities and forms of expression. It enables the performer to explore their vocal techniques while designing sounds, and to use evocative gestures, like in the sound tracing exercise, to then reproduce and control the designed sound. This process can be repeated until the musician is pleased with the result. Thus, in addition to involving different cognitive processes and activities, the embodiment of the interaction with miMic becomes an iterative process defined by the sound-interaction design loop. This last aspect is very important and relevant to the work presented in this thesis because the technology does not only afford certain modes of interaction and present the user a tool for sound and gestural design, but it also proposes a particular workflow that supports the sound design process and its exploration. In the same way, technological artefacts adopted and realised in my work, do not only aim to encourage a creative outcome, but also the process that leads to it.

A diverse approach is taken by Grégory Beller in his performance *Babil-on V2* (Beller, 2015). He captures the vocal sound through a wireless microphone and hand gestures with a custom-made system called SpokHands, originally ideated during the production of *Luna Park* Beller and Aperghis (2011). SpokHands is a system created with the aim of defining new approaches to vocal improvisation. It made of two wristbands embedding an accelerometer, tactile ribbons and piezoelectric sensors that allow the gestural control of real-time concatenative synthesis. In his performance, he records and plays back fragments of his voice, sometimes processed transformed, through a wide range of gestures such as percussion gestures, as if he had several percussions placed around him or on his hands, and different continuous movements.

With *Babil-on V2* he demonstrates how this approach can guarantee physical freedom to control sound processing. A system like SpokHands allows imbuing a stronger sense of physicality that transforms vocal sounds than any of the interfaces mentioned above, thus leading to the embodiment of the interaction with sound processes.

From the above reviewed examples, it emerged several factors that contribute to the impact of the interaction with an interface has on the performance. For example, the freedom of movement that an interface allows in relation to its shape, type of connection to the audio engine, the sensing modalities and the workflow they support. In the following section, these aspects have been considered looking at a broader range of interfaces.

2.4.2 Interacting with sound spatialisation parameters

Spatial sound opens a new dimension of creativity for composers, but it also raises several issues. Lyon (2014) highlights the problem around the need for a dedicated space to host multichannel audio systems. This is one of the primary issues that hold back many institutions to make these systems available to musicians. Another concern relates to the process of composing and the act of performing using sound spatialisation, about which Baalman (2010) writes that the use of a specific speaker set-up and the acoustic properties of the environment in which the system is used have the potential to cause transferability issues when reproducing the same performance in a different space. This technical issue can then alter the artistic intentions of composers and performers.

Something that is of my particular interest are the problems surrounding the interaction with sound spatialisation processes during live performance. The reasons behind such concerns on this particular topic related to the fact that controlling spatial sound effects is an example of interaction with a high number of audio processing parameters (i.e. gain, delay, filtering and reverberation) with the sole

scope of controlling the spatial position of the sound within the listening area. This observation on parameters' control can also be made about other audio elaboration techniques. For example, in controlling a harmoniser, a musician can control different parameters such as the balance between direct and effected processed sound, key and the number of voices. The attractiveness to controlling auditory spatial effects is related to the artistic possibilities that it gives to performers. To be precise, to the fact that it can extend musicians' creativity while preserving its timbral characteristic of the instrumental sound like pitch or, in the case of singing and intelligibility of the sung text.
With this in mind, below is a review of sound spatialising systems on the subject of interaction design.

Typical examples of systems presenting this issue are all fader-based interfaces, for instance, the ones utilised for controlling the BEAST (Wilson and Harrison, 2010), Gmebaphone (Clozier, 2001) and the Acousomium (Desantos et al., 1997). Although they offer fine control over the spatial characteristics of a sound in live performance, these interfaces do not provide the user with a "natural" and "intuitive" interaction with the system due to the high number of parameters requiring simultaneous control (Mooney, 2005). One of the first solutions to this problem was proposed in 1952 by Pierre Henry when he performed at Salle de L'Ancien Conservatoire in Paris using the Potentiomètre d'Espace. It was built by the French engineer, Jacques Poulin, to gesturally control the sound balance of fifty audio tracks using a quadraphonic audio system (Bates, 2009), through moving a reel across a space framed by metal loops (Bayle, 1993). Here the gestural interaction resembles the one for the Theremin, with the difference that the user holds parts of the interface and uses it to control spatial cues.
In more recent works, there is a change in the interaction design influenced by technological advancement. More recently, Ramakrishnan et al. (2006) realises Zirkonium to control the spatial location of the sound with a mouse or a joystick,

the position of the virtual sound source through a pointer tracker within an abstract representation of the acoustic space. The use of the joystick brings a novelty in the technological realm, but from the interaction perspective, as in the Potentiomètre d'Espace, the user keeps modifying the position of an object (the pointer) in relation to an abstract space (Graphical User Interface, GUI). Other examples of spatiality supporting this mode of interaction are Digital Audio Workstations (DAWs), which includes the plug-ins "Wave 360° Surround Tools" (Waves, 2015) and "Acon Digital Verberate Surround" (Acon Digital, 2014); and touch-screen based interfaces like "tactile.motion" (Johnson et al., 2014) and "ROTOR" (Reactable, 2016).
The human-computer interaction supported by these last examples of interfaces makes the control of sound spatial parameters easier compared to using a mixing interface. However, if using these technologies to perform spatialised music, these would still not enable the audience to appreciate all performative aspects involved in the execution of the piece. Those interfaces could amplify the problem if musicians play a musical instrument while controlling spatial cues. On this matter, Marentakis et al. (2008) argues that to enhance performability factors, gestural human-computer interaction provides visual feedback for the audience that enhances the perception of spatial audio cues. Referring to his work, it is possible to assume that the interaction with a sound spatialiser is not meaningful information for the audience unless the interaction itself and how it relates to the auditory feedback is shown to the audience.

Inertial Measurement Unit (IMU) based devices allow a free-space gestural interaction with sound parameters. In the case of wearing an IMU device, it allows one to consider our own body as the interface, thus, to embody modes of interactions, allow musicians to demonstrate the interaction with the computer through gestures visible by the audience. IMU devices can detect orientation, speed and acceleration of an object or body part. With the commercialisation of IMU-based devices – such as Nintendo's Wii Remote and smart-phones – and the lower fabrication costs, IMU devices have been used in musical applications. Valbom and Marcos (2005) realised

WAVE, an immersive virtual audio environment that senses hand and upper limb movements through six degrees of freedom (6DOF) sensors grabbed by the user with both hands. This interaction design resembles the one for the Potentiomètre d'Espace, with the difference that here the user can move without the constraints of a space framed by four metal loops. Nevertheless, due to the fact that the user has to hold the input controller with both hands, the interaction design with the system can still represent an issue for a musician who chose to play a musical instrument and to control an interactive system at the same time.

IMU devices have been used to explore the spatial characteristics of an auditory scene through head movements. Deldjoo (2009) took advantage of the IMU sensors of a Wii Remote to realise a head tracking system for controlling a 3D Audio Rendering; on the same line, Wozniewski et al. (2008) uses IMU sensors on a mobile phone for controlling sound spatialisation of a large-scale mobile audio environment for collaborative musical interaction. In these two works, the sound experience becomes closer to exploring sound in the real world through an interaction design that welcomes natural and already learnt body movements, such as twisting the head towards a sound that catches our attention.

An alternative approach is the extension of a musical instrument, making it become the interface for controlling sound processing parameter by mounting IMU devices on the body of the instrument. An application of this method can be found in the work of Luca Turchet who realised the Hyper-Hurdy-Gurdy (Turchet, 2016a) and the Hyper-Zampogna (Turchet, 2016b) mounting a 3-axis accelerometer on the body of the of a Hurdy-Gurdy and a Zampogna respectively.

Cameras represent an alternative to IMU devices for tracking body gestures. Lech and Kostek (2013) used a webcam to develop a system for mixing sound sources on the spatial auditory scene through gestural interaction with a GUI projected on a wall. In his work, modalities of interaction are limited to a combination of horizontal and vertical movements of the hands, due to tracking body movements from a bi-dimensional imagine. In the context of sound spatialisation, where the number of parameters to control at the same time is high, the use of a webcam can pose a limitation to the interaction design.

A solution to this technological limitation can be or use of systems built with depth-cameras, which enables to track a body in a 3D space. From a comparison of different of these systems (Vicon, Polhemus, Kinect, and Gametrak), Vigliensoni and Wanderley (2012) demonstrated that the Vicon is the system has the best performance. However, as it is more affordable and easier to use, the Kinect is more widely used and has inspired the production of interactive systems to manipulate spatial audio for musical and non-musical works. The Kinect has been widely used to realise mixing interfaces for establishing the sound spatial position of different audio files (Ratcliffe, 2014; Wakefield et al., 2017). Kronlachner (2013) developed externals for Pure Data that, together with the GEM library, have been used to control sounds and video projections in his multimedia art practice, by the IEM Computermusic Ensemble, and in *übersetzen - vertimas* (Kronlachner, 2012). The Kinect has also been used to track guitarists' gestures to create a hyperinstrumental performance in Graham and Bridges (2014). The Microsoft Kinect also embeds an array of microphones which has been used as a source localisation system to estimate the position of a musical instrument and use these data to then control a spatialisation system (Salvati et al., 2011).

Although the Kinect is a device that can support the interaction with interactive audio software, it constrains the user to interact with the computer within a limited physical space (Bullock et al., 2016; Vigliensoni and Wanderley, 2012). In reference to my research question: "how may disruption during a live performance with embodied human-computer interactions be reduced?", the use of a Kinect would not satisfy my aims. In the case of a musician moving on a large stage, using a Kinect could certainly disrupt the performance. Moreover, it requires being installed in a specific space and position in order to track our body accurately. Hence, the use

of these devices by non-experts in situations like a live performance, when the time and space for installing the system can often be limited, might be problematic.
Another one of the most used gestural devices in musical applications is the Leap Motion, an infra-red based camera device to track hand gestures. This last device has been used in different music applications such as to control a five-grain granular synthesiser and spatialisation parameters using a stage metaphor as interaction design (Hantrakul and Kaczmarek, 2014). A stage metaphor was also adopted by Ratcliffe (2014) for controlling amplitude and panning parameters in DAWs. Although a Leap Motion could facilitate the sound spatialisation of different sound sources, Ratcliffe mentions the difficulty of controlling a high number of sources.
Similar to the Kinect, the Leap Motion raises issues concerning the space of interaction. It successfully supports the interaction with the sound in a context in which the users' movements are in the range of 60.96 cm above the device. Thus, the Leap Motion is not a technological solution for my work, where the performer's hands movements are likely to be outside such range.

A different category of gestural interfaces with the potential to support design solutions to my research questions are glove based devices.
Based on the importance of gestural interaction as a way to communicate the audience information about the sound generation with clear and distinct gestures, Jessop (2009) created VAMP. It is a gesture-based wearable controller for transforming vocal sounds through a vocabulary of highly expressive gestures developed from choral conducting. However, VAMP suffers from usability limitations such as the need to be adjusted, and in some cases modified for each performer's arm physiology in order to be used exhaustively. Also, the fragile cabling around the arm, connecting the device to the computer, may inhibit the performer's freedom of movement.
To minimise these issues, Bokowiec (2011) created the Bodycoder. This is a more compact solution than the VAMP, but it nevertheless only partially solves this

issue, as the musician is still surrounded by cables going from the glove to a data transmitter on the waist belt. In some instances, this might result in disturbing the performer's body movements.

The "mi.mu" glove (Mi.Mu, 2016), also named the Gloves in Serafin et al. (2014), offers the most compact solution. In the mi.mu, the analogue-digital conversion of bend sensors – for tracking fingers' movements and an IMU monitoring the wrist's orientation – and the data transmission are all embedded in the glove. Thus, the electronics are not a limitation to the performance movements anymore. Moreover, it also embeds a LED to provide the user with visual feedback, which expands communication between the technology and performer (Jordà et al., 2007).

2.4.3 Multimodal approaches using biosensing technology

Multimodal approaches are those where the feedback is controlled and transformed through multiple parameters or interfaces (Tubb and Dixon, 2015). This use of interfaces is described by Tanaka and Knapp (2002) as *"complementary"*, in the way that each sensing modality and their features complement each other in the execution of a task. For example, when playing the clarinet, the blowing action enables the musician to produce sound and control the sound dynamic, and the fingering on the keys allows the tuning of the sound on a well-defined pitch. However, some interfaces are *"unidirectional"* so that they cannot ensure a *"free-standing mode of gesture-sound articulation"*. In the example of the clarinet, the musician-clarinet interaction is also defined unidirectional, as the blowing action is essential for playing the instrument. Blowing would allow the production of a sound even if not defined in the frequency domain. On the contrary, pressing the key without blowing will not produce a sound associated with the clarinet.

> The explicit goal [of multimodal interaction] is to integrate complementary modalities in a manner that yields a synergistic blend such that each mode can be capitalized upon and used to overcome weaknesses in the other mode.

(Oviatt, 2003)

Tanaka and Knapp introduce the concept of *"bidirectional complementarity"*, built upon Oviatt's goals of multimodal interaction. They propose using technology able to sense the isometric and isotonic activity of our body – for instance, a combination of electromyography (EMG) and IMU-based devices – for controlling and transforming sound through a *"bidirectional"* and *"complementary"* approach. While the position of a body limb might control a parameter, the muscles' activity of the same body limb performing the gesture could control other parameters of the sound transformation (Figure 2.2 and 2.3).

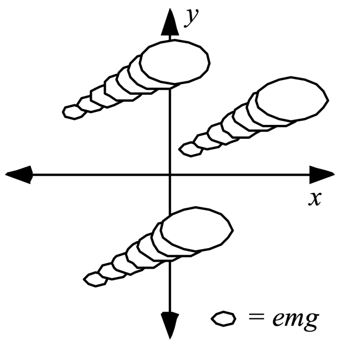


Figure 2.2: 'Bidirectional complementarity A: Position data complementing EMG gesture' (Tanaka and Knapp, 2002)

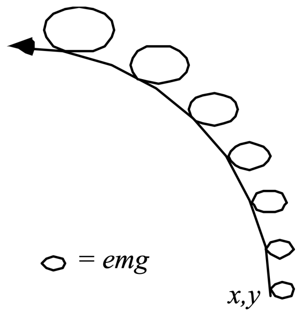


Figure 2.3: 'Bidirectional complementarity B: EMG data complementing positional displacement gesture' (Tanaka and Knapp, 2002)

Musicians' body movements are different, and often they are tight to their instrument. There are those musicians that when playing engage their body kinaesthetically with minimal muscular effort, those that do an intense work with their muscles without moving their limbs, and those whose movements are a combination of kinaesthetic and muscular activity. Adopting the Tanaka and Knapp approach, it would be possible to design modes of interactions that musicians can blend according to their creative and instrumental needs.

The first application of biosensing technology used in the field of computer music was made by Knapp and Lusted (1990) using the Biomuse system. It is a bioelectric controller to capture different biosignals such as electroencephalogram

(EEG), electrooculography (EOG) and electromyogram (EMG). These signals were then processed with a microcontroller and mapped into MIDI messages. In 1986, Chris Van Raalte ideated the BodySynth, a device similar to BioMuse that he built with the engineer Ed Severinghaus to translate EMG signals into MIDI messages, Marrin Nakra (2000) realised the Conductor's Jacket using biosensors made by Delsys Inc to capture conductors' biosignals and translate these into expressive musical gestures and Nagashima (2003) the built MiniBioMuse-II.

These technologies were used in musical performance mostly by their creator and a few other performers. BioMuse has been used in different performances by the Biomuse Trio (Lyon et al., 2014), and other performers like Atau Tanaka (Tanaka, 1993). The BodySynth has been used for performing by Van Raalte and by the artist Pamela Z in different performances (Van Raalte, 1998). This limited use is presumably attributable to the fact that these devices were produced in a small number. With the commercialisation of EMG sensing technology and lower fabrication costs, electromyography has started to be used for controlling interactive systems and software in musical applications by a more significant number of people.

At the time of writing, two of the commercially available EMG devices were the BITalino and the Myo.

BITalino is a low-cost toolkit designed for students, teachers, artists, researchers and corporate R&D to learn and prototype applications using body signals (Plux, 2014a) (Figure 2.4). Its users do not require any specific electrical knowledge. It allows capturing kinaesthetic data through an accelerometer and biodata such as EMG (electromyography), ECG (electrocardiography), EDA (electrodermal activity) and EEG (electroencephalography). BITalino also allows the measurement of lighting conditions (da Silva et al., 2014). The device can be wirelessly connected via Bluetooth and transfer data to a computer, mobile or microcontroller through dedicated APIs (Plux, 2014b).

Since its first release, BITalino has been used in different creative projects. The

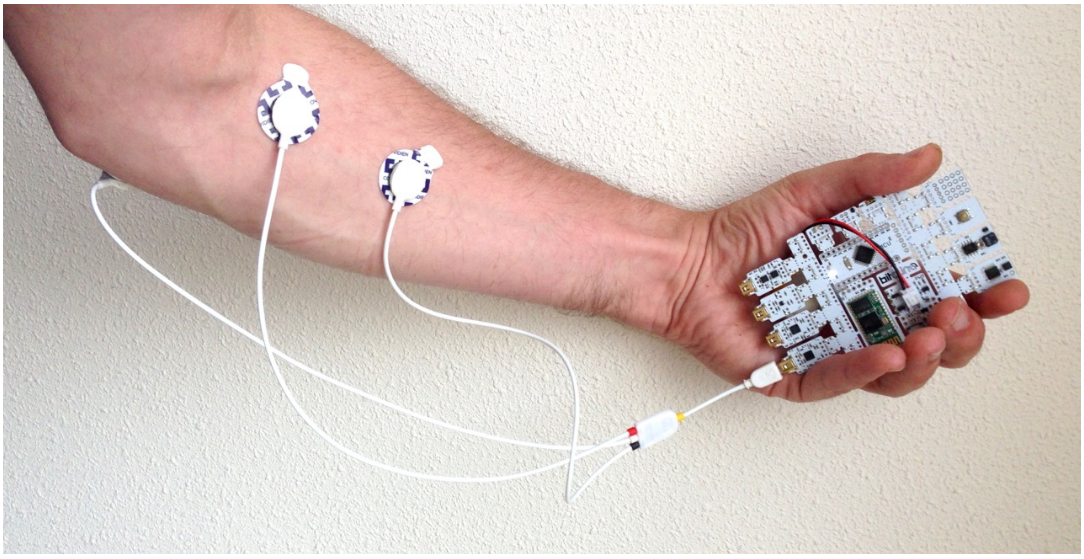


Figure 2.4: User wearing BITalino to track EMG. Adapted from Dertien (2016).

Glitch Chamber, by Protopixel (2014), is an interactive and immersive audio-visual space which reacts to the ECG activity of the person within the chamber. *ReactiFI* by Sad (2016), is an Augmented Reality (AR) environment based on the user's heart rate change. Cruz et al. (2015) created a Baymax doll (Disney's Big Hero character), whose lighting and sound can be controlled through the user's heart rate. BITalino supported all the above-cited works with successful results. Although its use does not require any specific knowledge, in the applications mentioned above, BITalino requires do-it-yourself (DIY) skills. This aspect is even more evident in the system developed by Pinto et al. (2016) to gather data from a swimmer in action. They had to find ways to attach the cables from the electrodes to the board, later both of those on the user's body. For this, they create a silicone case to make the board, the cables and all connections between the board and the electrodes waterproof. As also mentioned in the case of the glove-based technologies, the cabling has the potential to disrupt the execution of a permanence.

The Myo (Figure 2.5) is a wearable and wireless gestural controller, designed and produced by Thalmic Lab. It provides the user with spatial data and gestural data. Spatial data report the orientation of the armband in quaternions, which can be translated into rotation matrix or Euler angles, and the acceleration vector that describes the acceleration of the armband in time, as a three-dimensional vector (Thalmic Labs, 2015b). Gestural data give us the hand pose as well as the muscle activity of the arm. Spatial data are tracked by a nine-degrees of freedom (9DOF) IMU device placed into the main expandable casing (where the logo is, Figure 2.5). EMG data are extracted by recording the electrical activity of muscle tissue and the nerve cells that control them (motor neurons), using eight EMG sensors placed on the inner side of the eight pods composing the armband. Moreover, the Myo can provide haptic feedback to the user via vibration motors placed into one of the eight pods. As was the case with the LED on mi.mu gloves, this feature of the devices is extremely important to enhance the device-user interaction. The eight pods are linked to each other by a stretchable fabric, which allows adapting the armband to arms with different physiology. The monitored data are transmitted to the receiver device (e.g. laptop, tablet or smartphone) through Bluetooth protocol using a Bluetooth 4.0 Low Energy dongle provided by the manufacturing company.

Figure 2.5: Thalmic Lab's Myo

One of the first uses of the Myo in musical performance is exemplified by DJ Armin Van Buuren, who uses physical body movement as a way to control lighting effects (van Buuren, 2014). His technical team developed a system that allows him to control the orientation of moving headlights through a pointing gesture so that he could focus the light beam where he was pointing. In light of what discussed above regarding the control of sound spatialisation, in the same way, DJ Armin Van Buuren moved light beams towards the public, the pointing gesture could potentially be adapted to move the sound around the audience.

His work echoed in the DJ community and inspired the use of the Myo to create new software tools to control software and mix music through DJing scratching gesture. For instance, (Szanto, 2014) used the My to drive a Virtual 2D FX Table, Hammond (2014) created software to control for Serato Scratch Live, and Adesara and Ashank (2014) realised Myo Music Mixer (3M) who used the Myo in combination with the Leap Motion.

Interestingly, this is a case of hybrid *conscious/unconscious interaction*. The former refers to the conscious control of sound parameters, and the latter to the possibility of interacting with a system without consciously choreographing the sound control. The analysis of these two modes of interaction by Marshall et al. (2009), in the context of controls of sound spatialisation parameters, reports that *conscious interaction* should be well designed to not interfere with the musical performance, and the performer must adapt the interaction design to the performance without affecting instrumental technique. By contrast, *unconscious interaction* design frees the performer of the potential of any limitation as the system responds to any movement.

The scratching gesture, picked up using the Myo, allowed DJs to consciously interact with sound processing parameters, and at the same time, to focus their scratching technique on the turntable, and unconsciously interact with the machine. This analysis of the aforementioned work by DJs highlight the potential of the Myo to support different approaches to interaction design, as well as their combination.

The Myo has been used by the research and musician Atau Tanaka to compose and perform *Myogram* (Tanaka, 2015). Here the EMG signals are sonified and processed adopting various techniques. Together with complex mapping functions, he relates gesture trajectories with sound features in order to communicate a sense of closeness and direct body control.

Federico Visi uses the Myo in *Kineslimina*, a piece for viola and electric guitar, (Visi and Coorevits, 2015). Here we not only experience a music piece performed using an interactive system, but the simultaneous interaction between performer through the system. These interactions are also highlighted through lighting effects.

The work realised by Tanaka and Visi, prior and contemporary to this thesis, has contributed and informed my research. In earlier work, Tanaka (1993) expressed the idea that body movements lend themselves to the music and that they help to communicate the musicians' intentions, but at the same time it must be clear that those are not the cause of the sound. For example, a pianist could articulate a wide gesture, which is not sensed by the piano, in anticipation to play a note. However, the audience has a different experience of the performance because of that gesture. An example of such interaction can be found in another of Tanaka's pieces, *Suspension* (Tanaka and Nicolls, 2017). Also, Visi takes advantage of gestures that are already part of the musician's motor skills but pushes the boundaries of musicians' spaces during a performance (Visi et al., 2017).

Building on this body of work, I aim to take advantage of already-learnt gestures to reduce musician's effort to adapt to the interface and the disruption to the musical performance.

2.5 Mapping strategies

The design of musician-computer interactions is ultimately implemented through the adoption of strategies to map our actions into audio-visual feedback. To consider various aspects of mapping strategies is fundamental to put into practice interaction design solutions and to evaluate them during a musical performance. In computer music, this often translates into the mapping of input sensor data, from simple buttons to more sophisticated motion, ambient, touch pressure, into audio, visual or lighting processing parameters.

In the following subsections, I reviewed mapping strategies focusing on the interpretation of sensors data and what interaction with processing parameters they suit (Section 2.5.1), and approaches to implementing strategies (Section 2.5.2).

2.5.1 Sensor data and interaction

After comparing two different spatialisation systems, Marshall et al. (2009) identified various likely successful mapping strategies. They propose that commonly static parameters like room dimensions, wall reflectivity and speaker positions should be established before the performance and not associating them with any of the performer's actions. Source position, orientation volume and any other continuous controlled parameter should be controlled through continuously executed and sensed gestures. Thus, there are parameters that lend themselves to continuous control and those that do not. Consequently, the gestural interaction to control these needs to be considered appropriately, as well as thinking about what parameters work best in what contexts.

The relationship between spatialisation parameters themselves needs to be considered, such as the spatial position and orientation in each dimension of the virtual sound source. In reference to my work, gestural information should be considered similarly. For example, the position, orientation and posture of a musician should all be considered together. This can help the identification of essential sensor features to be mapped to feedback processing parameters.

In mapping gestures onto sound spatialisation parameters, it is also very important to chooses a sensing device that provides data with adequate resolution. The resolution of the control data should be determined by the driving spatial cues. For example, it is impossible to hear sound spatialisation cues on a centimetre scale, especially in the case of low frequencies (Saberi et al., 1991; Chandler and Grantham, 1992), there is no need of control data that changes the spatial position of the sound at that scale.

Marshall et al. also suggest that another aspect to consider when mapping sensor data to interactions with a system is the potential coordination and cognitive load limitations during a musical performance. In reference to the concept of *conscious* and *unconscious interaction* introduced in the previous section, *conscious interaction* requires a higher mental workload by the musician and degree of adaptation. Nevertheless, it can provide composers and performers with a more defined control of a system to create music. On the other hand, *unconscious interaction* allows the musician to be free from constraints and focus on the performance only. In this case, the interaction design and mapping strategies must be realised purposely for such performance; thus, the system is usable in a specific context only.

Differently, I aim to realise embodied design solutions that can enable *conscious interaction* with the feedback, but that also gives the opportunity to "wear" the interface and forget about the rules behind the interaction design. With this approach, performers will benefit from the possibility to control the audio-visual processes systematically, thus ensure the reproducibility of the performance, and at the same time the possibility to focus on the performance and on playing a second musical instrument.

2.5.2 Explicit and Generative Mapping

According to Hunt and Wanderley (2002), mapping strategies distinguished between explicit and generative. The difference between these two groups lies in the approach taken to it. Explicit mapping strategies define the relationships between input and output parameters. Generative mapping provides tools to create an internal adaptation of the system through the training of a model on the signals driving the parameters (i.e. neural network). Both of these approaches in the context of my research are reviewed in the following subsections.

2.5.2.1 Explicit mapping

Hunt and Wanderley, devised different approaches to explicit mapping indicating the number the input parameters mapped into a number of outputs feedback parameters. These approaches are so called *one-to-one*, *one-to-many*, *many-to-one* and *many-to-many*, outlined in Figure 2.6.

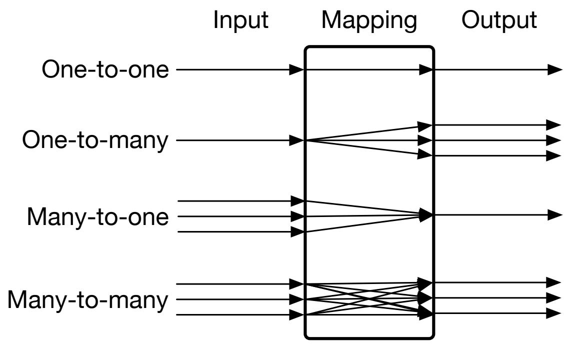


Figure 2.6: Mapping strategies diagram

They individuated fundamental aspects of these approaches and proposed a three-layer mapping method to use in conjunction with the ones just cited (Figure 2.7). In the first one, meaningful performance parameters are extracted, thus allowing *interface-specific* sensor data to be mapped into a set of abstract parameters (second mapping layer). These parameters perceptually relevant, as in Bates (1997) or Garrett and Goudeseune (1999), or derived from other interaction metaphors (Mulder et al., 1997) are then decoded and associated with audio engine parameters in the third layer.

By contrast, Wessel and Wright (2002) talked about expressive mapping strategies as a form of intimate interaction levels aiming at performer satisfaction. They argued that a *"high degree of control intimacy can be attained with compelling control metaphors"* as in the third layer of mapping defined by Hunt and Wanderley (2002). Wessel and Wright advocate the applicability of George Lakoff's work on embodied cognition supporting that the concepts of time and space are rooted in sensorimotor experience and proposed different metaphors to establish an expressive data input to sound output mapping. The metaphors they created were: (i) *drag and*

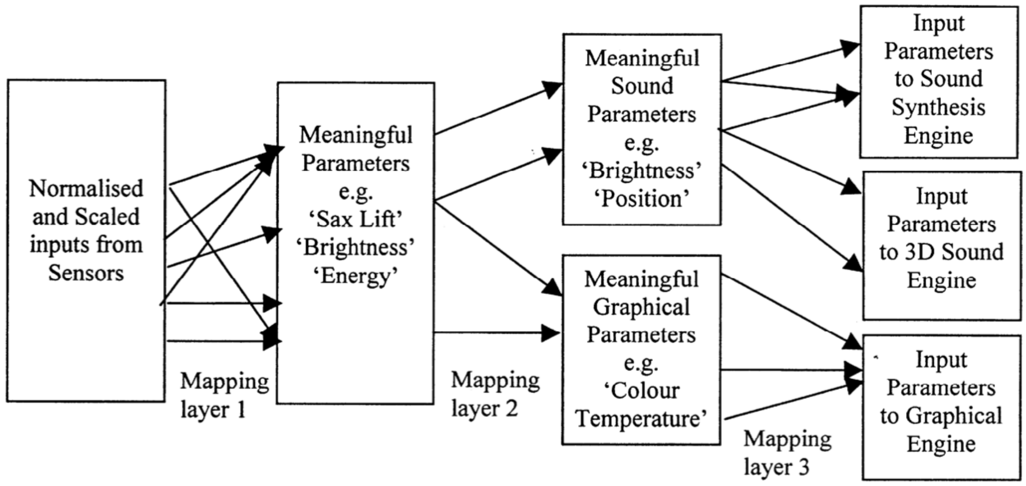


Figure 2.7: "Mapping layers within the RIMM project." adapted from Hunt and Wanderley (2002)

drop (the act of picking up an object and dropping it upon a process); (ii) *scrubbing*, (time-domain manipulation of sound without spectral changes); (iii) *dipping* (the control of multiple processes by which the computer processes different sound materials and the user establishes the presence of each material and parameters to process them); and finally, (iv) the *catch and throw* metaphor of intercepting the sound, processing it and throwing it back into the performance space.

Although these modes of interaction are close to a bodily experience of sound processes, they are best used in context-specific situations. For example, the *drag and drop* metaphor can be adopted to control spatial sound cues, as in the case of the stage metaphor, where an abstract visual representation of the sound is grabbed and moved to a different location of the auditory scene. In a situation where musicians can move freely, they would not have any issues. On the contrary, it can be difficult to perform those actions when playing a musical instrument.

2.5.2.2 Generative mapping

Generative mapping enables the adaptation of the functions to map input-output parameters in the context of the musical performance and the sensing modalities of the hardware. This approach can be implemented through the use of machine learning approaches, i.e. neural network (Hunt and Wanderley, 2002), or other machine learning approaches. These tools are powerful and offer a valid solution to this problem. However, it is not always the case that musicians have background knowledge of machine learning.

To facilitate the implementation of generative mapping approaches to musicians, Fiebrink et al. (2010) utilised Interactive Machine Learning (IML). IML is based on the interactive coaching in training of the model until the machine learning output does perform as desired. The workflow of interactive machine learning software is composed of three main stages: (i) recording the data set made of input examples and the respective labels, (ii) training the model, and finally (iii) executing it and then feeding it with newly generated data. This process fosters the creation models of input interfaces data-audio-visual processes, to evaluate its result and iterate this process until when the model output does satisfy the need of the musician.

In the context of my work, questions about how musicians interact with them, what are the benefits that they bring to the creative process, and the possibilities that they open when performing, are more important than the technological and implementation aspects behind those pieces of software. To evaluate these factors, Fiebrink et al. (2011) conducted a study where participants used motion-sensing technology to classify gestures and evaluate the models through cross-validation (a computational process in which the data set is split into two portions, one for training the model and one for evaluation) (Refaeilzadeh et al., 2009) – and by direct evaluation. This last one is a method that consists in observing a model's behaviour in real-time at any input data, via visualising the current outputs and the feedback that the results generate (Fiebrink et al., 2011). From this study, Fiebrink et al. observed that participants used cross-validation to evaluate the algorithm performances quickly; whereas direct evaluation was used to learn how to generate more effective training data, what were the easier types to build, and discovery of new problems which were not discoverable only through the use of cross-validation. Moreover, direct evaluation allowed participants to explore different ways of interacting during the training of the model. This approach makes training and evaluation an interactive process, whereby the user trains a machine learning model based on the ultimate input data-output auditory and visual feedback mapping realised by the machine learning algorithm (see Figure 2.8).

Direct evaluation is particularly powerful when developing an interactive audio system, as the user can, in real-time, evaluate the auditory feedback at, in the case of my work, different gestural interactions. At the same time, this way of evaluating a model might not highlight issues and potential improvements to the model, nor produce a complete quantitative evaluation of the model.

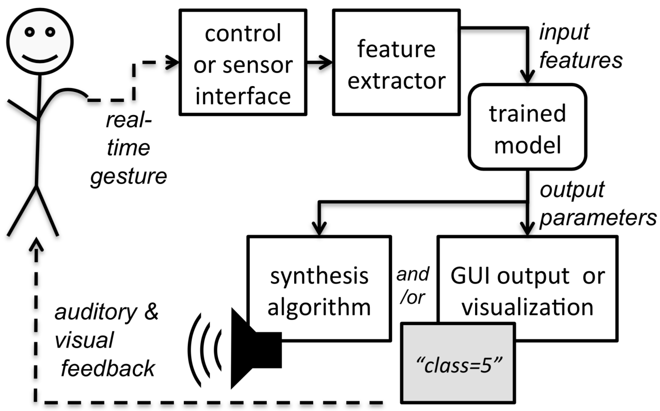


Figure 2.8: Direct evaluation model, adapted from Fiebrink et al. (2011)

The application of IML can lead to human-machine learning interaction that can be described as *co-adaptive learning* (Mackay, 1990), in which both machine learning models and humans are engaged in the model and adapt themselves to the end result. Another possible scenario is that during the iterative process of training the model, the goals evolve after adapting to the model: a phenomenon called *concept of co-evolution* by Kulesza et al. (2014).

As stated previously (Section 2.2.1), this phenomenon is particularly relevant to my work as it supports the iterative creative process of music-making. In the act of composing a musical piece, in which the musician writes on the score, plays it, and then evaluates whether the played music satisfies the artistic idea. The musician may eventually adapt the way the score is played, which is a co-adaptation of the playing technique to the musical objective represented by the score, or drastically rewrite the score towards a different musical objective. Similarly, in a rehearsal of a piece, the performer(s) may co-adapt their playing technique to achieve a certain musical result or impose themselves and so change the objective and thereby re-interpret the piece. Interactive machine learning, as well as facilitating the mapping of input data to output audio-visual feedback parameters, supports the musicians' workflow.

Different machine learning tools have been developed with these principles in mind. For example, Wekinator by Fiebrink et al. (2010), the Gesture Recognition Toolkit (GRT), by Gillian and Paradiso (2014), a C++ machine learning library designed purposely for real-time gesture recognition. Building on Gillian work, Bullock and Momeni (2015) developed ml.lib, a machine learning library for Max and Pure Data that implements GRT algorithms. Temporal models, for classifying time-varying gestures, have been added to the IRCAM MuBu Library for Max (Schnell et al., 2009; IRCAM, 2012) by Françoise et al. (2014).
Two of these, the Wekinator and ml.lib will be used and reviewed in the Pilot Studies (Chapter 4) and the two Case Studies (Chapter 5 and 6).

2.6 Chapter summary

This chapter set out the context for this thesis within the interdisciplinary field of Music and Human-Computer Interaction. It reviewed concepts of affordance and embodied interaction in the context of music, which allowed discussion of the objectives of this thesis, and on the interaction with sound and the parameters to control it electronically through corporeal experience.
Various technological artefacts for musical performance and the modes of interaction with them have been reviewed. This led to the conclusion that the use of multimodal technology combining IMU and biosensors are most appropriate to this research.

After considering different commercially available solutions, the Thalmic Lab's Myo was chosen as the input device. This choice was also made considering that at the time of writing, it was commercially available, and it is a device adaptable to the arm physiology of different users without any adjustment or custom support. From the name of the chosen input device and the primary audio processing technique involved, the interactive system realised with knowledge from this research work was named "MyoSpat".

Finally, mapping strategies and interactive machine learning approaches applied to music applications have been reviewed to lead to a deeper understanding of the software implementation of mapping gesture data to sound parameters.

Chapter 3

Methodology

3.1 Introduction

This research was conducted adopting a methodology that reflects its interdisciplinary nature, involving both Music and Human-Computer Interaction. The methodology outlined in this Chapter aimed to support the iterative work strongly influenced by musicians and their practice when composing, rehearsing and performing. Musicians' artistic practice was the main driver for the research. This entailed a compromise between musical and technical requirements when designing the MyoSpat system. A methodology for this kind of work should not aim only at the development and assessment of technological artefacts but should also take into account musical scenarios (composition, rehearsal and performance) and artistic objectives and practices (composition and performer intentions and the influence of instrumentation on interactions with the computer). Given the complexity of the context and the processes involved in musicians' activities, the adoption of a method for evaluating interaction design, musical and technological artefacts in the world of the creative process was essential. Therefore, the User-Centred Design method, with the musician at the centre of the design, was adopted.

The methodology outlined in the following sections does not encompass technical aspects such as: learning how to install and use the software, fixing problems with the graphical user interface, or potential issues in setting up a multi-channel audio system. This thesis instead focuses on musicians and the design of embodied interactions with sound, light, and visual processing parameters during music performance.

3.2 User-Centred Design (UCD)

UCD is "a broad term to describe design processes in which end-users influence how a design takes shape" (Abras et al., 2004). This method sees the application of an iterative design process formed of one or more cycles. The application of this method allows designing various artefacts around musicians' creative ideas and potential practical issues during musical practice (Monk, 2007). Moreover, it permits to challenge design assumptions and to assess a wide range of aspects related to interactions with technological artefacts in a real-world scenario. Finally, the UCD approach permits to elicit information that users (musicians) might not be conscious of, and similarly that the designer is unaware of.

The UCD method characterised by an iterative process. Each cycle of the process is composed of four main stages. The first one aims to understand the user and the context within which it operates. From these findings, in the second stage, it is possible to specify the user requirements and include them in the design process. In the third stage, a design solution is realised, which is then evaluated in the fourth stage. In this last stage, the system behaviours and the user experience are observed and analysed. Emerging issues from the design solution's evaluation go to inform any future iteration (see Figure 3.1).

Aspects of UCD and how it has been previously used in the field of music has been discussed in the Literature Review (Chapter 2). The following sections describe how the UCD is applied to this research and allowed answering the research questions.

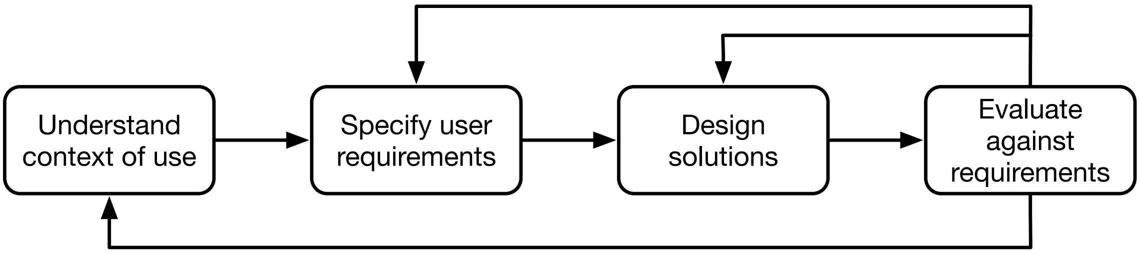


Figure 3.1: User-Centred Design cycle.

3.3 UCD Cycle

The UCD method has been applied in this thesis by contextualising the process described in the previous section. Here the musician, intended as the user, and the electronic music performance as the context of use are at the centre of the design. This UCD method was adopted to support the iterative process of designing modes of interaction with sound processing parameters and the MyoSpat system. Most importantly, through the progressive understanding of different musical contexts and varying needs.

Following an initial experimental phase made of five pilot studies, two cases studies (Voice Gesture and HarpCI) were conducted through two UCD cycles respectively. The UCD cycles include four main stages: (i) understanding the context of use, (ii) specification of user requirements, (iii) design solution and (iv) evaluation against requirements. The details of each stage are described in the following sections and summarised in Figure 3.2.

3.3.1 Understanding the context of use

Design solutions to answer the research questions were undertaken on the basis of the knowledge of the musician and the context of use, the creative processes involved in their art practice and surrounding aspects like instrumental technique, already learnt motor skills and use of technology.

These aspects have been continuously considered through the review of the existing literature. However, prior works do not always cover the important aspects of this

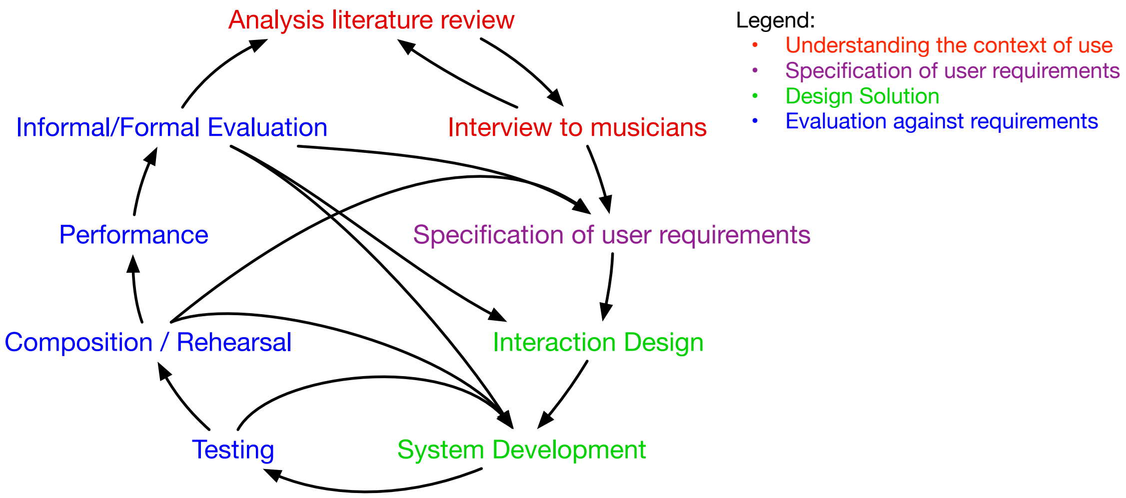


Figure 3.2: UCD cycle.

research. Consequently, semi-structured interviews were conducted to fill in the gaps in the literature on certain aspects of the research, such as the first case study, “Voice Gesture” (Chapter 5) with an interview to Beatboxers, and the second case study, “HarpCI” (Chapter 6) involving harpists.

3.3.2 Specification of user requirements

At this stage are defined user requirements concerning the interaction design with sound processing parameters. Here, choreographic aspects of music performance, musicians’ practice and needs that allow embodied interaction with processing parameters without disrupting the performance were individuated.

Specification of user requirements was delivered through the reasoning on the reviewed literature and early findings and intuitions from the interviews conducted in the previous stage, in function of the aims of each iteration (see Section 3.4).

3.3.3 Design Solutions

At this stage of the UCD cycle, modes of interaction were designed and implemented in the MyoSpat system. Thus, the MyoSpat system is an essential component of this research as it is the system which enables the exploration of the research questions.

Each UCD cycle iteration led to the development of a new version of the MyoSpat system listed with a version number, for example, $0.x$, where x is the iteration number. The MyoSpat audio-visual engine was developed using Pure Data and Max as Integrated Development Environments (IDEs). They offer a fast way for prototyping complex audio-visual signal processes and are used in a plethora of creative practices.

3.3.4 Evaluation against requirements

The application of theoretical intuitions and ideas in a real-world scenario is an essential activity of the UCD cycle. Design solutions were iteratively evaluated during the informally during the collaboration with musicians and in a controlled environment.

Collaborations with four performers and composers through their practice (composing, rehearsing and performing), were conducted to realise musical performances and explore new forms of interaction not exposed by the system through improvisation. Musicians continuously tested and provided useful information about the system and its critical, real-life applications, and consequently generated valuable artistic outcomes. The collaboration with performers and composers allowed the adaptation of the interaction design to various composing and performing aspects like technique, artistic idea and the creative process. Each concert and rehearsal informed decisions about interaction design as well as about technical difficulties encountered in the set-up of the MyoSpat in various concert venues.

In evaluating a digital musical instrument, interaction design and user experience in a real-world composing, rehearsal or performance setting can often be challenging. For this reason, in the second case study (Chapter 6), two user studies and two workshops in a controlled environment were conducted to complement observations made during the work with musicians and to evaluate the system performances.

Here, have been gathered information about the participants' views on strengths, weaknesses and potential improvements of the interaction design and the system, the system's performance, the user experience and the participants' perception of the audio-visual feedback. Details on the user studies methodology are described in Section 6.2.1.

3.4 UCD iterations

Pilot Studies (Chapter 4), comprising five informal studies, investigated different approaches to interaction design and gestural data mapping into auditory and visual feedback using the Myo and different audio software. These studies were also delivered with the aim to inform the initial development of the MyoSpat system.

In the first case study, Voice Gestures (Chapter 5), the UCD cycle is iterated five times with the aim of realising design solutions in two different contexts. In the first four iterations, with artistic output, the performance *VoicErutseG*, explores ways to manipulate and generate audio-visual media during a performance, in a context where the performer can move hands and arms freely (*unconstrained interaction*). Thus, a vocal performer was chosen as the subject of the study to perform a piece of music using MyoSpat system and a wireless headset microphone.
These four cycles also aimed at the exploration of various technological solutions to implement the MyoSpat system.

The fifth iteration, which context was set through the *Music Gesture Beatbox* performance, explored ways of interacting with auditory and visual feedback in a situation in which the performer's movements were constrained by instrumental elements (*constrained interaction*). Here a beatboxer was chosen to compose and perform a piece using the MyoSpat system with a wired microphone and a loop-station.

The second case study, HarpCI (Chapter 6), sees two iterations of the

UCD cycle. Those aim at extending the MyoSpat system to a wider range of instrumental performances, thus considering the constrains on performer movements imposed by the geometry of various musical instruments and their related playing techniques. The sixth and seventh iterations respectively led to the realisation of two performances: *The Wood And The Water* and *Star Cluster*. Each stage of the outlined methodology is summarised in Figure 3.3.

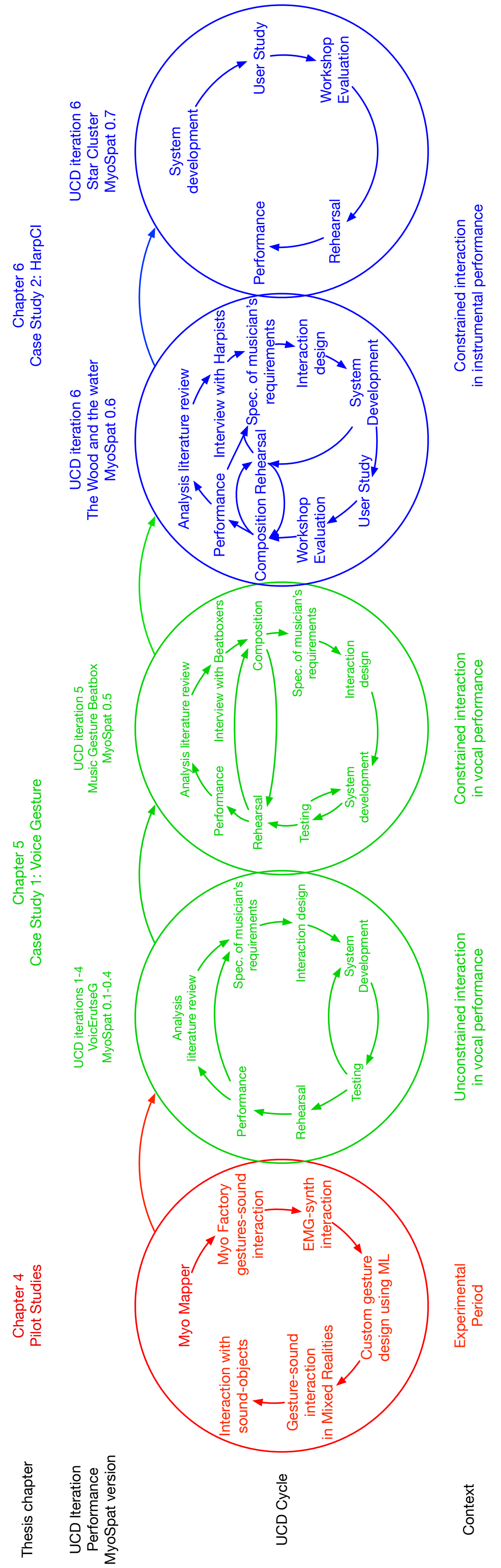


Figure 3.3: UCD cycle iterations.

3.5 Ethical considerations

All the research was undertaken in accordance with the ethical regulations and approved procedures of Birmingham City University (2014).

Chapter 4

Pilot studies

4.1 Introduction

This chapter describes five pilot studies that aimed to guide the initial phases of the designing new modes of interaction and the MyoSpat system development. Each study investigates different technological features of the Myo like the IMU and EMG sensing modality, factory software capabilities, and how to use these in conjunction with third party interactive music software.

The first pilot study explores the interaction with sound processing parameters through hand poses and arm movements using the factory gesture recognition software of the Myo. In the following study, it is described the use of the electromyogram signal to continuously control synth parameters. The third study investigates the possibility of realising a custom gesture library using Interactive Machine Learning. Building on the previous studies, the last two focus on the manipulation of sound through gestures as if it were a tangible object. The Myo Mapper, a software developed to enable and facilitate the use of the Myo in gesture-controlled interactive systems is also presented.

4.2 Methodology

Five pilot studies were conducted adopting an empirical approach. Each study was formed of three stages. In the first stage, ways of interacting with the sound parameters were designed and later implemented in the second stage through software development. Here, different strategies were adopted to investigate the capabilities of various tools matching the goals of this work. In the third stage, the capability and possible further developments of the interaction design and the developed software in musical performance were observed through the direct experience of the author.

Pilot studies were preceded by a period of development of the Myo Mapper software, which enabled the use of the Myo throughout this work.

4.3 Myo Mapper

At the time of writing, a number of software solutions to facilitate the use of the Myo in musical situations are available. One example is Leviathan by Precision Music Technology (2015), an application to easily control chords and effects parameters in Digital Audio Workstations (DAWs) through hand poses recognised by the Myo gesture recognition software. More control over Myo data mapping can be obtained by using Myo-Ableton (GonzaloNV, 2015), a Myo connector that maps Myo data to MIDI messages. Similarly, MyOSC by (Kuperberg, 2016) and myo-osc (Kamkar, 2015), are solutions to map Myo raw data into OSC messages. In addition to making Myo raw data available, Myo-maxpd (Caramiaux, 2016), an external for Cycling 74's Max, also gives access to connection and haptic feedback settings, like the Max external created by Jules (2016). Although these solutions have the full potential to support the creation of new interfaces for musical expression, they do not represent an easy-to-use solution for musicians that are not experts in interactive technology. Additional software would need to be developed for extracting high-level data features to facilitate the data mapping and use of machine learning for recognising non-factory hand poses and gestures. Due to the stochastic nature of the EMG signal, which makes EMG data hard to handle, the aforementioned software tools require pre-existent knowledge of how to design and implement their data feature extractors, calibration and scaling functions and means of communication to interactive machine learning software. They are, therefore, only realistically available to skilled developers.

By combining these aspects of interaction authoring into an easy-to-use workflow, Myo Mapper contributes to the work of the research community on the use of IMU and EMG based technology via the Myo. As such, it is useful to a wide range of users using different audio-visual software and interactive machine learning systems. The Myo Mapper software was developed as a bridge between the MyoSpat system and the Myo, but also to overcome some essential issues.

At the time of writing, Myo Mapper has been through three main iterations. This section describes the architecture of the third iteration and its implementation. Findings from an informal qualitative evaluation of the software are presented, together with a series of interaction design guidelines for the use of EMG data in creative applications drawn from the above experiences.
The source code and the binary files of the three versions of Myo Mapper are attached to this thesis in electronic format (see Appendix E, items no. 1, 2, 3).

The Myo Mapper software was developed using the C++ language, the JUCE framework (Roli, 2016) and the Myo Software Development Kit (SDK) (Thalmic Labs, 2015a). No other third-party software was used to realise this work.

4.3.1 Architecture and implementation

The back-end architecture of Myo Mapper is made of five main blocks: Myo communications, feature extractors, the OSC ports, shared spaces for storing application settings and a separate space for storing sensor data and extracted features. The GUI is made of three different windows: "settings", "calibrating and orientation" and "feature selection" (Figure 4.1).

The Myo SDK allows the application to communicate with the Myo through bindings to the `libmyo` C library. The entry point to the SDK is the Myo Connect application, which functions as a "hub", managing the connection between the computer and one or more Myos (Figure 4.1, a). The SDK gives access to the accelerometer, gyroscope, orientation and EMG data from the device and control over its vibrational motors (Figure 4.1, b).

To facilitate the mapping of the Myo data to audio-visual authoring environments and interactive machine learning software, Myo Mapper includes different feature extractors (Figure 4.1, c) that can be chosen in the "feature selection" window (Figure 4.1, e).

The GUI also enables the setting of the OSC communication between Myo Mapper and third-party applications and includes tools to visualise orientation data. These configuration settings (OSC ports, features, calibration and scaling parameters) are stored in a shared space to facilitate the communication between the GUI and the back-end (Figure 4.1, d). Myo Mapper includes an OSC receiving port, where OSC messages can be sent to control calibration and to scale features (Figure 4.1, g).

Gestural data is sent out through two independent OSC ports, *main* and *ml* (Figure 4.1, f). The *main* port is used to send Myo data to musical applications able to receive OSC messages, while the *ml* port concurrently sends OSC data to machine learning software. This second port was implemented to keep the end-user from having to use additional software to organise data features into a single feature vector to stream to the interactive machine learning software.

4.3.1.1 Feature extractors

From their evaluation of five EMG features – Mean Absolute Value (MAV), Variance (VAR), Willison Amplitude (WAMP), Waveform Length (WL) and Zero Crossing

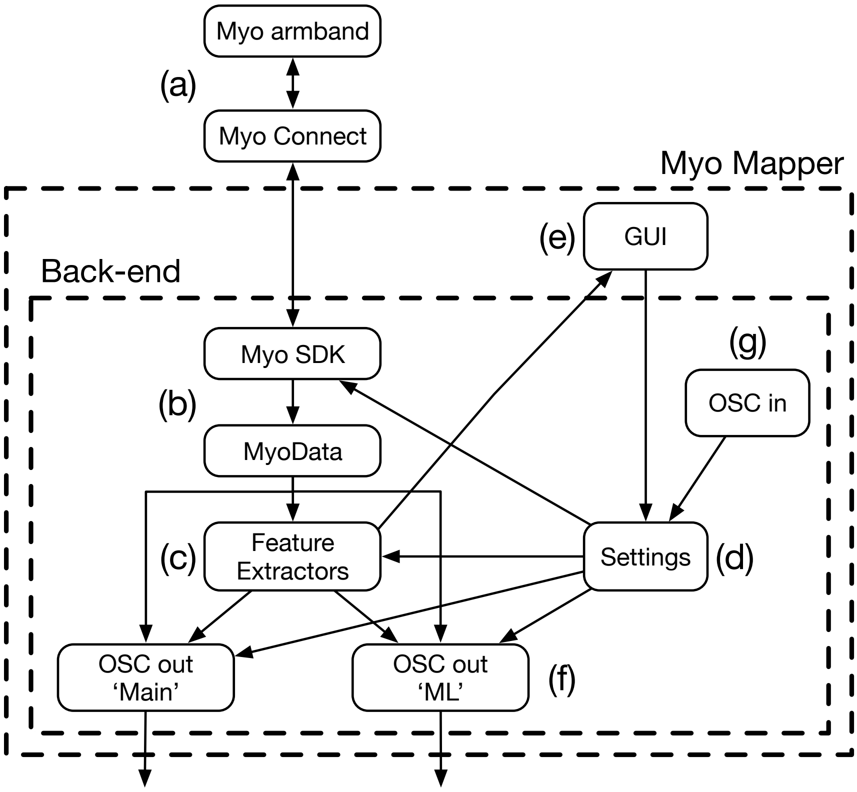


Figure 4.1: Myo Mapper architecture

(ZC) – Arief et al. (2015) recommended the use of the eight EMG signals' mean absolute value (MAV) for time series features extraction. Based on this knowledge, the MAV is here calculated as in Equation 4.1.

$$MAV = \frac{1}{N} \sum_{k=1}^{N} |X_k| \tag{4.1}$$

Arief et al.'s evaluation of EMG data series, the Zero-Crossing feature was found to be less useful than the others. On the other hand, the work by Roomi et al. (2010) and Boyali and Hashimoto (2016) on hand gesture-recognition through electromyographic analysis showed the importance of the zero-crossing (ZC) feature to observe spectral qualities of the EMG signal. Consequently, this feature was implemented into Myo Mapper. However, rather than using the equations identified by Roomi et al. or Boyali and Hashimoto, here Equation 4.2 from Arief et al. (2015) was implemented because it takes into account background noise present in the EMG signals generated by the Myo.

$$ZC = \sum_{k=1}^{N} sgn([X_k - 0.4][X_{k+1} - 0.4]) \tag{4.2}$$

where:

$$sgn(x) = \begin{cases} 1 \; x > threshold \\ 0 \; otherwise \end{cases}$$

Additional implemented data features were: minimum (MIN), maximum (MAX), absolute value (ABS), moving average (MAVG), first and second-order difference (FOD, SOD) based on Fiebrink's work with Wekinator and WekiInputHelper (Fiebrink, 2009). The MIN and MAX features were implemented to analyse the data range through which a gesture or pose is represented. The ABS is useful to observe EMG data that may be the negative component of a mirror image signal's positive component. FOD and SOD allow analysing the Myo data variation over time, for instance, in the case of orientation, the velocity and acceleration of arm movements. The MAVG can serve to smooth the EMG signals, so to make them usable as a control signal.

Feature extraction is useful to process sensor data for pose or gesture classification. For instance, the moving average of EMG absolute values (EMG ABS MAVG) over a window of 40 data points (Figure 4.2, red line) allows for the differentiation of the arm from its resting position (Figure 4.2, pink area) to performing in a fist pose (Figure 4.2, green area) by evaluating if the input value is greater than 0.05. With this method, gestures may be misclassified (as seen along the red MAVG curve in Figure 4.2), as a consequence of the time needed to calculate the data feature. However, as EMG data are streamed at a frequency of 200Hz, this classification latency is minimised. Figure 4.2 also shows that by using the absolute values of EMG raw data (Figure 4.2, blue line) the algorithm would have misclassified the input data more frequently – specifically, nineteen times using the EMG ABS as an input feature vector, opposed to nine times if using the EMG

ABS MAVG feature. An even less accurate response would have been obtained if using EMG raw data.

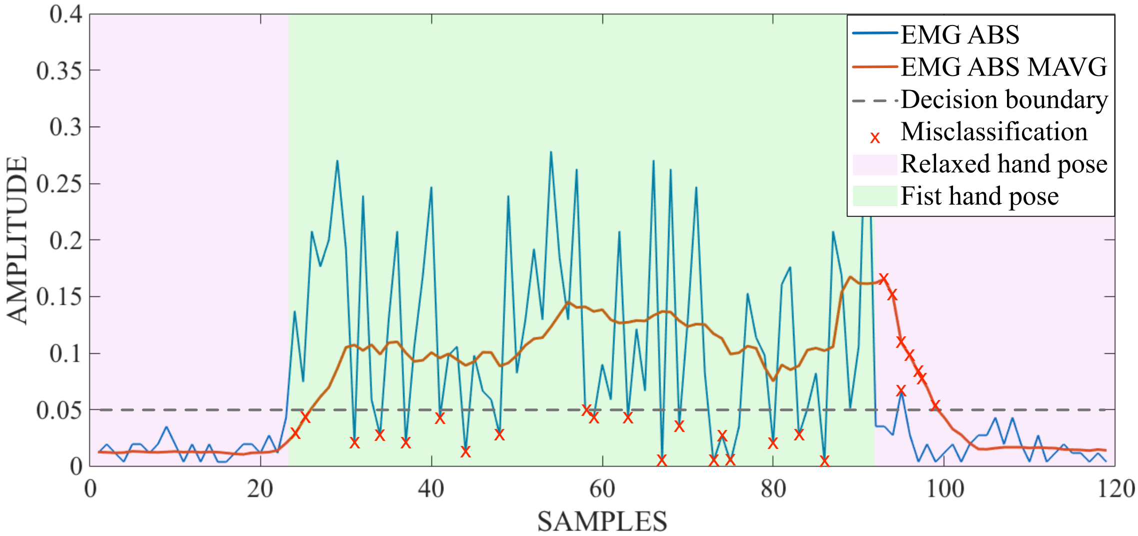


Figure 4.2: Comparison of EMG ABS and EMG MAVG feature data

4.3.1.2 Orientation data scaling

Orientation raw data from the Myo is represented in a range $[-2\pi, 2\pi]$. Myo Mapper scales this data to the range $[0, 1]$ to facilitate data mapping into audio-visual processing parameters.

The Myo suffers from yaw value drift by 3.7 deg/s before reaching a stable value (Nymoen et al., 2015). A similar issue was found in the roll value. To solve this issue, a *set origin* functionality that sets the current orientation data (yaw, pitch or roll) to a value of 0.5 was implemented. During the realisation of pilot studies (Chapter 4), it emerged that the variation in orientation data depends on the way the Myo is worn. Flipping or rotating the device on the arm changes the relationship between EMG channels and muscle groups and inverts the IMU. Without a software solution, users would have to take off the Myo and turn it around. To solve the issue, a *flip* function, $y = 1 - x$, on yaw, pitch and roll, was implemented. It was also observed that in some cases, arm movements produce a variation too small in the data to control audio parameters effectively. To address this, a range of functions were implemented to limit the values in input (*in min*, *in max*) and rescale the values in output (*out min*, *out max*).

4.3.1.3 OSC message streaming

The user can choose to stream Myo data and the extracted features to one or both OSC ports (*main* and *ml*). When selecting a feature in the feature selection window (Section 4.3.1.6), Myo Mapper combines the feature's data into a unique OSC message with tag `/myoX`, where `X` is the number of the selected Myo (i.e. `/myo1` for the Myo number one). The data included in the messages are arranged in the same order as they are selected. For instance, if streaming the raw orientation data and the EMG raw data, the OSC message will contain 11 floating-point values (i.e. `/myo1 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.9 0.11 0.12`), where the first three are the raw orientation data and the last eight the raw EMG data.

An OSC receiver port can be set through the "settings" window (see Section 4.3.1.4). Through this port, Myo Mapper can receive messages to configure remotely *set origin*, *flip*, *in min*, *in max*, *out min* and *out max* settings.

OSC messages' tag and type are shown via a tooltip display in the Myo Mapper software, and specified in Appendix B of this thesis and the online Myo Mapper Wiki page (Di Donato, 2016).

4.3.1.4 Settings window

The "settings" window comprises controls to set the OSC communication (ports' number and IP address). When more than one Myo is connected to the Myo Connect, it is possible to select which of the Myos' data series will be streamed.

4.3.1.5 Calibration and scaling window

The "calibration and scaling" window (Figure 4.3) embeds controls to recall the *set origin*, *flip*, *in min*, *in max*, *out min* and *out max* functions. When dragging the cursor over each button, a tooltip message containing instructions for controlling them via OSC will appear.

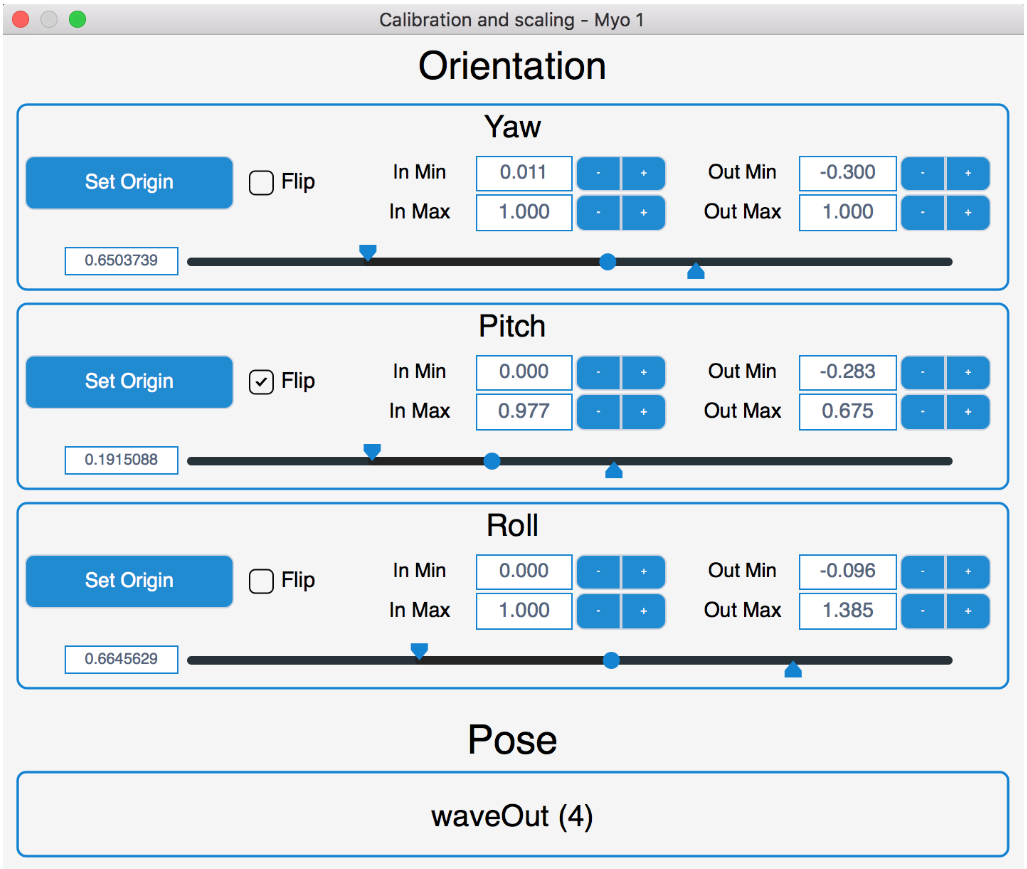


Figure 4.3: Myo Mapper "Calibration and scaling" window

4.3.1.6 Feature selection window

The feature selection window (Figure 4.4) allows the selection of one or more data features to stream through the *main* and *ml* OSC ports, respectively, using the *to main* and *to ml* toggles. The organisation of the features in a data tree is to represent the data processing chain of each feature. For instance, the moving average of the raw EMG zero-crossing rate can be streamed by selecting the fourth feature from the top in the EMG panel (Figure 4.4). A tooltip message is also implemented to inform the user about the data processing chain; referring to the above example, the tooltip would contain the text: `EMG raw -> zero-crossing rate`. Similarly, information relative to the OSC message of each feature (message tag, number of values and type, sender's port number and IP address) is shown in a tooltip window activated when hovering the cursor over a feature.

Finally, the Buffer Size value input boxes allow modifying the size of the buffer used to calculate the related feature. For example, in the case of the Moving Average feature, the buffer size refers to the number of samples over which the moving average has to be determined.

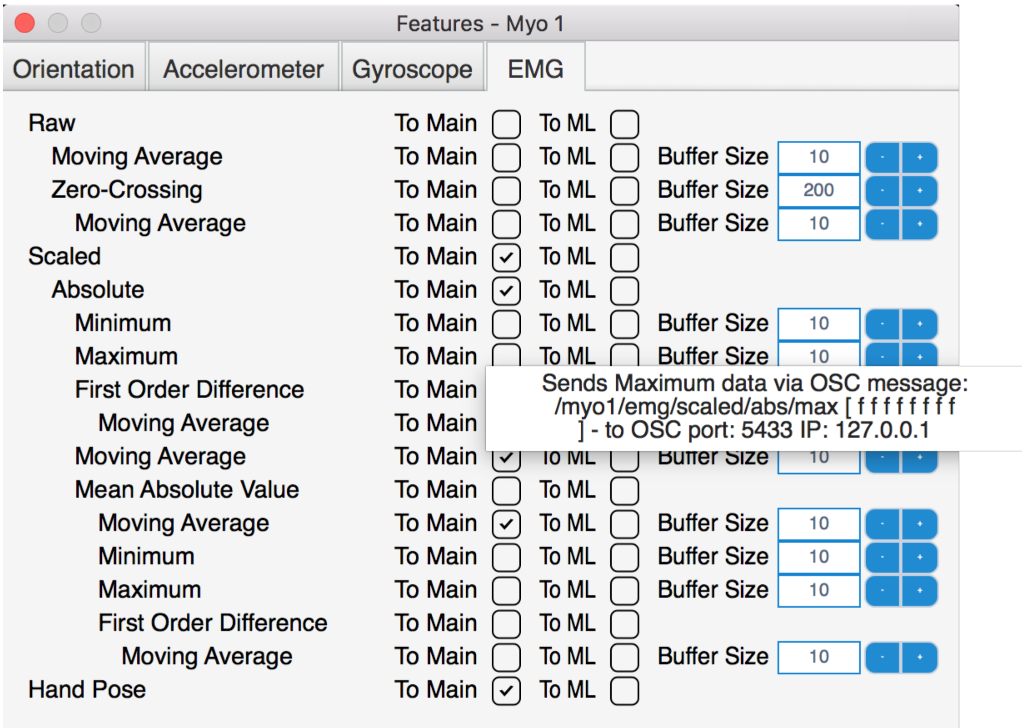


Figure 4.4: Myo Mapper "Feature selection" window

4.3.2 Informal evaluation

Myo Mapper was informally evaluated through a half-day workshop delivered at Goldsmiths, University of London.

Six students and two researchers from Goldsmiths' Computing, Music and Psychology departments took part in the workshop. Before the workshop, participants were given knowledge related to electromyography and its use in creative computing applications. All participants were aware of the Myo, and half of them had used it prior to the workshop. The Myo Mapper software was first presented and demoed to the participants. The demo included several interactive audio applications realised using Pure Data, Max and Wekinator. Successively, participants were left to explore the software during a hands-on phased of the workshop.

Problematic aspects of the GUI were highlighted. The organisation of labels in the feature selection window was not clear for some participants. Most of the participants asked several times about the OSC message's tag, type and number of values for each feature. Based on this feedback the verbose GUI tooltip was implemented.
Feature extractors were easy to use for all participants. In particular, the moving average of EMG absolute value was found to be very useful to filter out background noise. Through this feature, an undergraduate guitar student built an EMG data Neural-Network MLP classifier using Myo Mapper in conjunction with Wekinator to recognise three different plucking gestures using a moving average of EMG absolute values and the absolute value of gyroscope data. In fifteen minutes, the student created a model whose performance, evaluated through direct evaluation, was approximately 90% accurate.

A second participant, a PhD student in computing, aimed to recognise British Sign Language (BSL) gestures using Myo Mapper and Wekinator. However, due to the complexity and number of gestures she sought to classify, a robust machine learning model could not be built quickly. At the end of the workshop, participants commented on the software as being very useful and advised the implementation of additional features such as Root Mean Square (RMS) and Bayesian filters.

4.4 Pilot Study 1: Sound transformations through Myo's factory gesture library

The objective of this first experiment was to explore the Myo device and its potential for manipulating sound in real-time through hand poses and gestures using the Integra Live software as the audio engine.
Integra Live is a modular environment for interactive sound transformation and synthesis developed at the Royal Birmingham Conservatoire (Integra Lab, 2013). Integra Live offers an easy to use environment for controlling sound processes through an input device.

Accelerometer, gyroscope, orientation, and hand poses data were mapped into processing parameters of two different audio effects, taking advantage of Integra Live's inbuilt mapping functionalities. Both audio effects processed an audio file containing a pre-recorded wind sound.
The first effect, selected by performing a fist (Figure 4.5), consisted of a stereo granular synthesiser, based on asynchronous granular synthesis (Roads, 2012). The pitch shift of the audio grains was controlled by arm movements on the vertical axis as tracked by Myo's pitch value. The second audio effect, selectable by performing a spread finger pose (Figure 4.6), controlled the reproduction speed of the audio file by a twisting movement of the wrist, tracked by Myo's roll value. The output of both effects was then processed using a stereo panner controlled by linear mapping of Myo's yaw value.
The control of the pitch shifter by orienting the arm on the vertical axis was inspired by studies exploring the relationship between pitch and verticality, in particular referring to Wilkie et al. (2013) and their work on the application of Conceptual Metaphor Theory, where they use the conceptual metaphor *high pitch is up/low pitch is down*. An extended reflection on the relationship between the vertical position of the hand and the pitch is presented in Section 5.3. The twisting gesture of the arm,

to control of the playback speed parameter, is intended to emulate the action on a real knob. The control of the stereo panning through the horizontal movement of the hand aimed to establish a direct link between the sound's spatial position in the auditory scene and arm orientation on the horizontal plane.

The video recording no. 1 in Appendix F.1, demonstrates various audio effects that were triggered through hand poses from the Myo gesture library, as well as variations of audio processing parameters in various orientations of the arm. After testing this method of interacting with the computer, it was observed that even though the Myo factory gesture recognition software imposed modes of interaction, through the use of conceptual metaphors, it is possible to design interactions with the potential of being easy to learn and embodied.
From these and observations during the interaction design stage, it is possible to hypothesise that in a situation like this one, where the palette of recognised gestures is limited, it is probably best individuating the how to adapt those gestures to the affordance of GUI elements or to interaction metaphors.

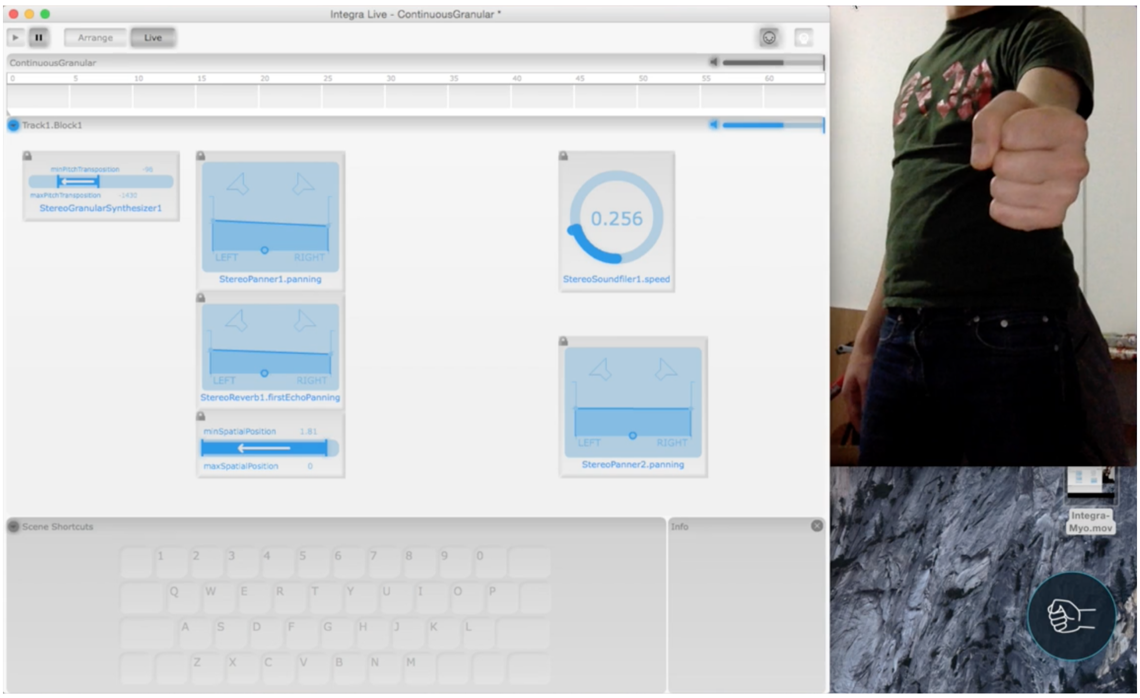


Figure 4.5: Fist pose. Screenshot from video recording no. 1 in Appendix F.1

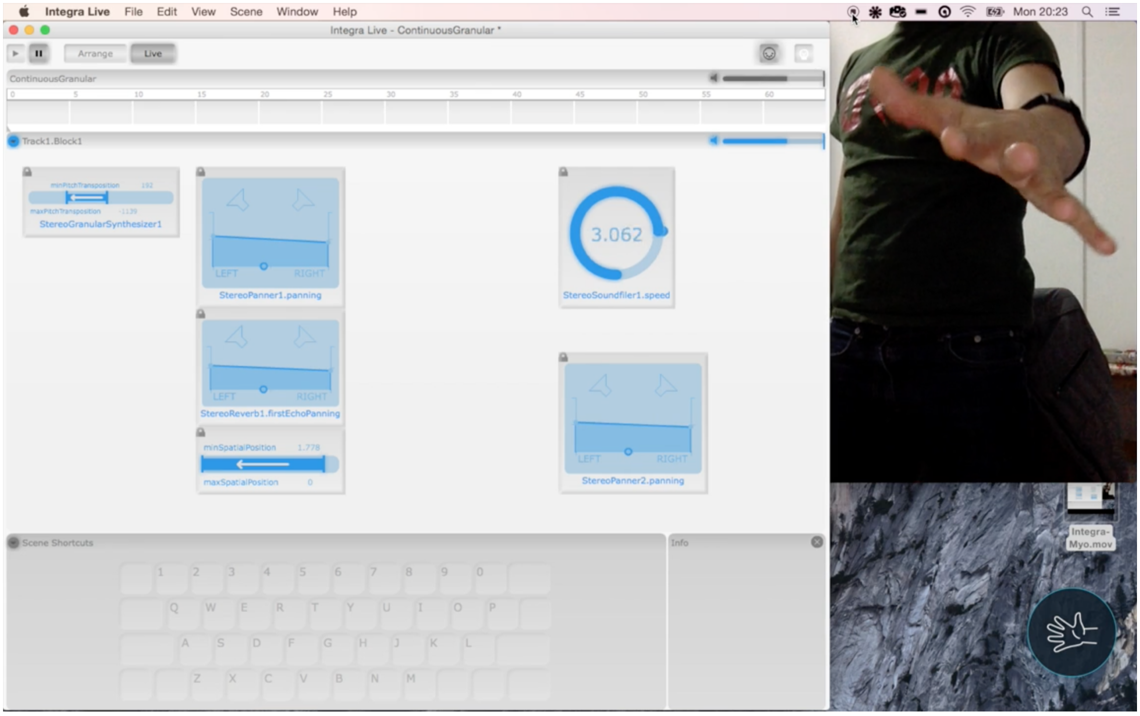


Figure 4.6: Finger spread pose. Screenshot from video recording no. 1 in Appendix F.1

4.5 Pilot Study 2: Interaction with synth parameters through muscles activity

This second study explored how raw EMG data can be used as parameters to control sound processes. An example can be found in the performance *Myogram* by Atau Tanaka (2015), who used the EMG signal to control audio transformations and its sonification as an audio source. A different approach was used by Pauletto and Hunt (2006). They used the EMG signal to prove its potential benefit as an auditory display. They mapped the amplitude of EMG signals to the amplitude of sine oscillators, with fixed frequency in a harmonic relationship.

In this case study, the approach of Pauletto and Hunt was adopted. Here, the absolute value of EMG data was directly mapped onto amplitude parameters of a bank of eight oscillators, with initial frequency and harmonic relationship of 200 Hz.

From this application were observe various aspects of our muscles activity, how

the Myo senses it, and how the audio feedback can differ with any muscle contraction or hand movement. When using EMG sensing capabilities of the Myo, it is important to maintain consistency in wearing the Myo at each use, in order to ensure the reproducibility of the interaction with sound parameters. The continuous control of amplitude parameters through muscle contraction provided an exploratory approach to both sound and gesture. During tests, it was observed that the audio engine would respond to any small contraction of the muscles and hand movement, and any of those gestures would produce a unique sound. This leads to the hypothesis that this mapping strategy could support the design of *unconscious interactions*.
This mapping approach has the potential to make a system vulnerable to any movement, which has the potential to ensure expressivity in the sound generation, but it is also important to considerate that this might lead to the production of unwanted sound timbres. This phenomenon was more evident when assuming different hand poses, as demonstrated in the video recording no. 2 in Appendix F.1. A comparison of Figure 4.7, 4.8 and 4.9 shows the differences in oscillators amplitude parameters (sliders at the bottom of Integra Live GUI) with different hand poses.

For this reason, the possibility to classify hand poses in order to have more conscious control over the sound will be explored in the third Pilot Study.

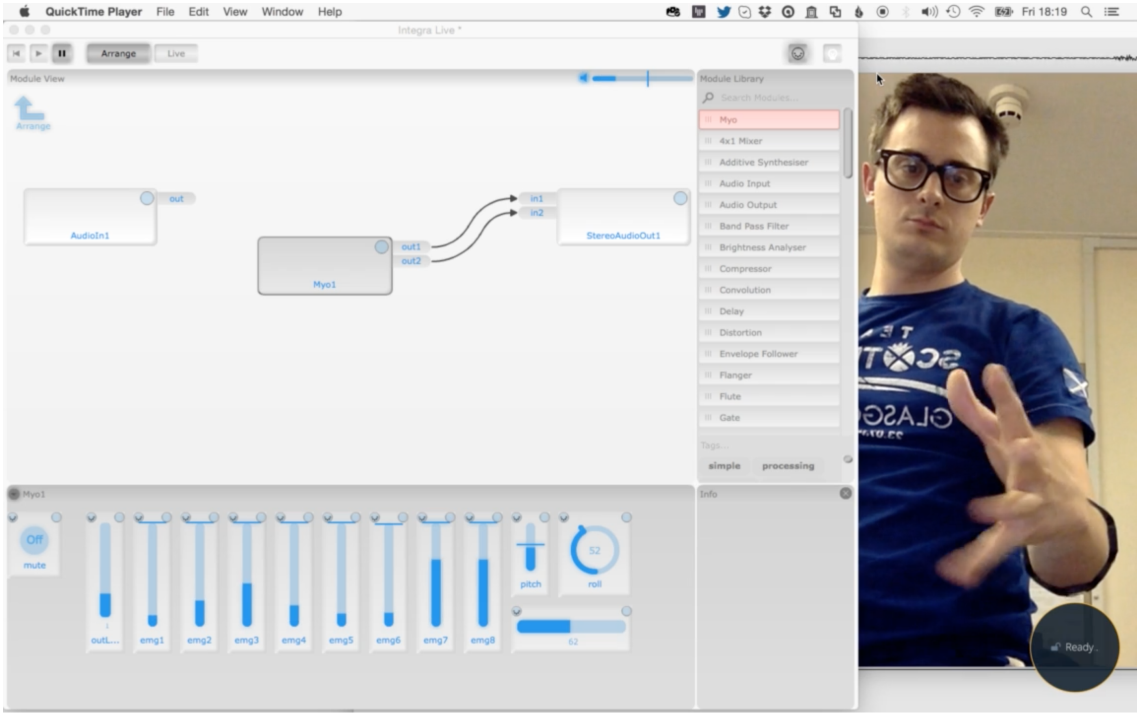


Figure 4.7: Hand pose (a) performed for engaging forearm muscles. Screenshot from video recording no. 2 in Appendix F.1

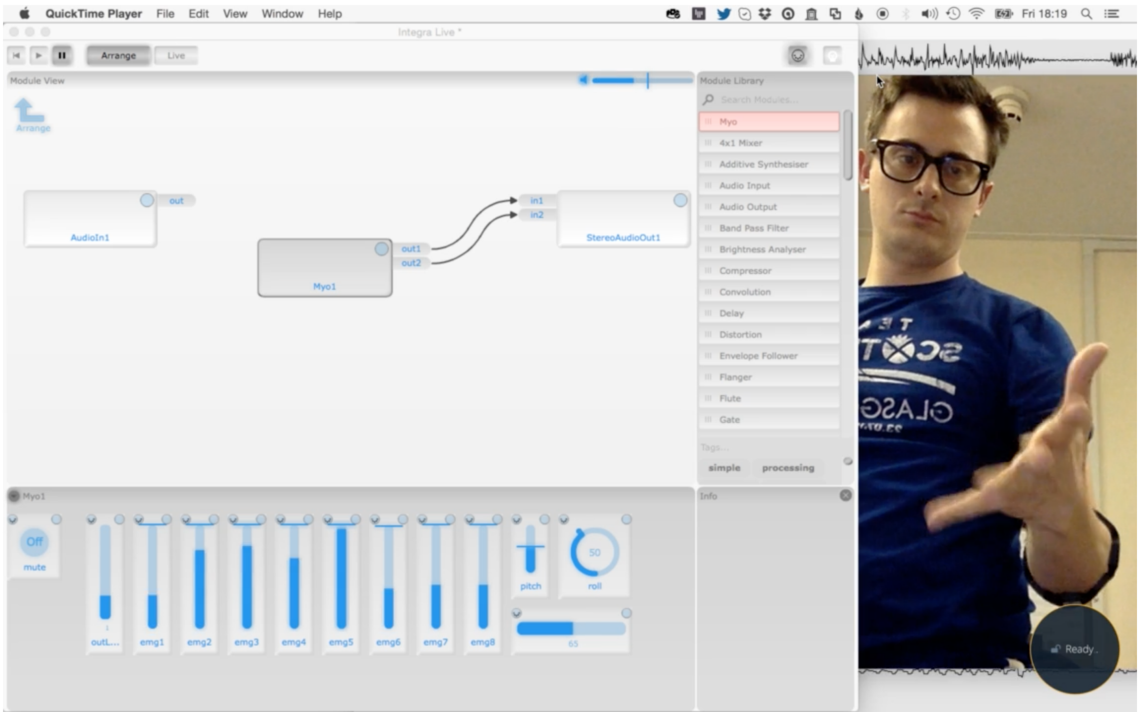


Figure 4.8: Hand pose (b) performed for engaging forearm muscles. Screenshot from video recording no. 2 in Appendix F.1

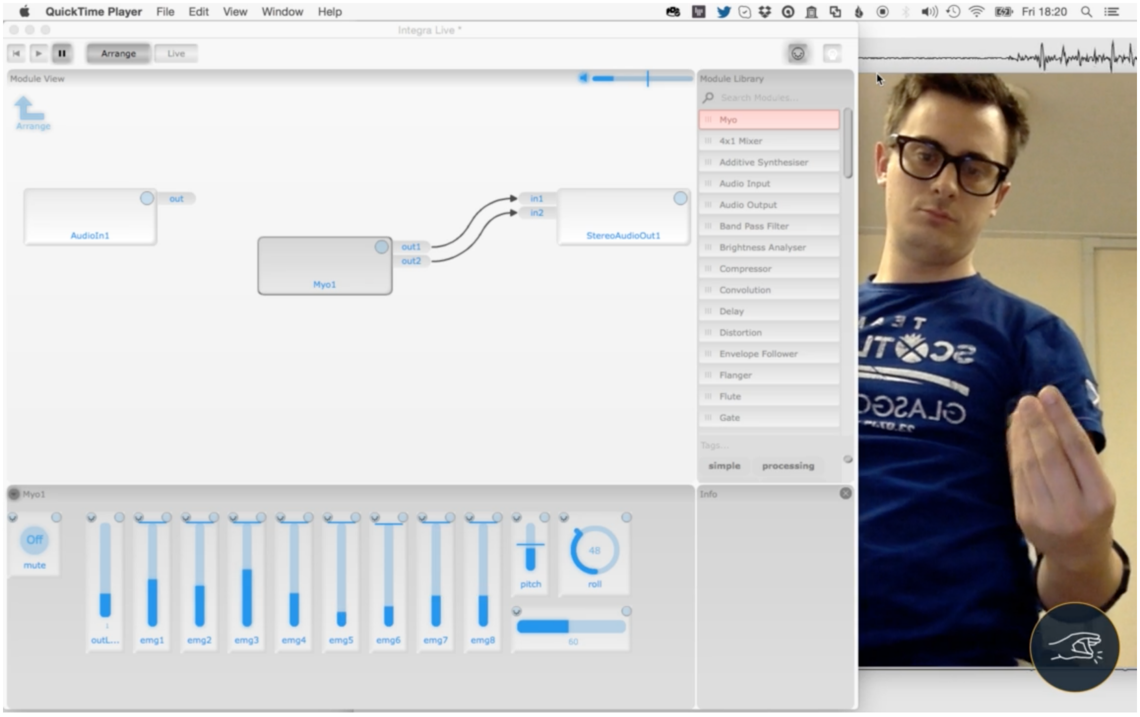


Figure 4.9: Hand pose (c) performed for engaging forearm muscles. Screenshot from video recording no. 2 in Appendix F.1

4.6 Pilot Study 3: Interaction with stereo panning effect parameters through hand gestures

During the previous pilot study, it was observed that we tense our muscles in distinctive ways when performing certain hand poses. In this third Pilot Study, three different poses were individuated empirically through observations of EMG absolute (ABS) values variations while improvising different movements. These were the *pointing* pose with the index finger (Figure 4.10, A). The second was the "L" letter in American Sign Language (ASL) (Figure 4.10, B), here called *L* pose. The last one was similar to the "i" letter in ASL, but also the ring and middle finger outwards (Figure 4.10, C), here called *three* pose.

Built on this knowledge, this pilot study aimed to explore the use of a custom gesture vocabulary, based on hand poses described above, to control the pan parameter of a stereo panner. At the recognition of each pose, the software, developed using Pure Data, triggered different panning settings. By performing the *pointing* pose the sound would be panned to the left channel, the *L* pose to the right and the *three* pose to the centre of the auditory scene.

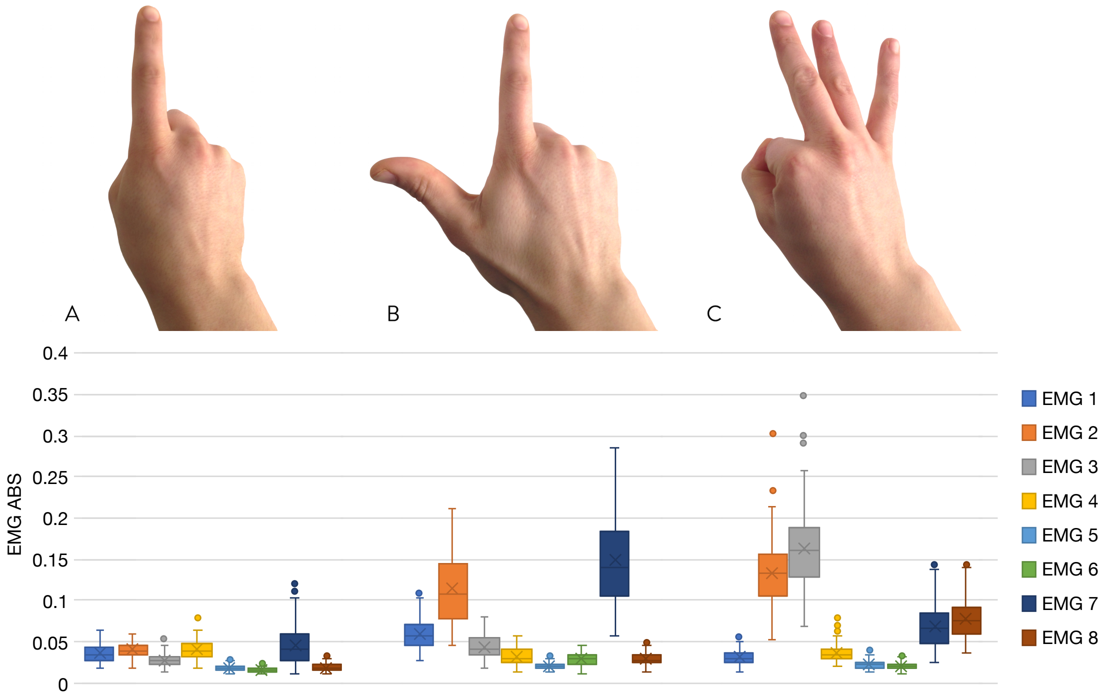


Figure 4.10: Hand poses selected from the second pilot study by observation of EMG ABS values. The chart summarise EMG ABS values from a 100 samples recording of the three poses.

The accomplishment of a successful result required the implementation of complex mapping functions, with eight EMG signals in input and the pan parameter as output. The complexity could rise exponentially, considering the stochastic nature of the electromyogram signal. As explored in Section 4.3, the adoption of interactive machine learning approaches can make the mapping process easier.

In this study, hand poses were classified using a Support Vector Machine (SVM). It delineates a linear path of separation, also called hyperplane, among data types (classes) by trying to maximise the distance between them considering data points closer to the margins dividing the two classes, also called support vectors (see Figure

4.11).

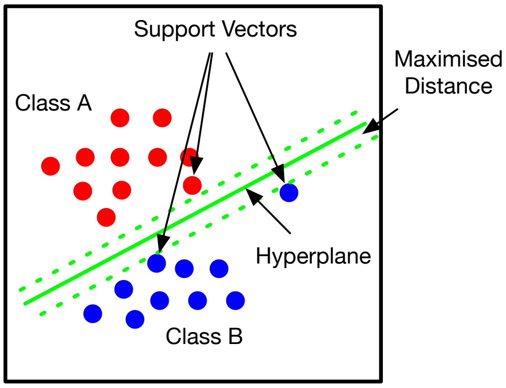


Figure 4.11: Support Vector Machine (SVM), non-linear classification.

In a non-linear classification problem, SVM algorithms use what is called a Kernel Trick, to map a high dimensional data space into a linear feature space by creating high dimensional models in which data are linearly separable (Antao, 2017). This process aims to improve the effectiveness of the hyperplane for separating feature data in different classes (Smola et al., 2000)(see Figure 4.12). A Kernel Trick can use different types of functions to divide classes, such as linear, radial basis functions, polynomial and sigmoid.

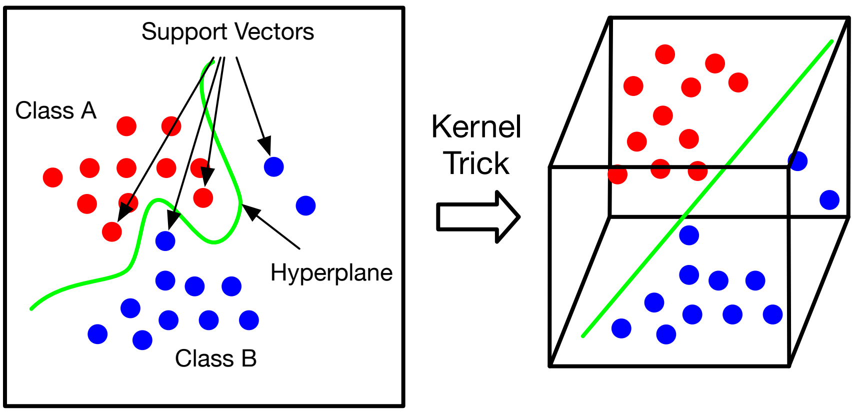


Figure 4.12: Support Vector Machine (SVM) Kernel Trick.

Prior the implementation of the system, considerations about the suitability of two software solutions for musical applications, namely Wekinator (Fiebrink and Cook, 2010) and ml.lib Bullock and Momeni (2015) (described in Section 2.5.2.2), for mapping Myo data to sound processing parameters were made.

If using Wekinator to classify Myo data, Myo Mapper data are sent first to Wekinator, and then Wekinator's output to the audio engine. Thus, Myo data have to go through two different environments before reaching the audio engine, developed using Pure Data (see Figure 4.13a). Alternatively, if using the ml.lib machine learning library for Pure Data, the gesture recognition process can be embedded in the audio engine. Hence, Myo data captured using Myo Mapper are sent directly to the audio engine (see Figure 4.13b). This second approach was adopted in favour of a simpler software architecture.

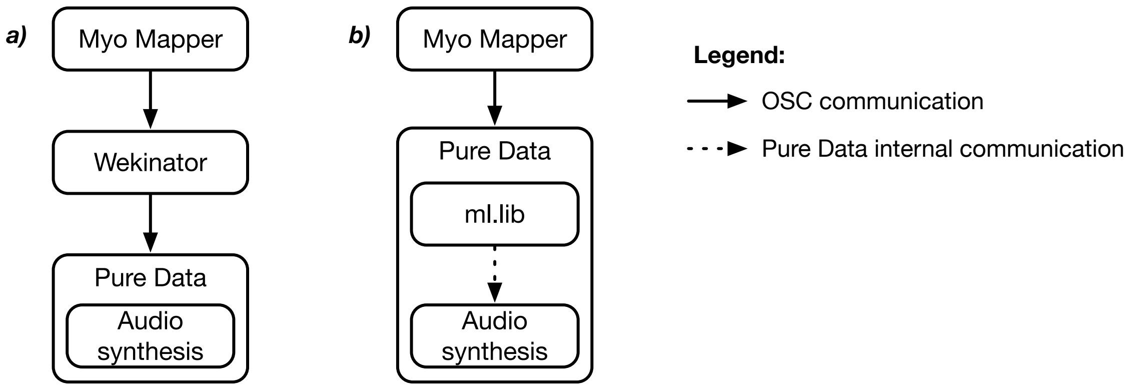


Figure 4.13: Comparison of gesturally controlled audio engines data flow using machine learning for gesture-sound parameters mapping layer. Diagram *a* illustrates the data flow if implementing gesture recognition using Wekinator. Diagram *b* shows the data flow when using ml.lib in Pure Data.

In this pilot study, a linear kernel function was utilised to solve a non-linear classification problem, and the machine learning model was trained with a feature vector composed of the moving average of EMG absolute values over fifty samples. All values were scaled within a range $[0, 1]$. This feature was used as it smoothes the EMG signals, so to make them usable as control parameters. The training set contained 339 samples (104 *pointing* pose, 111 *L* pose and 124 *three* pose).
The model was evaluated through direct evaluation (Fiebrink et al., 2011) and the result of accuracy was 95%. As observed in Section 4.3.1.1, the reason for the data misclassification could have been the time needed for the feature extraction process. Thus, to increase the responsiveness of the algorithm, the SVM was fed with the moving average of EMG data's absolute values over twenty samples. Additionally,

the SVM output data was filtered. The data filter would output a value every time a class had been detected consecutively three times, thus resulting in greater accuracy.
The output from the SVM was then mapped into panning parameters. A video demonstration of the system is included in this thesis as Appendix F.1 no. 3.

Three hand poses were classified with high accuracy (95%) by using an SVM algorithm. Observations were made in the potential to use machine learning for classifying hand poses. Although the gesture recognition algorithm approach was very robust, those poses could result in disrupting a performance where the musician plays a musical instrument. An attempt to train a model able to recognise a broader range of poses that was made. However, the results were not successful. From this pilot study, it was hypothesised that such an outcome was due to the EMG crosstalk between forearm muscles (Wehner, 2012), and the Myo's sensing capabilities.

4.7 Pilot Study 4: Interaction with sound in Mixed Reality

In this pilot study, modes of interaction with sound adopting the idea of sound physicalisation in the context of Mixed Reality are explored.

The reality-virtuality (RV) continuum is defined by Milgram et al. (1995) as a linear connection between the real and virtual environments included within a larger class: Mixed Reality (MR), as shown in Figure 4.14. Benford and Giannachi (2011) write about the idea of "trajectories" in MR, which are links and relationships, as well as boundaries between the two environments. From this idea, they created a conceptual framework for understanding cultural user experiences. In this pilot study, the idea of trajectories is used with the aim to design embodied interaction with audio parameters.

Tanaka (2012a) presents the idea of sound physicalisation. It refers to the

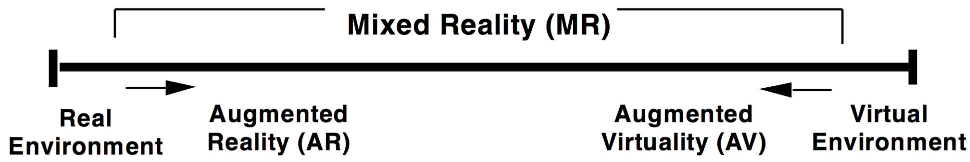


Figure 4.14: Simplified representation of an RV Continuum (Milgram et al., 1995)

translation of phenomena that are intangible in haptic terms, such as sound; in other words, the possibility to use scientific visualisation methods or the *"creative transduction of sound into physical phenomena"* as a way to represent intangible entities.

Based on the idea of trajectories of Benford and Giannachi (2011) and sound physicalisation in Tanaka (2012a), this pilot study investigates the potential of gestural interaction design as a method to initiate and experience sound's physicalisation and trajectories between the virtual and real world. The interaction with virtual objects and the auditory feedback are designed to mirror the interaction with the real-world correspondent object. Thus, I have used gestures that have the potential to summon familiar semantic signs, the so-called 'iconic gestures' by Kendon (2004). These gestures are also designed, taking into account the affordances of the real-world correspondent object, those interactions that determine the causality of a sound (see Section 2.3.1). Here the concept of embodiment through the action-perception loop is also relevant, as it is only through the action with the virtual object that we can reveal those affordances.

The interaction with an imaginary piece of paper occurred through a crumpling paper gesture, which controlled the generation of a crumpling paper sound (Figure 4.15). A real cardboard bin was translated into a virtual cavern through a reverberated crumpling paper sound generated whenever the imaginary piece of paper would have thrown into it, or the interaction with the paper would have happened inside the bin (Figure 4.16).

These modes of interaction are presented in the video recording no. 4 in Appendix F.1, in the following narrative. To demonstrate first the interaction with the real-world correspondent objects, a piece of paper is first scrunched up and throw it into the cardboard bin. Here, it is possible to hear a crumpling paper sound, when the paper is scrunched and when it is thrown in the bin. Next, an imaginary piece of paper is picked up in front of the camera and throw into the bin. At this moment, the interaction design (the crumpled paper gesture) and the auditory feedback (the scrunching paper sound and the reverberated version of it) give identity to the virtual objects: a piece of paper and a vast cavern. Finally, the imaginary piece of paper in the cavern is then looked for, picked up, and thrown back in front of the camera.

Myo EMG data were streamed to Integra Live, the audio engine. The Integra Live project contained an audio file playback module (Soundfiler), playing a pre-recorded crumpling paper sound in a loop, the amplitude of which was modulated through linear mapping of the EMG data's absolute value. As a result, the crumpling paper sound was produced every time the forearm muscles were engaged in the *crumpling* gesture. A reverb module was used to simulate the acoustic effect of an object thrown into a cavern. As the paper bin was placed on the floor, the user moved the arm pointing downwards to throw the object. The Myo pitch values, describing the arm orientation on the vertical axes, were mapped linearly to the input gain parameter of the reverb effect.

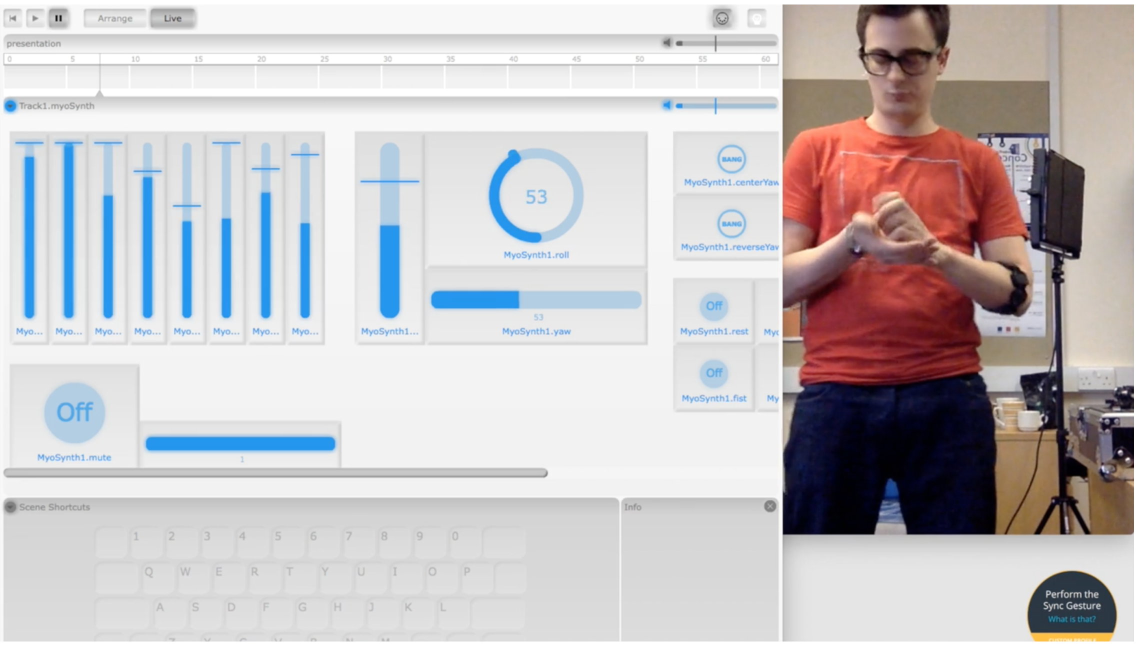


Figure 4.15: Crumpling gesture without paper. Screenshot from video recording no. 4 in Appendix F.1

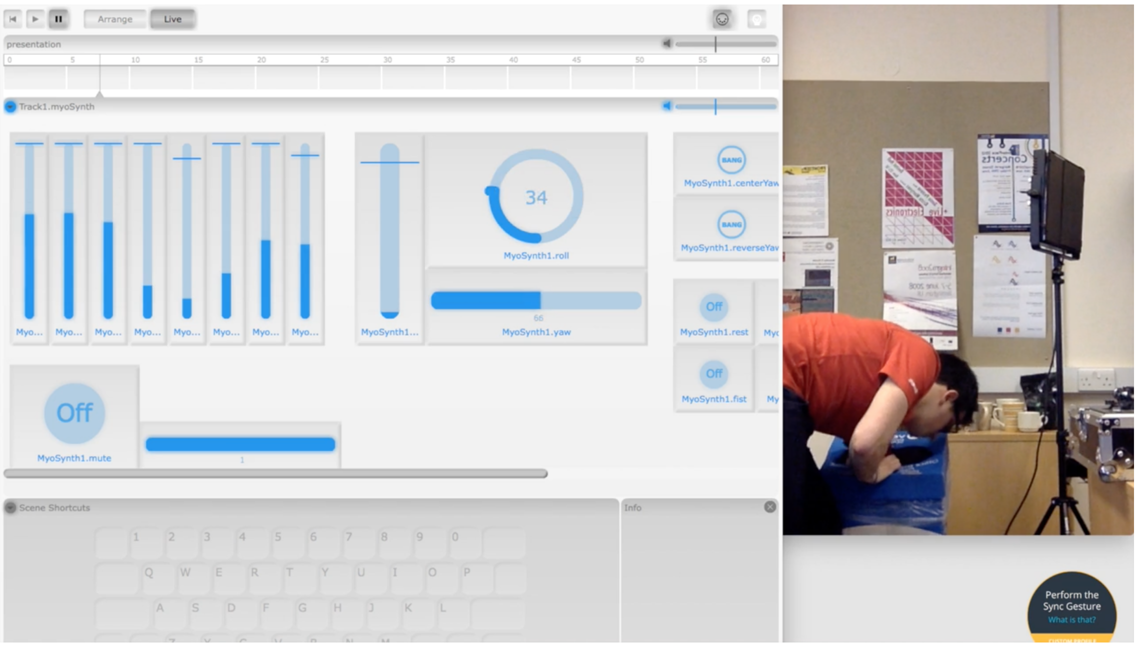


Figure 4.16: User looking for a piece of paper in the paper bin. Screenshot from video recording no. 4 in Appendix F.1

In this pilot study, principles of sound physicalisation resulted in fostering the interaction with real and virtual objects, thus in the generation of trajectories between the real and virtual world. Consideration of the affordances of the real-world objects allowed the design of embodiment and natural interactions with the virtual ones.

During testing, in the action-perception loop, I could have easily forgotten about the parameters or the mapping that would have generated the audio feedback, and associate a quiet crumpling paper gesture to a quiet crumpling paper sound, and vice-versa. Similarly, for the simulation of the cavern reverb, the combination of the bin's position and the mapping of pitch parameters allowed the interaction with the piece of paper into a cavern and removed the feeling of controlling the mix/wet signal ratio. Thus, it was a combination of real and virtual world factors that allowed the embodied interaction with sound.

4.8 Pilot Study 5: Interaction with "sound-object" in space

In this last pilot study, I explored the concept of sound physicalisation further, towards the idea of interaction with a "sound-object". With sound-object, I do not refer to Schaeffer's "object sonore" defined as a sound event over time that is perceptually detached from its source (e.g. the sound of a singer reproduced by a loudspeaker) (Schaeffer, 1966). Here a sound-object is the abstraction of looped audio file or live sound source positioned within a three-dimensional auditory scene as if it would be a tangible space with a focus on spatial characteristics of the sound, rather than the actual real-world correspondent object's sound.

Marshall et al. (2009) describe the idea of controlling sound spatialisation parameters by "steering" of a sound source within the space using a joystick, mouse, or any other gestural. Building on their work, modes of interaction that allow grasping, moving and throwing the sound-object within the virtual environment, or drawing trajectories which it will travel through were designed.

Building on those modes of interactions, I created a system with the Myo as the

input device and an audio-visual engine that renders the virtual environment and the drawn trajectory visually, as well as from the auditory feedback.
The system allowed the navigation of the virtual environment via a pointer represented by a 3D model of a hand (see Figure 4.17).

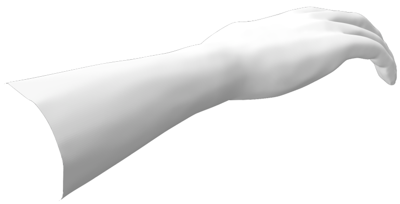


Figure 4.17: 3D hand model.

The sound-object is represented in the GUI by a sphere, here called "sound icon", whose texture is determined by the qualities of the post-processed sound. In the mapping of sound-texture features of the sound-object, I refer to the work of (Giannakis, 2006) concerning the relationship between sound and visual textures. Their findings report that there is a strong relationship between (i) sharp sounds and coarse textures, (ii) periodic sounds and non-granular simple textures, (iii) noisy sounds and granular and complex textures, and (iv) dissonance and repetitiveness in the texture.
Thus, to establish the texture of the sphere, the more noisy is the sound, the more granular and complex is the texture of the sound icon.

The GUI also represents the trajectory sound-object when in movement. In the physical world, a trajectory of moving objects can be observed simultaneously in the visual and auditory domain only under certain circumstances; for example, the condensation trail generated by an aeroplane and the Doppler effect. Even though unrealistic, examples of visual representations can be found in gaming applications such as Rory McIlroy PGA Tour in Figure 4.18.
Similarly, the sound icon shows a trace (Figure 4.19) that aims to represent the trajectory path of the sound spatialisation.

Figure 4.18: Rory McIlroy PGA Tour (Tiburon, 2015)

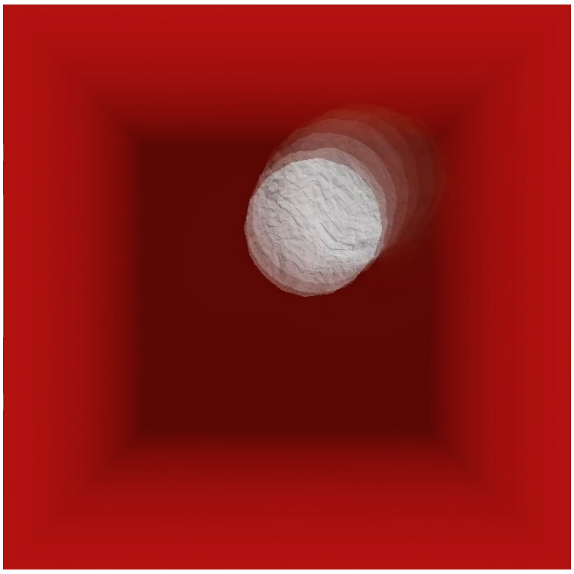


Figure 4.19: sound icon during movement

The interaction design consists of two modes, namely *moving* and *drawing*. The *moving* mode, activated by flexing the hand inwards (wave in pose), allows grabbing the sound icon, moving it and throwing it. The *drawing* mode instead can be selectable by flexing the hand outwards (wave out pose), and it allows the generation of automated sound trajectories.

The position of the pointer on the horizontal and vertical axes can be controlled by orienting the arm vertically and horizontally. By twisting the arm inwards, the pointer moves towards the far end of the room, and when twisting the arm outwards, the pointer moves towards the front of the virtual room.

To change the spatial characteristics of the sound icon, the user can interact with the system through five types of user interactions: *grab*, *drop* and *move*, *throwing* and *draw* gestures. The sound icon can be grabbed and dropped in any position within the virtual space. To *grab* the sound, the user moves the pointer next to the virtual sound icon and makes a fist pose (Figure 4.20). The sound icon can be then moved until a finger spread pose of the kind used in Section 4.4 is performed. Sound spatialisation trajectories can be generated by performing a *throwing* gesture as if the sound were a discrete sound source that can be grabbed and thrown away (Figure 4.21).

Recursive sound spatialisation trajectories can be generated in *drawing* mode through the *draw* gesture. This consists of moving the pointer while holding a fist pose (Figure 4.22). The user can then stop drawing by performing a *drop* gesture. Once the path is drawn, it is possible to *grab* and *drop* the sound icon on the drawn path (Figure 4.23 and 4.24) to then let the sound source move on along the drawn path on a loop (Figure 4.25).

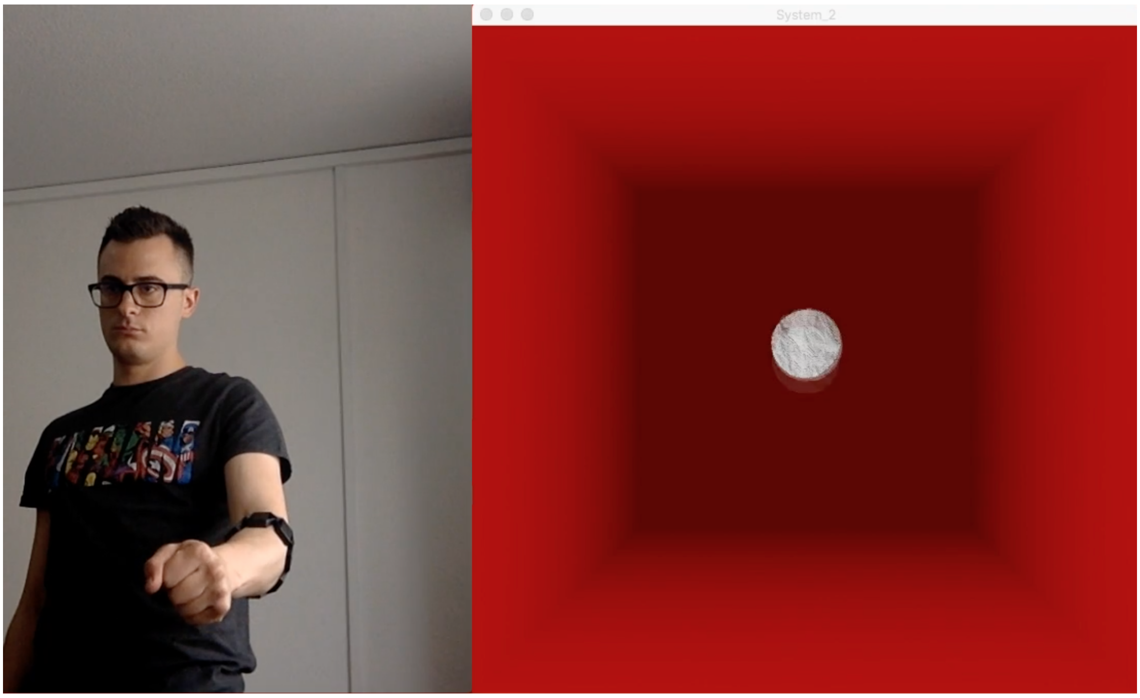


Figure 4.20: Grab gesture. Screenshot from video recording no. 5 in Appendix F.1

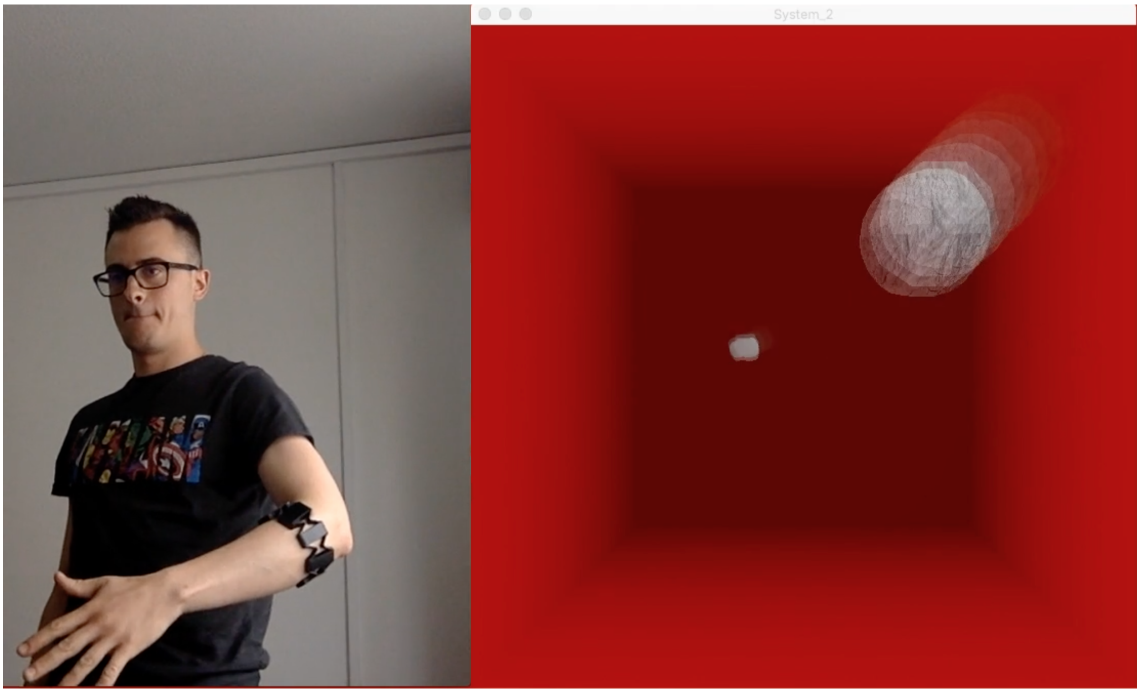


Figure 4.21: Throwing gesture. Screenshot from video recording no. 5 in Appendix F.1

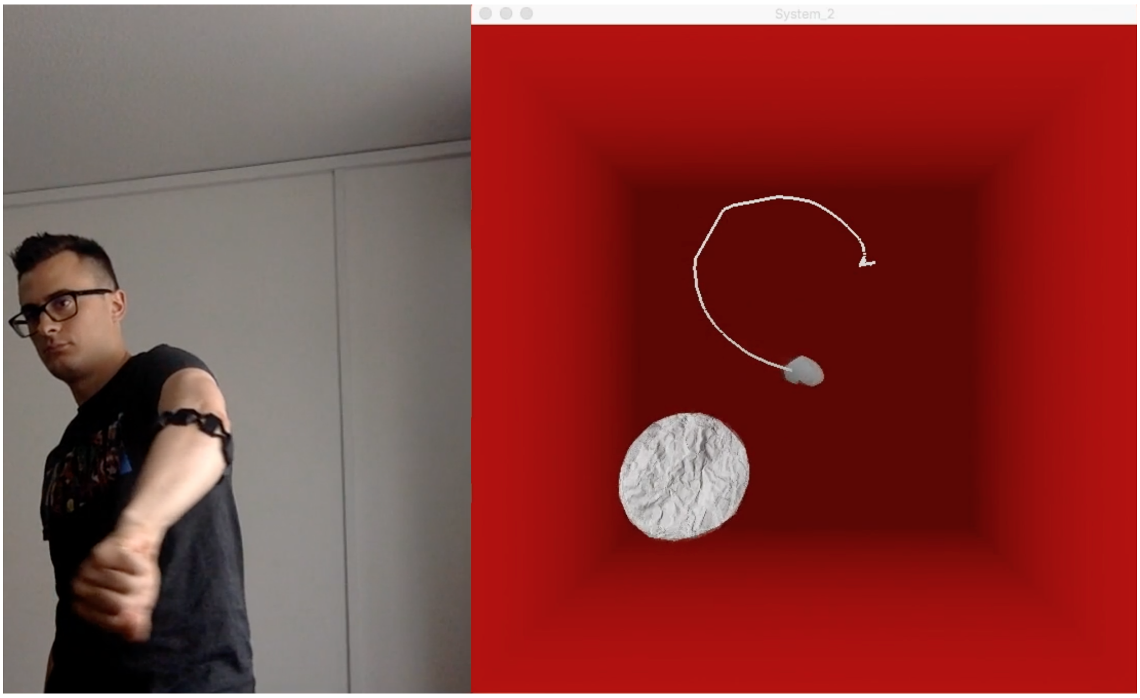


Figure 4.22: Draw gesture in. Screenshot from video recording no. 5 in Appendix F.1

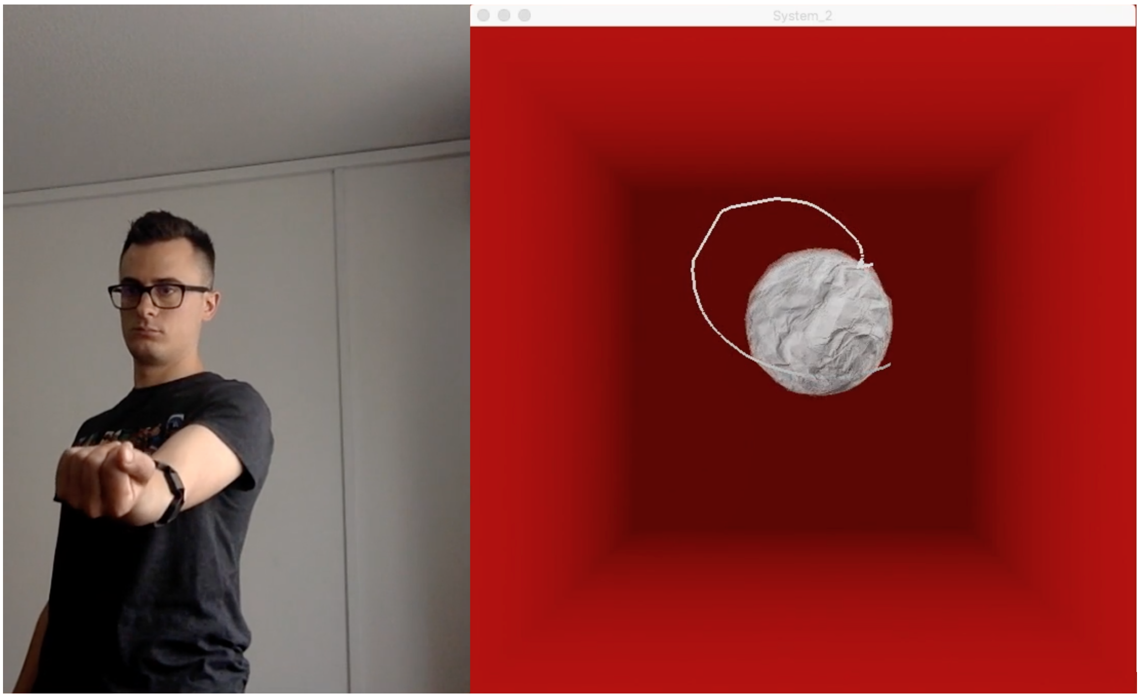


Figure 4.23: Move gesture in. Screenshot from video recording no. 5 in Appendix F.1

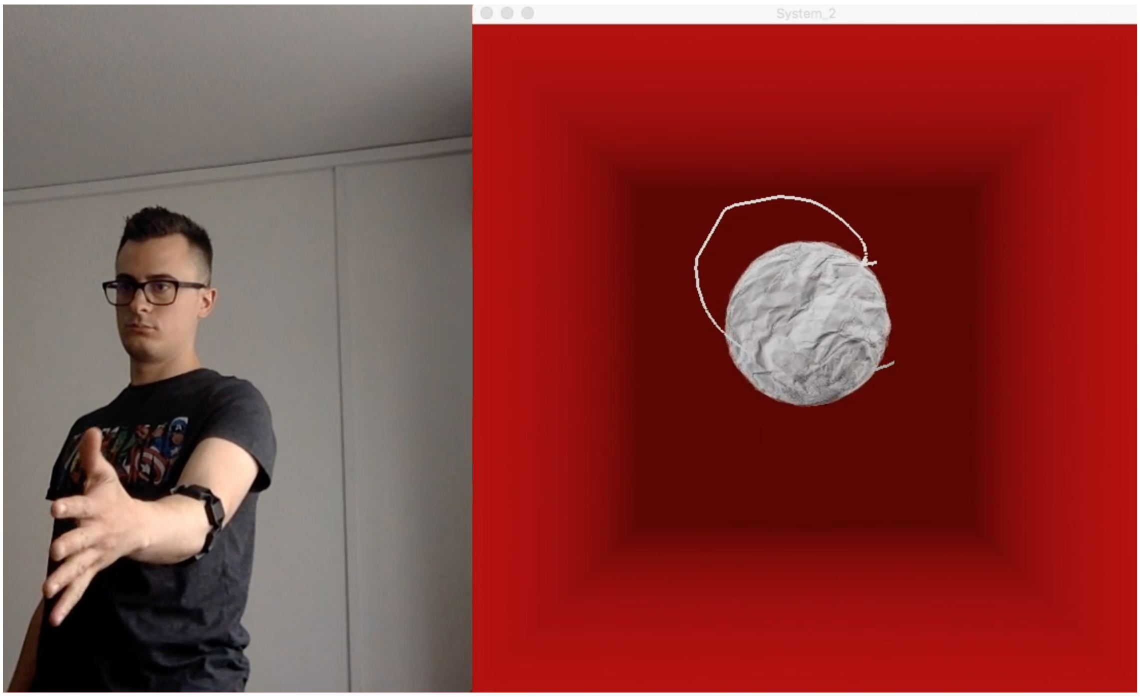


Figure 4.24: Drop gesture. Screenshot from video recording no. 5 in Appendix F.1

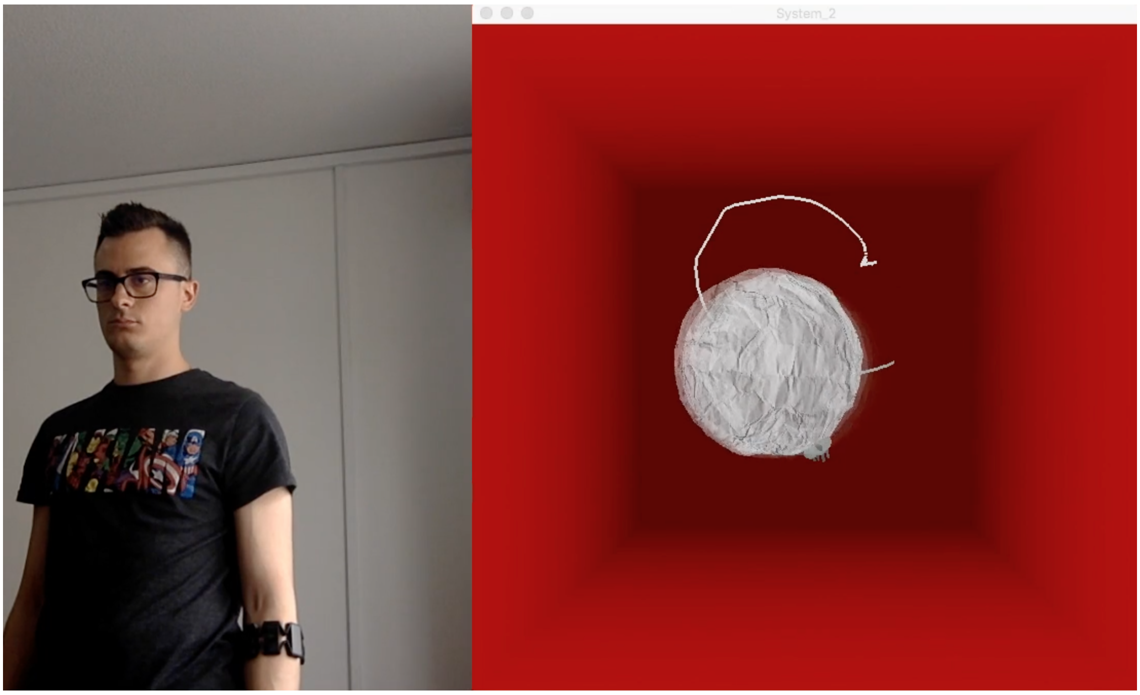


Figure 4.25: Sound icon moving along the drawn path. Screenshot from video recording no. 5 in Appendix F.1

All gestures are tracked through linear mapping IMU and EMG data. A *grab* gesture is detected when a fist pose is recognised, and at the same time, the pointer’s

coordinates are equal to those of the sound icon or within a margin of error of a quarter of the sound icon’s diameter. A *drop* gesture is detected if a finger spread gesture is recognised after a *grab* gesture was detected.
The *throwing* gesture is detected with a test on the average of the first order difference (FOD) of the accelerometer data and the EMG MAV value; if FOD is greater than 0.3 and MAV greater than 0.15, then a *throwing* gesture is detected. These thresholds have been calculated through a series of informal trials.
The movements of the pointer, on the horizontal and vertical axes, are defined through linear mapping of the yaw and the pitch respectively. The movement of the pointer on the depth axes is established with a test on the roll value; if the roll value is smaller than 0.5 (arm twisted inwards), the pointer’s coordinate value on the depth axes will decrease and vice-versa. Similarly, the sound icon’s movement, after a *grab* gesture is detected.
When the sound icon is thrown away, it moves along a linear path calculated at the moment the *throwing* gesture is detected. To be precise, the path’s origin coordinates are the same one of the pointer; and, the path’s end coordinates are calculated by adding to the pointer’s coordinates the result of the linear mapping of EMG MAV and the average of the absolute value of accelerometer data’s first order difference. The trajectory speed is calculated similarly. As a result, the trajectory travelled by the sound icon is towards the same direction of the *throwing* gesture and its length and speed are directly related to the level of the muscles’ engagement and acceleration of the arm in performing the gesture.
A video demonstration of the system is included in this thesis as Appendix F.1 no. 5.

During informal trials by the author, it was observed that the found interaction design and its implementation had the potential to offer a natural gestural control of sound spatialisation in a virtual environment. It allowed the interaction with sound as if it were an object. This idea of thinking sound and the interaction with it opened a new series of possibilities, which were later explored in the two case studies (Chapter 5 and 6).

The separation of modes of interaction, through the *move* and *drawing* mode allowed to implement a higher number of gestures compared to the earlier Pilot Studies. The most natural and intuitive gestures were the *grab*, *move* and *drop*, the which resembled the mouse interaction with desktop icons. The *throwing*, *drawing*, and the gestures to switch mode would instead require instructions and a steeper learning curve before they are easily used.

4.9 Chapter summary

Myo Mapper and five studies were outlined in this chapter. These explored various aspects of interaction design and the Myo.

Myo Mapper enabled users with different backgrounds to take advantage of feature extraction techniques and machine learning software for solving gesture recognition issues. It supported the sound exploration in each case study, and a high potential for supporting the creation of complex interactive artworks.

From an informal evaluation, it emerged that rescaling all values within the same range $[0, 1]$ facilitates user workflow. Smoothing EMG data using the MAVG feature makes the EMG data usable for controlling processing parameters. Raw EMG is a noisy and rich signal; thus, the effective use of feature extraction is fundamental in order to recognise its variations through a classification or regression algorithm. When using machine learning for recognising hand poses, it was observed that it is possible to achieve the desired result more easily when feeding the system with MAVG of EMG data, rather than any other feature listed in Myo Mapper. Moreover, results are more satisfactory when training machine learning with the EMG data in which MAVG was calculated to a sample size of 40 samples or above. However, this would have introduced a delay in recognising hand poses. To avoid this issue, the window size was then lowered when classifying new data. The MIN and MAX features appeared to be useful to observe the data range in which a pose or gesture occurs, to then adjust the data mapping accordingly. Linear mapping of orientation data has been successfully used to control the position and orientation of virtual and physical objects.

In the first case study, IMU and Myo factory gesture recognition outputs were used to track the up/down movement of the arm and associate this to the upward/downward pitch shift and the twisting movement to interact with an imaginary knob. This approach established a clear one-to-one sound gesture relationship but made use of the system to feel systematic and less exploratory.

In the second Pilot Study, it was investigated the potential of using muscle tension to interact with sound parameters without any restrictions in having to perform particular poses to control sound processes. Here the adoption of no coded gesture to control sound encouraged the exploration of both gesture and sound.

In the third case study, it was explored the possibility to extended the gesture vocabulary of the Myo after different poses were individuated as clearly recognisable by a machine learning algorithm in the second case study. This led to the adoption of machine learning applied to control audio panning settings in the fourth case study. Although results suggested that those gestures can be recognised reliably, this approach has the potential do disrupt a musical performance.

In the fourth study, the design of embodied interaction through the idea of sound's physicalisation and taking into account the sound affordances were explored in the context of MR. After the development and informal trial of a system that allowed the MR interaction with sound parameters, it was observed that this method allows drawing trajectories between the two ends of the MR continuum, here characterised by the auditory feedback, real and virtual objects.

In the fifth study, it was explored the interaction with a sound-object intended as an abstraction of the sound. Here the interaction design allowed to transform spatial characteristics of the sound by grabbing, moving, throwing sound-object or by drawing trajectories that it would travel through. In this and the fourth pilot studies was observed that through gestures that emulate the interaction with real-world objects, or that mirror the causality of the auditory feedback of such interaction in the real-world, fosters embodied and natural human-computer interactions.

Chapter 5

Case Study 1: Voice Gesture

5.1 Introduction

This first Case Study explored interaction design solutions in two contexts. In the first, performers could interact with the system without any constraint (*unconstrained interaction*). In the second scenario, the movements of the performers were partially constrained by a musical instrument or interface during a performance interaction with the system (*constrained interaction*).

The *unconstrained* and *constrained interactions* are explored through two musical performances, and the first five iterations of the MyoSpat system were developed. For *VoicErutseG* performed by Vittoriana De Amicis, the first four iterations of the MyoSpat system in the unconstrained context were used. The fifth iteration is used by Grace Savage for performing *Music Gesture Beatbox* in a constrained context, determined by the use of a hand-held microphone and a loop-station. The knowledge gathered in this chapter will inform later MyoSpat iterations that extend the interaction and the system design to support performances involving the use of different musical instruments.

In addition to resolving interaction design issues in this case study, the MyoSpat was also informally evaluated during composing and performing. While *VoicErutseG* explores the situation in which the performer has to rehearse and then perform already composed music, *Music Gesture Beatbox* aims to address issues when adopting the system during the music-making process.

In the subsequent section, the methodology adopted to attain the aims of this study is described. This is followed by an in-depth review of body movements during singing. This leads to the work being realised during the two performances from the musical, interaction design and system development point of view, illustrated in two sections (5.4 and 5.5). Finally, the Chapter Summary highlights significant findings.

5.2 Methodology

The aims of this study were achieved through two UCD cycles built on a similar model, but each of them focused on a musical performance: *VoicErutseG* and *Music Gesture Beatbox* (see Figure 5.1).

The following sections describe the implementation of different stages of the UCD method.

5.2.1 Understanding the musician and the context of use and specification of user requirements

The literature focusing on movements during singing was reviewed to gather what gestures are already part of a singer's vocabulary of gestures. Gestures identified from this review were later reconsidered in light of the importance of the gestures' visual feedback during verbal communication with others and when interacting with physical or virtual objects. In exploring the context of use, the literature on sound spatialisation methods was briefly reviewed, considering the artistic and technological objectives of this thesis. Later the MyoSpat system was developed to support the performance of *VoicErutseG*.

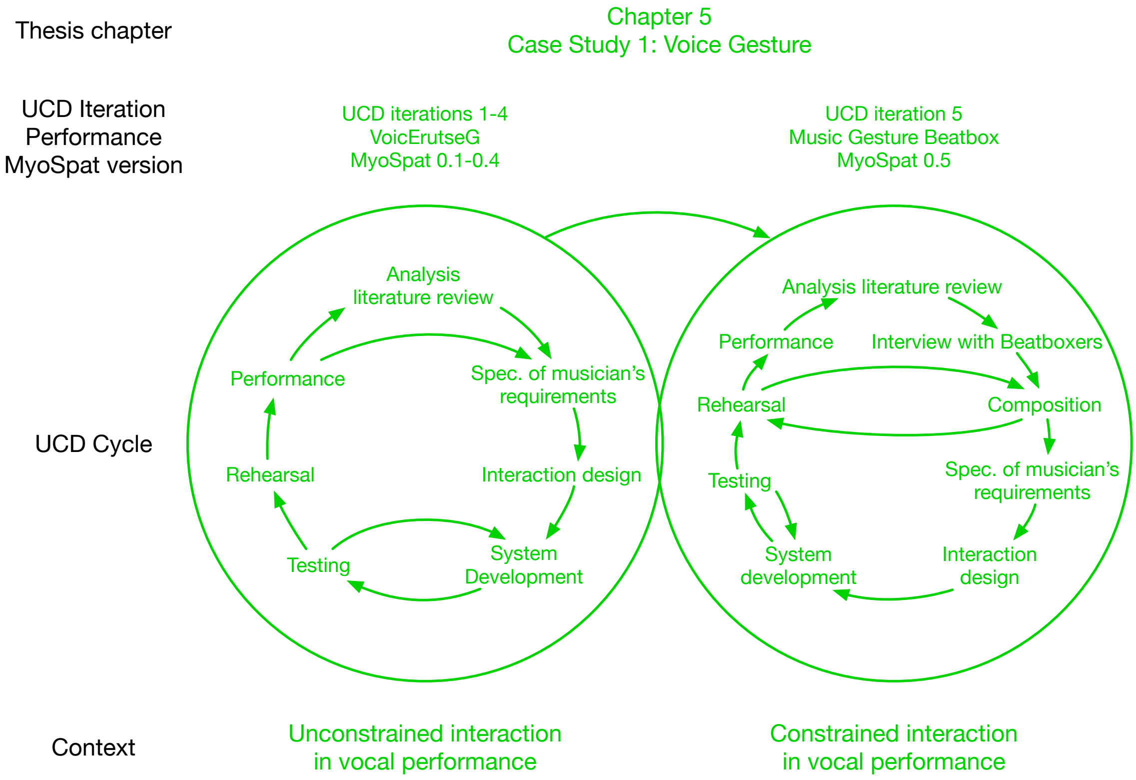


Figure 5.1: UCD cycles in Voice Gesture case study.

The challenge of implementing MyoSpat in the context of live performance with a beatboxer led to the reassessment of several aspects of the interaction design and further readings of the relevant literature. At the time of writing, the literature included studies on beatbox performance, for instance, focusing on real-time voice recognition (Stowell and Plumbley, 2010; Stowell, 2010; Hipke et al., 2014; Picart et al., 2015), functional endoscopic analysis (Sapthavee et al., 2014; Poliniak, 2013), paralinguistic studies (Proctor et al., 2013), historical, repertoire and pedagogical importance (Kuhns, 2014). However, none of the studies focused on hand gestures or body movements during beatbox performance. The findings from the interviews with eleven beatboxers, included in this case study, are intended as a contribution to the existing literature on beatbox performance.

5.2.1.1 Interviews with Beatboxers

The interviews were structured into two parts, preceded by background questions. The first part asked about movement and gestures, specifically about the beatboxers' definition of gestures, the importance of gestures in human beatboxing performances and the relationship with the musical results. The second part examined themes related to the use of software and hardware technology in live performance, its advantages and disadvantages, technological requirements to solve current related problems, and the use of sound spatialisation.

Background questions:

1. For how long have you been beatboxing?
2. Are you right-handed or left-handed?

Interview part one: gesture.

1. How do you define a body movement?
2. And a gesture?
   (a) Is it different than a movement?
3. Do you perform any movement or gesture during a musical performance?
4. If yes to question 3
   (a) When do you play gestures?
   (b) Which gesture do you play most?
   (c) Do you perform a specific beat per each gesture?
   (d) If yes to question 4c
      i. What gesture is it?
5. Have you ever choreographed your performances?
6. Some beatboxers suppose that gesture helps to conceptualise a sound and then produce it, to feel the beat or to keep the musical tempo. Do you agree or disagree and why?
7. Beatbox is performed in different locations: street, clubs, stages, and theatres.
   (a) Does the performing space influences your performance?
   (b) Does the size of the performing space influences your performance?

Interview part two: gesture and technology.

1. Do you use any technology in your performance?
2. If yes to question 1
   (a) Do you use mostly hardware or software?
   (b) Do you use a microphone?
      i. With which hand do you hold it?
      ii. Does the microphone affect your gestures?
3. Do you consider the interaction with digital technology suitable for a beatbox performance?
4. Does technology already satisfy all your needs?
5. Do you feel the needs of hands-free technology?
6. How would you interact with technology?
7. Does the technology you use to perform affect your choices in the performing space?
8. Do or would you make use of audio panning or sound spatialisation?
9. If using or would you like to use spatialisation, how would you relate the position of the sound to your movements while beatboxing?

Extracts from the Interview with beatboxers are then reported in different parts of this chapter to complement the literature review and support research findings. Full transcripts of the interviews are in Appendix G.1.

5.2.2 Design solutions

Designing solutions is the most crucial stage of the UCD cycle. In this first case study, findings from the literature review, Pilot Studies and the collaboration with performers merged towards the design of modes of interaction with sound processing parameters. The work on two performances motivated the design solutions: *VoicErutseG* (Section 5.4) and *Music Gesture Beatbox* (Section 5.5).
An experimental period of development, testing and rehearsal and was undertaken before each performance. In the case of *Music Gesture Beatbox*, this loop also included the activity of composing by the musician.

5.2.3 Evaluation against requirements

An informal evaluation of the work realised was conducted by gathering qualitative data through post-concert interviews with the performers and the audience when possible. Observations on the use of the system during a performance and its adaptability in different spaces were reported. Collected data would then inform the design of the next MyoSpat iteration.

Interviews and participants feedbacks were analysed utilising Thematic Analysis. This method allows for finding themes or patterns in qualitative data, according to Braun and Clarke (2006). Results have been found through inductive coding of the data. This is a "bottom-up" approach which consists in making the codes emerge from the data themselves, and without trying to fit it into a pre-existing coding frame. This approach has been already used in music applications. Tanaka et al. (2012b) used a Thematic Analysis methodology for qualitatively analysing a survey on the design of music software GUIs, with which they looked for recurrence of issues in participants' feedback, with interest in limitations of the touchscreen, lack of consistency in sensor input, latency, networked possibilities and toy-like music applications. Here, Thematic Analysis will be used to better understand gesture-sound relationships, musician-computer interactions, and their use of music technology during a performance. Elements of Grounded Theory Method, such as Open Coding, Axial Coding and Selective Coding methods in sequence, were used for the coding of the themes. Grounded Theory was first established by Glaser (1978), and later successfully and widely used for analysing qualitative data in the field of Human-Computer Interaction (Lazar et al., 2010).

In this thesis, the open coding method was firstly used to identify concepts in data by closely examining the dataset. Secondly, through axial coding, codes were clustered around the relationships between them. Finally, by applying a selective coding method, the most important information was highlighted. A spreadsheet was used to organise the data within the code and organise them within concepts and then categories. Codes are listed using the following structure, and the full code list can be found in Appendix G.2.

5.3 Body movements in vocal performance

Unlike other instrumental techniques, singing does not require macro gestures or movements to produce sounds. Sounds are produced by air coming from the lungs and tuned by our vocal folds together with a series of resonators that include the larynx, pharynx and mouth (Cronin, 2014), while abdominal muscles are involved in the diaphragmatic respiration (Shan and Visentin, 2012). Interestingly, after studying the gestures made during a performance, the music being performed and the performer's expressive intentions, Davidson (2001) demonstrated that body movements could contribute to the production and perception of sounds in vocal performance. These findings were later confirmed by Kurosawa and Davidson (2005) through an investigation of non-verbal behaviour in the performance of popular music. Ancillary gestures and stage presence have been considered an integral component of the musical message since the 17th century with Monteverdi (Carter, 2002). Interviewed beatboxers commented:

> [...] a lot of people go down a little bit when they are doing like a base sound, a lot of that it is like adding visual elements to the performance. Because it is something that people can because people can only listen to music right. So to add some sort of visual aspect to it.
>
> Interview with beatboxers, Participant 2 (2016)
> (extract from interview script in Appendix G.1.2).

> If you are using one speaker only, you feel like you have to do more with your voice, whereas when you are in concert environment, you have got two speakers so the sound is not gonna be an issue and you are more engaging and being more physical on the show, rather than being focused on the audio only which is the public basket setting.
>
> Interview with beatboxers, Participant 3 (2016)
> (extract from interview script in Appendix G.1.3).

So being able to do that while I'm moving helps the performance. It's not engaging if you beatbox a little bit, moving to one spot for using the loop-station and then start moving again afterwards.

Previous studies have shown that performing body movements when singing can facilitate the teaching and learning process (Hibbard, 1994; Halfyard, 2000; Davidson, 2001; Liao, 2002), and the interaction with co-performers or with the content of a song (Cook, 2000). These improvements are relatively subjective to each performer, and in some cases, body movements may even disturb the musical performance due to the performer's limits in coordination and cognitive load (Cook, 2017; Liao, 2002; Owens, 2005).

Five common gestures among singers were identified from the literature reviewed in this Section: *palette lifting*, *circling* or *pointing*, *opening/closing arms*, *beat* and *metaphoric* gestures.

5.3.1 *Palette lifting* gesture

The *palette lifting* gesture consists of a hand movement upwards or downwards with the palm facing the same direction of the movement (Davidson, 2001, 2007). The height of the hand specifies a direct relationship with the pitch of the vocal sound: high gestures for high-frequency sounds and low gestures for low-frequency sounds (Apfelstadt, 1988; Steeves, 1984). Grant (1987) and Ware (1997) also show that this gesture has the potential to improve the singer's intonation.

With work on the effect of gestures and movement on the intonation and tone quality of children's choral singing, Liao (2002) concluded that there is a solid relationship between pitch and hand position on the vertical axis during vocal performance. She also observed that the direction towards which the gesture is performed reflects melodic contour and emotional state. If the hands move upwards, it refers to happiness, and if instead, they move downwards, the movement indicates sadness. Moreover, the nature of the motion is directly linked to the dynamic level of the voice, the distance between the two hands reflects tone quality, and the movement now refers to vocal articulation (continuous and smooth movements are associated with a legato, while discontinuous movements are associated with staccato notes).

Similar behaviour is observed in beatbox performance, where several high hand gestures are performed when singing high pitch sounds and vice versa. To this extent, interview Participant 10 stated that they moved the whole body when performing low-frequency sounds (see video recording no. 29 in Appendix F.3).

5.3.2 *Circling* or *pointing* gesture

While working with five children, Liao and Davidson (2007) asked them to practice seven vocal patterns while performing hand gestures: pushing, a raising hand, side swinging, ear side circling, gathering, flicking. Children were free to make, and later to explore their gestures. From a conducted semi-structured interview, it emerged that *circular* or *pointing* gestures help singing the right pitch. *Circular* gestures are movements of the hand following the path of an imaginary circle. Similarly, the *pointing* gesture is a movement of the hand following an imaginary circle's path with a closed hand and the index finger extended. Liao and Davidson (2007) also found that if the gesture is made on a loop counter-clockwise, guiding the performer to sing a series of scales upwards, and downwards if the gesture is performed clockwise. If the gesture is made correctly, it can help the singer to produce a "smooth" and "round" sound. If not, the performer is more likely to sing with a more "throaty" and "constricted" sound. Finally, they also stated that *pointing* gesture could also improve the focus on musical performance.

5.3.3 *Opening/closing arms* gesture

Opening/closing arms gesture. The gesture of opening or closing the arms towards the audience is often related to vocal sound intensity. Singers often perform this gesture during the conclusion of a musical phrase. In the case of a crescendo, the performer opens her arms and closes them when performing a decrescendo. Gestures follow the dynamics of the sound, with more forceful gestures accompanying louder sounds and gentler gestures for softer sounds (Davidson, 2001). Similar results have been achieved by Turner and Kenny (2012), who, in a study on six popular Western singers, showed that there is a strong relationship between the quantity of movement and the intensity of the vocal sound. They found that when performers are moving on stage, they are likely to sing louder, and they are quieter when staying still. These movements are also found among beatboxers. The tenth interviewed beatboxer commented on performing wide gestures when beatboxing loudly, and the opposite when beatboxing quietly.

5.3.4 *Beat* gesture

Beat gestures are those that aim to highlight prosodic aspects of speech or the tempo of a song (McNeill, 2005a). These are commonly hand, head or feet gestures repeated recursively on the downbeat or upbeat of the music. Moreover, these gestures facilitate singing in the correct tempo during a performance (Liao, 2002; Davidson, 2001). Interviewed beatboxers also confirmed this aspect. They specifically mentioned using finger-pointing and striking gestures (Participant 5, 7 and 11 respectively) and shifting of the hand along the vertical axes (Interview Participant 10, 2016; video recording no. 27 n Appendix F.3). These gestures may also vary in relation to the genre of music being performed. Other performers tend to perform the *beat* gesture maintaining a fist, finger spread, or open hand pose.

> And see, I'm using my fingers here. Because I'm thinking the music in a different way, I send a different vibe, so I cannot do something like this [Participant moving weaving the fingers upwards]. I can do something like this using my finger if I'm doing something groovy and like party (see video recording no. 20 in Appendix F.3).
>
> Interview with beatboxers, Participant 7 (2016)

(extract from interview script in Appendix G.1.7).

All the interviewed beatboxers commented on making *beat* gestures in performance. A few also stated that they provide visual cues so that the audience can better appreciate their virtuosity. Beatbox Battle (2015) is a clear example of how both performers imitate the tempo of their beatboxing through waving the arm or the hand horizontally or vertically. The interviewed Participant 4 comments:

> [...] There are so many different ways that I could mean that [the music tempo] and so I could be like, [beatboxing]. Alternatively, I could be sort of [beatboxing]. So you really have to see them [beatboxers], and you have to see their gestures and the way that they move to understand what they are doing. [See video recording no. 6 in Appendix F.3]
>
> Interview with beatboxers, Participant 4 (2016)
> (extract from interview script in Appendix G.1.4).

While interviewing Participant 3, it emerged that mimicking the causality of a sound gesturally, beatboxers would try to represent spatial gestural characteristics. For example, beatboxers would produce a sequence of snare and hi-hat sounds by moving the hand in two different locations as if playing a virtual snare or a hi-hat. This observation is also shared in the work of Ehmann (1968), where he says that a virtual map of vocal sounds can enhance the learning and performing aspects of singing.

> [talking about the *pointing* gesture] It helps to coordinate what you are doing, so let's say you have a kick drum here, the hi-hat here and the snare here. [interviewer pointing at a different point in space] If you can do a rhythm on there you can do it with your voice, I mean if you can visualise it you can do it with your mouth. [...] if you cannot do a very quick hi-hat [referring to the drum's hi-hat sound], but then you play it virtually, you just start to do it subconsciously.
>
> Interview with beatboxers, Participant 6 (2016)
> (extract from interview script in Appendix G.1.6).

5.3.5 Metaphoric gestures

Gestures can also be defined as metaphoric gestures when representing the abstract meaning of a message as if it were an object or occupied a specific space (McNeill, 2005a). Metaphoric gestures are highly present in Beatbox performances (Kew, 2015). Findings from the interviewed beatboxers and the analysis of beatbox performances show that beatboxers tend to perform gestures that mimic the sound-producing object that they imitate in their singing or its sound qualities. A prominent example of these gestures is the performance by Thum (2013), where he sings and mimics through gestures the sound-producing object or its timbral characteristics. For instance, he mimics the action of playing the trumpet while producing a trumpet-like sound and the movement of a turntable cartridge as if it were moving across a vinyl, with the vinyl in a fixed position. The interview Participant 6 also describes a similar experience, for instance, when he mimics the sound of DJ scratching while performing a DJ scratching gesture. This is performed by waving the hand inwards and outwards consecutively, with the fingertips towards the ground (see video recording no. 16 Appendix F.3). These gesture results are in line with findings explored in the Literature Review (Section 2.3.1) by Caramiaux et al. (2011) that say that when enacting a gesture along with a sound in which we recognise its causality, we do tend to imitate that aspect of the sound.

Metaphoric gestures have the capacity to highlight the physical characteristics of a sound-producing object until they affect the identity of that sound. A demonstration of this concept is shown in *Pass the sound* by The Beatbox Collective (2012). This performance tells the story of a sound represented as a virtual object passed hand by hand across performers. The sound here gains its identity through gestural interaction with a virtual object, defined as exhibiting "external metonymy" in Cognitive Metaphor Theory (CMT) (Mittelberg and Waugh, 2009) or "pantomimed actions" with abstract objects in neuropsychology (Boyatzis and Watson, 1993). The subject of the performance is not the sound or the virtual object, but the combination of the vocal sound with the virtual object representing the sound.

From interviews, it becomes apparent how beatboxers tend to perform broad and relatively slow movements, often with their arms, to describe tempo, pitch or frequency. They enact faster movements, often with the fingers, to mimic the qualities of the sound that vary more rapidly, like the timbre modulation of a specific sound. For example, Participant 5 uses fingers to describe the tempo of a fast beatbox sequence, and the arm for a slow glissando (see video recording no. 7 in Appendix F.3).

This aspect can be related to the fact that larger joints are more suited to perform larger movements and vice versa (Rosenbaum et al., 1991).

5.4 VoicErutseG

5.4.1 Performance

VoicErutseG explores the possibility of extending existing works in the vocal solo repertoire through gesture-controlled sound spatialisation, paying careful attention not to disrupt the choreographic aspects and required vocal techniques of the original works. The performance aimed to inform and demonstrate the implementation of a natural and embodied interaction design for controlling the spatial position of the vocal sound, considering factors that might interfere with aspects of the performance and the concentration of the performer.

In this performance, I investigate how embodied interaction can facilitate rehearsals, avoiding potential issues of availability and learnability. For instance, the need of all due technology for rehearsing choreographic aspects of performance – such as a Myo, a multichannel audio system composed of a minimum of four loudspeakers installed in their practice space – and to learn new software. These problems are amplified when the performer in a different location, as in the case of the performer Vittoriana De Amicis, who was based in Italy, while this research was carried in the United Kingdom. I aim to find a solution to these issues through designing modes of interactions that allow performers to rehearse choreographic aspects of the performance and audio transformations without the system set-up.

VoicErutseG includes the performance of Luciano Berio's *Sequenza III* for voice (1965) and Cathy Berberian's *Stripsody* (1966). With *Sequenza III*, Berio aimed to explore extended vocal techniques. Choreographic cues are also an important aspect of the performance (Halfyard, 2000). For instance, hand gestures over the mouth are used to modify the sound of the voice (Figure 5.2). Apart from these instructions, Berio left the performer free to choreograph the piece as she wishes, yet based on a list of emotions, vocal behaviours, vocal flexibility and "dramaturgy" clearly indicated in the score (Berio, 1968). Through an analysis adopting Laban's taxonomy (dab, flick, float, glide, press, slash, thrust and wring) Halfyard (2000) highlights the importance of the gesture's *"force or weight (firm or gentle); its relationship with space, whether its trajectory is predictable and direct or unpredictable and flexible; and its relationship with time, whether its duration is brief or prolonged, resulting in gestures that are either sudden or sustained"*.

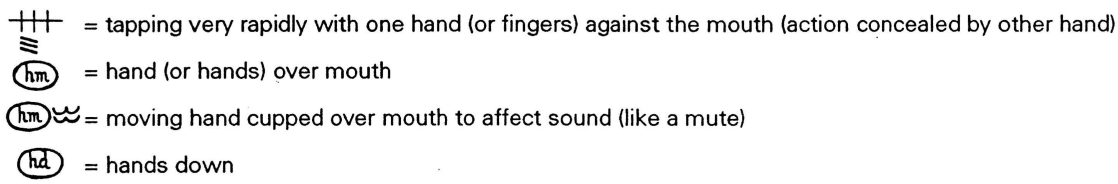


Figure 5.2: L. Berio's *Sequenza III*, notation of hand gestures (Berio, 1968)

Stripsody is a piece based on sounds that tell a comic story realised in collaboration with the comic designer Roberto Zamarin. The performer is asked to mimic "scenes" reported in the score through onomatopoeic sounds. These have to be accompanied by gestures "whenever possible" (Berberian, 1966). Berberian's performances were also known for their theatrical aspects (Meehan, 2011). In an interview for the TV program *From Monteverdi to the Beatles* in 1974, Cathy

Berberian says:

> I often say to myself that I'm not a singer with a talent for acting, but an actress who can sing.
>
> Varga (2013)

Sequenza III and *Stripsody* were chosen for their choreographic aspects, allowing the observation of MyoSpat's potential use in spatialising the vocal sounds of existing works from the repertoire.

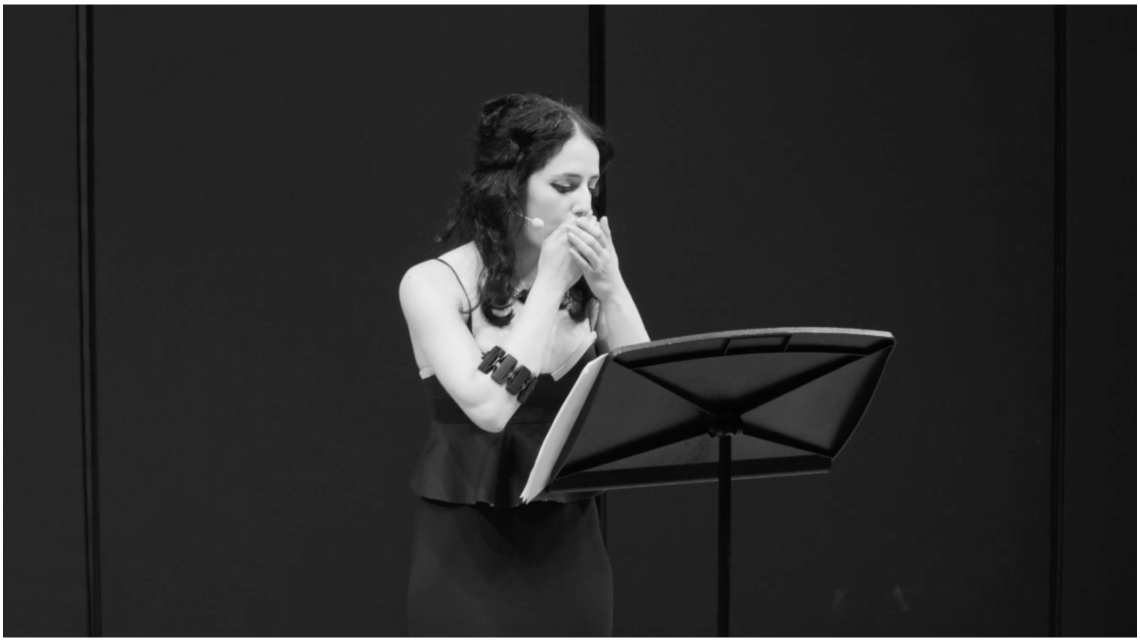


Figure 5.3: Vittoriana De Amicis, performing *VoicErutseG* at Frontiers Festival 2015, Birmingham, UK. Screenshot from video recording no. 1 Appendix F.2

5.4.2 MyoSpat v0.1-v0.4

5.4.2.1 Interaction Design

Gestures are an important factor in verbal communication. These can mislead our perception of the loudspeaker as in the ventriloquist effect (the act of making appear the voice from elsewhere, usually a prop or puppet) (Connor, 2000; Frassinetti et al., 2002), but also improve the communication of a message and the loudspeaker's intentions (Kendon, 1994; Gentner and Goldin-Meadow, 2003).

McNeill (2005b) proposed a categorisation of gestures, known as "McNeill's Gesture Dimensions", that groups them into iconic, metaphoric, deictic and beats gestures. Deictic gestures are those that we commonly use to refer to objects in space or a location while speaking. The most common is the *pointing* gesture, enacted by extending the index finger. This gesture, also called *deixis at phantasma* by Bühler (1982), has a significant role in verbal communication. After analysing a conversation between two people while working in an archaeological excavation, Streeck et al. (2011) revealed the importance of *pointing* gestures to refer to physical objects or subjects of the conversation in space. As well as between humans, the *pointing* gesture is often used to communicate with animals (Scheider et al., 2011).

The *pointing* gesture is also used for designing interaction with a computer. More recently, this gesture has been implemented to replace mouse-based human-computer interactions in various applications. Pajares et al. (2004) demonstrated that the *pointing* gesture could be easier to learn than a laptop's touchpad. The *pointing* gesture, for example, could be used in a camera-based system to interact with an adaptive virtual touchscreen (Jing and Ye-peng, 2013) or for air-typing (Ishijima et al., 2014). Findings from these applications showed that the *pointing* gesture is an embodied way of interaction with real or abstract subjects and objects around us.

From interviews with beatboxers and the reviewed literature by the author, it emerged that the *palette lifting* and *beat* gestures are the most common among singers. By contrast, the *opening arms* and *metaphoric* gestures are subjective and content related. Common to the *palette lifting* and the *beats* gesture is the *pointing* hand pose. Considering also findings from Marshall et al. (2007), who states that *"non-conscious gestures suit high-level or subtle spatialisation"*, the pointing gesture is a natural gesture that can describe spatial cues and that we perform unconsciously. This seemed well suited for the control of sound spatialisation parameters, whether the performer's gestures are performed consciously or unconsciously. These aspects of the pointing gesture make it the best design solution for supporting performer interactions with sound spatialisation parameters, without disrupting the musical performance.

The pointing gesture was redesigned taking into account the sensing modality of the Myo, which has a better capacity for recognising movement than hand poses in a complex context such as musical performance. In this case, the sound spatial position is not controlled through hand gestures with the finger-pointing at a specific location, but instead through the orientation of the forearm.

5.4.2.2 Audio engine

For the design and development of MyoSpat, the first consideration was a scenario in which the musician would interact with only one parameter in the context of electronic music: space. Later, the MyoSpat brought in additional audio processes, visual projections and lighting effects, as a result of working with musicians. The reason is being a spatialiser at the bottom of the MyoSpat, and the interaction design is that sound spatialisation is an example of sound processing technique where that requires the control multiple parameters and mapping layers are involved at the same time when modifying the spatial position of the sound. This choice was also conditioned by the fact that a spatialiser enables a sound to be processed while preserving its timbral characteristic. Thus, it gives a performer the possibility to extend their creativity without affecting other important features like pitch, dynamics or, in the case of singing, intelligibility.

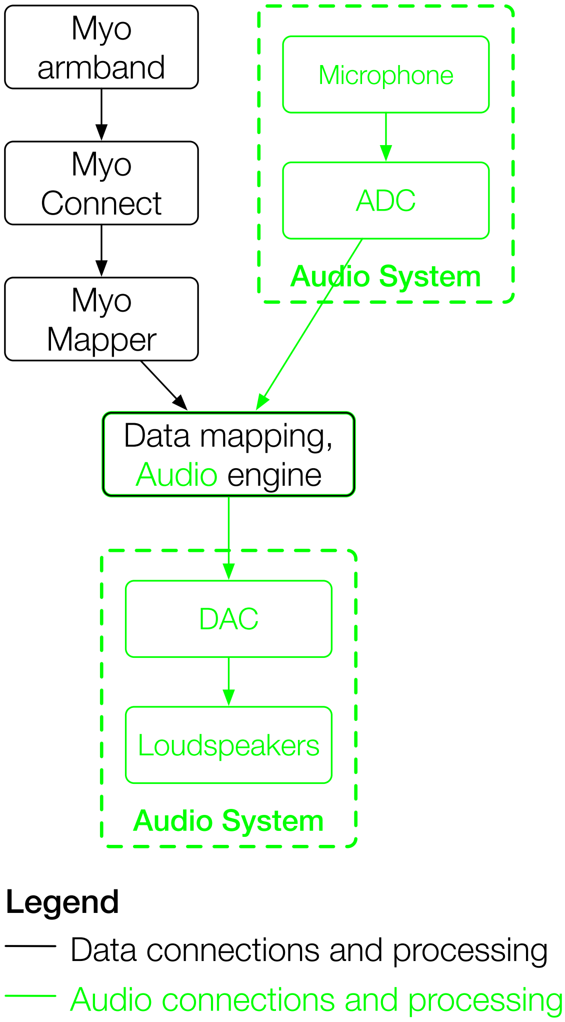


Figure 5.4: MyoSpat 0.1-0.4 architecture

A comparison of multichannel spatialisation algorithms, mostly used in audio IDEs (VBAP, IRCAM Spat, SSP, audioTWIST, Ambisonics, WFS and ViMiC) by Marshall et al. (2009), shows that all of them allow the user to establish the sound position. This characteristic is also shared with other systems not cited in the list above, such as DBAP (Lossius et al., 2009), BEAST (Wilson and Harrison, 2010) and Dolby Surround (Holman, 2008). To this extent, all the aforementioned had the potential to support the development of MyoSpat.

Spectral modifications are an essential aspect to consider when spatialising a sound, especially when it has to serve a musical purpose (Normandeau, 2009; Gupta et al., 2002). Timbral variations enable the imprinting of spatial information on a sound source diffused through a single loudspeaker. For instance, distance and room characteristics can be established by manipulating intensity, filtering and reverberation of a sound (Chowning, 1971; Giordano, 1996). Cognitive and psychological studies also showed that we tend to perceive verticality in relationship to pitch, placing high pitches high in space and low pitches low (Trimble, 1934; Rusconi et al., 2006). However, more recent linguistic studies showed that our perception of pitch is biased also by our linguistic backgrounds (Dolscheid et al., 2013). No spectral modification, however, can emulate the sound spatialisation on the horizontal axis. This is why MyoSpat should be used with a minimum of two loudspeakers, or a loudspeaker capable of diffusing the sound in a different direction on the horizontal plane.

The Icosahedral Loudspeaker Array, developed by Zotter and Sontacchi (2007) at the Institut für Elektronische Musik und Akustik - IEM (Graz, Austria), offers the representation of spatial sound through twenty loudspeakers mounted on the faces of an icosahedron (Zaunschirm et al., 2016). Through a user study, they showed that such a loudspeaker array could render both spatial position and distance. Another alternative to spatialising the sound on the horizontal axis is the adoption of a moving loudspeaker, such as the Leslie, built in 1940 by Donald Leslie.

The Leslie is composed of a woofer and a horn which can rotate on their horizontal axes at different speeds through two dedicated mechanisms. Specifically, Leslie added the "tremolo" effect to a Hammond Organ's sound, as a result of the acoustic Doppler effect generated by the rotations of the loudspeakers (Hammond, 2012). In the following years, the Leslie was replaced by guitar pedals and synthesisers which offered musicians more affordable and comfortable solutions to generate similar effects.

On the principle of moving speakers, Johnson et al. (2016) creates a mechatronic loudspeaker called "speaker.motion". This system was created with the aim of inspiring new spatial music aesthetics to use in performances and installations. It

embeds four loudspeakers which orientation can be set through MIDI messages to a custom hardware system. This system has the potential to foster expressive and engaging interaction with spatial sound and the acoustics properties of the physical environment in which the speaker is in.
K-array (2015a) released a similar technology, the KW8 Owl loudspeaker (Figure 5.5). It is a moving-head loudspeaker which aims to facilitate the sound positioning on stage and studios (InAVate, 2015). This loudspeaker is the first of its kind which is commercially available worldwide. The orientation of the loudspeaker can be moved over the full 360 degrees range. The KW8 Owl also embeds an HD camera mounted on the side of the loudspeaker to allow the user to visualise where the loudspeaker is *pointing*. The video feed and the possibility to pilot the Owl are available through the Owl Manager, an application realised using Max (K-array, 2015b). In addition to establishing the physical orientation of the loudspeaker, the Owl Manager enables the user to control audio parameters such as intensity, delay and filtering.

Figure 5.5: K-Array's KW8, Owl (K-array, 2015a)

The first four iterations of the MyoSpat explored different sound spatialisation solutions, intending to then choose and adopt one in future releases.
The spatialiser of MyoSpat 0.1 allowed the sound movement in a bi-dimensional

acoustic field through the control of a virtual sound source within an abstraction of the loudspeakers' position in the real world. The spatialiser was implemented by adapting work on Distance-Based Amplitude Panning (DBAP) for the distribution of the energies by Lossius et al. (2009), and findings from Giordano (1996) for the modelling and simulation of a sound source moving in the acoustic field. The audio engine was developed using Pure Data as IDE, and an octophonic audio system was used for rendering the sound in the acoustic space.

MyoSpat 0.2 was implemented following the same design adopted for MyoSpat 0.1. The only difference was the use of the IRCAM Spat's library for Max to implement the spatialisation engine, and the use of a quadraphonic sound system for the sound diffusion. IRCAM Spat was chosen as one of the spatialisation algorithms because it has been described as one of the most reliable solutions for reproducing sound localisation and a model of the listening area, using quadraphonic or octophonic audio systems (Marentakis et al., 2008). Moreover, it is easy to use, with its multidimensional control interfaces (Jot, 1999).

With MyoSpat 0.3, ways of mapping gestures onto parameters for a 3D sound spatialisation were explored. The three-dimensional sound was rendered using the High Order Ambisonics audio set-up system, and the audio engine was developed by adopting the High Order Ambisonics (HOA) library for Max realised by CICM (2014).

With MyoSpat 0.4, a solution to spatialise the sound using only one loudspeaker was explored. The system was implemented using the K-Array's KW8 Owl loudspeaker as the DSP engine to diffuse the sound and the Owl Manager software to control both the loudspeaker and the DSP within it.

5.4.2.3 Mapping strategy

Early tests were realised by mapping arm orientation data (pitch and yaw) to the virtual source position on the horizontal plane. Precisely, the gestural data range was linearly scaled into the virtual room size range.

However, taking into account *VoicErutseG*'s choreographic aspects – small and slow in *Sequenza III*, and fast and wide gestures in *Stripsody* – such an approach would have generated overly subtle spatialisation cues that are not perceivable by the audience in *Sequenza III* and rapid and complex trajectories in *Stripsody*.

To overcome this issue, the number of possible sound spatial positions was reduced to nine, established by performing a *pointing* gesture in nine different directions as summarised in Table 5.1.

Table 5.1: Correspondence virtual of source position from an audience perspective with the arm orientation

Position Number	**Virtual sound source's position**	***Pointing* gesture direction**
1	Behind	Up
2	Behind-Left	Up-Right
3	Left	Right
4	Front-Left	Down-Right
5	Front	Down
6	Front-Right	Down-Left
7	Right	Left
8	Behind-Right	Up-Left
9	Centre	Forward

After gathering and observing gestural data describing the *pointing* gesture in all nine directions, a discrepancy emerged between the pitch and yaw data and the intended spatial position of the virtual sound source, with the exception of position no. 9 (Figure 5.6).

Complex functions could have been applied to solve this data mapping problem. As discussed in the literature review (Section 2.4), interactive machine learning literature has the potential to simplify this process and allow to obtain a better result.

Here, the k-Nearest Neighbour (kNN) algorithm from the ml.lib machine learning library was adopted to map orientation data onto spatial parameters. kNN is a non-parametric approach which can be used for classification and regression. In

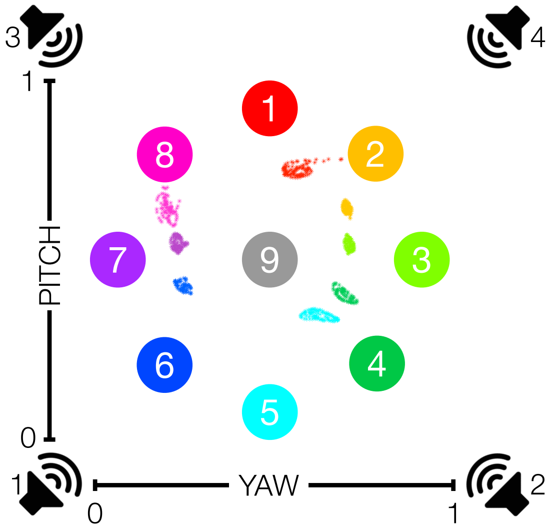


Figure 5.6: Comparison of Myo yaw and pitch data (∗) with the virtual sound source (circles) of the same colour

either situations, the input is composed by the k closest training examples in the feature space (Figure 5.7) (Peterson, 2009).

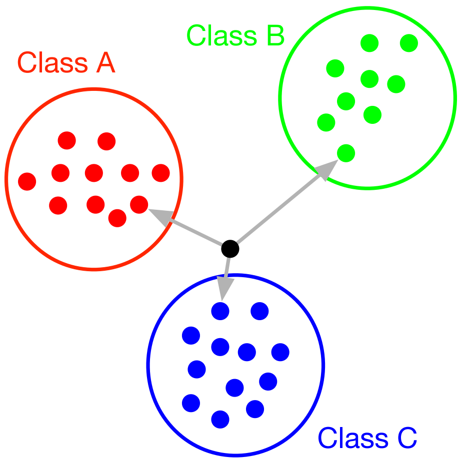


Figure 5.7: k-Nearest Neighbour (kNN) classifier

A kNN classifier was trained using 649 samples (81 samples per each class) of Myo yaw and pitch data (Figure 5.6). The evaluation of the model through direct evaluation reported an accuracy of 100% and 10-fold cross-validation with a Root Mean Square (RMS) error of 0. Even though the built model is overfitted, this is not considered an issue in this context. It was purposely created in this way to have a clear distinction between the nine *pointing* gesture directions.

Here, machine learning is not used in a way that it would satisfy formal and rigorous evaluation of the model, but in function to provide an expressive control of audio processing parameters and facilitate the musician-computer interaction. A similar approach is used across this thesis.

Sound spatialisation parameters were calculated taking into account the distance d_i from the position of the virtual sound source (x_{vs}, y_{vs}) to the position of the loudspeakers (x_{ls_i}, y_{ls_i}), using Equation 5.1, as in Lossius et al. (2009), where i is the loudspeaker's reference number.

$$d_i = \sqrt{\left(x_{ls_i} - x_{vs})^2 + (y_{ls_i} - y_{vs}\right)^2} \tag{5.1}$$

During empirical testing, it emerged that mapping distance data linearly into gain values – so that the closer the virtual source to the loudspeakers, the higher the value of the gain and vice versa – ensures a strong sound-gesture relationship as well as a noticeable spatial effect.
The spatialiser's Time Difference Of Arrival (TDOA), here also called "delay time" parameter, is calculated as $\frac{d_i}{c_\vartheta}$, where d_i is the distance calculated in Equation 5.1 and c_ϑ is the speed of sound at 20 Celsius degree (343.216 m/s) (Sengpiel, 2014). As in Giordano (1996), the low-pass filter's frequency cut-off (f), for simulating the air absorption effect, is calculated in relation to the distance of the virtual sound source from the loudspeaker with the equation $f = 10^{488} \times d_i^{-0.698}$, where d_i is the distance calculated using Equation 5.1.

To exclude any possibility that a linear mapping strategy would have been better than other approaches, in MyoSpat 0.2, the Myo yaw and pitch data $[0, 1]$ were linearly mapped onto horizontal and vertical Cartesian coordinates for establishing the position of the virtual sound source, within the range $[0, virtual\ room's\ width]$ and $[0, virtual\ room's\ depth]$.

To take full advantage of the High Order Ambisonic system, resident at

Conservatorio Santa Cecilia in Rome and the HOA library functionalities, the yaw and pitch $[0, 1]$ data were linearly mapped onto azimuth and elevation parameters of the virtual sound source $[0, 2\pi]$.

MyoSpat 0.4 was based on the use of a K-Array KW8 Owl as a unique loudspeaker for spatialising the sound. Given the similarities with the sound spatial parameters to control in MyoSpat 0.3 – azimuth and elevation of the loudspeaker, sound intensity – in MyoSpat 0.4 a similar mapping strategy was used. Yaw and pitch values were linearly mapped onto pan $[0°, 360°]$ and tilt $[0°, 180°]$ parameters of the loudspeaker, and roll value into sound intensity parameters $[0, 1]$. Data streaming to the KW8 Owl was managed using the Owl Manager through the DMX protocol.

5.4.2.4 System set-up

The MyoSpat system set-up varied across the first five iterations of the software. MyoSpat 0.1 used an audio system composed of eight loudspeakers placed as in Figure 5.8. In MyoSpat 0.2 the sound was diffused using four loudspeakers arranged as in Figure 5.9. MyoSpat 0.3 was implemented using the High Order Ambisonics audio system installed at Conservatorio Santa Cecilia in Rome. The system was composed of 19 Genelec 8050B loudspeakers on three circular levels, at about 1.5, 2 and 2.5 meters from the ground respectively above the audience area (Figure 5.10). The MyoSpat 0.4 system set up was composed of a single K-Array KW8 Owl unit placed at the height of about 2.5 metres from the ground above the user/performer (Figure 5.11 and 5.12).

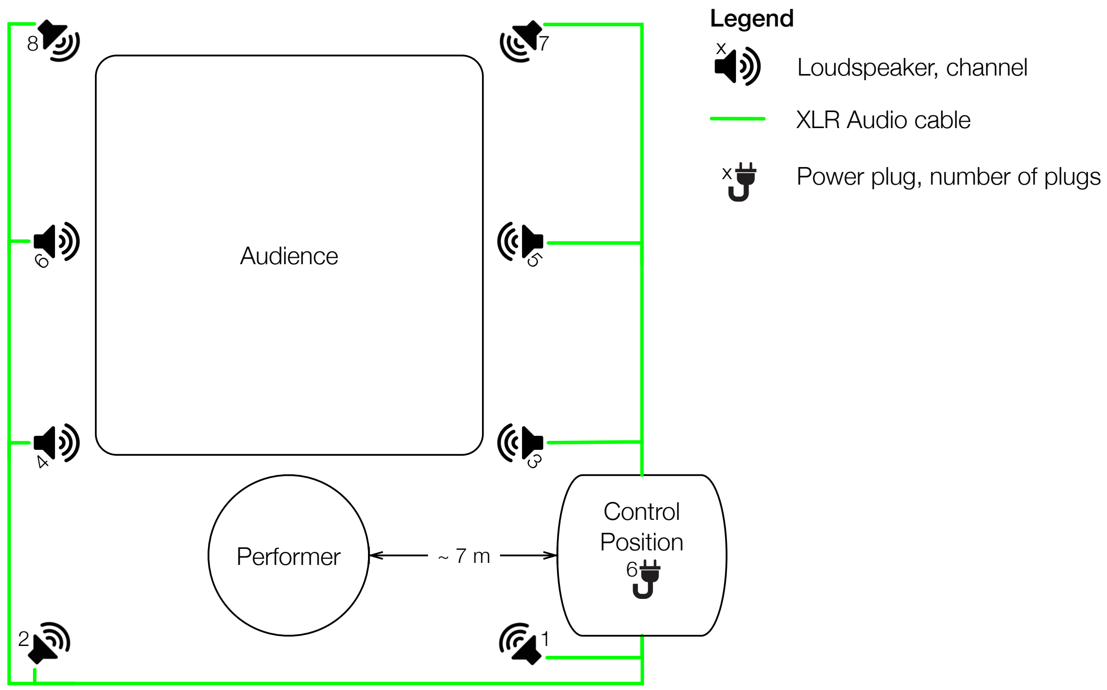


Figure 5.8: MyoSpat 0.1 system set-up

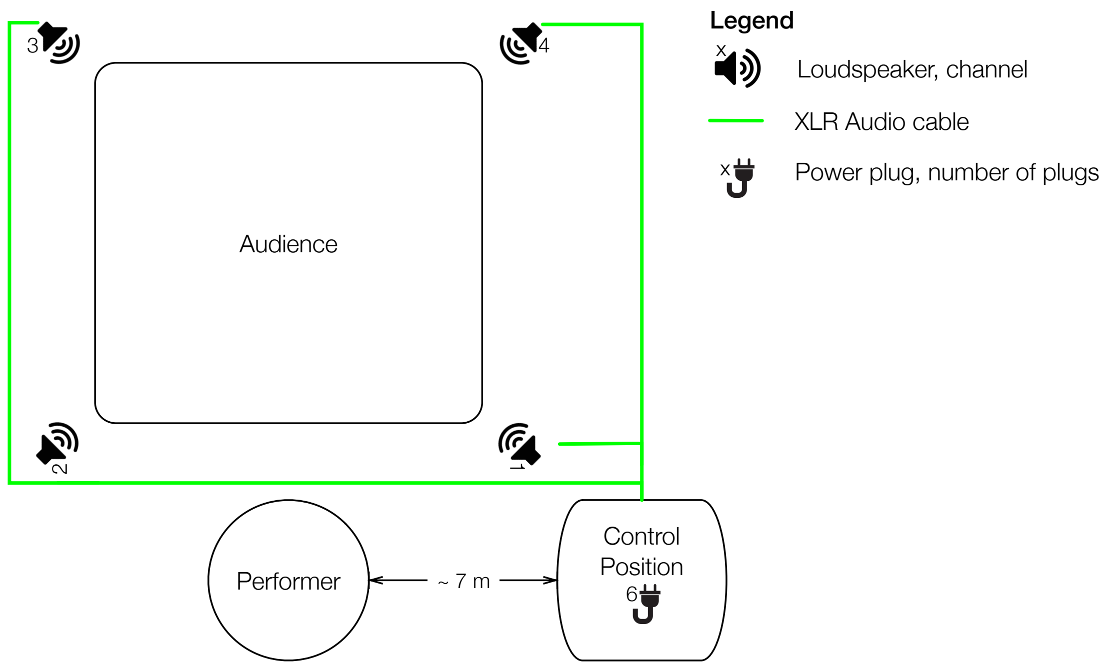


Figure 5.9: MyoSpat 0.2 system set-up

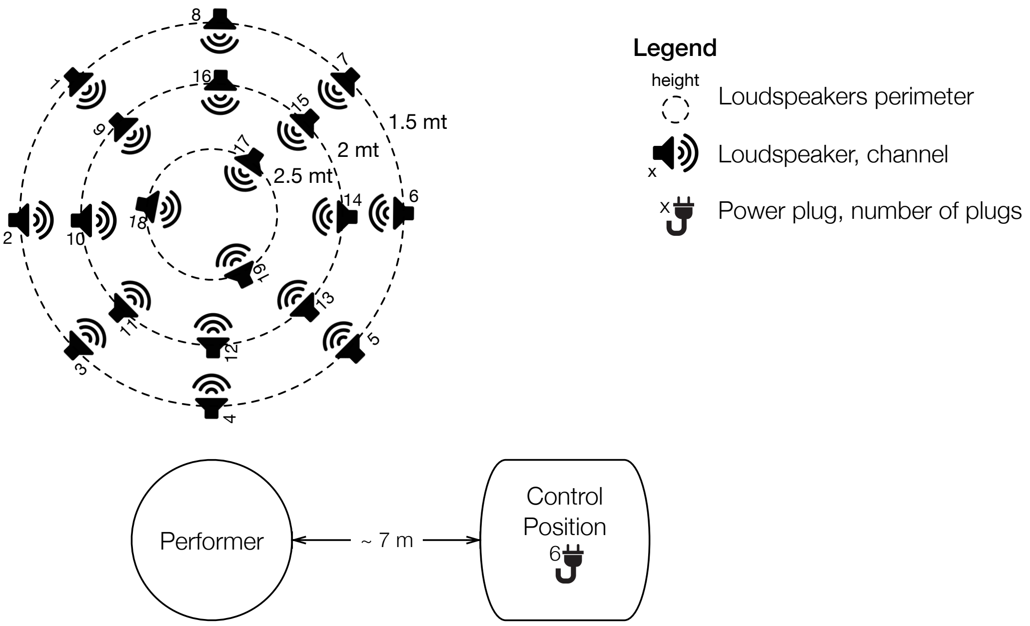


Figure 5.10: MyoSpat 0.3 system set-up

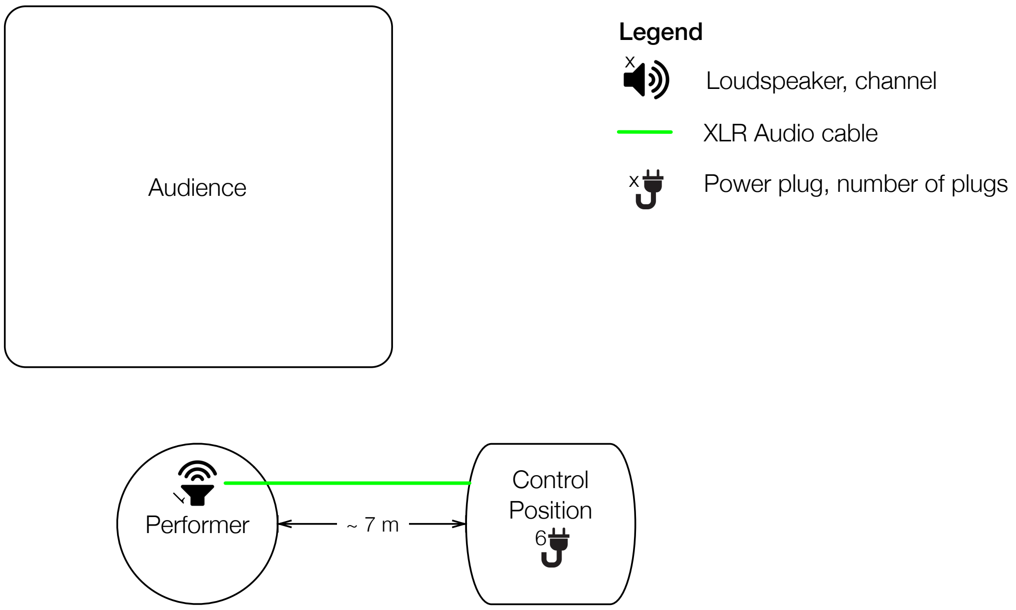


Figure 5.11: MyoSpat 0.4 system set-up

Figure 5.12: MyoSpat 0.4 system set-up. Screenshot from video recording no. 6 in Appendix F.1

5.4.3 Observations

Here are described observations gathered during the performance of *VoicErutseG* at Frontiers Festival 2015 at Royal Birmingham Conservatoire (Birmingham, UK) (Figure 5.3) (De Amicis and Di Donato, 2015a), EmuFest 2015 at Conservatorio S. Cecilia (Rome, Italy) (De Amicis and Di Donato, 2015b), and ElectroAQustica 2015 at Conservatorio A. Casella (L'Aquila, Italy) (De Amicis and Di Donato, 2015c). The recordings of some of the above performances are available in Appendix F.2, video recordings no. 1, 2, 3 and 4.

The interaction design, including the *pointing* gesture, allowed the performer to act on the position parameter of the virtual sound source relative to a predefined virtual room. The *pointing* gesture can be performed by orienting the arm towards the chosen position or movement of the sound source. This modality of interaction aimed to address two issues, (i) the design of a human-computer interaction that would be intuitive and learnable and (ii) that it would not disrupt the musical performance in its technical execution of the piece and choreographic aspects.

Although Vittoriana De Amicis did rehearse without the MyoSpat system, due to being based abroad, she did not find any particular difficulty when performing using the system. On the contrary, she commented that MyoSpat was easy-to-use and the modes of interactions intuitive and natural.

During the rehearsal prior to the performance, her comments focused on the excellent responsiveness with which the system reacted and the good relationship between arm orientation-sound position. Moreover, she commented that the way she performed was affected by the size of the room she was performing in and the audio system set up (quadraphonic, octophonic or ambisonic). She felt that the bigger the room and the farther apart the loudspeakers were, the less one could have perceived where the sound was from the stage.

Issues related to the performer's perception of the sound and on relationship sound-interaction design are due to coping with technical compromises in a real-world scenario. To optimise the audience's listening experience and to avoid any unwanted Larsen audio effects, the performer was always placed outside the listening area, and thus unable to perceive the spatialisation effects. This resulted in a decrease in the level of immersion and potential disruption of the performer's focus during the first performances.

Even though she was not performing at the acoustic sweet-spot when using the High Order Ambisonic, she experienced a higher level of immersion than in any other performance. The higher feeling of immersion compared from the previous set-ups was given by the sound diffusion of the loudspeaker array facing the performers. In reference to Figure 5.10, those loudspeakers are the ones on the semi-circle opposite the performer. From the audience point of view, on this occasion, the relationship between gesture and the spatial sound was stronger than in any other performance.

The venue size did not affect only the perception of the sound, but also the way the performer interacted with spatial audio processes. In the performance realised at EmuFest, where the performing space was significantly smaller than the one at

Frontiers Festival and ElectroAQustica, Vittoriana De Amicis performed the music with smaller movements.
Also, she commented that sound trajectories were more clear and defined in a smaller environment and with an audio system that could have delivered a higher level of immersion.

When the performer was making fast hand gestures, thus generating fast sound trajectories, the vocal sound was affected by a pitch bend effect. This unwanted effect was due to the movement of the Doppler effect, generated as a result of fast movements of the virtual sound source. This effect was underlined when the size of the virtual room was set to a value greater than five square metres. This issue was solved by implementing an interpolation algorithm to smooth the movement of the virtual source, thus reducing the Doppler effect.

The author's experience using the KW8 Owl loudspeaker allowed the exploration of the physical acoustic characteristics of the space with a fine degree of both responsiveness and resolution. On the other hand, it was observed that by moving the loudspeaker, the sound beam could be projected outside the listening area. This aspect could negatively affect the experience of the sound localisation. Due to the cost and the small number of available KW8 Owl loudspeakers, when this work was realised, this system was utilised in a laboratory environment only (see the video recording no. 6 at Appendix F.1) and discarded the idea of developing MyoSpat around it. However, taking into account findings from Johnson et al. (2016), who successfully evaluated the use of moving-head loudspeakers for musical performance, the K-Array KW8 Owl has the potential to be successfully employed for spatial sound in musical applications through an embodied interaction.

From the comments of the audience of the first two performances, where MyoSpat 0.1 (quadraphonic audio system) and 0.2 (octophonic audio system) were utilised respectively, an accurate rendering of the spatial audio was produced, while MyoSpat 0.3 (High-Order Ambisonic) gave them the sense of stronger spatial sound-gesture relationships. The audience of *VoicErutseG* at Frontiers Festival using MyoSpat 0.1, highlighted that when the performer was making small gestures, they were expecting a small spatialisation effect, but no spatial effect was perceived. This behaviour could have been seen as a weak mapping strategy, yet this was purposely implemented to reduce the gestural data discretisation, given the complex choreographic aspects in performing *Stripsody*.

Unfortunately, the MyoSpat 0.3 system's set-up is not affordable for the vast majority of users, and it is difficult to set up for non-specialist audio technicians. This is an issue present also in MyoSpat 0.4. For these reasons, those two MyoSpat versions were discarded for further development. Although MyoSpat 0.2 audio system it is not affected by affordability or availability issues, it was developed using proprietary software. This makes the MyoSpat system available only to those owning a dedicated software license. Thus, a decision was made to choose MyoSpat 0.1 as the basis for the development of all subsequent iterations.

5.5 Music Gesture Beatbox

5.5.1 Performance

Before discussing *Music Gesture Beatbox* by Grace Savage, it is important to introduce an important aspect of Beatboxing.

Beatbox is a form of art characterised by strong cross-contamination with the visual arts. A remarkable example is the work realised by Reeps One. As with many other artists such as Adam Matta (Lunches, 2013), Julia Ramos (McCormick, 2015) and Joaquin Manay (School, 2017), Reeps One imbues his musical ideas with his visual art practice and vice versa. For him, this is a natural process of being an artist. Drawing a series of lines comes as natural as putting together a series of beats (Flyotw, 2014). Reeps One also applies artistic concepts of Negative Spaces in his beatbox practice. This led him to consider the silence between each beat as important as the beats themselves (Idiosynphonic, 2012). During his live performances, he adopts cymatics (patterns generated by modal vibrations) as an approach for generating visual projections (Idiosynphonic, 2012; UKF, 2014), and obtaining an organic representation of the amplified vocal sounds through the loudspeakers. Reeps One has created a table that categorises sounds and represents them graphically (UKF, 2014).

Considering the importance of visual art in beatbox performance, the research and development work is focused on designing embodied gestural interaction with processes responsible for generating visual feedback and not audio-only. In addition to artistic objectives, visual feedback will also be used to increase the level of immersion, strengthen the sound-gesture relationship, improve usability and learnability (Mainsbridge, 2016), and in turn the comprehension of audio transformations by the audience through visual feedback (Jordà et al., 2007).

Music Gesture Beatbox is a performance divided into three parts, each one characterised by different audio-visual transformations.

In the first part, Grace Savage performs a beatbox routine and shows the process of layering vocal sounds. In this part of the performance, the artistic aim was to merge visual art and beatboxing. Inspired by Reeps One's art practice and interview Participant 8's comment, I decided to represent the process of creating and layering sounds, through an analogy of a painter brushing and layering different colours on a canvas.

> When I play gestures, it's like I'm painting, it is not that I'm doing painting gestures, but I try to make them understand the structure of what I'm doing. That's why I use my hands.
>
> Interview Participant 8 (2016)
> (extract from interview script in Appendix G.1.8).

In this first phase of the piece, Grace Savage draws circles on a screen through metaphoric painting gestures by throwing paint drops at the canvas and then brushing them away (Figure 5.13).

Taking into account findings from work on visualising the movements of "sound-objects" described in Chapter 4.8, here, "the drops of paint" (the circles), and their trace, represent the spatial position and trajectory of the sound respectively.

Figure 5.13: Grace Savage performing *Music Gesture Beatbox*, Scene 1. Screenshot from video recording no. 5 in Appendix F.2

In the second part of the performance, Grace Savage improvises on a video of Ella Fitzgerald singing *One Note Samba* by Antônio Carlos Jobim (1960). Savage's improvisation is visually represented through the manipulation of the *One note Samba* live performance recording. Here she uses different fragments of *One note Samba* several times. This led to considering the video playing speed as a central element of the transformations (reversing, fast-forwarding and playing the video at its original speed). Moreover, to represent her influence on the music, and to highlight the parts where the audience was hearing Grace Savage and not Ella Fitzgerald, the video would fade to a pre-recorded video of herself while beatboxing. The video would fade back to the *One Note Samba* performance when she would not sing (see Figure 5.14 and 5.15).

Figure 5.14: Grace Savage performing *Music Gesture Beatbox*, Scene 2, video reproduction speed control. Screenshot from video recording no. 5 in Appendix F.2

Figure 5.15: Grace Savage performing *Music Gesture Beatbox*, Scene 2, crossfade control. Screenshot from video recording no. 5 in Appendix F.2

In the last scene, to increase the degree of complexity of rhythmical patterns, Grace Savage uses a Boss Rc-505 Loop Station for layering vocal sounds and a band-pass filter to process her voice. Audio transformations are accompanied by effects including the continuous cross-fade between two recordings of her lips while beatboxing, and the control of the speed reproduction of these two videos through the Savage's body movement speed during the performance. Moreover, a colour filtering effect on the videos was applied to synchronise with the control of the band-pass filter (Figure 5.16).

Figure 5.16: Grace Savage performing *Music Gesture Beatbox*, Scene 3. Screenshot from video recording no. 5 in Appendix F.2

The lighting design, in line with the visual projections, supported and enhanced the relationship between the performer's gestures and audio transformations. The lighting cues were synchronised with the creation of a new circle in the first scene. The colours mirrored the colour of the circles, and the brightness was synchronised with the spatial position of the sound on the auditory scene (for example, if the sound moved towards the right, the right light became brighter than the others and vice-versa).

Music Gesture Beatbox has been recorded at Birmingham City University's Parkside Mediahouse (Birmingham, UK) (video recording no. 5 in Appendix F.2), and performed at Music Tech Fest Berlin 2016 (Berlin, Germany).

5.5.2 MyoSpat 0.5

5.5.2.1 Interaction Design

The interaction design was further extended as a result of the *Music Gesture Beatbox* performance. Three more gestures were added to the interaction dressing: *painting*, *brushing*, *playback* and *filtering* gestures, described here related to *Music Gesture Beatbox*.

5.5.2.1.1 ***Painting*** and the ***brushing*** gestures

In the first part of the performance, Grace Savage performs a beatbox routine and enacts *beat* gestures alongside. Given the importance of this and the frequency of this gesture in beatbox performance (Section 5.3), this was built on it to design the *painting* and the *brushing* gestures.
The *painting* gesture is performed as a beat gesture towards the screen as if throwing paint at it. This gesture is to create a circle on the screen. The *brushing* instead is to generate a trail of the same colour as the circle. The trail is composed of circles which progressively become smaller and increase in transparency. The *brushing* gesture in this case of the performance replaces the *pointing* gesture, as it also establishes spatial cues.

5.5.2.1.2 ***Playback*** gesture

A new gesture was also introduced in the second part of the performance, the *playback*. This served the purpose of controlling the video reproduction speed and the cross-fade between two videos. The playback speed was controlled by orienting the arm frontwards or upwards. If the arm was oriented frontwards, the video was reproduced at its original speed, and in reverse if the arm was oriented backwards. To highlight the relationship between the gesture and the video reproduction, Grace Savage performs the *playback* gesture by raising halfway and then increasing the angle of her arm by leaning her body backwards. This backwards movement can be considered an ancillary gesture to convey better the idea of controlling the video's reproduction speed.

With the same vertical movement, Grace also controlled the cross-fade between Fitzgerald's audio-visual recording and an image of herself with her voice as audio feed. This gesture might seem to conflict with the playback one, but here the same gesture is purposely used for controlling different parameters to supporting the creative idea.

5.5.2.1.3 *Filtering* gesture

In the last part of the performance, the interaction design here is to support the gestural control of a band-pass filter synchronised with colour filtering effect on a video and the audio-visual reproduction speed.

The literature review explored in Section 5.3 demonstrated a direct relationship between pitch and verticality when performing the *palette lifting* gesture. To this extent, Eitan and Granot (2006) argued that the pitch-verticality relationship is well-rooted in music through established music notation. This relationship has also been adopted to distinguish the pitches of a sound verbally through the words low and high (Hair, 1981). Studies concerning musical models and cross-domain mapping (Zbikowski, 1997a,b), and research conducted by Wilkie et al. (2009, 2010) on the relationship between musical language and interaction design, describe the validity of cross-domain mapping between sound and verticality. Based on this knowledge, and by working closely with Grace Savage on *Music Gesture Beatbox*, vertical movement of the arm was used to control the parameters of a band-pass filter, so that when she oriented the arm toward the loop-station, the sound would not be filtered, and when gradually raising the arm, the sound would be filtered from a low to high frequency. Referring to the type of audio-visual processing involved, the gesture that controls it is referred to as *filtering* gesture.

In this last section of the performance, the video content is a recording of Grace Savage's lips while beatboxing. The reproduction speed of the video is controlled by the speed at which any gesture is performed. This mode of interaction with video processing parameters had the ultimate goal to establish a strong relationship between the video content (lips movement) and the speed and movement of her gestures, and not to an effect applied to it.

5.5.2.1.4 ***Pointing* gesture**

The inclusion of new gestures demanded a review of the *pointing* gesture, which could conflict with the *playback* and *filtering* gestures. The *pointing* gesture was consequently redesigned to control spatial cues through the orientation of the arm along the horizontal axis only.

Variations in the acceleration of any gesture, similar to the *painting* gesture, were adopted as a way to trigger colour changing cues. The brightness parameter was controlled by the *pointing* gesture so that it would synchronise with sound spatialisation cues.

5.5.2.2 Audio-visual-lighting Engine

The MyoSpat 0.5 audio-visual engine was developed as a further iteration of the previous four versions. MyoSpat audio signal processing is composed of two types of audio processors: a spatialiser, and a band-pass filter. Visual projections include the generation of graphic elements and the manipulation of the speed of the video and colour filtering. The lighting data processing includes the control of colour and brightness changing cues. These data are then converted into a DMX signal for controlling lights, using an Arduino and a TinkerKit DMX Master Shield T140060.

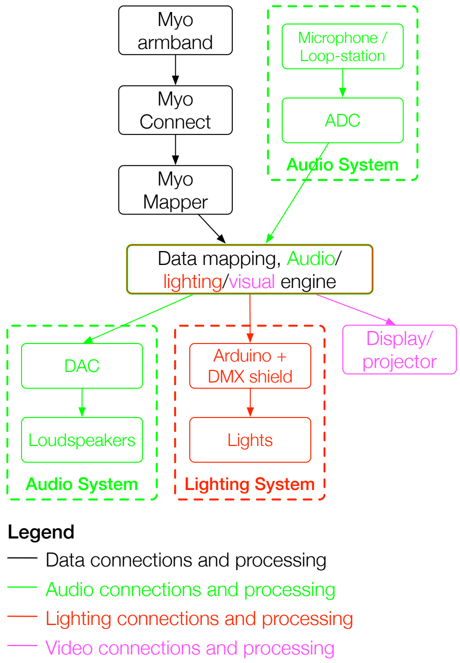


Figure 5.17: MyoSpat 0.5 architecture

In *Music Gesture Beatbox* performance the spatialisation algorithm used was MyoSpat 0.1, with the difference that here was implemented Cycling' 74's Max and a quadraphonic audio system. However, due to technical limitations, at Music Tech Fest Berlin 2016, the audio system had to be limited to a stereo loudspeaker pair.

In addition to a spatialiser, MyoSpat 0.5 also includes a band-pass filter implemented using Max's standard library.

The lighting system was composed of four 4 Par 86 x1W RGB LED lights. Cues to change colour and brightness were controlled using a serial data-DMX signal converter, realised using an Arduino board coupled with an Arduino TinkerKit DMX Master Shield T140060. The Arduino code for the serial to DMX conversion was realised using the DmxSimple Library by Stoffregen (2015).

In MyoSpat 0.5, the video rendering process was implemented using the OpenGL set of video processing extensions "Jitter" for Max.

5.5.2.3 Mapping strategy

Due to the inclusion of the audio filtering effect, substantial changes were made to the MyoSpat 0.5 mapping strategy. To support the new interaction design, pitch values were used for controlling band-pass filter parameters. Hence, in order to maintain a unique sound-gesture relationship, only the yaw values were mapped onto spatialisation parameters.

After evaluating Grace Savage's arm movements on the horizontal axis during rehearsals, it emerged that their amplitude was likely to be within a yaw data range $[\approx 0.3, \approx 0.7]$ ($\approx 45°$), thus significantly smaller than the range of the spatialisation parameters $[0, 1]$ (360°) (Figure 5.18). Consequently, the performer would have generated small changes in the spatial position of the sound. These limitations were imposed by the fact that she was not likely to turn her arm backwards unless the performance included choreography elements. It was also hypothesised that this phenomenon could have been exasperating if performers gestures were obstructed while playing a musical instrument such as a loop-station (in the case of Savage's performance).

From these considerations, yaw data were mapped using a sigmoid function (Figure 5.18, green line). This function allowed data compression into a range to control sound spatialisation trajectories through both small or wide gestures. Due to the experimental approach adopted, this function is here shown as a mathematical approximation with Equation 5.2, calculated by using fitting curve techniques implemented in the Matlab's Curve Fitting Toolbox (The MathWorks, 2014). Data reporting the function's goodness of fit are shown in Figure C.1 and Table C.1 in Appendix C.

$$yaw_{scaled} = \sum_{i=1}^{5} a_i \times sin(b_i \times x + c_i) \tag{5.2}$$

where:

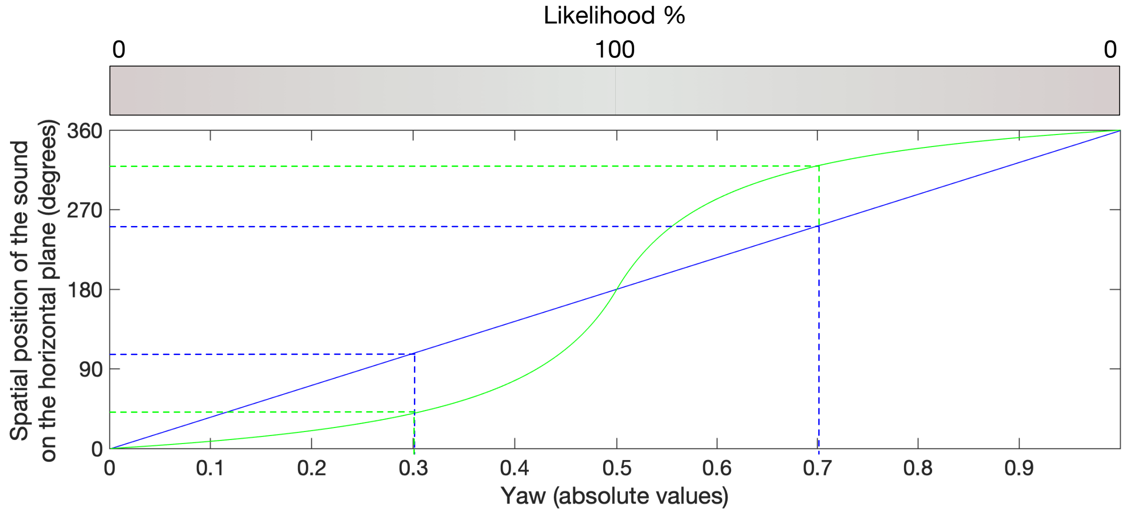


Figure 5.18: Linear (blue line) and sigmoid transfer function (green line) for Myo yaw data mapping into spatialisation parameters. From red to green is the likelihood of a gesture to occur in a region of the horizontal axis, so to explore a range of spatial sound parameters, with red less likely and green most likely

$$a_1 = 5.084 \; ; \; b_1 = 0.7936 \; ; \; c_1 = 4.966$$
$$a_2 = 4.966 \; ; \; b_2 = 0.8515 \; ; \; c_2 = -0.8806$$
$$a_3 = 0.07865 \; ; \; b_3 = 3.499 \; ; \; c_3 = 0.009737$$
$$a_4 = 0.02634 \; ; \; b_4 = 5.781 \; ; \; c_4 = -0.1182$$
$$a_5 = 0.01206 \; ; \; b_5 = 8.027 \; ; \; c_5 = 0.2973$$

Subsequently, data from Equation 5.2 (yaw_{scaled}) were mapped onto distance values of the virtual sound source from each loudspeaker ($d_{1,2,3,4}$) through four different functions (plotted in Figure 5.19), calculated through Equations 5.3, 5.4, 5.5 and 5.6, where $\Delta d_{1,2,3,4_{max}}$ is the maximum distance reachable by the virtual sound source. These equations were established through an experimental approach, based on their effect on the auditory experience. The transfer functions plotted in Figure 5.20 are obtained from this mapping.

$$d_1 = \begin{cases} [(yaw_{scaled} + 0.\overline{6}) - yaw_{scaled}] \times \Delta d_{1_{max}} & yaw_{scaled} <= 0.\overline{3} \\ yaw_{scaled} \times \Delta d_{1_{max}} & \text{otherwise} \end{cases} \tag{5.3}$$

$$d_2 = \begin{cases} (1 - yaw_{scaled}) \times \Delta d_{2_{max}} & yaw_{scaled} <= 0.\overline{6} \\ (yaw_{scaled} - 0.\overline{6}) \times \Delta d_{2_{max}} & \text{otherwise} \end{cases} \tag{5.4}$$

$$d_3 = \begin{cases} (yaw_{scaled} + 0.\overline{6}) \times \Delta d_{3_{max}} & yaw_{scaled} <= 0.\overline{3} \\ [1 - (yaw_{scaled} - 0.\overline{3})] \times \Delta d_{3_{max}} & \text{otherwise} \end{cases} \tag{5.5}$$

$$d_4 = \begin{cases} (yaw_{scaled} + 0.\overline{3}) \times \Delta d_{4_{max}} & yaw_{scaled} <= 0.\overline{6} \\ [1 - (yaw_{scaled} - 0.\overline{6})] \times \Delta d_{4_{max}} & \text{otherwise} \end{cases} \tag{5.6}$$

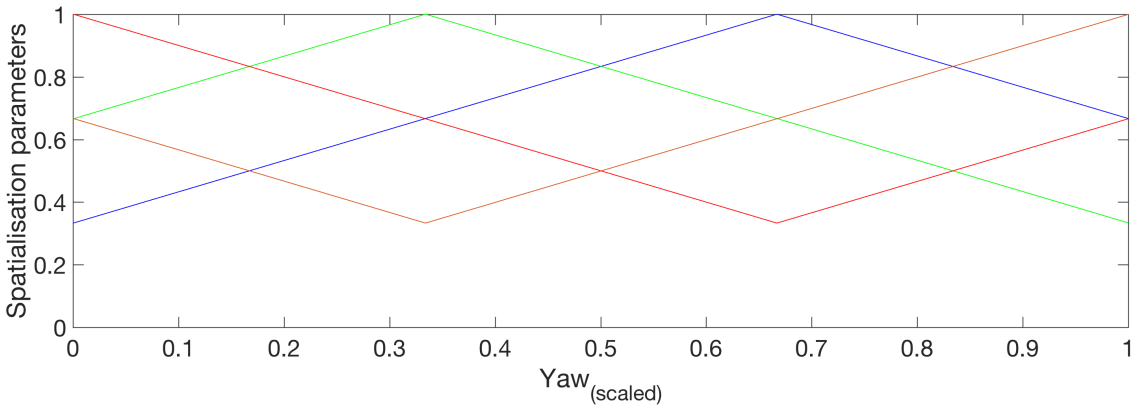


Figure 5.19: Transfer functions of Equation 5.3 (brown) Equation 5.4 (red), Equation 5.5 (green) and Equation 5.6 (blue) with results from Equation 5.2 as input data

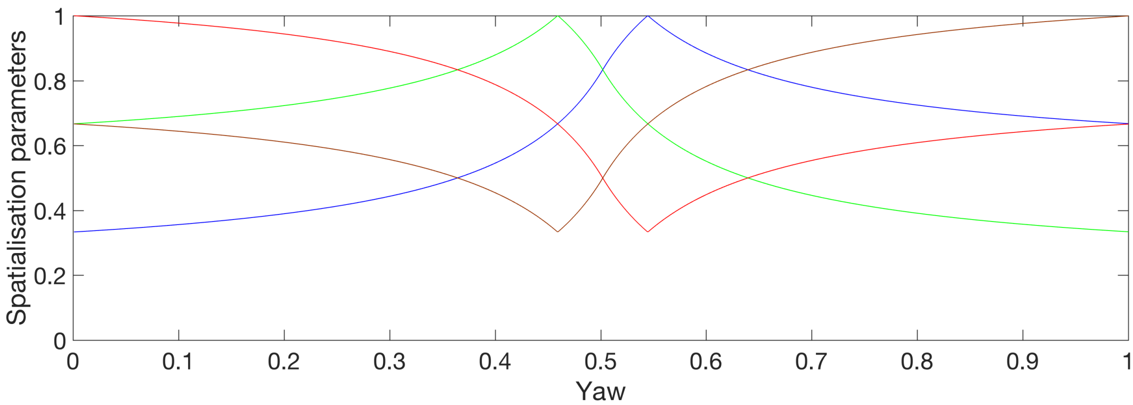


Figure 5.20: MyoSpat 0.5 spatial parameters' transfer functions

Parameters of the filtering effect are established through linear mapping of the Myo pitch value into band-pass filters parameters. Accordingly, when the performer raises their arm (pitch values from 0.5 to 1), the cut-off frequency goes from 1300Hz to 100Hz, the gain from 6 to 1, and the Q factor from 16 to 1. To obtain a smooth transition between the clean and processed audio signal when performing the filtering gesture, the gain of the unprocessed audio signal goes from 1 to when the pitch value goes from 0.5 to 0.85.

Regarding the visual projections, in the first scene of *Music Gesture Beatbox*, circles were generated on the screen when the Myo EMG MAV value was greater than 0.14. This threshold was identified during musical rehearsals with Grace Savage. The circles' position was established by linear mapping of the yaw and pitch values, and their diameter by direct mapping of the EMG MAV value. The bigger the EMG MAV value, the wider the diameter of the circle. Similarly, the length of the tail was directly proportional to the strength of the *painting* gesture. The colour of the circles was chosen randomly, and the brightness by mapping linearly the EMG MAV values $[0, 1]$ into brightness parameters $[0, 255]$ (see video recording no. 7 in Appendix F.1). Furthermore, each time a new circle was formed, some of the letters forming the text "Grace Savage" appeared and disappeared randomly.

In the second scene, the video reverse effect was controlled by a decision boundary algorithm. If the yaw value was greater than 0.5, the video was played at its original speed, and if the yaw value was smaller than 0.5, the video was played in reverse.

Finally, in the last scene, the amount of the colour filter effect was controlled through linear mapping of the pitch value, while the reproduction speed of the video was controlled by linear mapping of the gyroscope sensor's data average: the faster the movements of the performer, the faster the speed of the video.

The colour of the lights mirrored the circles' colour in the first phase of the performance and was randomly selected in the second and third part. To create a strong relationship between sound spatial cues and lighting effects, the brightness of the lights was calculated through linear mapping of each loudspeaker's sound spatialisation gain parameters to a pair of lights. Thus, the higher the gain value of a loudspeaker, the brighter the light placed in correspondence to that loudspeaker (see Figure 5.21 above). For example, if the sound was spatialised to be on the left side of the auditory scene, then the left lights would be brighter. This same mapping strategy was applied for the whole duration of the piece.

5.5.2.4 System set-up.

For MyoSpat 0.5, a stereophonic system was used to diffuse the sound, and four lights were positioned on the ground in front of the stage facing the performer (Figure 5.21). A projection screen or a video monitor at the back of the performer was used to project audio-visual cues.

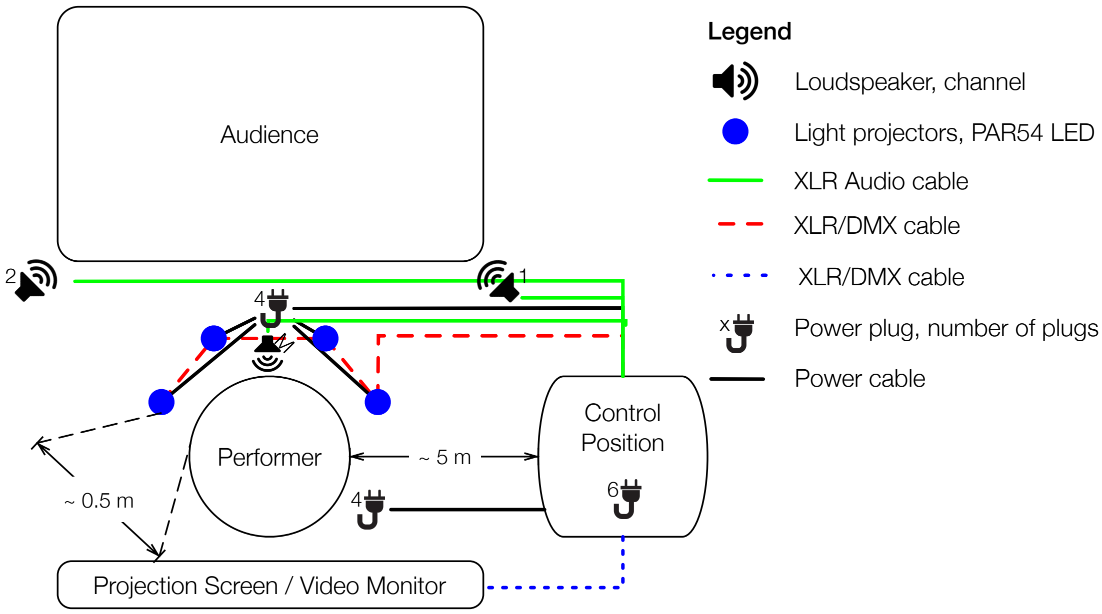


Figure 5.21: MyoSpat 0.5 system set-up

5.5.3 Observations

The interaction design played an important role not only in facilitating the musical performance but also in supporting the creative process. While Grace Savage was composing by adapting her ideas to the interaction design and the system's capabilities, the system was also being shaped to meet her artistic ideas. In this phase, artistic and technological objectives evolved together, co-adapted, towards the realisation of the *Music Gesture Beatbox* performance.

During this process, Grace Savage commented that MyoSpat is a tool that

facilitates the exploration of new vocal sonorities and choreographic elements of the performance. In composing the piece, visual and lighting feedback also represented a source of inspiration for sketching musical and choreographic elements.
However, she described the system as limited due to the small palette of available audio effects. Although the implementation of the filtering effect improved the creative use of the system, during rehearsals, she wished to experiment with a broader number of effects and to recall them through more hand gestures, in particular, finger movements. Similarly, concerning the gesture-audio effects mapping, Grace Savage manifested her wish to control loop-station parameters through gestures, thus extending MyoSpat interaction design to interface the system with third-party hardware and software.

The *painting*, *brushing* and the *playback* gestures fitted within the aims of the MyoSpat interaction design. They did not disrupt the musical performance and satisfied the artistic needs of the performer. However, the filtering gesture was not immediately easy to perform, as the performer needed to operate the loop-station and hold the microphone at the same time. After the first rehearsal, Grace Savage found ways to control the bandpass filter by having control of the processing of the loop-station and holding a microphone. This was easily achieved by adapting MyoSpat interaction design to the required actions of the loop-station and by synchronising interaction with both technologies.
During musical performance, Savage performed playback gestures differently than in rehearsals. She exaggerated this gesture by leaning backwards with the body to highlight the interaction and the rewind effects on the audio-video file. On this occasion, the system behaved consistently and gave her room for expression and improvisation.

During the performance, the system performed as expected. Gestures were recognised correctly when performing *Music Gesture Beatbox* at Music Tech Fest Berlin. On this occasion, MyoSpat 0.5 was utilised with an audio set-up resized to a stereophonic sound system. Even though this did not guarantee a level of immersion as high as if using a multichannel audio system, the performer commented on having still been able to express her artistic ideas through the interaction design and the visual feedback.

However, she found the interaction with the video difficult during the live performance in Berlin, as the projection screen was 10 meters to the left from the performer. This distance partially disrupted the choreographic aspects of the rehearsed performance. For instance, the relationship between the painting gesture and the visual feedback in the first part of the performance resulted to the performer being stronger in the studio performance than the live one, due to the wider distance from the screen.

On the contrary, comments from the audience reported that the visual feedback in the second and third part of the performance was stronger in the live performance, due to the bigger size of the projection screen. In the second and third part of the performance, the gestures were not spatially related to the visual feedback, but the musical composition. Since the screen was bigger than during rehearsal, it made the visual feedback stronger than in the studio performance.

In addition to strengthening the relationship between interaction design and audio-visual feedback during the live performance, lighting effects communicated a sense of immersion to both performer and audience, filling the gap between the projection screen and the performer.

5.6 Discussion

In the first performance, *VoicErutseG*, the interaction design consisted of the *pointing* gesture, prototyped in Section 4.4 to control audio panning parameters and in Section 4.8 to control sound spatialisation parameters. This knowledge has since been extended through the literature review of work on body movements during a vocal performance (Section 5.3) and cognitive HCI studies (Section 5.4.2.1) explored in this chapter, allowing the design of sonic interactions that allow the performer to control sound spatial cues of the voice without disrupting any aspect of the performance. Most importantly, the interaction design allowed the performer to control sound spatialisation parameters during a performance without any particular training. This facilitated the collaboration with performers when they were not able to rehearse with the system until the day of the performance.

The same interaction design was then adapted to the second performance, *Music Gesture Beatbox*. The adjustment of the interaction design allowed the introduction of new gestures, namely the *painting*, *playback*, *filtering* gestures. As these gestures overlapped with the *pointing* gesture, in the fifth version of the system the *pointing* gesture was redesigned to be performable by orienting the arm at 360° on the horizontal axis instead of orienting it at 360° on all axes. The introduction of these gestures was to satisfy Grace Savage's artistic needs in the three parts of the performance.

In contrast to *VoicErutseG*, in which the interaction design suited its performance in different venues, in *Music Gesture Beatbox* the gestural interaction had to be adapted to the technological requirements of the situation. In the performance at Music Tech Fest Berlin, the projection screen was far to the left of the performer and bigger. Thus, Grace Savage had to adapt her gestures to interact with the video. On the contrary, lighting effects were easily controlled, and they filled the physical gap between the performer and the projection screen. Although problematic during the musical performance, visual feedback was of vital importance to the audience for understanding the sound-gesture relationship. In the case of *VoicErutseG*, the gesture-sound relationship was not always clear. Members of the audience commented that they did not know who or what was controlling the sound spatialisation. From the comments of the audience experiencing *Music Gesture Beatbox*, it emerged that the inclusion of visual feedback and the filtering effect helped to blend the actions of the performer and the sonic result. From the musician's perspective, Grace Savage also mentioned the importance of visual feedback for enhancing learnability and usability factors.

The use of a bigger technological ecosystem for the *Music Gesture Beatbox* required a bigger effort from Grace Savage than Vittoriana De Amicis. For example, Grace Savage had to rehearse the filtering gesture carefully while holding the microphone and triggering samples with the loop-station. However, just as what happens when learning a traditional musical instrument, the playing of a digital musical instrument is also affected by a learning curve, during which users, designers and not-trained musicians should build a *"habituated mental model of its constraints"* and find the virtuosity within them (Magnusson, 2010).

The learning curve was not rigorously measured in this work, but from observations, it was evident that Grace Savage spent more time learning the system. However, it is important to remember that she used MyoSpat to compose a new piece of music, and this might have also impacted in the time spent using the system prior to the performance.

Although limiting the analysis to some aspects of the system, the adopted methodology supported the achievement of the objectives, such as the creation of an interaction design which would not disrupt the vocal musical performance regardless of whether the piece was composed using the system or not. In this case study, the system was utilised in two different musical contexts for obtaining a deeper understanding of its possibilities. These aspects will be better addressed in the next case study, where MyoSpat will be used more intensively and more diverse musical applications.

Among all the spatialisation systems, the one adopted in version 0.1 of the MyoSpat was able to render spatial audio and support the interaction design of the subsequent MyoSpat iterations. All MyoSpat iterations were used in musical performance with the exception of MyoSpat v0.3. Even though this one was not utilised in a real-world scenario, it was observed that the K-Array KW8 Owl speaker has the potential to open a series of as-yet-undeveloped interaction design possibilities and relationships between gestures and spatial sound parameters.

5.7 Chapter summary

The work presented in this chapter aimed at the creation of modes of interaction that would result in an embodied music performance. The goal of the interaction design was not to disrupt performance execution nor the creative process that led to its realisation.

Different gestures were defined in this Chapter. The *pointing* gesture, designed for *VoicErutseG*, aimed at giving the singer complete freedom of movement to control sound spatial parameters during the performance. The interaction design was also designed to allow the performer to rehearse vocal and choreographic aspects without the need of the MyoSpat system.
This gesture was chosen after reviewing works on singers' body movements during a performance, cognitive studies on gesture interaction during speaking and HCI literature about touchless interactions. It gave Vittoriana De Amicis natural control of spatial sound parameters, without the need to adapt the performance's already composed choreography to the interaction design.

In a performance of *Music Gesture Beatbox*, the interaction design aimed to support the constrained performer's freedom of movement by a handheld microphone, a loop-station and a projection screen. Four more gestures were added to the interaction design. The *painting*, *brushing* and the *playback* gestures were created to interact with video projections and lighting effects; the first for drawing on the screen and control lighting colour changing cues, and the second for controlling an audio-video file playback. The *filtering* gesture was controlling a band-pass filter's parameters and visual processing parameters. Due to the addition of these gestures, the *pointing* gesture was redesigned. Finally, variations in the acceleration of the arm were used to control lighting effects and visual feedback at defined points in the performance.

The interaction design was successfully applied in both performances, and the MyoSpat system supported the performance of *VoicErutseG* in three different venues, and *Music Gesture Beatbox* in a studio and live context. Nevertheless, due to technical reasons during the live performance, Grace Savage had to slightly adapt her performance choreography to the projection screen.

This first case study enabled the realisation of performances in which performers had total freedom of movement and performances where technological artefacts constrained performers' movements. The interaction design and enabling technology from this case study have been further developed in the next case study, focusing on instrumental performances where the performers' freedom of movement is constrained in various ways by the musical instrument.

Chapter 6

Case Study 2: HarpCI

6.1 Introduction

The second case study aims to extend MyoSpat interaction design to support instrumental performances. The role of the performer/composer and the tensions and limitations inherent in live audio processing are considered here through collaboration with Eleanor Turner, Principal Harp Tutor at Royal Birmingham Conservatoire at the time of writing. This case study consequently goes under the name "HarpCI", Harpist-Computer Interaction.

Controlling and transforming the sound of the harp in live performance poses several challenges for the player. The devices being used may impose uncomfortable body postures, and other issues may often arise concerning the wearability of the sensor and the size of the performance space. More generally, the movements needed to control sound processing may adversely affect those required to play the instrument (Berweck, 2012).

Solutions have been developed for bridging the gap between the performer's technique and the use of technology using sensing technology for controlling audio-visual processing through body gestures. Bevilacqua et al. (2006) developed a system that maps a violinist's bowing gestures into parameters for controlling the violin's live sound, by attaching two ADXL202 accelerometers at the base of the bow. Adopting a similar approach, Overholt (2005) created the Overtone Violin by embedding various types of buttons, knobs, a joystick and an accelerometer on the body of the instrument to control sound manipulations. Wu et al. (2015) realised a system to process vocals and synthesised sounds through "one-to-many" mapping strategies for connecting musicians, instruments and instrumental technique. The aforementioned works demonstrate existing attempts and approaches to extending the creative possibility of performing live electronic music. However, they also appear to be instrument-specific or strongly related to the instrumental technique. By contrast, MyoSpat aims to support musical applications outside the "sweet spot", and to support the performance of any acoustic or electronic instrument and the voice. Here the expression "sweet spot" does not refer to the optimal place for listening, but rather to the optimal condition in which the system interaction design is efficient, as shown in Figure 6.1 (Benford et al., 2005; Benford, 2010).

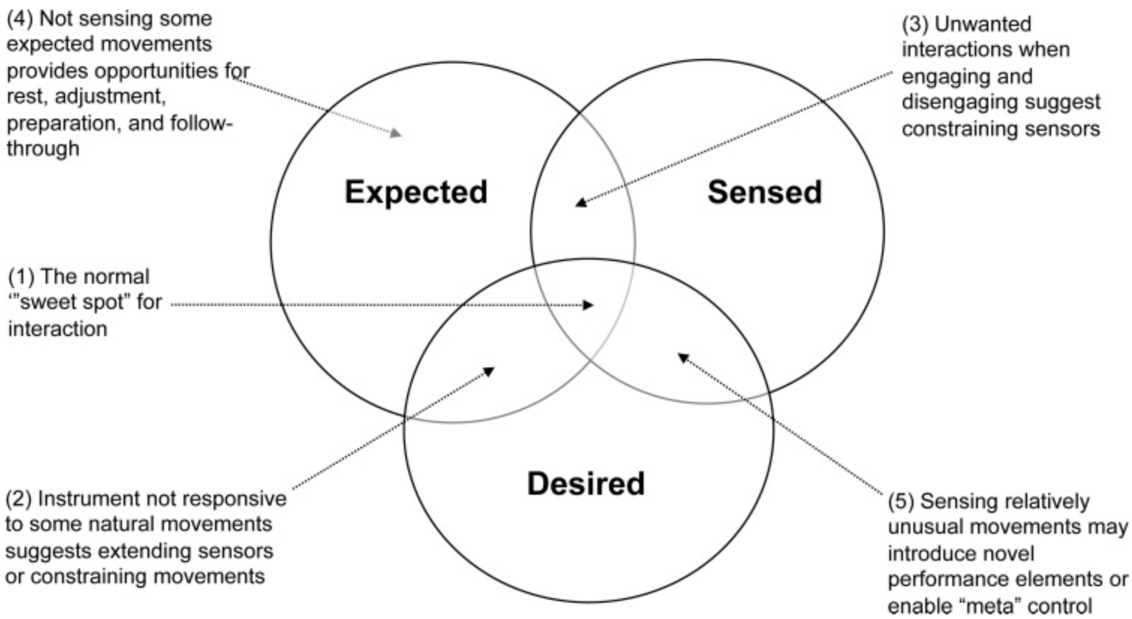


Figure 6.1: "Expected, sensed, and desired interactions for musical performance" (Benford, 2010)

6.2 Methodology

Like to the first case study, modes of interaction with audio-visual processing parameters are designed by absorbing knowledge gathered through a review of the literature on harpists gestures, and a constant feedback loop between system development and artistic creation.
At the time of writing, there was only a relatively small body of literature on the use of technology in harp performance. For this reason, a series of six interviews with professional harpists using technology in their performance practice was conducted in order to learn more. The interviews were semi-structured, with participants freely responding to questions based on various themes concerning harp technique and performance, live electronics, guitar pedals, audio programming environments and gestural controllers, and their general experiences of working in this area. Findings from the interviews were utilised to inform the MyoSpat system development.
Two iterations of the MyoSpat system were designed and developed to support the performances of *The Wood and the Water* and *Star Cluster*, (respectively MyoSpat 0.6 and 0.7).

In addition to observations during a musical performance, laboratory evaluations and workshops were conducted with musicians playing a different type of instruments. The evaluation focused on aspects of the system considering the diverse movement restrictions imposed by the geometry of the musical instrument played by the participant. Three evaluations of the system were carried out to measure the success of the MyoSpat 0.6 design and its implementation. The evaluations were conducted in a controlled environment in the form of a user study at Royal Birmingham Conservatoire, in a workshop delivered for Camac Harp Weekend 2017 at Cardiff Metropolitan University, and in another workshop at Southampton University.
A second user study was conducted at Berklee College of Music to evaluate the last MyoSpat iteration, 0.7.

The UCD cycle applied to this second study is summarised in Figure 6.2.

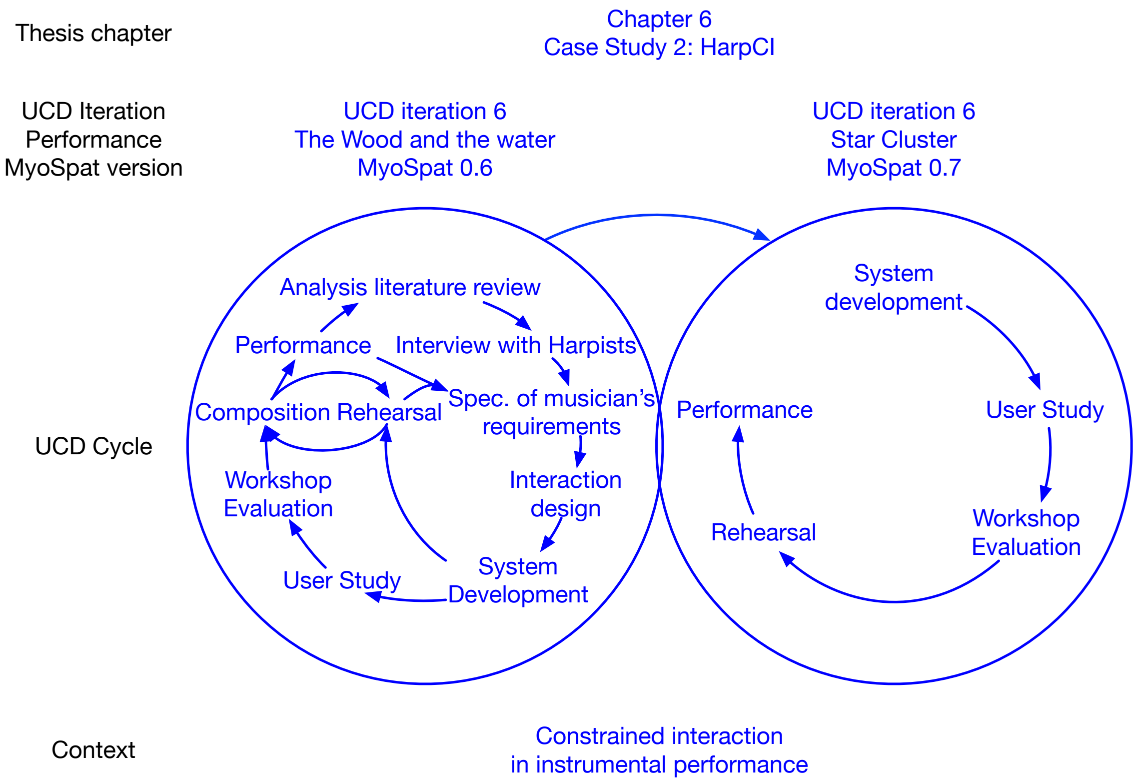


Figure 6.2: UCD cycles in Voice Gesture case study.

6.2.1 User Study

User studies were conducted following a mixed methodology that included: background questionnaires, task-based evaluation, semi-structured questionnaires and User Experience Questionnaires.

Participants were recruited through Birmingham City University's mailing lists, social networks and the author's website. Participants were recruited to represent the user population: musicians and technologists. No specialist training or capacity to play a specific instrument was required as the aim of the study was to observe the musician and the system behaviour in a wide range of situations.

6.2.1.1 Background questionnaire

The background questionnaire aimed at learning about the participants' professional background, their knowledge of audio processing. Specifically, sound spatialisation, and their experience with interactive gestural controllers, specifically the Myo.

The background questionnaire included the following questions:

1. Are you a musician?
   (a) If yes to question 1,
      i. In what role (i.e. composer, performer, conductor, technician, researcher, teacher)?
      ii. Do you play any musical instrument?
      iii. If yes to question 1(a)ii,
         A. What musical instrument?
         B. For how long have you been training your music skills?
   (b) If no to question 1,
      i. What is your background?
2. Do you know the principles of audio signal processing?
3. Do you know what sound spatialisation is?
   (a) If yes to question 3,
      i. Have you ever experienced sound spatialisation?
         A. If yes to question 3(a)i,
            I. In what occasion?
            II. Have you ever experienced it more than once?
4. Have you ever used interactive technology?
5. Have you ever used gestural controllers?
   (a) If yes to question 5,
      i. In what occasion?
      ii. Have you ever used the Myo?

6.2.1.2 Task-based evaluation

Orio et al. (2001) introduced the use of HCI methods for the evaluation of input devices used in musical performance. They presented the concept of "musical task" for evaluating input devices, also called controllers, during a musical performance.

A musical task involves the use of a controller to drive synthesis or processing parameters. They suggest designing tasks that allow evaluating the controller in each detail, even if the interaction and auditory feedback could be regarded as unmusical. Orio et al. take into account specific properties of the task, such as learnability (*the time needed to learn how to control a performance with a certain controller*), explorability (*exploration of the capabilities of the controller*), feature controllability (*the accuracy, resolution, and range of features perceived by the user when performing musical tasks*) and timing controllability (*temporal precision at which the musician can control the performance and its relationship with the tempo speed*). The task-based evaluation has been widely adopted in the NIME community as a method to evaluate interactive systems and tools and identify and expand their capabilities (Xambó et al., 2013).

In their framework for the evaluation of interfaces for musical expression, Lee and Wessel (1992), and successively Wessel and Wright (2002), highlight another essential aspect, the action-perception feedback loop to which the musician is exposed. In their framework, they also include a stage called "evaluation by perception" in which the musician evaluates and learns new possibilities afforded by the interface through the sensorimotor experience during playing (Figure 6.3).

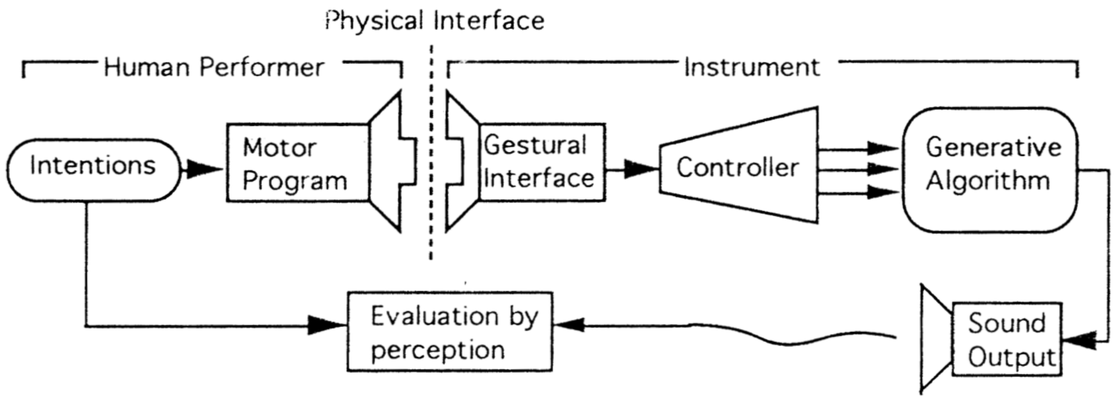


Figure 6.3: Lee and Wessel (1992) framework for electronic music instruments.

With this literature in mind, the present task-based evaluation looked at (i) the relationship between user interaction and audio-visual feedback, (ii) gesture recognition performances through direct evaluation, (iii) participants' perception of the audio-visual feedback and the relationship between intended and perceived interaction attributes, and (iv) participants' views on strengths, weaknesses and potential improvements of the system.

A demonstration of the MyoSpat system was delivered before presenting the tasks to the participants. The demonstration was conducted through verbal explanation first, and then by adopting a "learning through activity" approach, successfully used by different educational technologists (Roussou, 2004). This approach is built upon the principles of "learning-by-doing" and "hands-on" methods and encourages users to learn the potential use of interactive systems through active practice. Based on this knowledge, participants were asked to learn how to associate sound features with corresponding movement features while exploring the system. During the task-based evaluation, participants were also interviewed through a semi-structured interview. This allowed gathering qualitative data concerning their view on the audio-processing, the interaction design, the gesture-feedback relationship, potential issues, improvements and artistic use of the system.

6.2.1.2.1 Procedure

After demonstrating how to use the system, participants were asked to control and transform the sound and generate lighting effects through the related modality of interaction (see Section 6.5.2.1 for details about MyoSpat 0.6 and 0.7 interaction design). Participants were asked to utilise each audio-lighting effect five times, for a total of 25 tasks. Audio effects were presented randomly. They were chosen using an algorithm implemented in Pure Data to generate values from 0 to 4, with each value associated with an audio effect. Participants initially executed all tasks without visual feedback from the system, and then with visual feedback.

Tasks were executed in a free-space interaction context (without any constraint to the gestural interaction), and a pre-recorded audio file was used as an audio source for the MyoSpat. Later, where participants were musicians, they were asked to repeat the tasks while playing their musical instrument, thus feeding the MyoSpat system with the sound of the instrument. They were not asked to play anything specific nor with a particular instrumental technique. Participants were allowed to improvise or play any piece they preferred.

After each completed task, participants were asked (i) to rate positively or negatively whether their perception of the audio-visual feedback was as expected, (ii) to comment on the relationship between the interaction with the system and the audio and/or visual feedback, and (iii) any other observation relative to issues, strength and potential use of the system. The questions asked were:

- What do you think about the audio processing?
- How many times the system did not meet your expectations?
- Which of the gestural interaction was the easiest to perform?
- Which of the gestural interaction was the hardest to perform?
- What do you think about the sound-gesture processing relationship?
- Which of the performed tasks resulted in having the best relationship between the gestural interaction and the auditory feedback?
- If using lights, what do you think about the lighting-gesture relationship?
- If using lights, do lighting projections facilitate the use of the system?
- Would you use MyoSpat during a musical performance?
- Would you consider MyoSpat as an expressive tool to create music?
- Do you have any other comments?

The task-based evaluation procedure is summarised in Figure 6.4.

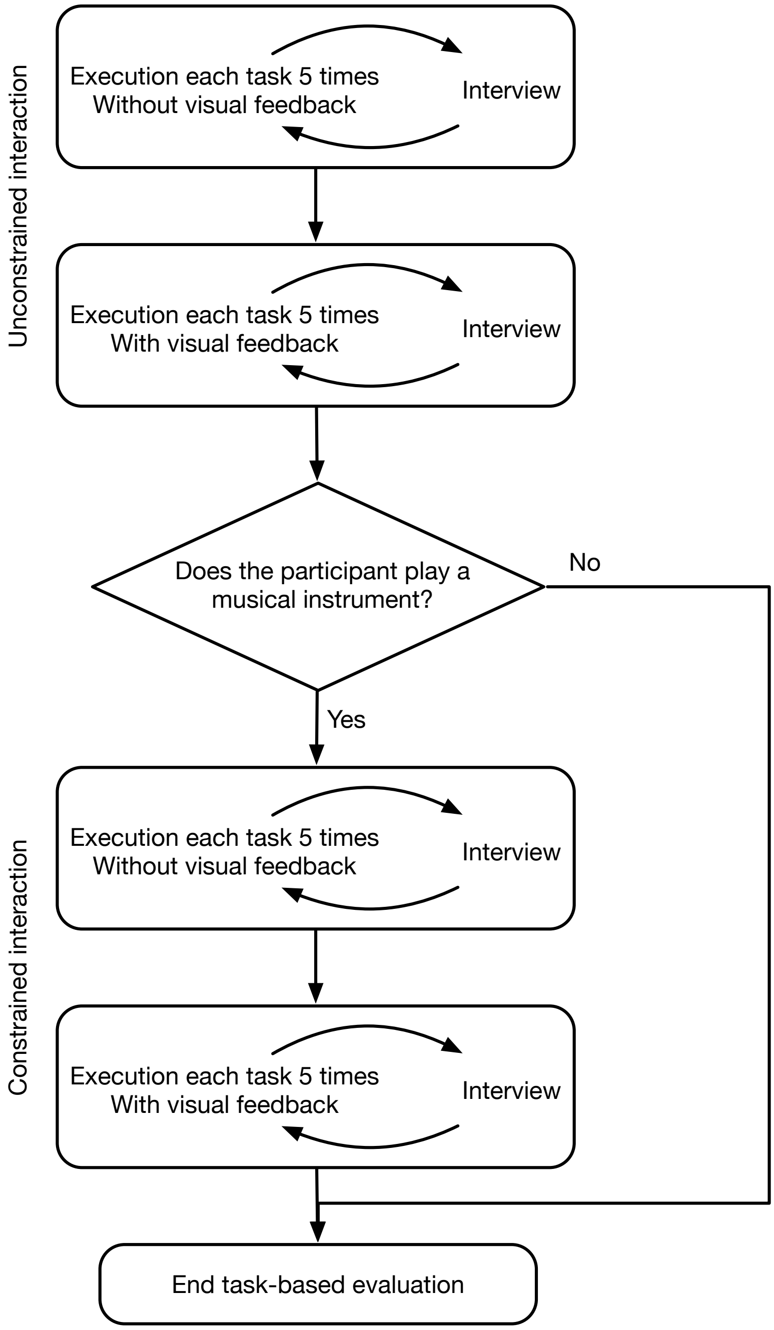


Figure 6.4: Task based evaluation procedure

6.2.1.2.2 Apparatus

To collect reliable data, MyoSpat was set up to ensure optimal usability conditions, and to reproduce potential issues that occur in live rehearsals and performances. For instance, the difficulty of having a quadraphonic audio system setup of identical loudspeakers in a quadratic perimeter, and the possibility for both performer and audience to appreciate the audio-visual feedback from a favourable position, such as *"the small area located on the symmetry axis between the loudspeakers —- the so-called sweet spot"* (Merchel and Groth, 2010). For these reasons, a quadraphonic system was used, composed of two pairs of different loudspeakers and positioned at slightly different heights. Loudspeakers were placed at the corners of a four-square metres space. Two Genelec 8030A loudspeakers were placed at the front-left and front-right of the participant, at the height of 1.6 metres from the ground. Two Genelec 8050A were placed at the rear-left and rear-right of the participant, at the height of 1.4 metres from the ground. Lights were mounted on a microphone stand and placed next to each loudspeaker. Both lights and loudspeakers were facing the user.

In the first phase of the task-based evaluation, a pre-recorded monophonic audio file containing the sound of continuous water flow was used as audio a source for the MyoSpat system. The file had a duration of one hour, with 44.1 kHz sample rate and 16-bit resolution. Through this approach, participants were able to test the system and evaluate audio-visual feedback through a comparable experience. The study could have also been extended to non-musicians to observe the system's behaviour and potential when used by people with no prior musical experience. A water sound file was chosen for its broad spectrum and random dynamic recursiveness, so as not to favour any audio effects over any others.

In the second phase, musicians were asked to repeat tasks while playing their musical instrument. The live audio signal was digitally converted using a Motu 828 firewire soundcard, with a sample rate of 44.1 kHz and 24-bit resolution. Electronic instruments were connected via Jack cable to the sound card. The sound from acoustic instruments was captured with an Audio Technica AT4050ST stereo condenser microphone positioned respectively to the sound radiation point of each musical instrument.

6.2.1.3 User Experience Questionnaire (UEQ)

To evaluate the user experience qualitatively, participants were asked to fill a User Experience Questionnaire (UEQ), at the end of the task-based evaluation. The UEQ is an easy to apply and reliable measure of user experience that can be used to complement data from other evaluation methods with subjective quality ratings (Rauschenberger et al., 2013; Holzinger, 2008; Schrepp et al., 2014). In this study, the outcome of the UEQ is to complement the task-based evaluation and the qualitative data from interviews.

The structure of the questionnaire allows the user to express feelings, impressions, and attitudes while using the product. The UEQ allows measuring standard usability criteria, like efficiency, controllability or learnability, and non-goal directed or hedonic quality criteria, like stimulation, fun-of-use, novelty, emotions or aesthetics. Specifically, it is possible to rate attractiveness, perspicuity, efficiency, dependability, stimulation and novelty. Attractiveness indicates an overall impression of the system. It answers the question of whether users like or dislike the product. Perspicuity tells if it is easy to get familiar with the system and if it is easily learnable. The efficiency factor shows whether the user employs unnecessary effort while using MyoSpat. The dependability factor rates the participants' feeling of being in control of the system. The stimulation factor evaluates how exciting and motivating the system is. The novelty factor rates the users' opinions about the system being innovative, creative and interesting. Attractiveness is a pure valence dimension. Perspicuity, efficiency and dependability are pragmatic quality aspects (goal-directed). Stimulation and novelty are hedonic quality aspects (not goal-directed).

Written and verbal instructions about how to complete the questionnaire were provided to the user. The UEQ models and spreadsheet were provided by the UEQ creators.

6.2.1.4 User Study Procedure

The user study procedure is outlined below:

1. Introduction to the workshop (5 minutes)
2. Background questionnaire (5 minutes)
3. MyoSpat demonstration (10 minutes)
4. Task-based evaluation (25/50 minutes)
   (a) Tasks with audio-file (25 minutes)
   (b) If the participant is a musician, tasks were repeated using a musical instrument (25 minutes)
5. Free exploration (30 minutes)
6. User Experience Questionnaire (10 minutes)

6.2.1.5 Data analysis

Different approaches were used to analyse user studies data. A light statistical approach was used for quantitative data and thematic analysis for qualitative data. While statistics from quantitative data aimed to compare objective results between iterations of the system, qualitative data gave insights about the subjective participants' opinion.

Participants' answers to the background questionnaire were clustered and reported using tables (see Table A.1 and A.2).

Interviews and feedback from participants were analysed as in the first case study, utilising Thematic Analysis and elements of the Grounded Theory Method. Codes found among the interview are listed in Appendix D.3.

Quantitative data relative to the performance of the gesture recognition system (how many times the gesture recognition process outputs the correct result) and the user's perception of the feedback (how many times the user perceives the audio-visual feedback as expected), were analysed by drawing statistics. Data were grouped by user interaction and statistics were calculated in percentages using Table A.3 and summarised using A.4 and A.5 in Appendix A.

UEQ results were calculated using the UEQ Data Analysis Tool provided by Hinderks (2015). This last one consisted of a spreadsheet template that allows a statistical analysis of completed UEQ forms by the participants.

6.2.2 Workshop

MyoSpat was also evaluated during workshops delivered by Eleanor Turner and the author to assess harpists' user experience when improvising with the system. Due to the workshop schedule, which included a series of activities focused on the use of different ways of processing the harp sound, improvisation with the system was limited to five minutes per each participant. The behaviour and comments of the attendees were observed and noted, and they completed a User Experience Questionnaire at the end of the workshop.

Workshops were realised at Royal Birmingham Conservatoire, the Camac Harp Weekend 2017 and the University of Southampton.

6.3 Interviews with harpists

Six professional harpists working with technology took part in the interviews: Audrey Harrer, Arnaud Roy, Hélène Breschand, Jennifer Ellis, Sofia Asunción Claro and Úna Monaghan. A common theme to emerge from the interviews was that of learnability with regard to technology and electronic systems, demonstrating that a high number of skills and techniques are required to approach this discipline with confidence. The extracted codes from the interview are listed in Appendix H.2.

Most harpists commented that technology requires "too much time" to be learnt and requires time away from music-making; in some cases, the interviewee felt the time invested in learning technology for a limited number of performances was not well spent.

> [...] technology can push you away from being creative because as a

> result, many stimuli, so many instruments [...]. I have to constrain myself to learn another technology because my main goal is to create music and no to learn new tools. [...] I was very happy to learn the MIDI harp because I was one of the firsts to do this. I was very proud of it, spent much time to learn the technology and learn how to play. But, in my live show, I played the MIDI harp maybe four or five times. So it's not much compared to the time I spend to create the pieces. It's a form of research, so you have to be patient.
>
> Arnaud Roy (2017)
> (extract from interview script in Appendix H.1.5)

Harrer, Breschand and Ellis highlighted the importance of practising technology to embed it cohesively in their musical practice, considering the study of technology equal to that of any other instrumental technique. Asunción Claro shared this sentiment saying: *"My experience with electronics and technology was very pleasant. I did enjoy it. I never had problems because I was practising to have things under control"*. Similarly, Harrer emphasised the importance of practice in a comment: *"Using technology is like learning anything else. It is like learning a passage on the piano. It's a matter of practice. It's just something that develops through practice. It's like anything."*

Most of the interviewees had the experience of using guitar pedals in their practice, with the above observations describing how study and practice were vital to confidently developing the muscle memory needed to precisely trigger effects. Breschand considered them *"a musical instrument"*, *"an extension of the harp"*. The freeze pedal was the most used one among the interviewees. As the harp sound is characterised by a short sustain, being able to make the sounds indefinitely longer is particularly appealing to harpists. To overcome this limitation of the harp sound, Arnaud Roy adopted an original approach to the freeze pedal in *The Families* (Lumeris Theme) (Roy and Bouchet, 2017), where he combined the use of the pedal with audio transformations driven by the Leap Motion gestural controller (Leap Motion, 2014). Specifically, he used the freeze pedal to create and consequently manipulate "pad" sounds using the Leap Motion as an input device and Live as an audio engine.

The majority of interviewees described pedals as reliable devices that make music technology more accessible. Ellis underlined the fact that harpists are very skilled at moving their feet, and potentially already have the skills required to integrate guitar pedals into performance confidently and with reduced practice time. However, the use of guitar pedals with a pedal harp can potentially disrupt the performance, as the feet are normally required to perform chromatic changes on the instrument. This is one reason why Ellis believes that gestural technology could be a solution for designing a closer interaction with the audio processing. Roy supported this idea when talking about his experience with the Leap Motion:

> I think the Leap Motion is great because it adds signage for the audience. It is something special when I move my hands the sound changes. So I think this actually is very interesting for people, the Leap Motion. [...] I think it's a very interesting way to perform, to have good control of your sound, and it's very accurate. I think the Leap Motion is accurate so I can do whatever I want with it. Sometimes it's a little bit buggy, a little, but I feel comfortable with it. [...] When I started to play with the Leap Motion, I was amazed by the way my hands were controlling sound. It's kind of magical, is like you are feeling the sound, like you are a sculpture, you know? And you can make a face or something like a landscape. It's really magical, so I think it is a really enjoyable performance for the performer and also for the audience because the audiences feel that when you are moving your hand the sound changes. I think, maybe, another idea is to have lights. People love that.
>
> Arnaud Roy (2017)
> (extract from interview script in Appendix H.1.5)

Roy acknowledged the importance of communicating to the audience the action and response between physical gesture and sonic result through the performer's gestural interaction with the technology used. Other interviewees shared the idea that gestural technology represents a form of visual feedback for the audience to experience and engage with the musical performance. This is how Asunción Claro described it:

> [talking from the audience point of view] I think it is interesting to listen [to sound transformations as] the result of my own movement, [it] is quite unexpected that because you move because you do some movement, the sound will change.
>
> Sofia Asunción Claro (2017)
> (extract from interview script in Appendix H.1.4)

In response to *"Do you design gestures considering the visual impact it would have on the audience or do you design them to facilitate the musical performance?"* Monaghan said:

> Both of them. I work on all these things in parallel. If I want something to work with a gesture and that's going to happen, I'm aware of the visual feedback of the audience. Sometimes, I push the button on the motion sensor, and I want the audience to see it, sometimes I don't want to show that, so I don't make it obvious. But other times I want them to know what's going on. So, I make it more obvious. [...] It is always a parallel process between what is practical, what is compositionally desirable, what is desirable for the sound. [...] I'm thinking of composition, performance and practicality of the software and coding in parallel. I work with these strands as important as the others.
>
> Úna Monaghan (2017)
> (extract from interview script in Appendix H.1.2).

In Roy's live performance of *FlyByNo The Families* (Lumeris Theme), he used a Leap Motion, placed on a laptop, and controlled it with his left hand (Figure 6.5). From an interaction point of view, the Leap Motion does not solve all the issues that harpists encounter during a live electronic performance. For example, having to move the hands away from the instrument, thus interfering with the continuity of the performance, or having to adapt the instrumental technique in response to the demands of the device's motion capture capabilities.

Ellis noted her experience of this issue whilst using IRCAM's MO (Modular Musical Objects device) (Bevilacqua, 2012), for performing *Weav-Weav-Weaving I, II* and *III* (Ellis, 2014). Ellis took an empirical approach to develop how she could effectively use the device. Attaching the device to her left arm using a smartphone

Figure 6.5: Arnaud Roy performing *FlyByNo The Families* (Roy and Bouchet, 2017)

support band, Ellis used the Max software to develop meaningful mappings between movement and audio processing for *Weav-Weav-Weaving*. Monaghan describes a similar approach to develop gesture-controlled audio processing, having used IRCAM devices as well as X-IO Technologies X-IMU device (x-io Technologies, 2016).

A preference for using the left hand to control audio manipulations emerged among the harpists interviewed, whether it is turning a dial on a guitar pedal or using a gestural device. With the harp leaning on the right shoulder, the left hand has more freedom to move and is used to dampen the strings. Both Roy and Monaghan discuss this.

This is the chosen approach for establishing mapping strategies as it facilitates the interaction with the electronic audio processing in a way that is sympathetic to standard harp technique. Roy's work with the French harp manufacturer Camac on a new MIDI harp examined ways to manipulate the sounds through an interaction design built on instrumental technique. Roy did, however, report how the MIDI harp presents difficulties in creating meaningful mappings, and how unnatural it is to manipulate sounds or play virtual instruments with it. Comparing the MIDI harp to gestural controller technology, he described the latter as "more easy to set up, more easy to use and more enjoyable"; adding how attaching the Leap Motion to the body of the harp would facilitate a method of gestural interaction that does not require the hands to move away from the instrument. Camera-based devices were also used by Asunción Claro. She performed using a camera to control and transform the sound in *Sound Shapes* by Graugaard (2008). Although she performed the piece without encountering any technical problem, she also described the vital need of a skilled technician to set up the interactive system.

In addition to choosing the right technology to support the gestures inherent in performing a specific musical work, mapping strategies play an equally important role in creating coherent connections between action and response when using interactive technology (Hunt et al., 2002). In her *Weav-Weav-Weaving I*, *II* and *III* for live electronics, Ellis demonstrated her use of the same devices to manipulate the sound of the harp with different interaction design and mapping strategies for each work. The technology here becomes composition-specific, adapting to the type of harp and the required instrumental technique for each work, especially concerning movements of the feet when playing the pedal harp.

Developing strong, lucid connections between instrumental technique, gesture, technology and musical output emerged as a recurring theme throughout the conducted interviews. An approach that considers these aspects can create meaningful musical experiences for performer and audience alike.

6.4 Body movements in harp performance

Jensenius et al. (2009) divide musical gestures into sound-producing, sound-facilitating, communicative and sound-accompanying. This taxonomy was used to analyse the harpists' gestures in performance. Sound-producing gestures are those responsible for the production of the sound. Communicative gestures express the intentions of performers to the audience or co-performers. Sound-facilitating gestures enhance the production of sound, for instance, to keep the tempo, to reach the strings or to control a note. Sound-accompanying gestures do not play any role in the production or control of sound. Although distinguishable from their relationship with the sound, it is also important to take into account body postures and the context in which gestures are played (Leman, 2009; Visi et al., 2014) and the fact that the body posture and the performed gestures are cross-related (Desmet et al., 2012).

Results from a gestural interaction analysis during harp performance carried out by Chadefaux et al. (2012), showed that the global posture of harpists is very stable during playing and that the relationship between arm and limbs is constant. Furthermore, they found that the shoulder represents a fixed point from where sound-producing gestures have their origin, and the hands perform more accentuated movements. They concluded that most of the harpist's interpretation is delivered through movements of the hands. In a later study, Chadefaux (2013) analysed and compared shoulder, elbow and wrist movements among three participants playing Debussy's *Danse Profane* (1904). From this analysis, she estimated that the shoulder is the joint that performs most of the work during performances, secondly the elbow and ultimately the wrist. Studies from Le Carrou et al. (2008), who analysed the finger-string interaction during harp performance, and later findings by Chadefaux et al. (2013), who developed a model of the harp plucking, show the importance and commonalities of hand and finger gestures among performers.

6.4.0.1 Sound-producing gestures in harp performance

The primary sound-producing gesture in harp performance is the *plucking* gesture. It can be divided into three temporal stages: stick, slip and free oscillation phase (Chaigne and Kergomard, 2016). In the stick phase, the strings are pulled until a certain threshold is reached. Then they are released (slip phase) and left to oscillate freely (free oscillation phase).

Another sound-producing gesture is the brushing technique, which involves the performer sliding the fingers very fast across a group of notes. The hand should assume a standard position and each finger should close after release. This is a standard technique used for playing an ascending or descending Aeolian chord, like the one in Figure 6.6.

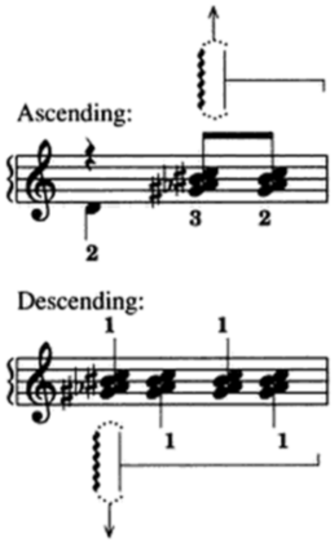


Figure 6.6: Example of ascending and descending aeolian chord (Kondonassis, 2006)

6.4.0.2 Sound-facilitating gestures in harp performance

The most commonly employed sound-facilitating gesture in harp performance is the *raising* gesture. It is performed by gradually closing the hand from the plucking position into a fist while raising the wrist. This gesture can be enacted using the French or Salzedo methods (Huang, 2011). Renié's *raising* gesture (French method), is performed by an outward wrist movement, either with or without the arm. This gesture facilitates the control of the sustain (Huang, 2011). Salzedo's *raising* gesture is executed with the elbow at a fixed point, raising the closed hand with complete control. The Salzedo method also includes indications concerning facial expression: during the rising gesture the performer should not look at the hands but in the direction of the soundboard or the strings, and should listen accurately until the sound is extinguished; to this extent, the raising gesture improves tone, technique and musical interpretation. This movement also enhances muscular and mental relaxation and physical respiration during a performance (Dennett, 2015). The *raising* gesture helps to decrease the noise produced by the premature replacement of the fingers on the strings (Kondonassis, 2006). Salzedo (1952) also points out the importance of the aesthetic aspect of this gesture:

> Since the instrument [the harp] does require necessary gestures, these gestures should be correct not only functionally but aesthetically as well. Music is meant to be heard, but also to be looked at - otherwise radio would long ago have supplanted the concert stage which, fortunately, it has not. Like the orchestral conductor, the harpist must learn how to externalize music. This he does through his gestures, which should be inspiring and not dry; these gestures should emphasize and not negate the intent of the music. On this basis, I give my students training, which permits them to play aesthetically as well as musically. I explain to them that the harpist should consider himself like the orchestra while his gesture-making hands are the conductor who calls forth his playing and depicts phrases through motion.

Salzedo then continues commenting on the importance of the *raising* gesture:

> The basic harpistic gesture is the raising of the hands slowly and with complete control. Once this ascending gesture has been mastered, the perfectly controlled hands are then at your service for all kinds of touches. When the hands react sensitively to the various rhythmical and emotional requirements of music, there is no difficulty in rendering a composition as intended by the composer. This slow, controlled raising of the hands grows out of complete relaxation - first mental, then muscular aesthetically accomplished. For example: to prevent tenseness, some harpists recommend dropping the hand wrist up. This is an ugly gesture. I cure tenseness by stressing no particular gesture as law, but by inviting mental relaxation and, for a while, avoiding loud or fast playing.
>
> Salzedo (1952)

The importance of sound-facilitating gestures is further confirmed by Chadefaux et al. (2012), demonstrating through kinematic analysis of the harpists' gestures that a high level of hand and arm activity occurs after the plucking action.

6.5 The Wood and the Water

6.5.1 Performance

The Wood and the Water is a piece for electric harp and live electronics composed and performed by Eleanor Turner using MyoSpat. Turner aimed to express and communicate some of her personal experiences through an original type of musical poetry. The first step was writing down the foundation poem on paper to establish what she wished to express. Then, using British Sign Language (BSL), it evolved into a more descriptive and expressive poem that she could sing and play on the harp. The simplest musical gestures in the piece are BSL signs that begin on the harp with the plucking of the strings to create the sound and are completed away from the strings. In fact, they often continue for a long time away from the strings and even away from the instrument - above, around, behind, underneath and on the side of the harp, enabled by the MyoSpat sound spatialisation and delay.

Aside from those exact signs that create the poem, the music sets the scene of walking through a forest, hearing the feet crunching through the leaves, atmospheric sounds coming from all around and being alone with one's thoughts. A connecting musical motif takes us further on our walk through the forest; a break from the signed poem and complex electronic effects. The most intense moment in the piece uses spatialisation together with the gesture-controlled effects and the signed poem; all brought about by the discovery of a pool of water calling for the author's honest, personal reflection. The rhythmical spoken word and music passage that follows is Turner's impassioned response to this challenge and is dense with electronic effects and the complexity of words and music angrily spilling out all over each other.

The *Wood and the Water* has been recorded in studio (Figure 6.7; video recording no. 6 in Appendix F.2), and performed in different music festivals and conferences, such as Birmingham Harp Day 2016, Camac Harp Weekend 2017, Audio Mostly 2017 and Shanghai's Electronic Music Week 2017. The video recordings of selected performances of the piece are included in this thesis as Appendix F.2, no. 7 and 8. The musical score is also attached to this thesis in Appendix I.

Figure 6.7: Eleanor Turner performing *The Wood and the Water*. Screenshot from video recording no. 6 in Appendix F.2

6.5.2 MyoSpat 0.6

6.5.2.1 Interaction Design

Five new gestures to interact with sound and lighting processing parameters were designed from considerations upon harpists gestures during a performance. The gestures were: the *crumpling*, *extend*, *lower*, *throwing* and *clean* gestures. These, and the *pointing* gesture borrowed from previous iterations of the system, formed the MyoSpat 0.6 interaction design.

6.5.2.1.1 *Crumpling* gesture

The *plucking* gesture is the physical body movement responsible for excitation of the harp strings, which then causes the production of the harp sound. This gesture is shared with the instrumental technique of other chordophone instruments. In addition to the gestural aspects, the sound produced by chordophone instruments and their synthesis has been a subject of interest in the research community.

Karplus and Strong (1983) realised one of the earliest examples of the synthesis of a plucked string sound. Later, Cook and Trueman (1998) used function transfers from plucked sounds of guitar strings to allow the composition of music in a virtual environment. Karjalainen et al. (1993) proposed new methods for improving guitar sounds synthesis. The Digital Musical Instruments (DMIs) community has created augmented string instruments by combining both gestural interaction and sound synthesis inspired by the action of plucking strings. For instance, Kapur et al. (2004) created the Electronic Sitar Controller based on the sitar's instrumental technique and a representation of the sitar sound among other audio processing techniques. Other examples can also be found in consumer digital musical instruments, for instance, Jamstik+ (Zivix, 2017), Artiphon (Artiphon, 2017), MI Guitar (Artiphon, 2017) and the Guitar Hero Controller (Activision, 2017).

Given these considerations of the *plucking* gesture in playing the harp and any other chordophone instrument, and in the wider field of DMIs, this gesture was taken as the starting point for designing a new interaction modality.
After exploring the Myo's muscle sensing capacity for controlling audio parameters (Section 4.5), the activity of the forearm muscles during the *plucking* gesture was chosen to manipulate audio and visual effects. Inspired by the initial experiment in Mixed Reality (Section 4.7), EMG data were used to control the gain factor of an amplitude modulation (AM), and the delay time of a delay effect, connected in series with the AM effect, so that audio processing was controlled by the contraction of forearm muscles as a consequence of the *plucking* gesture. Due to the similar audio processing techniques adopted for the initial experiment, this gesture was called the *crumpling* gesture.

With the *crumpling* gesture, the user can also control the brightness of the lighting system, so that it mirrors the sound intensity.

6.5.2.1.2 *Extend* gesture

The *raising* gesture is considered one of the most important gestures in harp performance for its functional and aesthetic aspects. It helps the performer to control the sustain of the sound, and it is partially responsible for the rhythmic and expressive aspects of the performance. However, the sustain of the harp cannot be finely controlled or extended after the string has been plucked. Findings from the interviews with harpists show that this limitation of the instrument can be resolved by adopting guitar pedals, such as the freeze guitar pedal.

Gestures similar to the harpist's *raising* gesture can be observed in the performance of other musical instruments. Davidson (2007) discusses the lifting of the hand after playing a note in piano performance, and the relationship of this vertical movement with the musical expression of the piece. It commonly occurs in moments of the piece where there is "space" (rest and held notes). However, it is also executed at various salient points during a performance. This means that the performer voluntary enacts this movement and then continues supporting the idea that such gestures complement the musical expression. In the case of electronic music, this gesture has already been used to control sound transformation. For example, Nicolls (2010) uses the lifting gesture to elaborate the sound through processes controlled using the Myo; and, in *Kineslimina* by Visi and Coorevits (2015), the violist controls electronic sound transformation, also tracked with a Myo, with a raising gesture after bowing the strings. This interaction design can also be applied to a percussionist, who hits the percussion and then moves her hands away from the instrument, or even to a trombone player using the plunger mute.

The intention here was to augment the performer's level of control over sustain and decay, allowing her to extend the sound of the harp to an unnatural length. A long-tailed reverb is applied to the sound, and the wideness of the rising gesture controls the dry/wet ratio parameter of the reverb. Wideness is represented by the angle formed by the wrist, the performer and the harp in the horizontal axis (Figure 6.8). Thus, the further the hand moves from the body of the harp, the longer and more distant the resulting sound. Inspired by the audio processing characteristics associated with this user interaction, namely to extend the duration of the sound, this gesture was called the *extend* gesture.

When performing the *extend* gesture, the musician can also control the lighting changing cues so that it becomes blue.

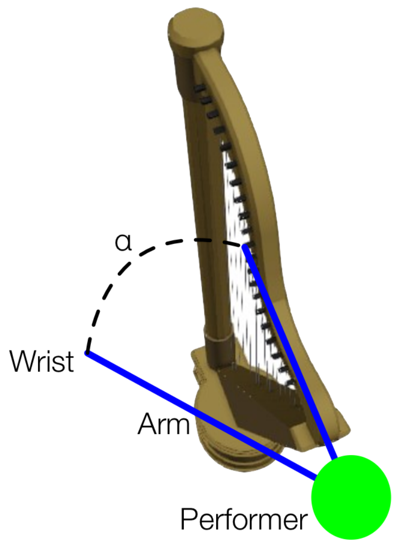


Figure 6.8: Angle (α) formed by the harp and the performer's arm

6.5.2.1.3 ***Lower*** gesture

As argued by O'Modhrain (2011), in order to allow performers to communicate their musical intention successfully through their instrument, its design has to satisfy their needs. Accordingly, the filtering gesture was reconsidered to support a language closer to what the musicians wanted. The filtering that was previously used to isolate and highlight a portion of the audio spectrum was translated into a pitch shift to one octave lower. This choice was also influenced by artistic work in collaboration with Eleanor Turner.

Adopting the same interaction design approach used for the *filtering* gesture, this was redesigned with a downward movement of the arm. This gesture was designed to control the dry/wet ratio parameter of the pitch shifter. Due to the nature of the audio manipulation and the interaction design, this gesture was named the *lower* gesture.

In addition to controlling audio processing parameters, the *lower* gesture allows the musician to change the lighting colour to green synchronously.

6.5.2.1.4 *Throwing* gesture

To broaden sound spatial possibilities and taking into account the successful interaction design implemented in Section 4.8, the *throwing* gesture was here replicated to allow the performer to generate sound spatialisation trajectories. When performing the *throwing* gesture, as well as when changing the spatial position of the sound through the *pointing* gesture, the user can change the lighting brightness as described in Section 5.5.2.3. The variation in the lighting intensity aims to represent the spatial position of the sound visually.

6.5.2.1.5 *Clean* gesture

Considering the potential need for playing a musical part without digitally processing the sound, a gesture to obtain such a sound was designed. This consisted of moving the arm frontwards or closer to the chest. To create a strong analogy with the absence of an audio effect, the corresponding lighting colour is white. Adopting the taxonomy commonly used in the field of audio processing, this gesture was called *clean* gesture.

Figure 6.9 summarises the arm movements required to process the audio and lighting signals.

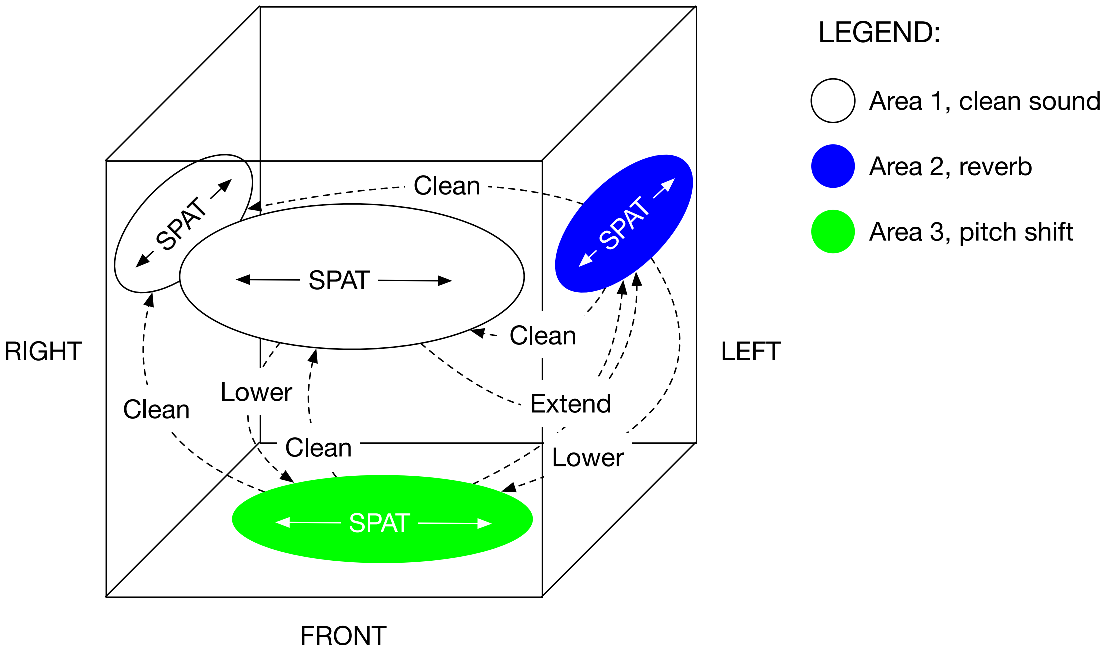


Figure 6.9: MyoSpat 0.6 interaction design

6.5.2.2 Audio-visual engine

With regard to *The Wood and the Water* and *Star Cluster*, the importance of the decision taken by both composers to use sign language to express their artistic ideas should be recognised. Contemporary to this work, Abreu et al. (2016) confirmed the potential use of EMG data and the Myo device as a method for recognising sign language. However, the primary goal of this user study and MyoSpat 0.6 was to include in the interaction design harpists' gestures analysed in Section 6.4 and to make possible for performers of other instruments to interact with MyoSpat using the same gestures (the MyoSpat 0.6 interaction design is outlined in Section 6.5.2.1). For this reason, the specific gestures associated with sign language are not included in the discussion.

On the contrary, several audio effects such as pitch shift, reverb, and amplitude modulation chained to a delay were implemented to answer the artistic needs of the composer and performer.

The audio and lighting engine, implemented using Pure Data (Pd), consists of a data receiver, two signal routers (a and b), four audio processors (reverb, pitch shifter, AM+Delay effect and spatialiser), a generator of sound spatialisation trajectories, three mapping processes (a, b and c) and a serial port. The architecture of the system is outlined in Figure 6.10. Taking into consideration the technical details and their effect on the interaction with the system, the use of visual projections was discarded in this and later iterations of the system. Only the lighting system was used to provide visual feedback.

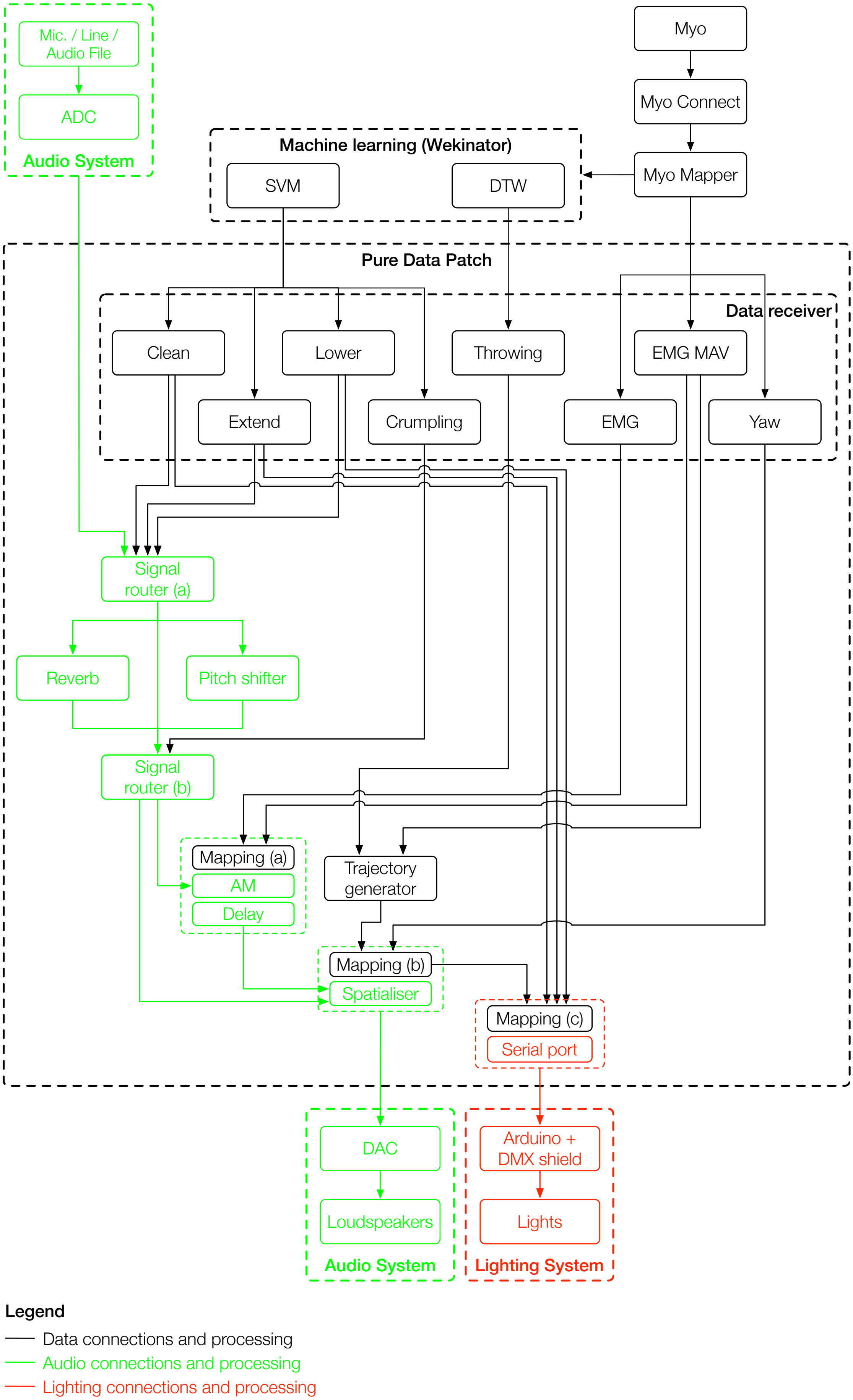


Figure 6.10: MyoSpat 0.6 architecture

MyoSpat features two audio signal routers in series. The first router, signal router *a*, controls the level of the dry, reverberated and pitch-shifted signal through mapping the gesture-recognition algorithm output relative to the *clean*, *extend* and the *lower* gesture. The audio signal is sent to the signal router *b*, which controls the amount of signal to be processed with an amplitude modulator and delay line when the *crumpling* gesture is performed. Finally, the audio signal is sent to the spatialiser.

The pitch shifter is an implementation of Integra Live's Pitch Shifter module (Integra Lab, 2016a). Based on the auditory perception of the audio effect relative to the user's interaction with the system during informal tests and rehearsals with the harpist Eleanor Turner, the pitch parameter was set to a fixed value of one octave lower.

The reverb is an implementation of Integra Live's Reverb module (Integra Lab, 2016b). As for the pitch shifter, the reverb settings were left static. After informal tests, these settings fit the creative requirements of the composer/performer. These are summarised in Table 6.1.

Table 6.1: Reverb effect settings

<table>
<tr><th colspan="2">Parameter</th><th>Early Echoes</th><th>Late Diffusion Reverb</th></tr>
<tr><td colspan="2">Level (dB)</td><td>-3</td><td>0</td></tr>
<tr><td colspan="2">Cut-off Frequency (Hz)</td><td>15000</td><td>5266</td></tr>
<tr><td colspan="2">Min Delay Time (ms)</td><td>20</td><td>25</td></tr>
<tr><td colspan="2">Max Delay Time (ms)</td><td>75</td><td>110</td></tr>
<tr><td colspan="2">Min Damping (dB)</td><td>-6</td><td>-6</td></tr>
<tr><td colspan="2">Max Damping (dB)</td><td>-2</td><td>-2</td></tr>
<tr><td rowspan="6">Decay Time</td><td>Range</td><td colspan="2">Milliseconds</td></tr>
<tr><td>High</td><td colspan="2">20</td></tr>
<tr><td>Mid High</td><td colspan="2">18</td></tr>
<tr><td>Mid</td><td colspan="2">16</td></tr>
<tr><td>Mid Low</td><td colspan="2">14</td></tr>
<tr><td>Low</td><td colspan="2">12</td></tr>
</table>

MyoSpat 0.6 included the same spatialiser as in MyoSpat 0.5, described in Section 5.5.2.2, but this time implemented using Pure Data. The choice of using a free and

open-source interactive audio software, such as Pure Data, was made in order to give access to MyoSpat to a wider user base, considering affordability issues.

The AM+Delay effect is composed of four amplitude modulators chained to a delay line. Each of the four audio processors routes the outgoing audio signal to the four channels of the spatialiser (Figure 6.11).

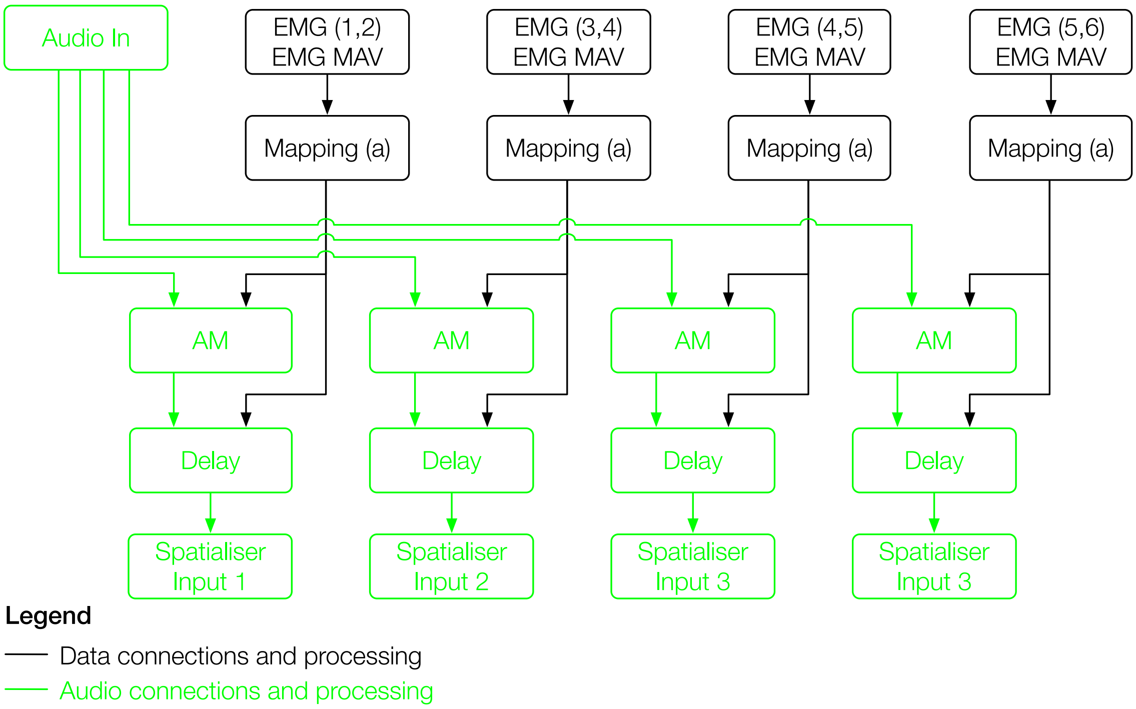


Figure 6.11: MyoSpat 0.6 AM+Delay effect architecture

Data from "Mapping (c)" (Figure 6.10) were converted first into serial messages and then into a DMX signal for the lighting system; as in previous iteration of the system (Section 5.5.2.2).

6.5.2.3 Mapping strategy

In previous case studies, ml.lib and Wekinator were successfully used to map Myo data to audio and visual parameters. From previous experiences using both ml.lib and Wekinator, the latter offered a higher degree of freedom in building and editing a model. Thus, Wekinator and many-to-many mapping strategies were used to map gestural data into sound and lighting processing parameters.

From an observation of the *clean*, *extend* and *lower* gestures, it emerged that the beginning and the end of each gesture could have been considered as the poses for training the machine learning algorithm. These poses would consist of keeping the orientation of the arm towards one of the areas of interaction outlined in Figure 6.9.

Once the model is trained, the output of the algorithm can then be used to control the amount of signal to process through the signal router *a*.
To solve this pose-recognition problem, a Multilayer Perceptron (MLP) regression algorithm was utilised. MLP is a feedforward artificial neural network model that maps input data to output data in a linear or non-linear way and can be used for both classification and regression (Witten and Frank, 2005). An MLP includes multiple layers of nodes, with each node fully connected to the next one (Figure 6.12).

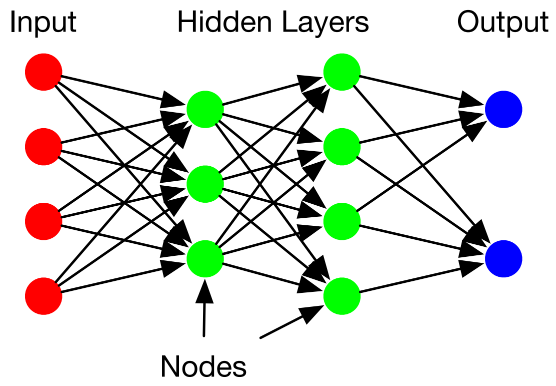


Figure 6.12: Multilayer Perceptron (MLP)

Continuous values from the MLP regression algorithm relative to the different models were directly mapped into parameters for controlling the signal router *a*. The model was trained using the yaw, pitch and EMG MAV features from Myo Mapper as an input feature vector; the training data set was composed of 3,327 samples. Models used for the *clean*, *extend*, *lower* and *crumpling* gestures were built using two, two, one and one hidden layers respectively, with each hidden layer composed of three nodes. This number of layers was chosen as it was the minimum number of layers that allowed to obtain a model with acceptable accuracy, and the number of nodes as by default in the Wekinator software. Here the degree to which a model is acceptable was established by the performer during rehearsal. This choice was also informed by a detailed 10-fold cross-validation carried out using Weka Explorer 3.8.1 (Witten et al., 2014). This reported RMS error of 0.0019 for the *clean* gesture, 0.0048 for the *extend* gesture, 0.0038 for the *lower* gesture and 0.1217 for the *crumpling* gesture. Detailed data of the model's evaluation are shown in Table D.6 in Appendix D.1.

No further mapping strategies were implemented for controlling the reverb and pitch shift, but linear mapping strategies were used to control the AM+Delay effect. EMG data were linearly mapped into amplitude modulation's gain control values $[0, 1]$, delay's time in milliseconds $[0, 4000]$ and feedback $[0, 1]$.

The *crumpling* gesture was recognised by distinguishing two different levels of muscles contraction: low and high muscle contraction. These two poses were then used to control the signal router *b*.

Sound spatialisation trajectories were generated by the *trajectory generator*. This is a Pure Data patch for overriding the spatialisation parameters using a quasi-random function, called each time the *throwing* gesture occurs. This gesture was recognised using a Dynamic Time Warping (DTW) algorithm. DTW is an algorithm that attempts to find the best alignment between two time-dependent data series (Figure 6.13)(Müller, 2007). This is obtained by first calculating the *local cost measure* of each sample of the data series (Figure 6.14). Afterwards, it is possible to establish the degree of match between the data series by calculating an overall value of the cost. The lower the cost, the more probable that the data series are similar.

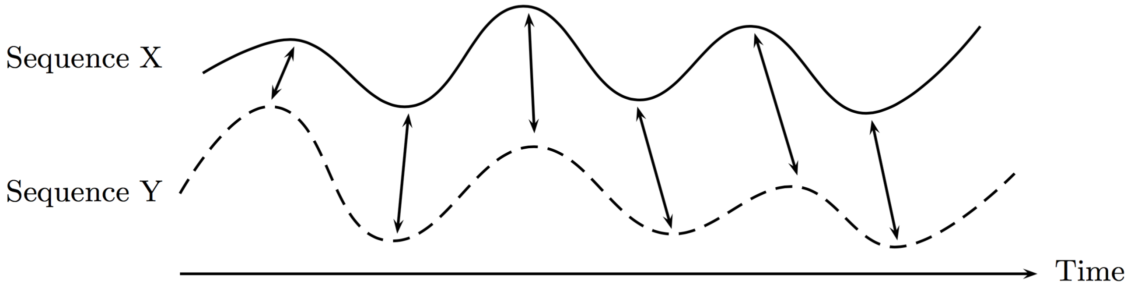


Figure 6.13: Time alignment of two time-dependent sequences. Aligned points are indicated by the arrows (Müller, 2007)

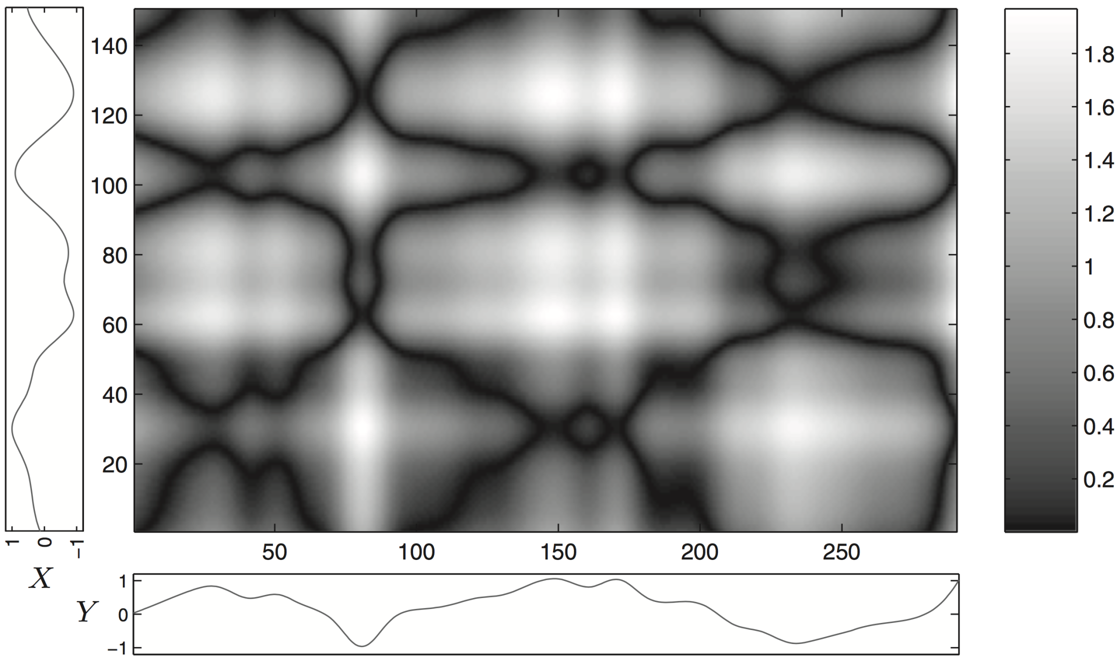


Figure 6.14: Cost matrix of the two real-valued sequences X (vertical axis) and Y (horizontal axis) using the Manhattan distance (absolute value of the difference) as local cost measure c. Regions of low cost are indicated by dark tones and regions of high cost are indicated by light tones (Müller, 2007)

A DTW was used through the second instance of Wekinator, as it is not possible to run a DTW and an MLP in the same instance. The model was trained with fifty samples of *throwing* gestures and five samples of the *clean*, *extend* and *lower* gestures using the Gyro ABS values as feature vector data. The DTW, evaluated through direct evaluation (Fiebrink et al., 2011), performed $\approx 95\%$ accuracy. This evaluation was made through informal trials by the author, in which the *throwing* gesture was recognised nineteen times out of twenty. A second evaluation, in which trials were performed by the harpist Eleanor Turner, reported accuracy of $\approx 85\%$ (seventeen out of twenty recognised trials).

The following mapping strategy was adopted to relate the complexity and duration of the sound trajectories to the applied force in performing the *throwing* gesture. The EMG MAV value, at the moment when the gesture was detected, was mapped linearly to several trajectory segments, whose origin was the target value of the previous one, and its target value was randomly generated (Figure 6.15). The first segment has its origin at the virtual source position at the moment when a *throwing* gesture is detected. The duration of each segment is established through linear mapping of the EMG MAV to create segments with a maximum duration of one second each.

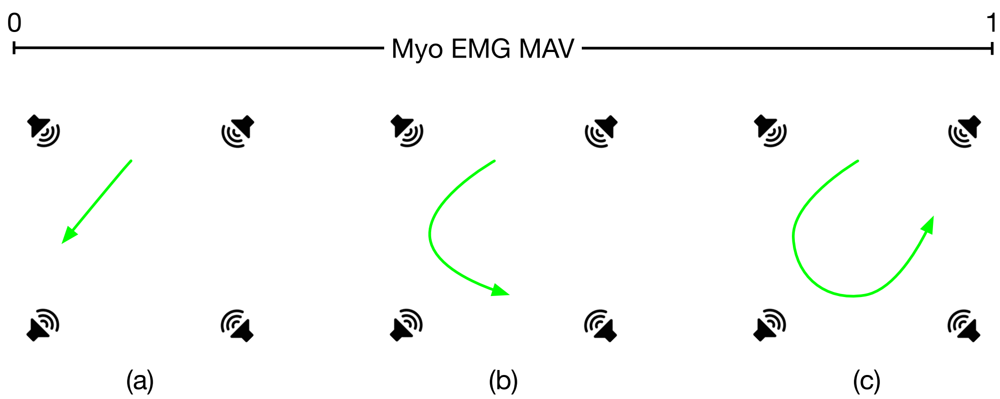


Figure 6.15: Examples of automated spatial trajectories generated by the *throwing* gesture using MyoSpat 0.6

The spatialiser's gain factors $[0, 1]$ were mapped linearly to brightness parameters $[0, 255]$ so that each light projector was strongly related with the sound coming out from its associated loudspeaker. The colour of the lights was calculated by mapping MLP output into colour-changing cues. To be specific, the *clean*, *extend* and *lower* gestures were mapped to a range $[0, 255]$ for controlling the white (W), green (G) and blue (B) components of the four lights respectively.

No synesthetic study was conducted to establish the gesture-colour and gesture-brightness mapping strategy. The colour mapping was established in collaboration with the harpist Eleanor Turner through an experimental approach.

6.5.2.4 System set-up

MyoSpat 0.6's system was composed of a quadraphonic audio system of full-range loudspeakers and the same lighting system used for MyoSpat 0.5. The system was installed in different configurations in relation to its purpose: performance, rehearsals, user study and workshops.

The performance set-up in Figure 6.16, first designed and used for the premiere of *The Wood and the Water* by Eleanor Turner, aimed to put the audience in an optimal position to appreciate the performance. Here the audio system is arranged around the audience and in a semicircle configuration in front of the performer. To allow the performer to appreciate spatial cues in the sound, the set-up includes a minimum of two wedge monitors in front of her.

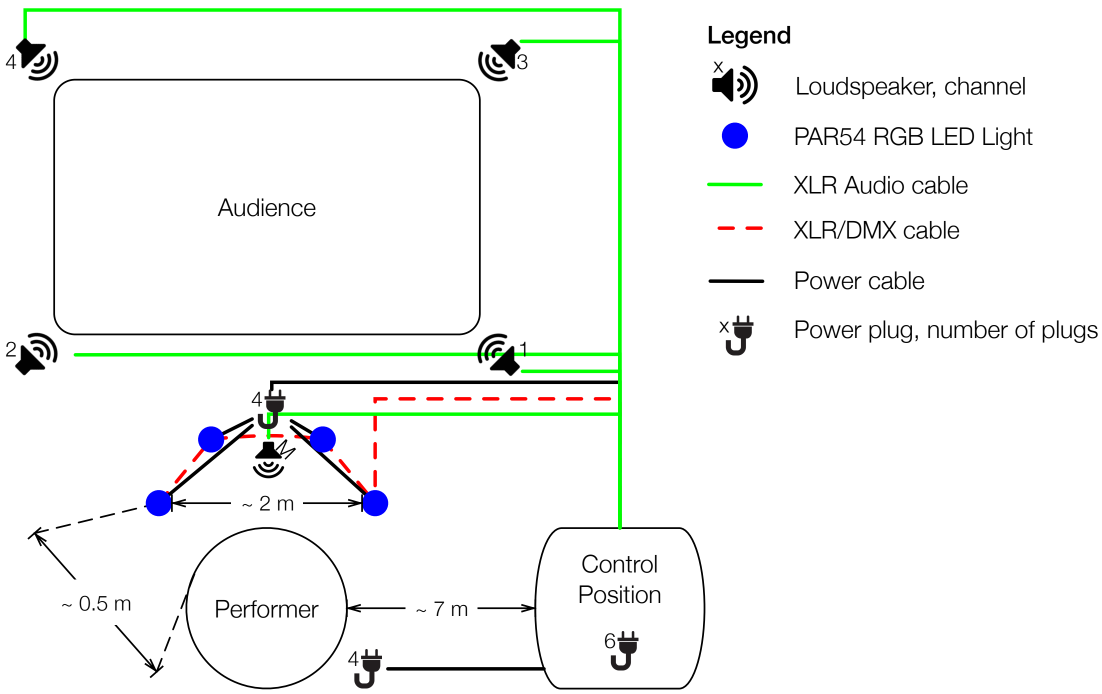


Figure 6.16: MyoSpat 0.6 performance set-up

The rehearsal set-up (Figure 6.17) aims to offer optimal conditions during rehearsal. It allows the performer to have the same auditory experience as the audience during a performance. By enabling to build a mental image of the sound they project to the audience, this set up helps the performer adjusting the instrumental technique and the interaction with the system in relation to their objectives. With this in mind, a quadraphonic audio system is set up around the performing space and the lighting system in a semicircle in front of the performer.

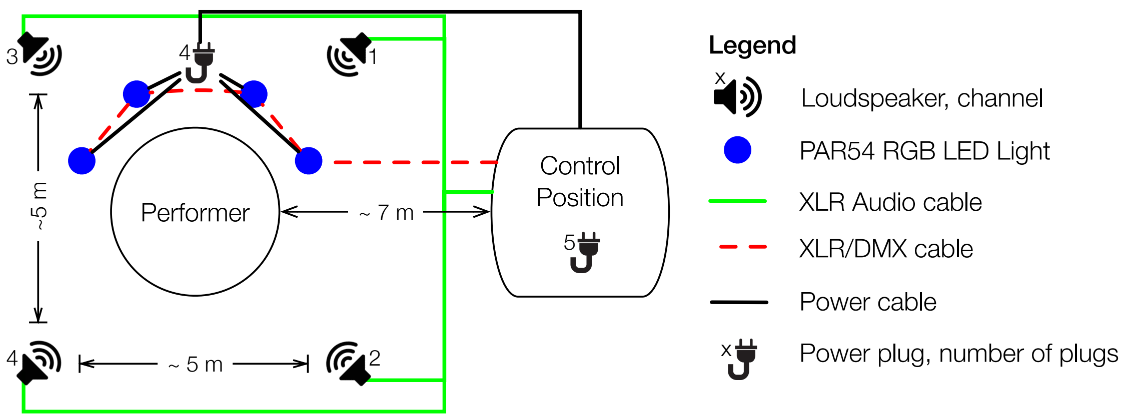


Figure 6.17: MyoSpat 0.6 rehearsal set-up

The user study configuration outlined in Figure 6.18, aims to set the optimal condition for user study participants to explore the MyoSpat system, to test its limits, and to appreciate the relationship between the gestural interaction and the audio-visual feedback.

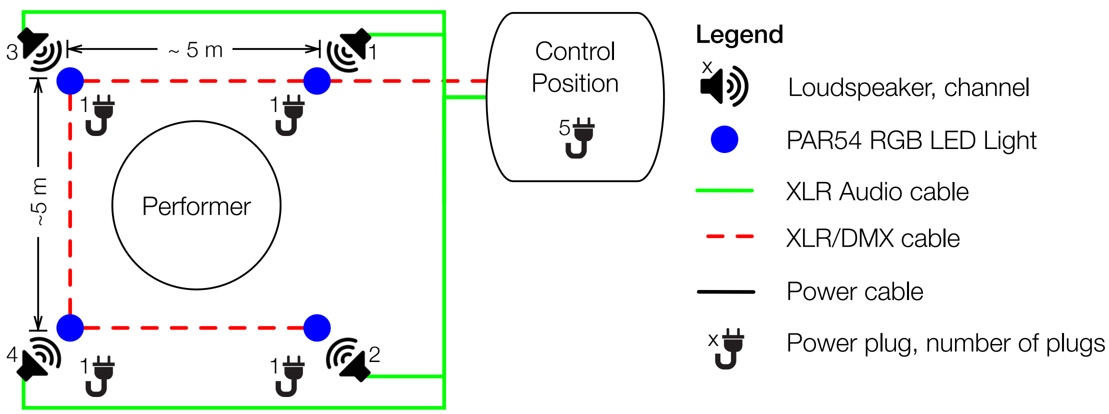


Figure 6.18: MyoSpat 0.6 user study set-up

The workshop configuration mirrors the performance set-up with the addition of a projector screen or video monitor (Figure 6.19).

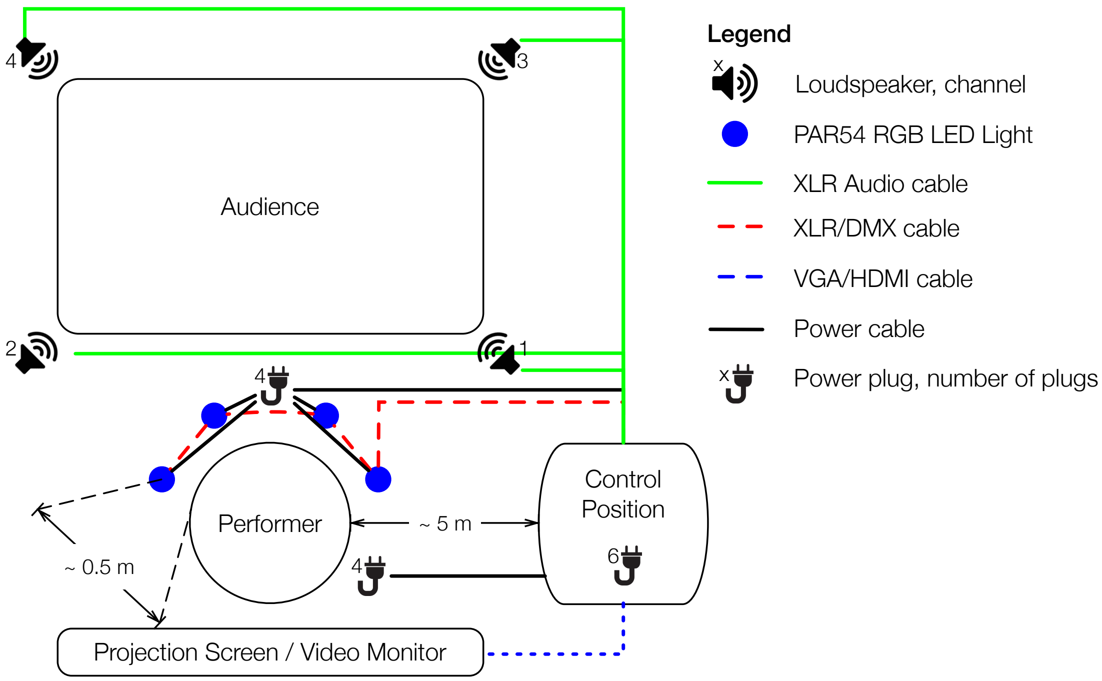


Figure 6.19: MyoSpat 0.6 workshop set-up

6.5.3 User Study

MyoSpat 0.6 system was evaluated through a user study conducted at Royal Birmingham Conservatoire, adopting the methodology described earlier in this chapter (Section 6.2). The results of each of the analyses are presented, followed by a discussion of the different outcomes.

6.5.3.1 Participants

Eleven participants took part in the user study: one engineer, one dancer and eight musicians, of which two guitarists, two keyboard players, a harpist, a percussionist, a saxophonist and a singer. All musicians completed the user study using MyoSpat while playing their musical instrument. Nine participants were aware of the concept of audio processing, particularly sound spatialisation. After explaining what sound spatialisation is, all participants confirmed having experienced it in cinemas. Nine participants had experienced sound spatialisation in music concerts through a multichannel diffusion. Only four participants had direct experience of sound spatialisation through music production, compositional works or sound art installations. Only one participant had never used interactive gestural controllers. Three participants used such devices in gaming applications, five in musical performance only and two in both musical performance and digital arts (visual art, sound design, sound art installations). Eight participants already knew the Myo. However, only three had used it before the user study. Data are resumed in Figure 6.20 and outlined in detail in Table D.1 and D.2 in Appendix D.1.

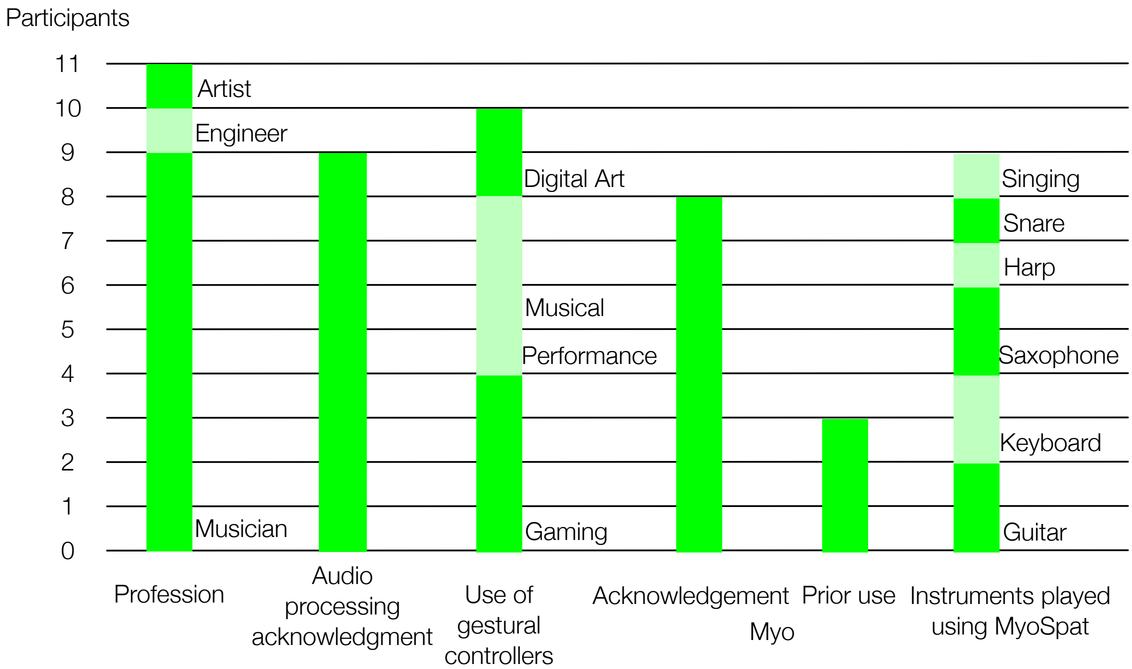


Figure 6.20: MyoSpat 0.6 user study Participant information

6.5.3.2 Results

6.5.3.2.1 Task-based evaluation

The task-based evaluation lasted an average of 39'13" (min 24'59", max 62'32"). Participants spent an average of 12'31" (min 7'48", max 20'39") manipulating the pre-recorded sound file, and an average of 18'51" (min 10'30", max 20'39") using MyoSpat while playing their musical instrument. Before starting the task-based evaluation, they spent an average of 4’54” (min 2’30”, max 10’18”) learning and practising with the system. Detailed data are outlined in Table D.3 in Appendix D.1.

Comparing the gesture-recognition performance’s data with the participants’ perception of the audio-visual feedback (Table 6.2 and Figure 6.21), it emerged that although the gesture-recognition algorithms were accurate for 94% of the time, only 88% of the participants perceived the audio-visual feedback as expected. Specifically, the MLP recognised the *lower*, *extend*, *clean* and *pointing* gestures 100% of the time; and 95% of the time the *crumpling* gesture. With the *throwing* gesture, the model performed with 75% accuracy. The participants’ perception of the audio feedback scored an average of 5% less than the gesture-recognition algorithm. From Figure 6.21, it is possible to compare both sets and conclude that the *clean* and *pointing* (100%), *lower* (99%) and *extend* (98%) were the gestures recognised with the higher accuracy, against the *crumpling* (91%) and finally the *throwing* gesture (68%). Detailed data are shown in Table 6.2.

Table 6.2: MyoSpat 0.6 task-based evaluation result

Evaluation	**Gesture**						
	Clean	**Lower**	**Extend**	**Crumpling**	**Throwing**	**Pointing**	**AVG**
Gesture Recognition Accuracy (%)	100	100	100	95	75	100	94
Participants’ Perception Response (%)	100	98	96	87	62	100	89
AVG	100	99	98	91	68	100	91

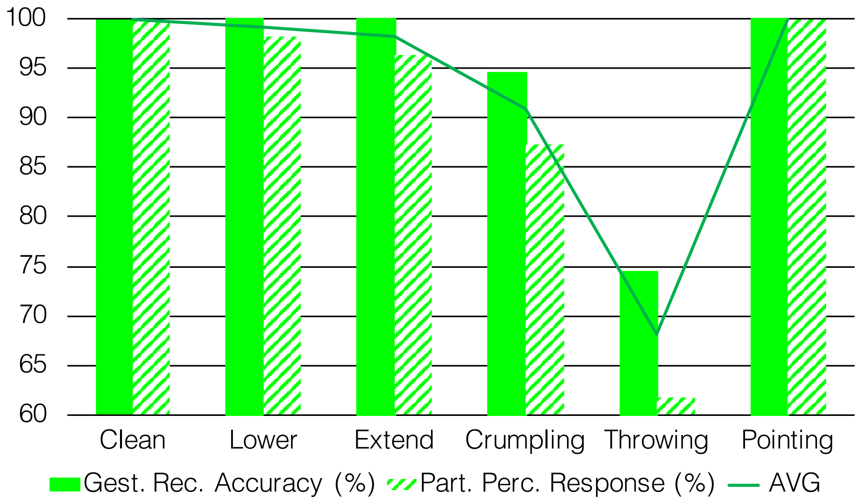


Figure 6.21: MyoSpat 0.6 task-based evaluation result

6.5.3.2.2 Semi-structured interview

The Myo was seen as a comfortable device to wear and perform with only two of the participants communicating discomfort. Although the Myo is entirely built with non-allergic materials, a participant reported a feeling of itchiness; and another one commented that the Myo caused tiredness. Myo can be worn on either arm when not playing a musical instrument. However, to take full advantage of the MyoSpat's interaction design in performance, it is better to wear the Myo on the same arm whose hand is most engaged in playing the musical instrument. All participants started using MyoSpat wearing the Myo on the left arm. Six participants felt the necessity to swap and wear it on the right arm. In all cases, this decision was due to the instrumental technique. After swapping, four participants preferred to wear the Myo on their left arm, four on both arms and three of them on the right arm. Interestingly, two participants, both guitarists, decided to wear the Myo on two different arms. Participant 1 was more comfortable with the armband on the left arm because he felt the need to have a higher degree of freedom with the right hand when plucking the strings. Participant 9 instead commented that wearing the Myo on his left arm was less convenient. As he was playing bar chords, this would inevitably tense the muscles and so trigger the AM+Delay effect even when this was not desirable (see video recording no. 1 in Appendix F.4).

MyoSpat 0.6 was successfully used by singers or performers playing very different instruments, such as snare drum, guitar, and saxophone. Six musicians interacted with the system by taking the hands off the instrument. Two keyboardists could not trigger the reverb effect comfortably due to the instrumental technique. After an average of five minutes of practice, six instrumentalists found custom ways to trigger effects while playing. The rest needed an average of ten minutes. Importantly, the unconscious interaction with the system was commented as natural and embodied. This behaviour was first noticed by Participant 7 (keyboard player), who obtained a tremolo effect on the keyboard's sound by modifying the hands' pressure on the keyboard keys after hitting them (see video recording no. 2 in Appendix F.4). The AM+Delay effect was used by three musicians to create more interesting rhythmical patterns (see video recording no. 3 in Appendix F.4) and one used it to overlap different sounds (see video recording no. 4 in Appendix F.4). Seven participants used the spatialiser delay to bend the pitch of the musical notes.

The interaction-feedback relationship was very strong for five participants. Among all gestures, five participants judged the *lower* gesture as having the strongest relationship with the auditory and visual feedback, followed by the *crumpling* gesture (four participants) and the *pointing* gesture (two participants). Interestingly, participants took advantage of the pitch shifter to process the pre-recorded sound file of a flowing river through gestures that mimic the interaction with water in a bathtub or a big bowl (see video recording no. 5 in Appendix F.4). All participants defined the *throwing* gestures as having the weakest relationship with the audio-visual feedback.

All participants confirmed that the visual feedback enhanced both user interaction and the level of immersion, and at the same time, increased their awareness. However, all participants suggested a more gradual transition between the colours relative to each audio effect. Three participants commented that the visual feedback was an element of distraction from the auditory perception. The visual feedback helped six participants to understand automated trajectories better; one participant used it as a guideline for navigating through the system; six participants suggested a more noticeable visual elaboration. Among all lighting effects, four participants appreciated the one triggered by performing the *throwing* and *crumpling* gesture. Interestingly, Participant 6 suggested the use of red, green and blue to relate to the different areas.

Seven participants interacted with the system through non-predefined gestures. Two triggered the pitch shifter by standing up while playing their musical instrument (see video recording no. 6 in Appendix F.4). One tried to trigger the pitch shifter by changing the body posture (see video recording no. 7 in Appendix F.4). Participant 7 sets off the pitch shifter gesture by raising the shoulder and straightening his arm when playing (see video recording no. 8 in Appendix F.4). The same effect was triggered by guitarists (Participant 9 and 7) by tilting their body towards their left; such behaviour was noticed in two different cases, where the Myo was worn on different arms (Figure 6.22 and 6.23). Participant 11 triggered the pitch shifter by lifting the elbow. Three of the participants panned the sound purposely through hand gestures only by taking advantage of the spatial cues generated through the crumpling effect. Participants sent the delayed signal towards a specific loudspeaker. Five musicians used the delay effect to simulate a tremolo through the AM+Delay effect. MyoSpat was also used by Participant 5 to sonify dance moves through sudden contractions of muscles, also known as "locking" or "popping" technique in Hip Hop dance (Engel, 2001) (see video recording no. 9 in Appendix F.4).

Figure 6.22: Participant 1 tilting his body to the left to trigger the pitch shifter

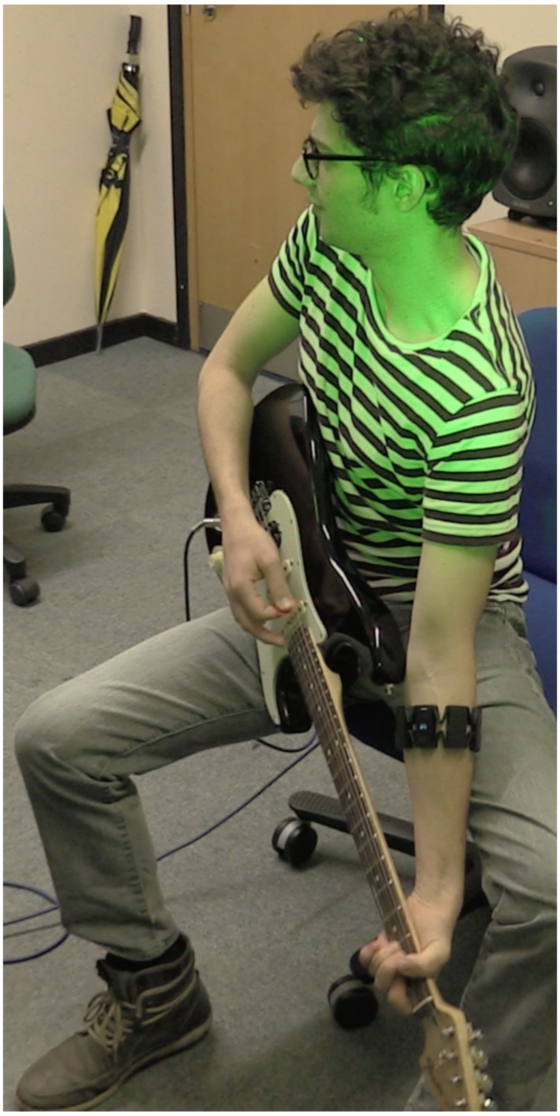


Figure 6.23: Participant 9 tilting his body to the left to trigger the pitch shifter

Most of the participants suggested a more extensive palette of gestures, and one suggested a different way to implement the *extend* gesture, as it imposes a spatialisation effect on the sound. When feeding the system with the audio file, two participants suggested improving the AM+Delay effect control and audio processing, as they could not clearly perceive when the sound was no longer processed. However, this reaction by the participant was observed only with randomly dynamic sounds characterised by a broad spectrum, such as the water sound recording of the user study. When the processed sound was constant and spectrally defined the participants felt in control of the effect.

Most participants suggested improvements to the audio processing, for example, to make panning effects more evident and sound trajectories generated through the throwing gesture clearer. Three participants suggested having continuous control of the pitch shift's frequency by moving the arm on the vertical axis. Finally, while most of the participants suggested a smoother transition between audio-visual effects, four defined the pitch shifter as very precise and discrete.

6.5.3.2.3 User Experience Questionnaire (UEQ)

The results of the UEQ, summarised in Table 6.3 and Figure 6.24, show that the user experience was positive in all UEQ categories. The lowest evaluation was made for perspicuity and dependability, which respectively scored 1.909 and 1.864. However, answers related to these two categories were slightly inconsistent. All the other categories scored between 2 and 3. Results show that the best MyoSpat quality is stimulation (2.682), followed by attractiveness (2.621), novelty (2.591) and efficiency (2.114). The scales of the UEQ grouped into pragmatic quality (perspicuity, efficiency, dependability) and hedonic quality (stimulation, originality), confirm that the user had a better experience when free to explore the system, rather than when performing specific tasks (Table 6.4 and Figure 6.25)

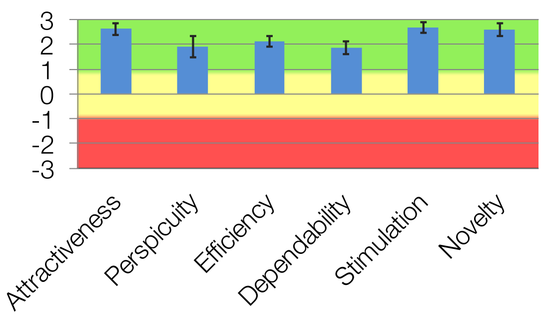


Figure 6.24: UEQ scales, MyoSpat 0.6 evaluation

Table 6.3: UEQ scales, MyoSpat 0.6 evaluation

UEQ scales	
Attractiveness	2.621
Perspicuity	1.909
Efficiency	2.114
Dependability	1.864
Stimulation	2.682
Novelty	2.591

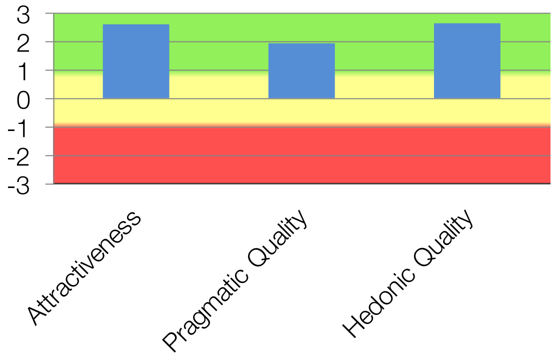


Figure 6.25: UEQ grouped scales, MyoSpat 0.6 evaluation

Table 6.4: UEQ grouped scales, MyoSpat 0.6 evaluation

Pragmatic and Hedonic Quality	
Attractiveness	2.62
Pragmatic quality	1.96
Hedonic quality	2.64

6.5.3.3 Discussion

The interaction design was described as easily learnable by all participants, and most of them found it adaptable to their instrumental technique with some practice. The participants felt confident using the system after an average of 4'54" (min 2'30", max 10'18"). Considering that only three participants had used the system before, they learned to control audio-lighting effects in a relatively short time.

The input device was comfortable to wear in different situations. Thanks to its flexibility and Myo Mapper's easy calibration process, MyoSpat was used in performance by six different instrumentalists. Ordered from the most performable to the least performable: voice, saxophone, guitar, snare drum, harp, keyboard.

Reflecting the varying degrees of freedom of interaction that the system affords, participants took different approaches when interacting with it. They also manipulated the sound through gestures and body postures not included in the system's interaction design. This aspect seems to confirm that the system is open to exploration and leaves the musicians to adapt it to their technique. It is worth noting that five participants responded that the system was limited and that it might impose body postures and instrumental techniques potentially detrimental to the musical performance. However, the same participants did not interpret the system limitations as constraints, but as barriers and rules to break and be creative with. Particularly, Participant 11 (saxophone player) said:

> [...] when the fingers are moving, it causes certain things to happen, and I'm one who doesn't move much, so in this case, the system imposes a performing style that you don't necessarily want but exploring the three spaces is very interesting. However, I would like to have a smoother transition between the effects, for the reason why I like to get to the spaces in between. Because the three spaces are so well defined, apart from the drop in pitch, the reverb, the delay and the trajectories there is not much. However, it's a good sign because you can easily understand and control what you want to do. But by the way, I perform, the places in between would be the interesting spaces that I want to get to, but I can't.
>
> [...] My approach to electronic systems is to have something very limited, which then gets incorporated with other things. The saxophone is very limited, it's a metal tube, and actually, it does one thing, part of the fun is to find out where the boundaries are. With three effects it's enough, there is interesting stuff there, and you can also start to test where the edges are, where breaking points are, to test the misbehaviour.
>
> Participant 11

From this case study, it emerged that lighting effects enhance the level of immersion and raise the accuracy of the interaction with the system. Similar results were also found in the previous case study (Chapter 5). The lighting system also guides users through the exploration of the system. The downside is that visual feedback has the potential to make the user unfocused towards the auditory one.

Quantitative analysis of the task-based evaluation showed consistent behaviour of the system at each user interaction. The system performed positively 93.5% of the time and the auditory feedback met the participants' expectations 86.9% of the time.

Most participants first interacted with the sound by engaging the forearm muscles through the *crumpling* gesture. By performing the *lower* gesture, participants changed the pitch of the water one octave lower, so that it sounded like an underwater current. In this situation, participants interacted with the sound as if they were immersing their hand, splashing or moving water in a bathtub or a bowl in the real world (see video recording no. 5 in Appendix F.4). Similar sonic interaction design has been investigated previously by Boyer et al. (2015), who created a system that allows interacting with a virtual water surface through hand gestures, and generated similar audio feedback as if touching a real water surface. The Leap Motion's hand-tracking capabilities allowed Boyer et al. to successfully explore the perception of audio virtual surfaces (AVSs) through hand gestures. Their observations confirm the validity of their concept of sound-oriented tasks, through which they show the potential to place the sound at the centre of the interaction within the sensorimotor loop.

Participants also took advantage of the AM+Delay effects, triggered by engaging the forearm muscles and the spatialiser's delay line, through the *pointing* gesture, to obtain a tremolo and vibrato effect of the sound. In these cases, they obtained such audio feedback by performing the same gesture of a string or guitar player when attempting to play vibrato (see video recording no. 2 in Appendix F.4). Interestingly, this type of interaction with similar audio effects is also embedded in different devices created concurrently with this research work, such as the Roli's Seaboard keyboard, which allows the user to control audio parameters through the fingers' pressure on the keys (TechCrunch, 2014), the Enhancia, a IMU-based device (Enhancia, 2018) and Touché, a touch-sensitive gestural MIDI controller (Expressive E, 2017) or the Ondes Martenot as discussed in Chapter 2. The same gesture was performed not only by guitarists but also by keyboard and saxophone players.

The *lower* and *clean* gestures were described by all performers as easy to learn and perform, intuitive, natural and highly related to the auditory feedback. Such movements were easily adapted to different instrumental techniques. The *extend* and *throwing* gestures were considered as the least related to the audio and visual feedback. Although the reverb effect was performed with an adaptation of the *extend* gesture to the instrumental technique, it did not require any particular effort by most participants, and it was easily recognised by the system. The *throwing* gesture was described as having a strong conceptual relationship with the associated audio feedback. Nevertheless, when participants performed the *throwing* gesture, the gesture recognition algorithm had an accuracy of 74%, and the audio feedback met the expectations of the participants, only 62% of the time. The *throwing* gesture was recognised at each attempt only with Participant 11, who also perceived every time the sound spatialisation trajectory (see video recording no. 10 in Appendix F.4). Due to gesture-recognition accuracy and the implemented spatialiser, the relationship between the *throwing* gesture and its related audio feedback was considered the weakest. The *crumpling* gesture, on the other hand, was detected 94% of the time and the related audio feedback was perceived clearly by the participants 87% of the time. All other effects were described as having the strongest relationship with the triggering gestures.

When using MyoSpat with the support of visual feedback, participants commented that the relationship user interaction-feedback was more robust and coherent. They felt more conscious and in control of the system, immersed within the multimedia environment and guided by the system itself.

MyoSpat was considered by all performers as a potentially creative tool for musical performance. In fact, each of the effects made performers go beyond the sonic limitations imposed by the physicality of their instrument and technique. MyoSpat gave performers the possibility to produce new sounds through the direct interaction with the instrument, in addition to the interaction with the system's input device. The pitch shift of one octave lower allowed them to extend the harmonic dimension of their music. The AM+Delay effect was employed for the production of more interesting rhythmical patterns and tremolo effects. The reverb extended the sound resonance of each produced sound, while the delay time of the spatialiser's delay line was used to generate vibrato effects.

6.5.4 Workshop

Workshops were delivered at Royal Birmingham Conservatoire, Cardiff Metropolitan University for the Camac Harp Weekend 2017, and at Southampton University, adopting the methodology described earlier in this chapter (Section 6.2). The results of each of the analyses are presented, followed by a discussion of the different outcomes.

6.5.4.1 Participants

Compressively, twenty-six harp students with age between fifteen and forty-five (AVG 21.6) participated at the three HarpCI workshops. Five of them had prior experience in processing harp sound using guitar pedal effects, yet none of them had any experience with wearable technology, nor controlling audio-visual feedback through gesture sensing devices.

6.5.4.2 Results

6.5.4.2.1 Observations

Participants to the workshops confirmed that MyoSpat is easily learnable and applicable to most harp techniques. However, the *crumpling* and *throwing* gestures were found to be relatively difficult to adapt to the participants' playing technique. In particular, performers were not able to finely control the contraction of the forearm muscles and the resulting audio effect. Interestingly, the limitations imposed by MyoSpat's interaction design on the one hand, and the harp characteristics on the other, pushed performers to manipulate the sound by taking advantage of the harp geometries, for instance, contracting forearm muscles as a consequence of the pressure imposed by the hand on the body of the harp.

Participants felt all gestures were easy to perform while playing the harp. Most harpists highlighted that when they performed the crumpling gesture, they felt like playing their own body in conjunction with the harp.

> This makes me more conscious of my instrumental technique.
>
> Workshop attendee (Birmingham, 2017)

> This armband gives me the possibility to feel the sound through my muscles.
>
> Workshop attendee (Cardiff, 2017)

> I feel like playing my body, like if I could control the harp sound through my inner self. This makes me have a new intimate relationship with the instrument and the music.
>
> Workshop attendee (Southampton, 2017)

6.5.4.2.2 User Experience Questionnaire (UEQ)

Results from the UEQ, outlined in Figure 6.26, showed that the overall user experience among workshop attendees was positive. They had an excellent general impression of the system and its usability while playing the harp. The system was considered interesting and exciting, as well as being a novelty to performers. However, the data also shows that the system was not very clear to attendees and that they did not feel in complete control. It must be noted that these last observations from the questionnaire may not be valid, as the values of perspicuity and dependability were inconsistent with the dataset.

Results grouped by pragmatic and hedonic quality showed in Figure 6.27, confirm the behaviour of the workshops' attendees. Results from the UEQ show that the pragmatic quality of the system was lower than the hedonic quality. In other words, the user experience was not as enjoyable during task performance as when improvising using the system.

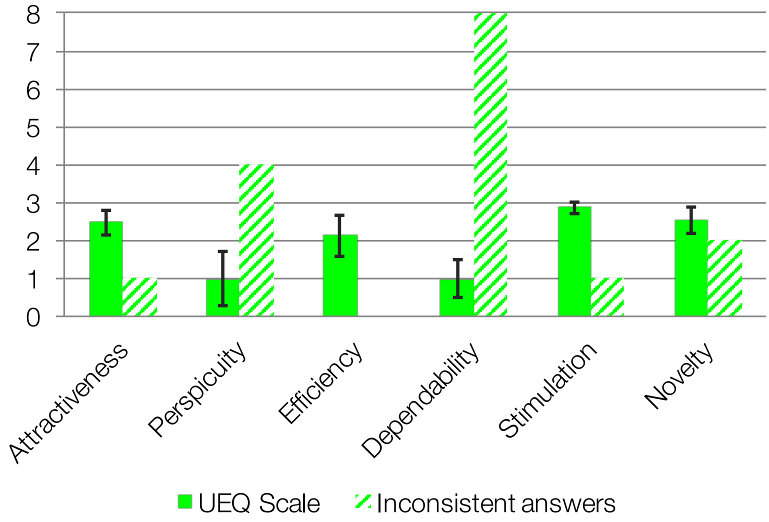


Figure 6.26: MyoSpat 0.6 evaluation among harpists, UEQ scales and inconsistent answers count

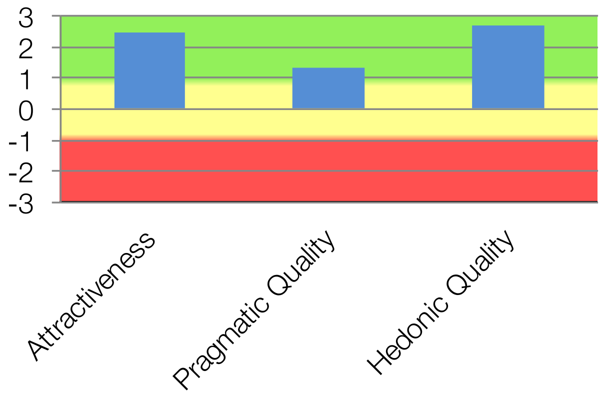


Figure 6.27: MyoSpat 0.6's evaluation among harpists, UEQ pragmatic and hedonic quality

6.5.4.3 Discussion

In taking advantage of the body of the harp for controlling audio parameters, participants exhibited different behaviours compared with Eleanor Turner. While they used the harp body to help them tense their muscles, Eleanor Turner found solutions for spatialising the harp sound in any direction by moving the arm above the body of the instrument (see Figure 6.28). This unexpected behaviour of the participants can allude to the fact that modes of interaction implemented in MyoSpat inspired participants to perform as if their limbs didn't simply extend, but became a complementary part of their musical instrument.

Figure 6.28: Eleanor Turner spatialises the sound overcoming the limitations imposed by the physicality of the harp

The feeling of *"playing the body"* (Workshop attendee. Southampton, 2017) when using EMG based technology as in MyoSpat was also shared with Eleanor Turner, who after composing *The Wood and the Water* said:

> When I first met Balandino, I had been playing the electric harp for twelve years, making looped pieces using effects, but it did not feel real. Compared to the responsiveness and varied timbres of the acoustic harp, I was severely limited with my basic set-up. I wanted to create relevant music to express myself and something of the world today, and I had a dream of being like Björk; collaborating with sound designers and creating an other-worldly amplified show with emotional impact. Until I tried MyoSpat, that concept could not have been further away. MyoSpat has transformed my experience of playing the electric harp. [...] Working with Balandino has given me a direct means of communicating with the audience through a more physical performance style and the use of multiple effects and lights, intrinsically linked with the musical creation.
>
> Eleanor Turner (Birmingham, 2017)

Similar comments were also made by pianist Sarah Nicolls about *Suspensions* by Tanaka and Nicolls (2017) says:

> Working with the EMG sensors with Atau Tanaka was challenging as they require certain physical tension — muscles need to be contracted and tensed whilst playing. A new gestural language evolves around the performance, and it feels as if one is playing two instruments at once almost: the external piano and the internal instrument of the body.
>
> Nicolls (2011).

Unlike when performing composed pieces, participants had difficulties in repeating improvised movements. The fact that participants spent a short time using the system had a relevant impact on the reproducibility of particular sounds and lighting effects. It is then possible to conclude that MyoSpat is a system that does facilitate the interaction with sound and lighting parameters, but at the same time requires practice to be able to replicate sonorities and lighting effects systematically.

6.6 Star Cluster

6.6.1 Performance

Star Cluster, composed by Kirsty Devaney, is a piece for harp, loop station and MyoSpat. It was performed by Eleanor Turner and recorded at Royal Birmingham Conservatoire. The video recording of *Star Cluster* is included in this thesis as Appendix F.2, no. 9.

This piece was commissioned by the Future Blend Project Competition 2015/2016. Initially, *Star Cluster* was not composed for MyoSpat. It was, in fact, for harp and electronics realised using a series of guitar pedals such as delay, reverb, filter bank and loop station (Devaney, 2016). Approximately one year after its composition, the composer Kirsty Devaney and the performer Eleanor Turner aimed to enhance the live performance aspects by reducing the number of pedals, to clarify the performer's intentions through gestural interaction, and to bring an additional layer to the performance through lighting effects.

6.6.2 MyoSpat 0.7

In this version of the system, no changes were made in the interaction design. Improvements were made to the spatialisation algorithm and, through new mapping strategies, to the processing of the audio signal. The lighting system was also updated with new mapping strategies and the use of different lights. In addition to discussing these changes, an evaluation of the seventh MyoSpat iteration is also included, the result of a user study conducted at Berklee College of Music, Valencia Campus (Valencia, Spain).

6.6.2.1 Audio-visual engine

In version 0.7, the gesture recognition algorithm is implemented using the ml.lib library for Pure Data (Figure 6.29), with the result that the load on the OSC network is reduced and the internal data flow within the audio-visual engine is smoother.

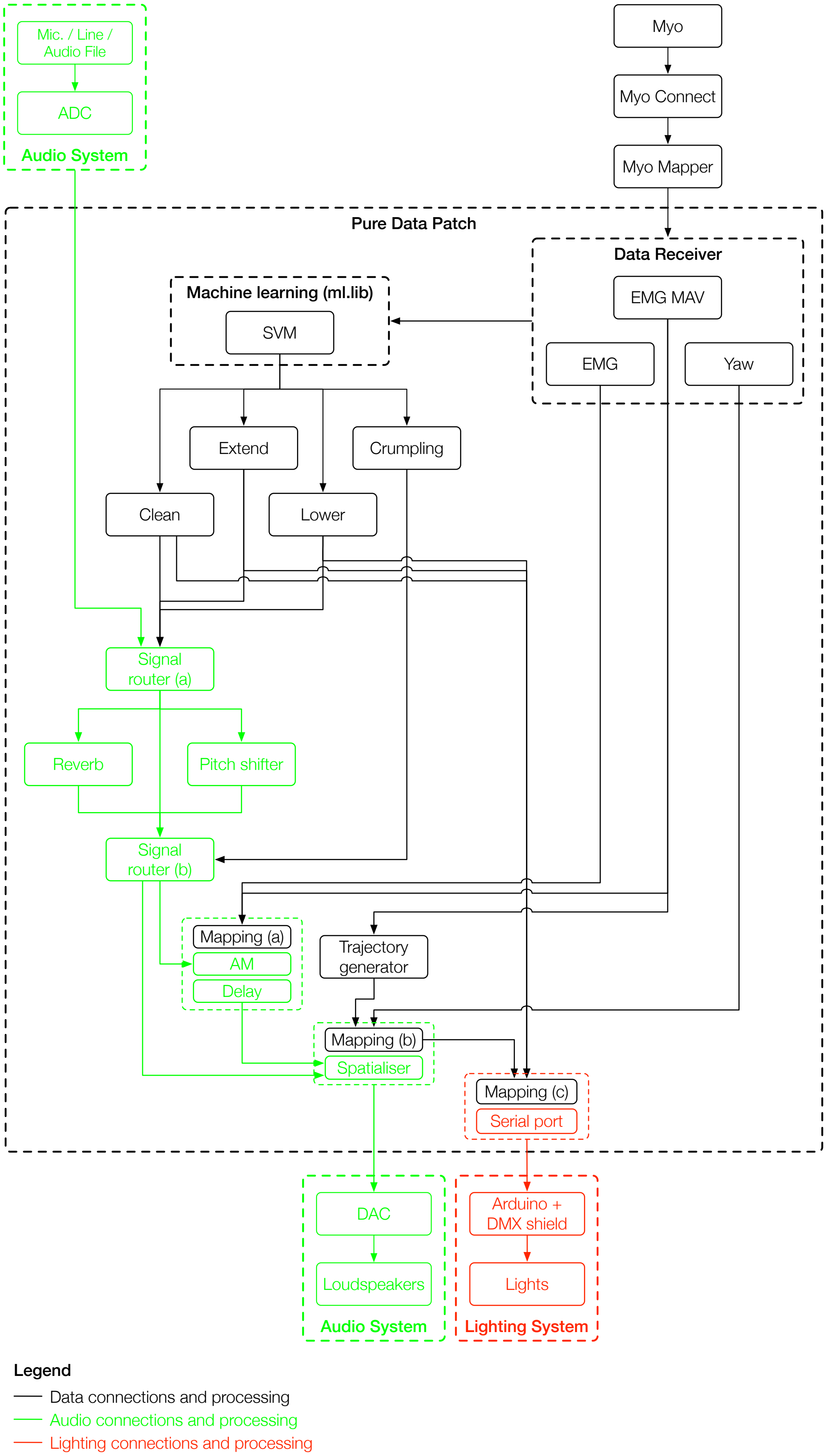


Figure 6.29: MyoSpat 0.7 architecture

Audio synthesis changes were made only on the spatialiser. The gain control of each output channel was replaced by an Ambisonics algorithm, implemented following an experimental approach to distribute sound energies within the listening area. Moreover, a reverb effect was added to the spatialiser algorithm to inform the listener of the virtual room characteristics and enhance the precedence effects of the direct signal. The reverb follows the same approach used by Chowning (1971), where the reverberated signal is complementary to the direct one. Both clean and reverberated audio signals were rendered using the same High Order Ambisonics Library used in MyoSpat 0.3 but compiled for Pure Data.

6.6.2.2 Mapping strategy

The ml.svm Pure Data object that implements the SVM algorithm, already mentioned in Section 4.6, was used for recognising three classes of gestures: the *clean*, *extend* and *lower* gestures. The models were trained with a data set of 908 samples per each gesture using the same approach used in MyoSpat 0.6. An evaluation of the model through direct evaluation reported an accuracy of 100%, and 10-fold cross-validation with a Root Mean Square (RMS) error of 0. In MyoSpat 0.7 the *crumpling* and *throwing* gestures were recognised and controlled through direct mapping.

The evaluation of MyoSpat 0.6 showed that the reverb effect did not give enough room for the sound exploration. Thus here, the continuous values from the SVM algorithm were linearly mapped also into early echoes and late diffusion parameters of the reverb, so to allow a more expressive control over the reverberated sound.

In response to comments from participants to the MyoSpat 0.6's evaluation, where they mentioned their difficulty in perceiving spatial sound effects clearly, a different mapping strategy was used for calculating the virtual source position.
Here, yaw data were linearly mapped into azimuth values for driving the Ambisonic algorithm using a sine function so that it would move within a fixed circular

perimeter. The mirrored data were to control the reverberated signal. This method allowed to establish the amount of reverberated sound in relationship to the virtual sound position within the virtual room (see Figure 6.30).

MyoSpat 0.6's evaluation showed that the sound trajectories generated as a consequence of the throwing gesture were not perceived as expected. Taking advantage of the new mapping strategy, the generated trajectories were also simplified. Their path now follows a circular trajectory whose direction (clockwise or counterclockwise) is established by the arm's orientation on the horizontal axes at the moment of performing the *throwing* gesture, and whose length is established by the maximum speed of the arm at the same moment.

Figure 6.31 (a) shows the trajectory resulting from a slow *throwing* gesture towards the left, Figure 6.31 (b) a medium-fast *throwing* gesture towards the left and Figure 6.31 (c) a fast *throwing* gesture towards the right.

The *throwing* was detected with a decision boundary algorithm fed with gyroscope data, using the value 0.3 as the threshold. Consequently, a test on the yaw value, at the moment of the gesture detection, decided on the direction of the trajectory. If the yaw value was greater than 0:5, then the trajectory was clockwise and vice-versa.

The duration and speed of the trajectory's path were established through the linear mapping of gyroscope sensor data.

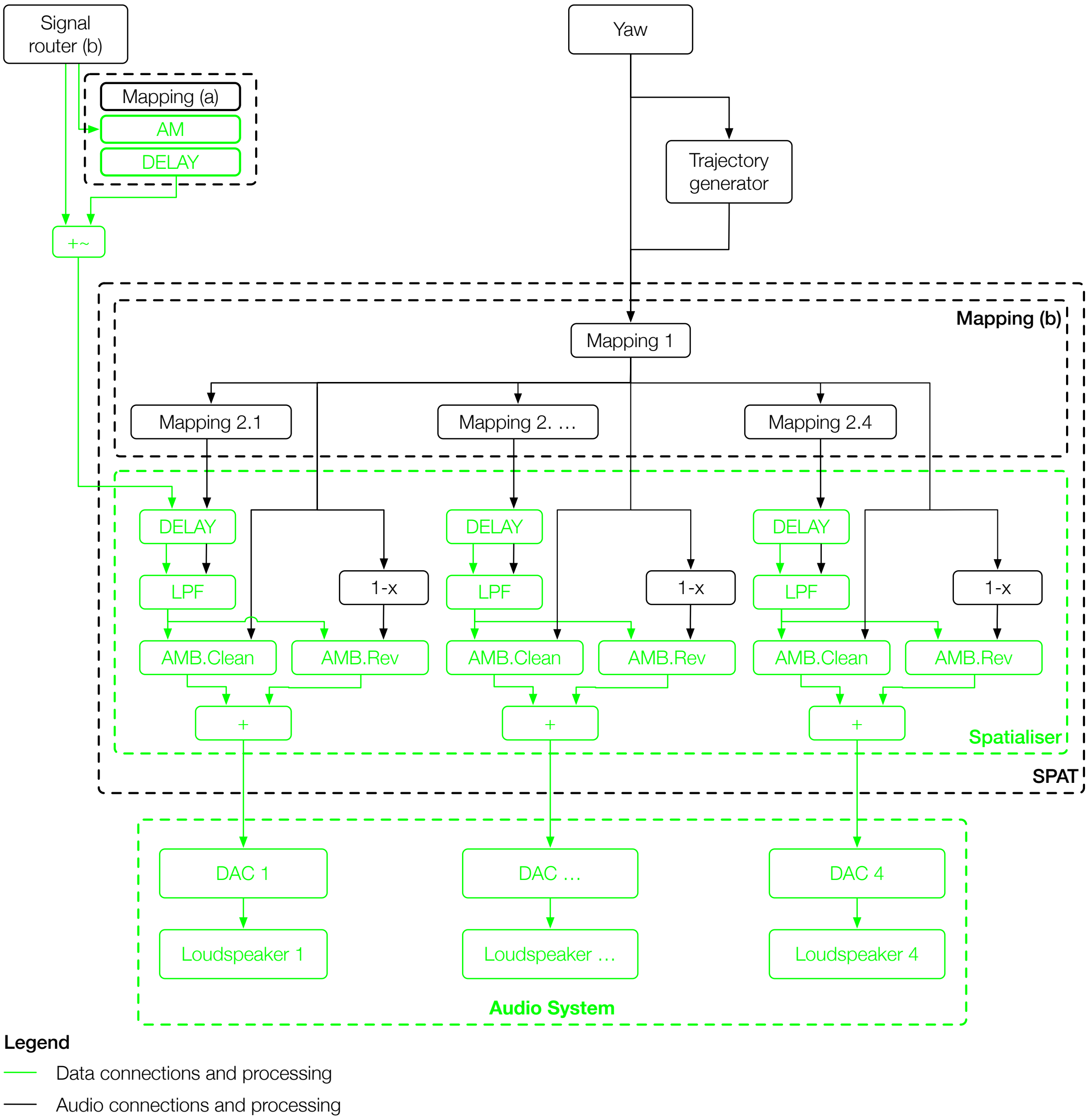


Figure 6.30: MyoSpat 0.7 spatialiser architecture

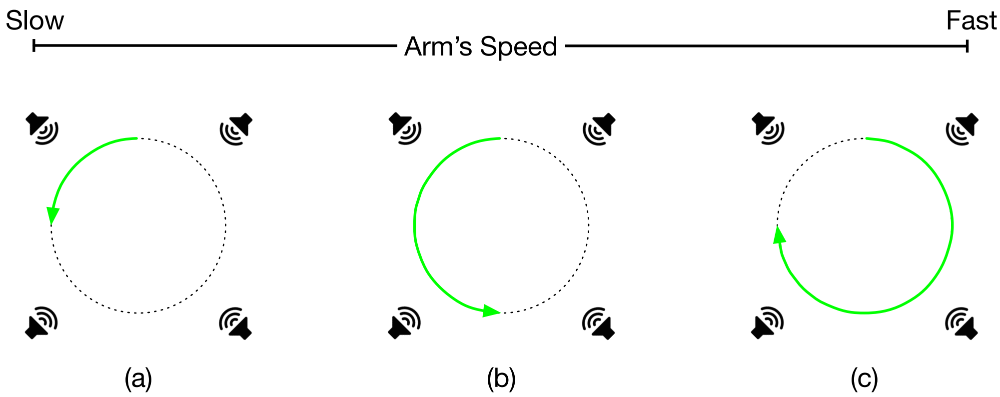


Figure 6.31: Examples of automated spatial trajectories generated performing the *throwing* gesture using MyoSpat 0.7

The AM+Delay effect rendered using MyoSpat 0.6 was not always perceived by the user. Thus, the mapping strategy was modified to make the audio processing more gesture related and prominent.

Here, the *crumpling* gesture is recognised whenever the EMG MAV value is higher than a threshold of 0.14; this threshold was established from findings that emerged from the work with Grace Savage (Section 5.5.2.3), and empirically when working with Eleanor Turner. To give finer control of audio processing, the same approach used for calculating gain factors of the amplitude modulation in MyoSpat 0.6 was used to establish also the feedback of the delay.

To strengthen the relationship between sound and lighting when the AM+Delay effect occurs, in MyoSpat 0.7 the brightness of each light was calculated by mapping linearly into brightness values the distance from the virtual sound source to the associated virtual loudspeaker and the envelope of the audio signal; in this way, the lighting effects represent the spatial position of the sound and its dynamic variations visually, as a consequence of the AM+Delay effect.
The transition time between cues for brightness and colour changes was also reviewed to improve the perception of the virtual sound source movement. The time is calculated independently for each light considering the TDOA value applied to the audio signal diffused by the loudspeaker associated to the light; this ensures that the cues for brightness and colour changes are temporally synchronised with the spatial effects of the sound.

6.6.2.3 System set-up

The MyoSpat 0.7 configuration remains the same as in the previous iteration. However, the previous lights (Par 86 x1W RGB LED) were replaced with four Par 54 x3W WRGB LED to achieve more defined lighting projections in terms of colours and brightness. These lights were chosen as they could be controlled over DMX and were affordable to an average consumer audience. Another important aspect behind the choice of these type of lights is that at the time of writing, they were affordable and easily available worldwide. This means that the system can be installed anywhere without the need to transport the lights.

A case in point is the performance of *The Wood and the Water* performance at Shanghai's Symphony Hall for the Electronic Music Week in 2017. The lighting set-up for this performance was realised by lighting designer Emanuela Mentuccia (Figure 6.32). Her design was based on MyoSpat 0.6's system set-up in the user study configuration to increase the level of immersion and to reinforce the sound-gesture relationship even more. She improved on the previous configuration by mounting the lights above the performing area in order to avoid the production of shadows on the scene. Profile lights with a Frost 201 filter were also used to ensure clearer visibility of the performer at any moment of the performance.

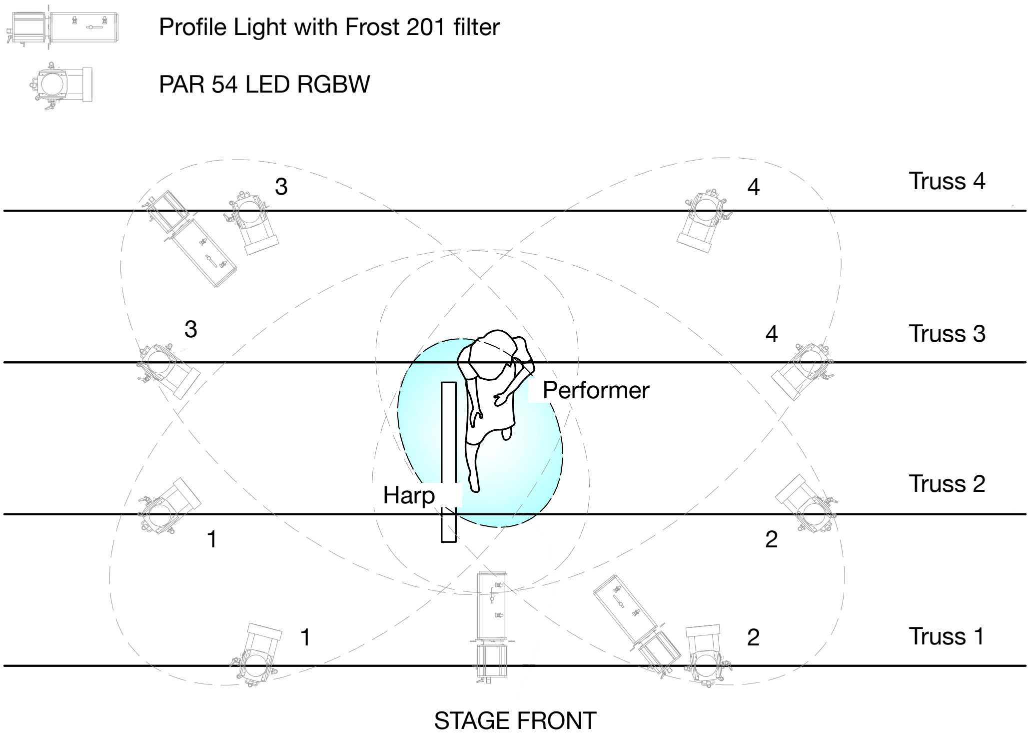


Figure 6.32: MyoSpat 0.7 lighting design

6.6.3 User Study

MyoSpat 0.7 system was evaluated through a user study conducted at Berklee College of Music, Valencia Campus, adopting the methodology described earlier in this chapter (Section 6.2). The results of each of the analyses are presented, followed by a discussion of the different outcomes.

6.6.3.1 Participants

Nine participants – seven musicians, one engineer and one programmer – took part in the user study. Four of the musicians joined in both parts of the study (with and without instrument). They were one singer, one singer/electronic musician, one DJ and one pianist. All participants were aware of the concept of audio processing, and all of them confirmed that they had experienced sound spatialisation in cinemas. Only two participants experienced sound spatialisation as audience members in concerts with multichannel diffusion. The same participants had direct experience of performing with sound spatialisation. Although three participants already knew the Myo, only one participant had used it before the user study. Data are summarised in Figure 6.33 and outlined in detail in Table D.7 and D.8 in Appendix D.2.

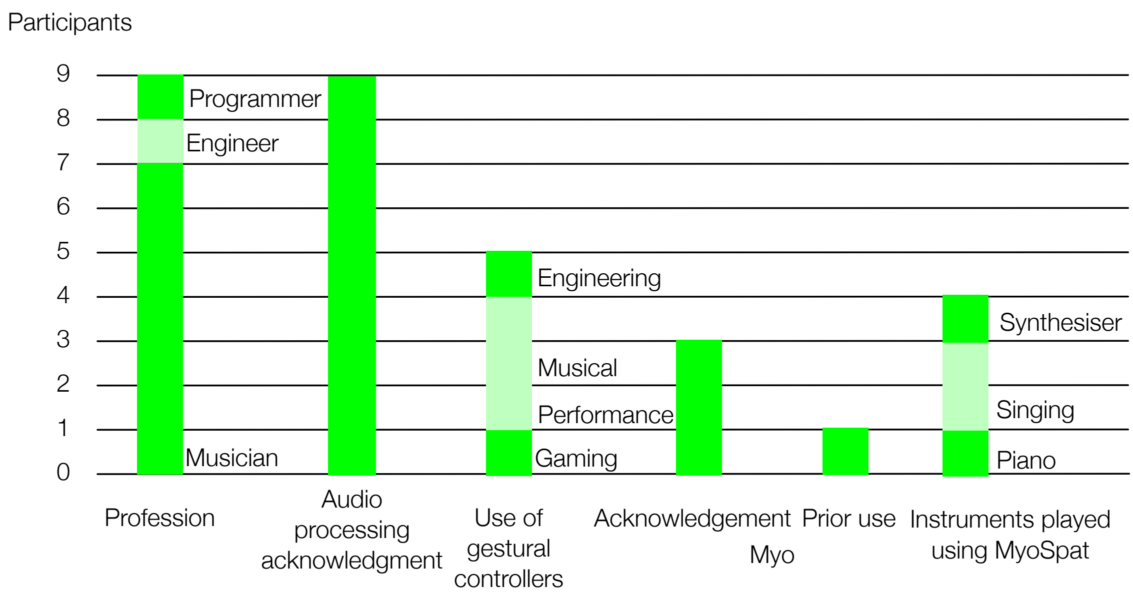


Figure 6.33: MyoSpat 0.7 user study Participant information

6.6.3.2 Results

6.6.3.2.1 Task-based evaluation

The task-based evaluation lasted an average of 44'23' (min 27'36", max 63'04"). Participants spent an average of 23'50" (min 15'11", max 31'45") manipulating the pre-recorded sound file, and an average of 16'44" (min 13'40", max 21'15") using MyoSpat while playing their musical instrument. Before the task-based evaluation began, they spent an average of 4'46" (min 3'33", max 6'16") for learning and practising the system. Detailed data are available in Table D.9 in Appendix D.2.

Comparing the gesture recognition accuracy with the participants' feedback perception (Table 6.5 and Figure 6.34), it emerged that gestures were recognised 99% of the time, and participants perceived the audio and lighting feedback as expected 97% of the time. To be precise, the *lower*, *extend* and *pointing* gestures were recognised with an accuracy of 100%, and the *crumpling* gesture 97%. Participants perceived the audio-visual feedback as expected 98% of the time when performing the *lower* gesture, 100% for the *extend* gesture, 99% for the *clean* gesture, 91% for the *crumpling* gesture, 92% for the *throwing* gesture and 100% for the *pointing* gesture. Comparing both series of data, it is possible to conclude that the last MyoSpat iteration reached a consistently high level of gesture recognition accuracy and rendering of the audio and lighting feedback in all predefined gestural interactions.

Table 6.5: MyoSpat 0.7 task-based evaluation result

Evaluation	**Gesture**						**AVG**
	Clean	**Extend**	**Lower**	**Crumpling**	**Throwing**	**Pointing**	
Gesture Recognition Accuracy (%)	99	100	100	97	98	100	99
Participants Perception Response (%)	99	100	98	91	92	100	97
AVG (%)	99	100	99	94	95	100	98

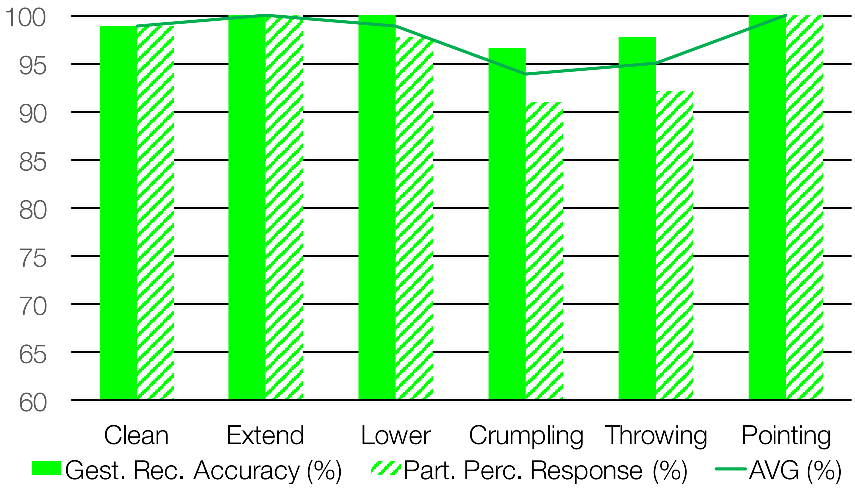


Figure 6.34: MyoSpat 0.7 task-based evaluation result

6.6.3.2.2 Semi-structured interview

All performers commented that the system and the interaction design were easy to learn. Most participants described MyoSpat as a reliable system for audio-visual control and transformation in live performance, as it responded as expected while performing. In particular, Participant 1 was able to perform her vocal pieces and feel in control of the system all times.

Half of the participants commented on having perceived a smooth transition between the various effects. Similarly to the previous user study, participants transformed the pre-recorded audio file as if the sound were a tangible entity. This way of interacting with the sound was found once again to be natural and embodied. Interestingly, while performing gestures that would generate a "splash" sound, Participant 7 said: *'This gesture in real life would have created a splash, but this has not happened'*. It is remarkable that with this comment, Participant 7 did not talk about the interaction by referring to the system or sound control parameters, but to the sound itself as if it were part of the MyoSpat ecosystem; he described the interaction with the sound itself, not through the system.

Participant 9 found the transition between the *lower* and other gestures faster than others. Only one participant (Participant 5) found MyoSpat's interaction design as limiting the movements of the user.

Through the new mapping strategy, all participants perceived both aspects (path and duration) of the sound trajectories generated by the *throwing* gesture. However, two participants (Participants 7 and 9) found these spatial effects less evident among the other effects. Participant 7 hypothesised that the cause of this was due to the audio file's timbral content. By contrast, the AM+Delay effect was perceived and controlled successfully by all participants except one. Most of the participants found that the modulation of brightness intensity, generated through the *crumpling* gesture, was strongly related to the gestural interaction. All participants thought that lighting effects enhanced the perception of the sound effects and made the user experience more immersive. Participant 8 suggested that he would have learned the system quicker if the training phase had started with the lighting system on. However, most of them explicitly stated that even without the lighting system on, they clearly perceived the sound transformations. Half of the participants preferred not to have lighting effects, as they distracted them from the perception of the audio transformations.

Participant 1 proposed the implementation of a higher pitch shift and a freeze effect, while Participant 9 suggested a Wah-Wah or volume effect, associated with the orientation of the arm in respect to the vertical axis. Most participants suggested a smoother transition between effects, stating that visual feedback helped them have more stable control over the effects crossfade.

Participant 5 requested an extension of the interaction design and tried to reverberate the signal through a different interaction than the one proposed by the author. He also suggested the implementation of more complex lighting effects, as in his view sound transformations were significantly more elaborated than the lighting ones. To this extent, Participant 8 suggested the inclusion of a smoke machine in the lighting system.

6.6.3.2.3 User Experience Questionnaire (UEQ)

Chart 6.35 and 6.36, and Table 6.6 and 6.7 show that all participants had a very positive experience using MyoSpat 0.7. Moreover, data show that no particular UEQ scale prevailed the others.

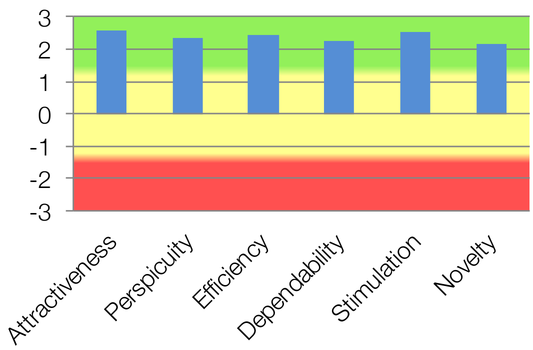


Figure 6.35: UEQ scales, MyoSpat 0.7 evaluation

Table 6.6: UEQ scales, MyoSpat 0.7 evaluation

UEQ scales	
Attractiveness	2.593
Perspicuity	2.333
Efficiency	2.444
Dependability	2.250
Stimulation	2.528
Novelty	2.167

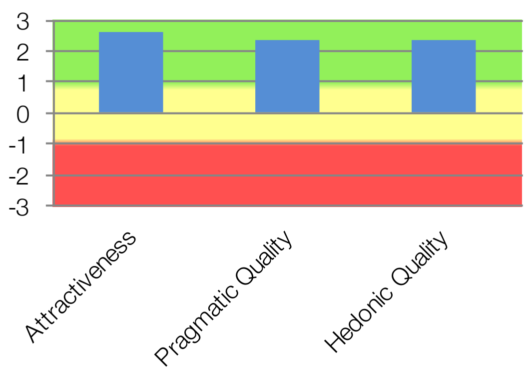


Figure 6.36: UEQ grouped scales, MyoSpat 0.7 evaluation

Table 6.7: UEQ grouped scales, MyoSpat 0.7 evaluation

Pragmatic and Hedonic Quality	
Attractiveness	2.59
Pragmatic Quality	2.34
Hedonic Quality	2.35

6.6.3.3 Discussion

Comparing quantitative data about the gesture recognition accuracy, and the statistics concerning the perception of the auditory and visual feedback, overall, MyoSpat 0.7 improved in respect to the previous version (Figure 6.37). Noticeable is the improvement of the gesture accuracy of the *throwing* gesture and the perception of the triggered audio effect by those that took part in the evaluation of the system.

Most participants felt comfortable using the system. They agreed that it

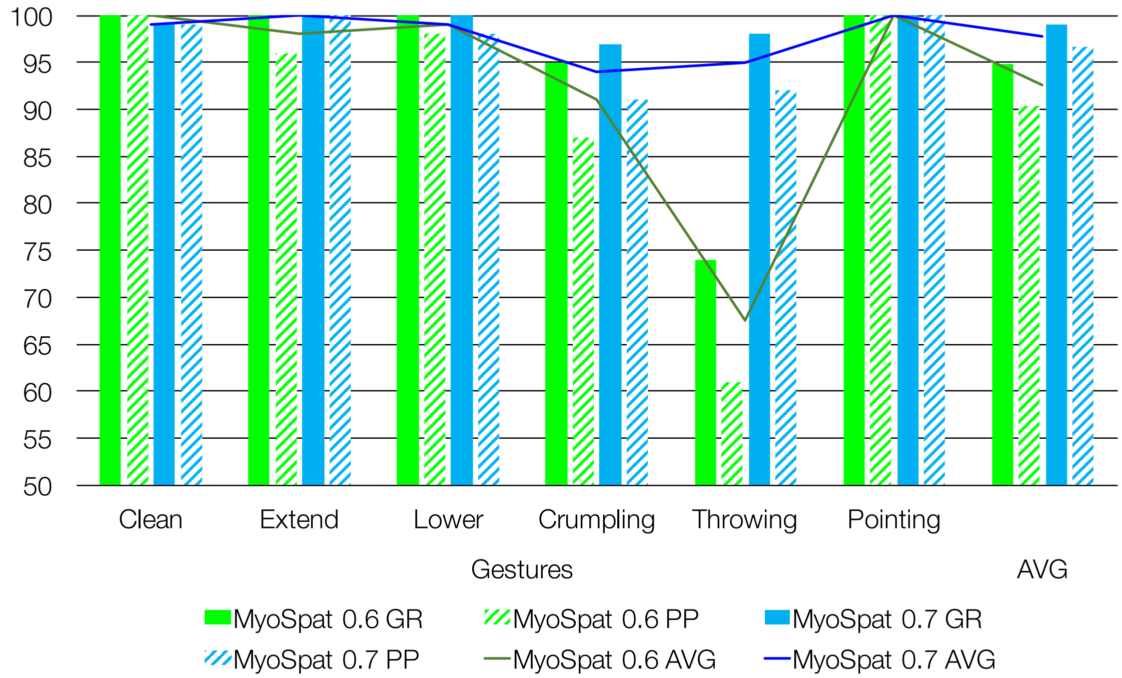


Figure 6.37: Comparison of task-based evaluation results from MyoSpat 0.6 and MyoSpat 0.7 user studies. GR = Gesture Recognition, PP = Participant Perception.

can support a live performance and stimulate musical creativity. Interestingly, behaviours similar to those of the MyoSpat 0.6 user study were observed. Users controlled and transformed the sound of water flow as if they were interacting with water. In both user studies, the lighting effects enhanced usability, and a participant suggested that it had the potential to improve learnability.

The lighting system also enhanced the control of the transition between different audio effects. However, issues concerning the transition between effects experienced in the previous iterations of the system were not completely resolved.

Comparing UEQ results from the evaluation of the 0.6 and 0.7 versions (Figure 6.38, Table D.10 in Appendix D.2), it is possible to see that this latest version improved considerably the user experience. Most importantly, there is no relevant difference between pragmatic and hedonic quality, which gives reason to believe that the system offers the same degree of control of the sound and lighting effects in the performance of both composed and improvised music.

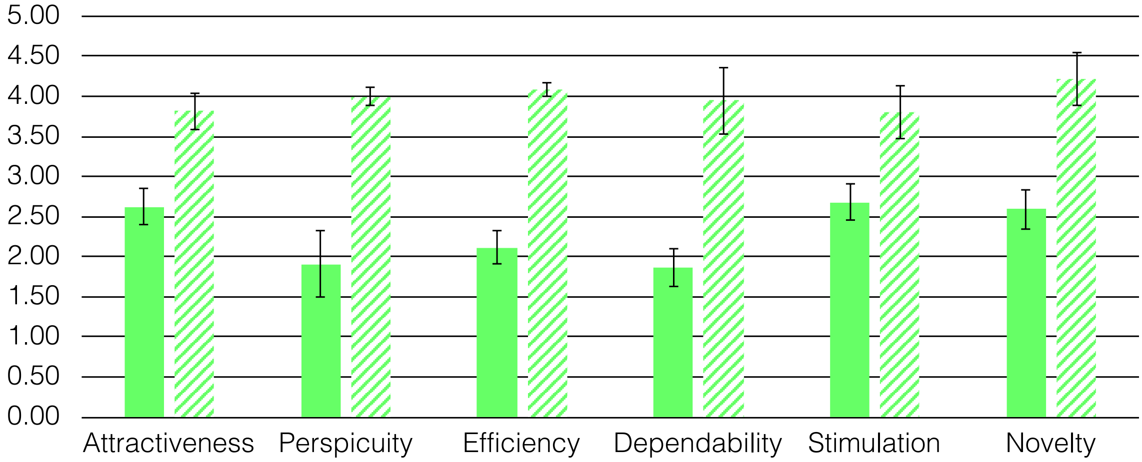


Figure 6.38: Comparison of UEQ scales results from MyoSpat 0.6 (solid green) and MyoSpat (green stripes) 0.7 user studies

By comparing UEQ scales of data gathered during the task-based evaluation and the workshop, shown in Figure 6.39 (detailed data is outlined in Table D.4 and D.5 in Appendix D.1), a significant decrease of perspicuity and dependability factors in the workshop situation emerged. These two properties refer specifically to the learnability and degree of control of the system respectively.

From these findings, it is possible to conclude that the system has the potential to provide a better user experience during improvisation than performance or composition. Examples of MyoSpat's use for improvisation are shown in the video recording no. 11 in Appendix F.4. However, it has also been observed that answers relative to dependability factors were inconsistent (Figure 6.26). Moreover, the composition and performance of *The Wood and the Water* and performance of *Star Cluster* proved that MyoSpat is able to support both composition and performance. Another aspect that emerges from comparing the user study and workshop situations is that the user needs an average of 39'13" to become familiar with the system and feel in control of the interaction.

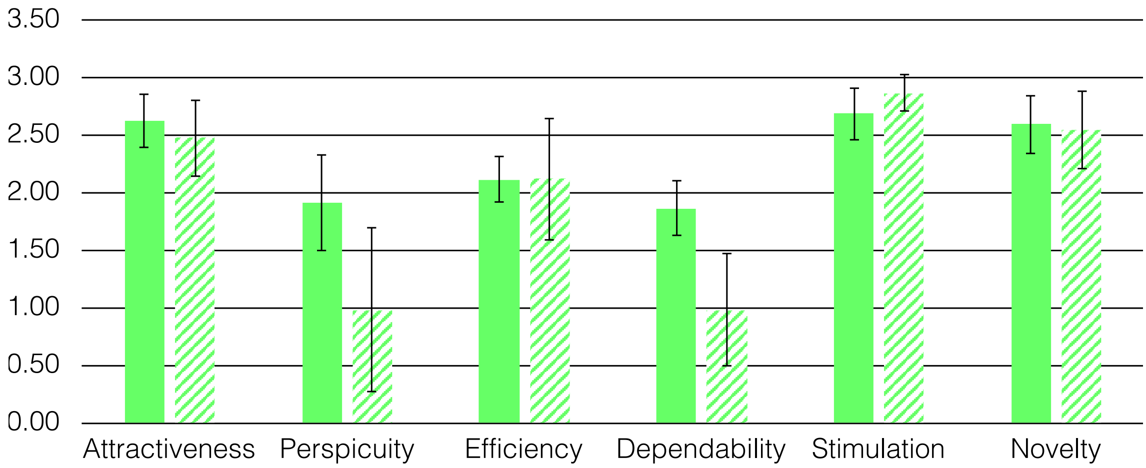


Figure 6.39: MyoSpat 0.6 comparison of UEQ scales resulted from data gathered through the user study (solid green), and evaluation among harpists (green stripes)

6.7 Chapter summary

Findings from this case study showed that the interaction design realised through a UCD approach, has the potential to empower performers to express themselves through a natural and embodied control of audio and lighting transformations. The six gestures (*pointing*, *crumpling*, *extend*, *lower*, *throwing* and *clean*) and the ability of MyoSpat to detect them with a high degree of reliability and map them to meaningful musical parameters expands the creative possibilities open to performers exponentially, without disrupting their established relationship with the instrument, their instrumental technique and their connection with the audience.

In this chapter, the use of MyoSpat in different situations – performance, workshop, and in the composition process – was discussed. Results showed that the interaction design implemented through the system is easily learnable without diminishing the creative process. The system limitations were seen by musicians as elements to play with, and inspiration to drive their choreographic and playing technique beyond those limits. A particular example of this during a performance was the use of BSL, in both *The Wood and The Water* and *Star Cluster*. On these two occasions, the performer and composer aimed to express artistic intentions through gestures not codified by the system. Although the interaction design did not include some of the gestures used in these two performances, it supported the creative objectives of the performer as well as those of the composer.

The evaluation of the system demonstrated the validity of using standard instrumental gestures and techniques as the starting point to design interactions between performers and technology. Performers were in direct control of the sound transformations and processing, and more likely to regard technology as an extension of their instrument and their performance practice, rather than as an external appendage.

Chapter 7

Conclusions

7.1 Summary and Contributions

In this thesis, I have explored three main research questions: (i) how can musicians most effectively control and transform auditory, visual and lighting effects during a live performance through gesture? (ii) What interaction design considerations should be addressed that allow performers to interact simultaneously with an instrument and audio-visual-lighting processing? (iii) How can disruption during a live performance with embodied human-computer interactions be reduced?

With this thesis, it has been argued that through embodied interaction design, audio, visual and lighting processes can be controlled effectively. In using this approach, we can design interactions that minimise the disruption of the musical performances and at the same time, introduce challenges that stimulate the creativity of the musician. Many factors contribute to the creation of embodied interactions with auditory and visual mediums.

From this work, it has emerged that it is possible to interpret and encode into the interactive system itself the intentions of the musician. More to the point, the identification of specific human behaviours during a performance and in daily life can drive the design of embodied interactions with sound transformation parameters.

Merging artistic and technological considerations and the continuous iterative process of music-making and system development have been the strengths of this work and adoption of the User-Centred Design (UCD) method has greatly facilitated this process. This approach allows one to embrace and include the creative and artistic process, allowing artistic intentions, and the resulting artistic experiences and technological artefacts to be fully captured in the iterative development. Moreover, the methodology for this research permits the gathering and inclusion of musicians testimonies about their gestures and the use of technology in their artistic practice.

The interviews with beatboxers, harpists, and all studies and workshop participants revealed various vital aspects to consider when designing modes of interaction, such as the importance of the musical and broader artistic context in which such interactions are situated. Having a deep understanding of the musicians and their relationship with both their instrument and other art forms leads to the formulation of gestural interactions that can become embodied, as they relate to the musicians' artistic expression and cultural background.

Different aspects of embodied interaction design were explored in the Pilot Studies (Chapter 4), and the two UCD cycles: Case study 1 - Voice Gesture (Chapter 5) and Case study 2 - HarpCI (Chapter 6). These are summarised and discussed in the following sections reflecting the approach taken in this research: exploration of musical interaction design using EMG and IMU-based devices, followed by observations of interaction design in two music performance contexts, one where the performer's movements are free from any constraint (*unconstrained interaction*), and the second one dealing with instrumental-constrained performance (*constrained interaction*). Finally, the impact of Myo Mapper in the music community is presented.

Limitations and questions that remained unanswered in this thesis are described at the end of this Chapter.

7.1.1 Exploration of musical interaction design using EMG and IMU-based devices.

Different modes of interactions were explored in five pilot studies. The first study focused on ways to control dials similar to physical knobs, taking advantage of the up/down-high/low pitch and left/right arm orientation-left/right panning gesture-sound relationships. The interaction design was realised considering the affordances of the GUI controls (dials and sliders) and the sensing capabilities of the Myo. It was later implemented combining data sources and the Myo factory gesture recognition library and IMU sensors. From this first pilot study, it becomes apparent that the Myo is a reliable device and with potential in to drive interactive audio processes, but with a factory gesture recognition software that supports modes of interactions that are systematic and not expressive.

In the second pilot study, electromyogram signals from the forearm were sonified in a similar manner to Pauletto and Hunt (2006) to evaluate the potential of linear mapping strategies for the control of audio engines and to better understand the nature of the signal. From here, it emerged that EMG is a very rich signal that empowers the user to create potentially infinite combinations of gesture-related sonic responses. At the same time, fixed sound-gesture relationships were found. This allowed focussing on particular poses to be recognised in the following pilot study.

In the third pilot study, three different poses (*pointing*, *L* and *three* pose) were successfully recognised using a Support Vector Machine algorithm, trained with eight channels of EMG absolute data from the forearm, to recall three audio panning settings of sound on the left, centre and right channel. It was observed that tracking a wider range of hand poses is a significant challenge. Moreover, imposing such modality of interaction might lead to the design and development of interfaces that are instrument oriented. The use of hand poses to control audio processes could support specific instrumental techniques, which can be adopted for one or a small number of instruments.

In the fourth case study, Tanaka's principles of sound physicalisation and Benford and Giannachi's concept of trajectories between realities were adopted in order to design embodied interactions with sound. This interaction design was implemented in a system that enabled the user to generate sounds by taking advantage of the sound affordances, thus referring to the sound-producing object and action. For example, the generation of crumpling paper sounds through crumpling paper gestures. This pilot study highlighted the importance of the action-perception loop to understand and experience the interaction with a musical system.

In the fifth study, interaction techniques explored in the previous pilot studies were combined to support interaction with a sound-object. The sound-object here is intended as looped or continuous sound with a spatial position in the auditory scene and visually represented with a sound icon in a 3D GUI environment. The sound icon is a sphere, with time-varying characteristics informed by sound time and frequency domain properties. The interaction design included the possibility to grab the sound object with a grab-like gesture and to move it, throw it, or to draw a line in the 3D environment and later make the sound object travel through the drawn line in a loop. From the last two studies, it was observed that if designing interactions with sound processing parameters to emulate the actions with a real-world object causing auditory feedback similar to one of the synthesised sound, the interaction with sound processing parameters becomes embodied and natural.

Different approaches to interaction design using IMU and EMG devices were identified in these studies. These ranged from the control of sound with sound parameters to interaction with sound as if it were a tangible entity to be shaped and manipulated like a real-world object. Among all gestural interactions, the crumpling, grabbing, dropping, and throwing gestures exhibited the most robust gesture-sound relationship. The grabbing, dropping and throwing gesture have the potential to be applied in a broader range of applications. Regardless of the sound source – in the case of the last pilot study, a looped guitar sound – the interaction design allowed the control of spatial parameters ad supported the gesture-sound relationship. By contrast, the crumpling gesture is a mode of interaction that can be utilised in a context-specific situation, for example, to generate a crumpling paper sound.

The design and development process of a system, mapping strategies and other technological implementations, although vital, are only the means to bring a technological artefact to life. It is by focusing on modes of interaction and the gesture-sound relationship that we can avoid building machine-oriented systems that do not take into account the performer's experience and their mental image of the sound. At this stage of my work, this argument was only speculative, as based on my own informal evaluations. Aspects of interaction design and mapping strategies were investigated in greater details and evaluated formally in the later case studies.

7.1.2 Unconstrained interaction

After having investigated different ways of interactions using the Myo in different contexts, I started working in the context of music performance.

In the first Case Study, Voice Gesture, this work was framed in a context where the performer could move her hands freely, without any constraint. The musical scenario for this first case study was singing, a practice where the musician is free from instrumental constraints, apart from her own body. After a review of body movements during a performance, behavioural and HCI works, the *pointing* gesture was found to be the most appropriate gesture to control different spatialisation algorithms based on VBAP (quadraphonic audio system), the IRCAM SPAT (octophonic audio system) and High Order Ambisonics. This gesture consisted of orienting variously to interact with the position of the virtual sound source. A kNN algorithm was successfully utilised to facilitate the gesture-sound parameters mapping.

Spatialised versions of the *Sequenza III* by Luciano Berio and *Strypsody* by

Cathy Berberian were performed by Vittoriana De Amicis using MyoSpat at Royal Birmingham Conservatoire (Birmingham, United Kingdom), Conservatorio S. Cecilia (Rome, Italy) and Conservatorio A. Casella (L'Aquila, Italy), using the three different audio spatialisation set-up.
From the performances, it emerged that the *pointing* gesture – here intended as when the whole arm points and not just the index finger - allows the control and embodiment of spatial parameters of the sound without disrupting the musical or choreographic aspects of the pieces. The fact of not using specific hand poses enables the performer to move the hands freely as the composer intended while controlling the spatialiser and maintaining a strong gesture-sound relationship. In relation to the reviewed literature in Section 2.5.1, this mode of interaction can be considered *unconscious*, in the sense that the performer unconsciously drove the spatialiser. The fact that sound spatialisation parameters can be controlled unconsciously so that the system adapts to any of the performer's movements allows the performer to rehearse the piece without the system.

7.1.3 Constrained interaction

The interaction design and the MyoSpat system were then extended to embrace constraints imposed by a wide varied of musical instruments and musical needs.

The first case study, Voice Gesture, focused on vocal performance, constrained by a microphone and vocal audio processors (a loop-station in this case). The interaction design and the MyoSpat system were reviewed to support the composition and performance of the piece *Music Gesture Beatbox* by Grace Savage. The interaction design included new gestures: the *painting* gesture to generate drawings on a projection screen, *playback* gesture to control the reproduction of an audiovisual file, and *filtering* gesture to control a band-pass filter. The MyoSpat system was further developed to enable interaction with visual and lighting effects. In contrast to an unconstrained interaction context, modes of interactions were

designed around the performance *Music Gesture Beatbox* adopting a *conscious* interaction design approach. The gestural interaction with parameters took into account Grace Savage's artistic ideas and the influence of visual art in beatboxing. This approach resulted beneficial to the musical performance in the sense that it ensures a high degree of reproducibility, and facilitates establishing strong links between performing aspects, the music and the cultural background. However, this method risks confining the application of the interaction design uniquely for specific musical performances, such as *Music Gesture Beatbox*.
The MyoSpat system sufficed the requirements for performing at Music Tech Fest Berlin and fostered the embodied interaction with sound, visual and lighting processing parameters. It could be used to perform other pieces of music, but in that case, the MyoSpat system should be further developed for every single musical performance.

The last case study, entitled HarpCI, aimed at breaking any dependence imposed by performance requirements. Here, the interaction with sound and lighting effects parameters was designed to support the composition of new pieces and for the performance of already composed music. MyoSpat was successfully utilised by Eleanor Turner for composing *The Wood And The Water* and performing it at the Electronic Music Week at the Shanghai Opera House (Shanghai, China), and for a studio recording; and to perform *Star Cluster* composed by Kirsty Devaney and as a tool for teaching how to compose and perform electronic music.
HarpCI was first positioned in the context of harp performance and later extended to a wider range of instrumental performances. This approach was taken to accommodate the interaction design to any electronic music performance in which the performer seeks an embodied interaction with the machine.

The interaction design implemented in the last case study was composed of the *pointing* gesture and five more to control audio and lighting processes: (i) the *crumpling* gesture to control an amplitude modulation chained to a delay, (ii) the *throwing* gesture to trigger automated sound spatialisation trajectories, (iii) the *extend* gesture to control parameters of a reverb effects, (iv) the *lower* gesture to drive a pitch-shifter and, (v) the *pointing* gesture to continuously control sound spatialisation cues.

The fusion of gestures expressing music metaphors (i.e. *extend* and *lower* gestures) and the ones that do not (i.e. *crumpling*, *throwing* and *pointing* gestures) permit the realisation of an interaction design that is non-instrument-specific, and easy to learn and use during a musical performance. This approach to interaction design can be considered a limitation by some users, but can however encourage musicians to discover each nuance of the system through a "musical multi-modal approach". The interaction design can introduce challenges that can result in generating interest from the musicians, rather than problems and limitations.

The relationship between the gestures, the audiovisual processors, and the different gestures themselves can also allow the creation of polysemous interaction design. For example, the *lower* gesture was perceived as interaction with a pitch shift effect when feeding the system with an instrumental sound; but when the system was fed with the sound of a river's flowing water sound, the musicians perceived the interaction as being with a virtual bathtub or bowl filled with water. Thus, the interaction became natural and embodied when the interaction and the generated feedback fit equally within the same context and informed each other.

In analogy with the concept of bidirectional complementarity proposed by Tanaka and Knapp (2002), the present work shows the synergy between the audiovisual feedback and the modality of interaction that provides an embodied and natural experience, and not only between the single components of the interactive system itself – such as input device, mapping strategy and the audio/visual/lighting processing.

With this research, it is also argued that a successful interactive musical system should be like a musical instrument. It should not only make the user forget about the interaction with the computer, but it should also promote interaction with the sound as if it were a physical and tangible entity so that it becomes embodied and natural. Interfaces for musical expression should empower users to interact with the audiovisual medium, not with a sophisticated and complex machine.

Solutions explored in this work included the implementation of an interaction design that produces a similar auditory and visual feedback when interacting with real-world objects, and that takes into account the musicians' behaviours and expressivity, for example, the *throwing* and *pointing* gestures. These findings are in line with Caramiaux et al. (2011) who demonstrated that if the sound-producing cause is identified, we tend to describe the action that produces such sound in a real-world scenario.

This work showed that the use of electromyography as a way of capturing gestural data can be novel to many musicians, yet it allows interactions that became embodied to the point that musicians "felt the sound through their muscles" (Workshop attendee, Cardiff, 2017). The use of biosignals empowers musician to transform the sound through a variety of physical action, mid-air or on the instrument. This last aspect allows them to blend the interaction with their instrument and with the computer and made their bodies an augmentation of the musical instrument. More importantly, the interaction design, the input device and the mapping strategy can support the creation of a system that amplify the artistic intentions of the musicians and do not disrupt the musical performance.

Findings from this work suggest that a limited number and relatively popular audio processing techniques, as implemented and evaluated in MyoSpat, can make a system intuitive and easily learnable, at the same time, it allows users to explore the temporal and dynamic properties of the sound freely. Visual projections and lighting effects are the "fifth element" of a system. They strengthen the sound-gesture relationship and allow performers and the audience to immerse themselves in the performance. Visual cues can provide musicians with a deeper understanding and continuous control of the system, for example, in informing them if the sound is being processed by the system, and through which signal processing technique.

7.1.4 Creative projects realised by the community using Myo Mapper

The Myo Mapper software was developed to enable musicians and researchers to take advantage of the Myo's potential for creative applications. This technological artefact has played a vital role in supporting the use of the Myo for this research work. In addition, Myo Mapper has inspired musicians and researchers in the production of different creative outputs in the field of music and dance performances, Virtual Reality (VR), robotics applications and gesture recognition. Its principal contribution was to facilitate the implementation of mapping strategies, thus bridging the gap between the interaction design and the audiovisual engine in an interactive system.

Myo Mapper was used in the realisation of *Fantasie pour violoncelle* by Villegas Curulla (2016), for cello and electronics. The piece explores the field of noise and saturation by means of extended notation and live processing with Cycling 74's Max. Here, the right arm of the cellist controlled filtering, pitch-shifting and time-stretching audio effects using the Myo and Myo Mapper for the mapping of gestures to parameters. Scaled data from the accelerometer and gyroscope modified time stretching and granulation factors (Villegas Curulla, 2018). In *X/Centris Delirium Machine: Les bruits de l'esprit* by Lagacè and Dagher (2017), the instrumental and vocal sounds produced by the performers were transformed by mapping Myo data.

Myo Mapper has been adopted for realising interactive dance performances. In *Nu Body* by Sabri (2017), orientation data from the Myo's IMU data were mapped to parameters for controlling sound spatialisation parameters, changing cues for the lighting colours and visual projections. A Neural-Network Multilayer Perceptron

(MLP) classifier was implemented using Wekinator for the mapping process and Pure Data as the audiovisual engine (Dooley and Di Donato, 2016). T. Michailidis and the author presented an interactive system aiming to translate the performer's gestures into haptic feedback, sensed by another performer using vibrational motors (Michailidis and Di Donato, 2016; Michailidis, 2017). This system aimed to empower dancers to inform each other of their artistic intentions during a performance. Here, yaw and pitch's FOD and EMG ABS data of the Myo worn by a performer were mapped to parameters for controlling the intensity of vibrational motors of other Myos worn by the second performer. Interesting observations emerged from this work; for example, how the electromyography's inability to monitor involuntary movements can be exploited as an element of creativity in dance performance. Dancers can control an interactive system differently when they are interacting or guiding each other's bodies, for instance, when a dancer moves another dancer's arm. Furthermore, orientation data scaling controls give the possibility to adapt different mapping strategies to audiovisual software. The same controls can also be beneficial in adjusting the data mapping process in live performance.

At the DigiDance workshop by TaikaBox (2016), dancers controlled sound and video projections. Myo Mapper was used to streaming yaw, pitch and EMG ABS data to Isadora (TroikaTronix, 2016) and Live (Ableton, 2017). Arganini (2018a) used Myo Mapper in *Follow Me* (2018) to transform his dance movements into parameters for modulating the amplitude of an audio file using Live. In this performance, gestural data were linearly mapped and classified using machine learning software (Arganini, 2018b).

Myo Mapper was used by John Collingswood and Tanja Råan, in collaboration with the industrial robotics company Probot LTD, to explore choreographies for humans and robots. Myo Mapper has also been used for controlling the movements of a robotic arm with a light mounted on its extremity through linear mapping of orientation data (TaikaBox, 2017). The outcome of this work informed the creation of an interface to map arm gestures into robotic arm movements through linear mapping of gyroscope sensor data from two Myos into parameters for controlling the movements of two robotic arm joints (TaikaBox, 2018).

7.2 Limitations and Open Questions

This thesis concludes with a description of the limitations of this research, and the new short- and long-term perspectives that it generates.

Recalling the topic of affordance explored in the literature review and the evidence gathered through this research, it is clear that musicians learn about the modalities of interaction with the sound, visual and lighting effects through the bodily experience of the music and by moving while playing. Since our bodies are the interface to interact with a non-tangible or visible entity (audio and lighting processing parameters), we have no idea of how to interact with the feedback until we experience it. To this extent, I ask myself, what if the system was not first demonstrated to them? Would the musicians still have the same experience? What are the issues that they could face? Further research concerning the affordance of the interface and how it presents itself to the musicians is undoubtedly needed.
Interaction design limitations also included the dependability between the *extend* and *pointing* gestures. As a side effect, spatial cues generated through the *pointing* gesture can also be masked by the reverb effect when performing the *extend* gesture. New work should consider avoiding dependent gestures and interaction designs that affect different controls at the same time.
All interaction modalities here described did not consider the importance of micro-gestures during a performance. A user study participant highlighted the potential of using the MyoSpat during a dance performance, and Eleanor Turner to use it through BSL gestures. Even though this is a fertile terrain of exploration, it was not further investigated. Finally, in relation to the interaction with lighting

effects parameters, no synesthetic study was conducted to establish the sound-gesture-colour or brightness relationships. These were established in collaboration with the musicians adopting an experimental approach. This left space for future work in this field.

Although the MyoSpat system was successfully used in various music performances, this technology was not assessed adopting a rigorous scientific approach. This brings us to the formulation of empirical conclusions concerning the usability and learnability of the system and its limitations, noted during its use in different contexts. MyoSpat requires technical knowledge of Pure Data and the OSC communication protocol. From experience gained with Myo Mapper, future software should be developed to be a cross-platform application and have a simple GUI to manage both input device and audio-lighting processes. On the hardware side, MyoSpat should be able to generate and render the audio, lighting and visual feedback through consumer as well as more professional level audio, lighting and projection systems.

Using off-the-shelf technology brought many advantages, but it meant that the research depended upon the decisions of the industries producing such hardware or software. Most of all, at the time of publication of this thesis, both software and hardware components of the MyoSpat system, were no longer available. The HoaLibrary for Pure Data, utilised for spatialisation algorithm, was not supported on machines with macOS 10.14.2 as Operating System, and the Myo was discontinued after completing this research. Thus, MyoSpat will need further development to welcome different EMG and IMU-based devices.

Appendix A

MyoSpat evaluation tools for data gathering and analysis

Table A.1: Template for gathering participant information (part 1)

Participant	**Profession**					
	Musician				**Other**	
	Y/N (0/1)	**Role**	**Instrument**	**Years of Experience**	**Y/N (0/1)**	**Profession**
1						
2						
3						
...						
TOTAL						
AVG (%)						

Table A.2: Template for gathering participant information (part 2)

Participants number	**Spatial audio**		**Interactive technology**			**Myo**		**Task-based evaluation**	
	Acknowledgement (Y/N, 1/0)	**Experience**	**Use of gestural controllers (Y/N, 1/0)**	**Device**	**Purpose**	**Previously used (Y/N, 1/0)**	**Acknowledgement (Y/N, 1/0)**	**With instrument (Y/N, 1/0)**	**Instrument**
1									
2									
3									
...									
TOTAL									
AVG (%)									

Table A.3: Template for gathering task-based evaluation data

User Interaction	**Attempt**	**Performance (Pos./Neg., 0/1)**	
		Gest.-rec. accuracy	**Part.'s perc. response**
Clean	**1**		
	2		
	3		
	4		
	5		
	AVG (%)		
Extend	**1**		
	2		
	3		
	4		
	5		
	AVG (%)		
Lower	**1**		
	2		
	3		
	4		
	5		
	AVG (%)		
Crumpling	**1**		
	2		
	3		
	4		
	5		
	AVG (%)		
Throwing	**1**		
	2		
	3		
	4		
	5		
	AVG (%)		
Pointing	**1**		
	2		
	3		
	4		
	5		
	AVG (%)		

Table A.4: Template for analysing task-based evaluation data

Participant	**Gesture recognition accuracy and participants' perception of the feedback response (%)**													
	Clean		**Extend**		**Lower**		**Crumpling**		**Throwing**		**Clean**		**AVG (%)**	
	Gest.rec. acc.	**Part. perc.**	**Gest.rec. acc.**	**Part. perc.**	**Gest.rec. acc.**	**Part. perc.**	**Gest.rec. acc.**	**Part. perc.**	**Gest.rec. acc.**	**Part. perc.**	**Gest.rec. acc.**	**Part. perc.**	**Gest.rec. acc.**	**Part. perc.**
1														
2														
3														
...														
AVG (%)														

Table A.5: Template for resuming task-based evaluation data

Evaluation	**Gesture**						**AVG (%)**
	Clean	**Lower**	**Extend**	**Crumpling**	**Throwing**	**Pointing**	
System Performance (%)							
Participant's perception (%)							
System-Perception AVG (%)							

Appendix B

Myo Mapper OSC communication

Table B.1: Myo Mapper 3 OSC messages list (part 1)

OSC tag	Myo parameter	Number of values	Value Type	Range
/myoID/orientation/quaternion	$Quaternion_{x,y,z,w}$	4	float	$[-2\pi, 2\pi]$
/myoID/orientation/raw	$Orientation_{yaw,pitch,roll}$	3	float	$[-2\pi, 2\pi]$
/myoID/orientation/scaled	$Orientation_{yaw,pitch,roll}$	3	float	$[0, 1]$
/myoID/orientation/scaled/velocity	$Orientation_{yaw,pitch,roll}$'s Velocity	3	float	$[0, 1]$
/myoID/orientation/scaled/acceleration	$Orientation_{yaw,pitch,roll}$'s Acceleration	3	float	$[0, 1]$
/myoID/acceleration/raw	$Acceleration_{x,y,z}$	3	float	$[-16, 16]$
/myoID/acceleration/fod	$Acceleration_{x,y,z}$'s first order difference	3	float	$[-16, 16]$
/myoID/acceleration/scaled	$Acceleration_{x,y,z}$'s scaled;	3	float	$[0, 1]$
/myoID/acceleration/scaled/fod	$Acceleration_{x,y,z}$ scaled's first order difference;	3	float	$[0, 1]$
/myoID/acceleration/scaled/fod/mavg	$Acceleration_{x,y,z}$ scaled first order difference's moving average;	3	float	$[0, 1]$
/myoID/gyro/raw	$Gyro_{x,y,z}$	3	float	$[-2000, 2000]$
/myoID/gyro/raw/fod	$Gyro_{x,y,z}$'s first order difference	3	float	$[-2000, 2000]$
/myoID/gyro/scaled	$Gyro_{x,y,z}$'s scaled	3	float	$[0, 1]$

Table B.2: Myo Mapper 3 OSC messages list (part 2)

OSC tag	Myo parameter	Number of values	Value Type	Range
/myoID/gyro/scaled/abs	$Gyro_{x,y,z}$ scaled's absolute value	3	float	$[0, 1]$
/myoID/gyro/scaled/fod	$Gyro_{x,y,z}$ scaled absolute value's first order difference	3	float	$[0, 1]$
/myoID/gyro/scaled/fod/mavg	$Gyro_{x,y,z}$ scaled absolute value first order difference's absolute value	3	float	$[0, 1]$
/myoID/EMG/raw	EMG	8	int	$[-127, 128]$
/myoID/EMG/raw/zcr	EMG's zero crossing rate	8	int	$[0, \infty]$
/myoID/EMG/raw/zcr/mavg	EMG zero crossing rate's moving average	8	int	$[0, \infty]$
/myoID/emg/scaled	EMG scaled	8	float	$[0, 1]$
/myoID/emg/scaled/min	EMG scaled's minimum	8	float	$[0, 1]$
/myoID/emg/scaled/max	EMG scaled's maximum	8	float	$[0, 1]$
/myoID/emg/scaled/fod	EMG scaled's first order difference	8	float	$[0, 1]$
/myoID/emg/scaled/sod	EMG scaled's second order difference	8	float	$[0, 1]$
/myoID/emg/scaled/abs	EMG scaled's absolute value	8	float	$[0, 1]$
/myoID/emg/scaled/abs/mav	EMG scaled's absolute value mean absolute value	8	float	$[0, 1]$

Appendix C

MyoSpat mapping strategy

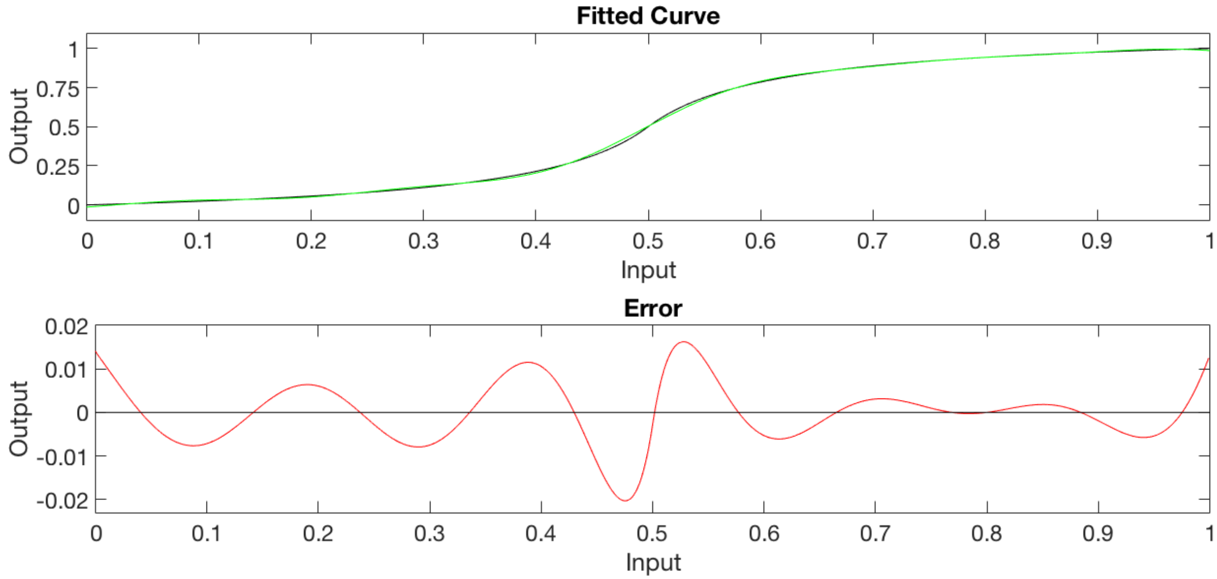


Figure C.1: Fitted curve and error estimation of Equation 5.2

Table C.1: Equation 5.2's goodness of fit

SSE	**R-Square**	**Adjusted R-square**	**RMSE**
0.0458	0.9997	0.9997	0.006812

Appendix D

MyoSpat's evaluations result

D.1 MyoSpat 0.6

Table D.1: MyoSpat 0.6 user study, participant information (part 1)

Participant	Profession						Spatial Audio	
	Musician				Other		Acknowledgement (Y/N, 1/0)	Experience
	Y/N (0/1)	Role	Instrument	Experience (Years)	Y/N (0/1)	Profession		
1	1	Performer, producer	Guitar	7	0		1	Cinema, concerts, ATMOS system
2	1	Performer	Percussion	20	0		1	Cinema, concerts
3	0				1	Manufacturing	0	Cinema
4	1	Performer, composer, teacher	Clarinet, saxophone, guitar, piano	20	0		1	Cinema
5	0				1	Art historian, sculptor, dancer	1	Cinema, concerts
6	1	Performer, composer, teacher, researcher	Keyboard, piano, guitar, bass guitar, drum	22	0		1	Concerts, cinema, music production, performance
7	1	Performer, composer, teacher, researcher	Piano, guitar	25	0		1	Concerts, cinema, composition
8	1	Composer, performer, teacher	Guitar, voice, recorder, harmonica	19	0		1	Cinema, concerts
9	1	Performer, songwriter	Guitar	12	0		1	Cinema, concerts
10	1	Performer, composer, teacher	Harp	27	0		1	Cinema, concerts, performance
11	1	Performer, Researcher	Saxophone	23	0		1	Cinema, concerts, performance, composition
TOTAL	**9**				**2**		**10**	
AVG (%)	$81.\overline{81}$			$19.\overline{4}$	$18.\overline{18}$		$90.\overline{90}$	

Table D.2: MyoSpat 0.6 user study, participant information (part 2)

Interactive technology				Myo		Usability Study	
Participant	**Use of gestural controllers (Y/N, 1/0)**	**Device**	**Use**	**Previously used (Y/N, 1/0)**	**Acknowledgement (Y/N, 1/0)**	**With instrument (Y/N, 1/0)**	**Instrument**
1	1	Wii Remote	Gaming	0	0	1	Guitar
2	0			0	1	1	Snare
3	1	Kinect	Gaming	0	0	0	
4	1	Wii Remote	Gaming	0	1	1	Saxophone
5	1	Myo, Kinect	Gaming	1	1	0	
6	1	Kinect, Leap Motion, Wii mote, Wii Balance Board, smartphone, Arduino	Digital art, musical performance	0	1	1	Keyboard
7	1	Wi mote, Myo, custom built interface, Arduino, X-OSC, webcam	Digital art, musical performance	1	1	1	Keyboard
8	1	iPad, touch OSC, MIDI controllers	Music performance	0	0	1	Singing
9	1	iPad, midi controllers	Music performance	0	1	1	Guitar
10	1	Myo, Kinect	Music performance	1	1	1	Harp
11	1	IMU devices	Music performance	0	1	1	Saxophone
TOTAL	**10**			**3**	**8**	**9**	
AVG (%)	$90.\overline{90}$			$27.\overline{27}$	$72.\overline{72}$	$81.\overline{81}$	

Table D.3: MyoSpat 0.6 user study, duration

Participant	**Duration (h:m:s)**						
	Introductory Questionnaire	**Training**	**Usability Test**			**UEQ**	**Total**
			Soundfile	**Instrument**	**Total**		
1	00:04:00	00:02:30	00:14:54	00:21:42	00:36:36	00:04:09	00:47:15
2	00:03:17	00:05:06	00:09:53	00:15:48	00:25:41	00:03:47	00:37:51
3	00:02:31	00:03:38	00:09:57	*	00:09:57	00:08:53	00:24:59
4	00:02:47	00:05:11	00:08:50	00:42:14	00:51:04	00:03:30	01:02:32
5	00:02:06	00:06:01	00:18:41	*	00:18:41	00:03:01	00:29:49
6	00:03:10	00:02:30	00:09:40	00:16:57	00:26:37	00:03:18	00:35:35
7	00:02:29	00:04:22	00:19:06	00:10:56	00:30:02	**	00:36:53
8	00:03:08	00:08:03	00:07:48	00:12:51	00:20:39	00:04:01	00:35:51
9	00:03:52	00:10:18	00:09:33	00:16:35	00:26:08	00:01:27	00:41:45
10	00:02:03	00:02:44	00:08:35	00:10:30	00:19:05	00:02:26	00:26:18
11	00:03:33	00:03:36	00:20:39	00:22:03	00:42:42	00:02:41	00:52:32
TOTAL	**00:32:56**	**00:53:59**	**02:17:36**	**02:49:36**	**05:07:12**	**00:37:13**	**07:11:20**
AVERAGE	**00:03:00**	**00:04:54**	**00:12:31**	**00:18:51**	**00:27:56**	**00:03:43**	**00:39:13**
MINIMUM	**00:02:03**	**00:02:30**	**00:07:48**	**00:10:30**	**00:09:57**	**00:01:27**	**00:24:59**
MAXIMUM	**00:04:00**	**00:10:18**	**00:20:39**	**00:42:14**	**00:51:04**	**00:08:53**	**01:02:32**

Legend:

- $*$ = Participants that did not complete the user study by playing a musical instrument.
- $**$ = Users who complete the UEQ online.

Table D.4: Comparison of UEQ scales results from the MyoSpat 0.6 user study and the workshop

UEQ scale	MyoSpat 0.6 - Usability Study						MyoSpat 0.6 - Workshops					
	Mean	**STD**	**N**	**Confidence**	**Confidence Interval**		**Mean**	**STD**	**N**	**Confidence**	**Confidence Interval**	
Attractiveness	2.62	0.39	11	0.23	2.39	2.85	2.47	0.61	13	0.33	2.14	2.80
Perspicuity	1.91	0.70	11	0.41	1.50	2.32	0.98	1.32	13	0.72	0.26	1.70
Efficiency	2.11	0.34	11	0.20	1.91	2.32	2.12	0.97	13	0.53	1.59	2.64
Dependability	1.86	0.41	11	0.24	1.62	2.11	0.98	0.89	13	0.49	0.50	1.47
Stimulation	2.68	0.37	11	0.22	2.46	2.90	2.87	0.28	13	0.15	2.71	3.02
Novelty	2.59	0.42	11	0.25	2.34	2.84	2.54	0.62	13	0.34	2.20	2.88

Table D.5: Two Sample T-Test of UEQ scales resulted from the evaluation of MyoSpat 0.6 in the user study and among harpists

Alpha level: 0.05

Attractiveness	0.4823	No Significant Difference
Perspicuity	0.0405	Significant Difference
Efficiency	0.9952	No Significant Difference
Dependability	0.0052	Significant Difference
Stimulation	0.1958	No Significant Difference
Novelty	0.8086	No Significant Difference

Table D.6: MyoSpat 0.6 machine model 10-fold cross evaluation at different hidden layers

Pose	**No. hidden layers**	**Correlation coefficient**	**Mean absolute error**	**Root mean squared error**	**Relative absolute error (%)**	**Root relative squared error (%)**	**No. Samples**
Clean	0	0.7967	0.235	0.2877	54.8446	62.1559	3327
	1	0.9748	0.0321	0.1035	7.4891	22.3607	3327
	2	**1**	**0.0019**	**0.0046**	**0.4514**	**0.9864**	**3327**
	3	0.9927	0.0065	0.0556	1.5205	12.0205	3327
Extend	0	0.7113	0.2632	0.3351	62.4057	72.9788	3327
	1	0.9178	0.1009	0.1823	23.9154	39.6949	3327
	2	**0.9999**	**0.0022**	**0.0048**	**0.5166**	**1.0436**	**3327**
	3	1	0.0009	0.0019	0.2235	0.4159	3327
Lower	0	0.8991	0.1721	0.2167	36.2493	44.4591	3327
	1	**1**	**0.0017**	**0.0038**	**0.361**	**0.7834**	**3327**
	2	1	0.0005	0.001	0.1072	0.215	3327
	3	1	0.0002	0.0004	0.033	0.0803	3327
Crumpling	0	0.8108	0.2356	0.2962	47.1186	59.2317	3327
	1	**0.9708**	**0.0522**	**0.1217**	**10.4469**	**24.3291**	**3327**
	2	0.9666	0.0529	0.131	10.5836	26.1919	3327
	3	0.9634	0.0526	0.1375	10.5092	27.4995	3327

D.2 MyoSpat 0.7

Table D.7: MyoSpat 0.7 user study, participant information (part 1)

Participant	Profession						Spatial Audio	
	Musician				Other		Acknowledgement (Y/N, 1/0)	Experience
	Y/N (0/1)	Role	Instrument	Years of Experience	Y/N (0/1)	Profession		
1	1	Performer	Singing	17	0		1	Recording DAW
2	1	Performer	Synthesisers, drum, guitars, singer,	15	0		1	Surround sound, Cinema
3	1	Performer	Piano, guitar	20	0		1	Malton, Emisphere,
4	1	Performer	Bass guitar, synthesisers	6	0		1	Cinema
5	0		Singing	3	1	Engineer	1	Cinema
6	0		Drum	0.5	1	Programmer	1	Cinema
7	1	Performer	Guitar	25	0		1	Cinema
8	1	Performer	Computer music, drum,	7	0		1	Live performance
9	1	Performer	Piano, guitar	5	0		1	Cinema
TOTAL	**7**				**2**		**9**	
AVG (%)	$77.\overline{7}$			$10.9\overline{4}$	$22.\overline{2}$		**100**	

Table D.8: MyoSpat 0.7 user study, participant information (part 2)

Participant	**Interactive technology**			**Myo**		**Usability Study**	
	Prior use (Y/N, 1/0)	**Device**	**Use**	**Previously used (Y/N, 1/0)**	**Acknowledgement (Y/N, 1/0)**	**With instrument (Y/N, 1/0)**	**Instrument**
1	0	None	None	0	0	1	Singing, Loop-station
2	1	Leap Motion	Musical performance	0	1	1	Drum machine
3	1	Kinect	Engineering, video processing	0	1	0	
4	1	iPad	Musical performance	0	0	0	
5	0	None	None	0	0	1	Singing
6	1	WiiMote	Gaming	0	0	0	
7	0	None	None	0	0	0	
8	0	Mobile devices	Musical performance	1	1	0	
9	1	None	None	0	0	1	Piano
TOTAL	**5**			**1**	**3**	**4**	
AVG (%)	$55.\bar{5}$			$11.\bar{1}$	$33.\bar{3}$	$44.\bar{4}$	

Table D.9: MyoSpat 0.7 user study, duration

Participant	**Duration (h:m:s)**						
	Introductory Questionnaire	**Training**	**Usability Test**			**UEQ**	**Total**
			Soundfile	**Instrument**	**Total**		
1	00:03:15	00:04:20	00:21:40	00:17:28	00:46:43	00:04:17	00:51:00
2	00:02:30	00:03:40	00:19:25	00:13:40	00:39:15	00:04:30	00:43:45
3	00:03:44	00:03:33	00:31:45	∗	00:39:02	00:03:28	00:42:30
4	00:03:43	00:04:11	00:28:05	∗	00:35:59	00:04:48	00:40:47
5	00:04:54	00:05:42	00:21:24	00:14:32	00:46:32	00:05:50	00:52:22
6	00:03:34	00:04:29	00:15:11	∗	00:23:14	00:04:22	00:27:36
7	00:04:03	00:05:19	00:23:16	∗	00:32:38	00:03:45	00:36:23
8	00:04:19	00:06:16	00:25:58	∗	00:36:33	00:05:23	00:41:56
9	00:04:34	00:05:24	00:27:43	00:21:15	00:58:56	00:04:08	01:03:04
TOTAL	**00:34:36**	**00:42:54**	**03:34:27**	**01:06:55**	**05:58:52**	**00:40:31**	**06:39:23**
AVERAGE	**00:03:51**	**00:04:46**	**00:23:50**	**00:16:44**	**00:39:52**	**00:04:30**	**00:44:23**
MINIMUM	**00:02:30**	**00:03:33**	**00:15:11**	**00:13:40**	**00:23:14**	**00:03:28**	**00:27:36**
MAXIMUM	**00:04:54**	**00:06:16**	**00:31:45**	**00:21:15**	**00:58:56**	**00:05:50**	**01:03:04**

Legend:

- ∗ = Participants that did not complete the user study by playing a musical instrument.

Table D.10: Comparison of UEQ scales from the evaluations of MyoSpat 0.6 and 0.7 through user study

UEQ Scale	MyoSpat 0.6						MyoSpat 0.7					
	Mean	**STD**	**N**	**Confidence**	**Confidence Interval**		**Mean**	**STD**	**N**	**Confidence**	**Confidence Interval**	
Attractiveness	2.62	0.39	11	0.23	2.39	2.85	3.81	0.35	9	0.23	3.59	4.04
Perspicuity	1.91	0.70	11	0.41	1.50	2.32	4.00	0.18	9	0.12	3.88	4.12
Efficiency	2.11	0.34	11	0.20	1.91	2.32	4.08	0.13	9	0.08	4.00	4.16
Dependability	1.86	0.41	11	0.24	1.62	2.11	3.94	0.65	9	0.42	3.52	4.37
Stimulation	2.68	0.37	11	0.22	2.46	2.90	3.81	0.51	9	0.33	3.47	4.14
Novelty	2.59	0.42	11	0.25	2.34	2.84	4.22	0.51	9	0.33	3.89	4.55

Table D.11: Two Sample T-Test of UEQ scales from data gathered after the MyoSpat 0.7 user study and workshop

Alpha level: 0.05

Attractiveness	0.0000	Significant Difference
Perspicuity	0.0000	Significant Difference
Efficiency	0.0000	Significant Difference
Dependability	0.0000	Significant Difference
Stimulation	0.0001	Significant Difference
Novelty	0.0000	Significant Difference

D.3 Codes from interviews with MyoSpat evaluations' participants

1. Played musical instrument
   (a) Guitar
   (b) Saxophone
   (c) Keyboard
   (d) Harp
   (e) Snare
   (f) Voice
   (g) Piano
   (h) Synthesisers
2. Visual feedback
   (a) Strength
      i. Improvement of the user interaction with the system
      ii. Improvement of learnability factors
      iii. Increase of the awareness' level
      iv. Increase of the immersiveness' level
   (b) Issues
      i. Decrease of the focus on the auditory feedback
   (c) Suggestions
      i. Different colour for visually representing the MyoSpat gestures
      ii. Use of primary colours
      iii. Light on the musical instrument
3. Input device
   (a) Wearability
      i. Preference to wear Myo on the left, right or both arms
   (b) Issues
      i. Myo armband causes tiredness
      ii. Myo armband causes itchiness
4. User interaction
   (a) MyoSpat gesture library
5. Embodied interaction
   (a) Interaction with pre-recorded sound

   (b) AM+Delay effect and spatialiser delay line used to generate a tremolo effect through interaction with the musical instrument

6. Use of the system during musical performance
   (a) Interaction with the system through an uncomfortable body posture
   (b) Use of the system during musical performance
   (c) Using MyoSpat without contact with the musical instrument
7. Suggestions
   (a) Improvements to the gestural data mapping process
   (b) A wider palette of gestures to interact with the system
8. Interaction through non pre-coded gestures
   (a) Generation of spatial trajectories by performing the *crumpling* gesture
   (b) Attempt to pitch shift the sound an octave higher by arising the arm
9. Audio feedback
   (a) Auditory perception
      i. Noticeable perception of the audio processing
      ii. Unclear perception of AM+Delay effect on the sound
      iii. Unclear perception of automated sound spatial trajectories
      iv. Enrichment of perception of spatial cues by means of the visual feedback
   (b) Suggestions
      i. Wider palette of audio signal processing techniques
      ii. To make the presence of audio effect more noticeable
   (c) Creative use
      i. AM+Delay effect used for creating more rhythmical patterns
      ii. AM+Delay effect used for layering sounds
      iii. AM+Delay to sonify dance moves
      iv. Pitch shift triggered by standing up from a seated position while playing
      v. Spatialiser's delay line used for bending the instrumental pitch
   (d) Gesture-Feedback Relationship
      i. Relationship between one of the gestures and audiovisual feedback is the strongest among all gestures
      ii. Relationship between one of the gestures and audiovisual feedback is the weakest among all gestures
   (e) System
      i. Strength
         A. Usability in musical performance
         B. Usability in dance performance

Appendix E

Software

1. Myo Mapper 1
   USB hard drive's file path: `BDDPHD18/E-Software/1-MyoMapper-1`
   URL: `https://github.com/balandinodidonato/MyoMapper/releases/tag/1.4`.

2. Myo Mapper 2
   USB hard drive's file path: `BDDPHD18/E-Software/2-MyoMapper-2`
   URL: `https://github.com/balandinodidonato/MyoMapper/releases/tag/2.5.13`.

3. Myo Mapper 3
   USB hard drive's file path: `BDDPHD18/E-Software/3-MyoMapper-3`
   URL: `https://github.com/balandinodidonato/MyoMapper/releases/tag/v3.1.9`

4. Myo Mapper repository
   USB hard drive's file path: `BDDPHD18/E-Software/4-MyoMapper-Repository/MyoMapper-master.zip`
   URL: `https://github.com/balandinodidonato/MyoMapper`

5. MyoSpat repository
   USB hard drive's file path: `BDDPHD18/E-Software/5-MyoSpat-Repository/MyoSpat-master.zip`
   URL: `https://github.com/balandinodidonato/MyoSpat`

Appendix F

Audiovisual Materials

F.1 Software demonstrations

1. Pilot Study no.1
   USB hard drive's file path: `BDDPHD18/F-Recordings/F1-SoftwareDemo/1-PilotStudy-1.mp4`
   URL: `https://youtu.be/tpU_cEjm7KI`

2. Pilot Study no.2
   USB hard drive's file path: `BDDPHD18/F-Recordings/F1-SoftwareDemo/2-PilotStudy-2.mp4`
   URL: `https://youtu.be/3eM-xOLLWo8`

3. Pilot Study no.3
   USB hard drive's file path: `BDDPHD18/F-Recordings/F1-SoftwareDemo/3-PilotStudy-3.mp4`
   URL: `https://youtu.be/GlURLxzvkis`

4. Pilot Study no.4
   USB hard drive's file path: `BDDPHD18/F-Recordings/F1-SoftwareDemo/4-PilotStudy-4.mp4`
   URL: `https://youtu.be/KSCMFQtRaF0`

5. Pilot Study no.5
   USB hard drive's file path: `BDDPHD18/F-Recordings/F1-SoftwareDemo/5-PilotStudy-5.mp4`
   URL: `https://youtu.be/uXBOrchC26s`

6. Live Sound Spatialisation Using Myo and K-Array KW8
   USB hard drive's file path: `BDDPHD18/F-Recordings/F1-SoftwareDemo/6-KArry.mp4`
   URL: `https://youtu.be/-P7YU6NRte4`

7. Music Gesture Beatbox Drawing scene
   USB hard drive's file path: `BDDPHD18/F-Recordings/F1-SoftwareDemo/7-Drawing_scene.mp4`
   URL: `https://youtu.be/i36sAPyVFQM`

F.2 Performances

1. *VoicErutseG* (*Sequenza III*), live at Frontiers Festival 2015 (De Amicis and Di Donato, 2015a)
   USB hard drive's file path: `BDDPHD18/F-Recordings/F2-Perfomances/1-VoicErutseG_FrontiersFestival_Sequenza.mp4`
   URL: `https://youtu.be/xUo1TyymAQc`

2. *VoicErutseG* (*Stripsody*), live at Frontiers Festival 2015 (De Amicis and Di Donato, 2015a)
   USB hard drive's file path: `BDDPHD18/F-Recordings/F2-Perfomances/2-VoicErutseG_FrontiersFestival_Stripsody.mp4`
   URL: `https://youtu.be/r7_NN8ppu3g`

3. *VoicErutseG* (*Sequenza III*), live at EmuFest 2016 (De Amicis and Di Donato, 2015b)
   USB hard drive's file path: `BDDPHD18/F-Recordings/F2-Perfomances/3-VoicErutseG_EmuFest_Sequenza.mp4`
   URL: `https://youtu.be/pQOY-YsKHPY`

4. *VoicErutseG* (*Stripsody*), live at EmuFest 2016 (De Amicis and Di Donato, 2015b)
   USB hard drive's file path: `BDDPHD18/F-Recordings/F2-Perfomances/4-VoicErutseG_EmuFest_Stripsody.mp4`
   URL: `https://youtu.be/gdhnKbxUkUg`

5. *Music Gesture Beatbox*, live at Music Tech Fest 2016 (Savage and Di Donato, 2016)
   USB hard drive's file path: `BDDPHD18/F-Recordings/F2-Perfomances/5-MusicGestureBeatbox.mp4`
   URL: `https://youtu.be/DRFqCXpvfW0`

6. *The Wood and the Water*, studio version (Turner and Di Donato, 2016)
   USB hard drive's file path: `BDDPHD18/F-Recordings/F2-Perfomances/6-TheWoodandTheWater_Studio.mp4`
   URL: `https://youtu.be/gYu4Za-1E48`

7. *The Wood and the Water*, live at Audio Mostly 2017 (Turner and Di Donato, 2017a)
   USB hard drive's file path: `BDDPHD18/F-Recordings/F2-Perfomances/7-TheWoodandTheWater_AudioMostly.mp4`
   URL: `https://youtu.be/3n_2y3OerVQ`

8. *The Wood and the Water*, live at Electronic Music Week 2017 (Turner and Di Donato, 2017b)
   USB hard drive's file path: `BDDPHD18/F-Recordings/F2-Perfomances/8-TheWoodandTheWater_EMW.mp4`
   URL: `https://youtu.be/p0_VR9N4x8w`

9. *Star Cluster* (Devaney et al., 2017)
   USB hard drive's file path: `BDDPHD18/F-Recordings/F2-Perfomances/9-StarCluster.mp4`
   URL: `https://youtu.be/9ToP33Ki2SE`

F.3 Interviews with beatboxers

1. *Participant 4 (a)*
   USB hard drive's file path: `BDDPHD18/F-Recordings/F3-InterviewWithBeatboxers/1-Participant_4a.mp4`
   URL: `https://youtu.be/GGqON4K7Dkw`

2. *Participant 4 (b)*
   USB hard drive's file path: `BDDPHD18/F-Recordings/F3-InterviewWithBeatboxers/1-Participant_4b.mp4`
   URL: `https://youtu.be/rnwgzXyv7c8`

3. *Participant 4 (c)*
   USB hard drive's file path: `BDDPHD18/F-Recordings/F3-InterviewWithBeatboxers/1-Participant_4c.mp4`
   URL: `https://youtu.be/JnSPnw2lx7Y`

4. *Participant 4 (d)*
   USB hard drive's file path: `BDDPHD18/F-Recordings/F3-InterviewWithBeatboxers/1-Participant_4d.mp4`
   URL: `https://youtu.be/HD386M7_qQA`

5. *Participant 4 (e)*
   USB hard drive's file path: `BDDPHD18/F-Recordings/F3-InterviewWithBeatboxers/1-Participant_4e.mp4`
   URL: `https://youtu.be/g-SMS6Vjwhc`

6. *Participant 4 (f)*
   USB hard drive's file path: `BDDPHD18/F-Recordings/F3-IntervwWithBeatboxers/1-Participant_4f.mp4`
   URL: `https://youtu.be/LIrJpMdpcrY`

7. *Participant 5 (a)*
   USB hard drive's file path: `BDDPHD18/F-Recordings/F3-InterviewWithBeatboxers/2-Participant_5a.mp4`
   URL: `https://youtu.be/s1OPjA_tzHo`

8. *Participant 5 (b)*
   USB hard drive's file path: `BDDPHD18/F-Recordings/F3-InterviewWithBeatboxers/2-Participant_5b.mp4`
   URL: `https://youtu.be/MdWjNw-UCfo`

9. *Participant 5 (c)*
   USB hard drive's file path: `BDDPHD18/F-Recordings`

/F3-InterviewWithBeatboxers/2-Participant_5c.mp4
URL: https://youtu.be/o_M2iBNIER4

10. *Participant 5 (d)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/2-Participant_5d.mp4
URL: https://youtu.be/kpp8Xrp6M44

11. *Participant 5 (e)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/2-Participant_5e.mp4
URL: https://youtu.be/vPyf-v805N0

12. *Participant 5 (f)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/2-Participant_5f.mp4
URL: https://youtu.be/tzrtUvlwS7Q

13. *Participant 5 (g)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/2-Participant_5g.mp4
URL: https://youtu.be/5xZuvTD8S8U

14. *Participant 5 (h)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/2-Participant_5f.mp4
URL: https://youtu.be/eO_Nz26WcWM

15. *Participant 6 (a)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/3-Participant_6a.mp4
URL: https://youtu.be/XcQcQQUVsq0

16. *Participant 6 (b)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/3-Participant_6b.mp4
URL: https://youtu.be/TA6aw8TP1Sg

17. *Participant 7 (a)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/4-Participant_7a.mp4
URL: https://youtu.be/RCxQbCJ9Cek

18. *Participant 7 (b)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/4-Participant_7b.mp4
URL: https://youtu.be/18TFvWUhd38

19. *Participant 7 (c)*
USB hard drive's file path: BDDPHD18/F-Recordings

/F3-InterviewWithBeatboxers/4-Participant_7c.mp4
URL: https://youtu.be/9qix2AXO5wU

20. *Participant 7 (d)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/5-Participant_7d.mp4
URL: https://youtu.be/v-I4x9gaztg

21. *Participant 7 (e)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/5-Participant_7f.mp4
URL: https://youtu.be/27kiDutv91s

22. *Participant 7 (f)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/5-Participant_7g.mp4
URL: https://youtu.be/Xm1CSqjlBVc

23. *Participant 8 (a)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/5-Participant_8a.mp4
URL: https://youtu.be/Xw6J947vYT8

24. *Participant 8 (b)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/5-Participant_8b.mp4
URL: https://youtu.be/rDDRMlQ4D14

25. *Participant 8 (c)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/5-Participant_8c.mp4
URL: https://youtu.be/1O9gEV3jm1o

26. *Participant 10 (a)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/6-Participant_10a.mp4
URL: https://youtu.be/Hw_HkZXfgWY

27. *Participant 10 (b)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/6-Participant_10b.mp4
URL: https://youtu.be/hoiX1SpjVcQ

28. *Participant 10 (c)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/6-Participant_10c.mp4
URL: https://youtu.be/tfNaEYKwL4E

29. *Participant 10 (d)*
USB hard drive's file path: BDDPHD18/F-Recordings

/F3-InterviewWithBeatboxers/7-Participant_10d.mp4
URL: https://youtu.be/5uFpg35L57Q

30. *Participant 10 (e)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/8-Participant_10e.mp4
URL: https://youtu.be/i6P5DZD5gBY

31. *Participant 10 (f)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/8-Participant_10f.mp4
URL: https://youtu.be/iDYGNDffcRg

32. *Participant 10 (g)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/8-Participant_10g.mp4
URL: https://youtu.be/PY-gfJSh3FM

33. *Participant 10 (h)*
USB hard drive's file path: BDDPHD18/F-Recordings
/F3-InterviewWithBeatboxers/8-Participant_10h.mp4
URL: https://youtu.be/60bXg3Tkw28

F.4 MyoSpat evaluation

1. *Guitarist playing the guitar using MyoSpat0.6 with the Myo on the left arm*
USB hard drive's file path: BDDPHD18/F-Recordings/F4-MyoSpatEvaluation/
1-MyoSpat06-GuitraMyoLeftArm.mp4
URL: https://youtu.be/kyzlbJITpGA

2. *Tremolo effects generated using MyoSpat 0.6*
USB hard drive's file path: BDDPHD18/F-Recordings/F4-MyoSpatEvaluation/
2-MyoSpat06-Tremolo.mp4
URL: https://youtu.be/tTYUbqV-2Oc

3. *Crumpling gesture on Sax using MyoSpat 0.6*
USB hard drive's file path: BDDPHD18/F-Recordings/F4-MyoSpatEvaluation/
3-MyoSpat06-CrumplingOnSax.mp4
URL: https://youtu.be/ainhq4-M0cU

4. *Dealy to build up harmonic elements using MyoSpat 0.6*
USB hard drive's file path: BDDPHD18/F-Recordings/F4-MyoSpatEvaluation/
4-MyoSpat06-Harmonics.mp4
URL: https://youtu.be/dMzfQ_d6sSI

5. *Tangible interaction using MyoSpat 0.6*
USB hard drive's file path: BDDPHD18/F-Recordings/F4-MyoSpatEvaluation/
5-MyoSpat06-TangibleInteraction.mp4
URL: https://youtu.be/0iBEj_DawqI

6. *Lower gesture by standing up from a seating position while MyoSpat 0.6*
USB hard drive's file path: `BDDPHD18/F-Recordings/F4-MyoSpatEvaluation/6-MyoSpat06-LowerGestureStandingUp.mp4`
URL: `https://youtu.be/mUyOymFQMsQ`

7. *Lower gesture while changing the posture of their body using MyoSpat 0.6*
USB hard drive's file path: `BDDPHD18/F-Recordings/F4-MyoSpatEvaluation/7-MyoSpat06-LowerGestureBodyPoture.mp4`
URL: `https://youtu.be/kKMn7oqmJ4w`

8. *Lower gesture while playing the keyboard using MyoSpat 0.6*
USB hard drive's file path: `BDDPHD18/F-Recordings/F4-MyoSpatEvaluation/8-MyoSpat06-LowerGestureKeyboard-7.mp4`
URL: `https://youtu.be/nPsRK-KRVOM`

9. *Dancer using MyoSpat 0.6*
USB hard drive's file path: `BDDPHD18/F-Recordings/F4-MyoSpatEvaluation/9-MyoSpat06-Dancer.mp4`
URL: `https://youtu.be/QiRRpSge7xA`

10. *Sax player generating sound trajectories using MyoSpat 0.6*
USB hard drive's file path: `BDDPHD18/F-Recordings/F4-MyoSpatEvaluation/10-MyoSpat06-Trajectories.mp4`
URL: `https://youtu.be/yjpTl8Gmqwk`

11. *Music improvisations using MyoSpat 0.6*
USB hard drive's file path: `BDDPHD18/F-Recordings/F4-MyoSpatEvaluation/11-MyoSpat06-Improvisations.mp4`
URL: `https://youtu.be/8Wl1boPo8Zg`

Appendix G

Interview to Beatboxers

G.1 Scripts

G.1.1 Participant 1 - Grace Savage

Interviewer How would you define a movement?

G. Savage An action through space

Interviewer And a gesture?

G. Savage It has more emotion attached to it or emotional intention. Movement can be applied to any object, that's a movement, but it doesn't mean anything. With a gesture, you try to explain something.

Interviewer When you do beatboxing do you employ some acting skills as well?

G. Savage It comes to the confidence I think. When I first started beatboxing, I was so scared and nervous to get a microphone, it's scary to do. So your safeness it's to look at the ground, do not look at the audience, focus on the beat you are doing, and then it will be over. The more you become comfortable with performing and confided with your own skills and abilities, the more you start to look at a new style, you start to relax. It is more engaging for an audience. Acting gave me more confidence. Acting is a great base for anyone, even if you are into theatre. It helps to get up and to articulate things confidently. I think that it inevitably that had an influence on my skills and all aspects of my of performing. Still, it is not something I consciously though thought of that.

Interviewer You worked on a show called *Blind*. When you worked on it, was there a choreographer directing you on stage or you were everything by yourself.

G. Savage Blind was my solo show, so it was just me on stage, but I worked for a theatre company called the Paper Birds. They are a sort of experimental feminist theatre group from Leeds. So most of the time, it was me and Jemma (the director) rehearsing in the room and when we were testing ideas and stuff. We spent a lot of time just messing around with voice and emotion so she would lie down on the floor and read out a word. I had to react with a sound, just on impulse. So before we had that natural story and structure of idea of what we wanted, we were just playing around ideas of body and the voice.

Interviewer What do you mean when you say "playing around ideas of the body"?

G. Savage I went through a range of emotions and different sounds, for example, crying or laughing. If you just hear it, you can't really tell the difference. So we did this routine where I was scared of something. I was making scary sounds, and suddenly it will turn into laughter, and then will turn into shock. We did this five times and recorded it. I looked like a mental patient but, the sounds mixed with the body gestures all the way, all your body tenses up, suddenly is really readable. If you just ear the sound isolated without the body movements, you may not be able to please it so much.

Interviewer Was all linked, the sounds and the movements?

G. Savage Yeah.

Interviewer You worked on the production of a second show: *Life in progress*. Was there a choreographer for the dancer and you where beatboxing or you had a part on stage as well?

G. Savage That's the show I've been touring the whole year. There are four different pieces. I'm in the first piece by a choreographer called Akram Khan. He uses a lot of Indian classical stuff with contemporary dance. There was an Indian percussionist a violinist and me sort of making vocal sounds and using a loop-station. So, I'm very much behind the stage, behind a sort of see-through fish curtains and she is on stage dancing.

Interviewer If you were the choreographer, would you have done something different?

G. Savage There is an interesting thing that we have done before. Sometimes the movement instructs the music. You see the movement, and you go "how can I interpret that with my sounds?". Other times, the musicians come up with the music, and the choreographer goes, "right, how do we fit to dance to that sound and sometimes is a bit of friction?" Because musicians are like: "I can't do anything you don't come with the choreography", and the dancer will say "I can do nothing if you don't come with the music". This, sometimes, creates friction in that way. But this show they already rehearse few ideas, and they had an idea that they want to be technology versus evolution and humans and not technology taking over. It has got this epic sort of landscape of the heart of the universe and scuttling like a little animal. They used to have a microphone that sort of moved and followed her around and produced the electronic sounds. I made sort of electronic sounds. So yes, they both inform each other.

Interviewer In terms of choreography and performance, have you ever planed a choreography for one of your shows?

G. Savage Yes, in another show, which has got nothing to do with emotions or these intentions of the scene. It is literally: you see this line, and you go and stand here. You stand on a line, and you put that cup over there, and there is a stage manager who writes those things down and has diagrams known as the actors go there the actor goes there. But some directors are very intent on making sure that the blocking is really rigid, and the same every time. Other directors would say if you are not feeling moving over there on that line and you feel like you wanna walk that way this time, and be in the moment you can. So there are kinda two ways of directing.

Interviewer When you say, the actor doesn't really feel about moving there, towards left or right whatsoever. Why would you move there or somewhere else?

G. Savage You end doing a show, and you do the same routine, the same thing, eight times a week. Instead of the same line, you do the same movements, and the job of an actor/performer is to make it seems as it is fresh, it is original the first time. That can be difficult if you do the same show for three or four months. I totally get why some actors just like comfort. Like, look I know this routine, I can do this same routine for 10 years, but you need to keep it fresh you need to change it a little bit. You need to warn the other performers about you are gonna do that. Some actors are just all over the place, and it is not fair, it's a selfish way of acting because you have to be really giving to the other persons. So tonight, I'm supposed to bang this phone on the table, but I might check it over there instead. Or I may smash it or I may I don't know. Otherwise, it loses the authenticity of the action.

Interviewer Is it something that comes out from your position on the stage, the other performers' position, or it may be from what you are singing?

G. Savage It depends a lot about how you are feeling on that day. And that's another thing, you could, you have to go and perform all day every night, no matter what's happening, you know. Your mum could have just died, I know that's really extreme, and horrendous things could happen to you, but you still have to go on stage and perform. All those kinds of things can affect and also the energy of the other people performing on stage can influence. The audience has a huge influence. For example, after you do the show over and over again, one night, the audience can hate it. Another night they could love it, or they can be really quiet, they can be really loud, and performers tend to take it really personally.

Interviewer Or is not like, I feel that this sound should go along with my movements, while I'm moving fast or slow.

G. Savage It depends when you are talking about acting or beatboxing or singing because they are all completely different.

Interviewer In *Blind*, did you feel like some sounds were going out better performing lying down on the floor or running or doing something else?

G. Savage There is one thing at the very end of the show, I find my voice in a cheese way. I beatbox on the microphone thinking all the sounds I have heard throughout my life all, mashing them into 30 seconds of really intense beatboxing, and throwing myself around. Meanwhile, the base is really loud. There is no musical structure to it, but I just let loose. Sometimes, I really go for that, and It's exhausting. Even though it's 30 seconds, sometimes, I fall to the floor because I know that's my cue and then the lights go out, and I drop to my knee. And that feels a bit forced. I've done that show a hundred times or something. So that's hard to have real and genuine passion every time. If I really commit myself to it, and I'm going to put myself in this physically and vocally 100%, I always get better results.

Interviewer Did more acting influence beatbox or the other way around?

G. Savage I think the acting influenced the beatboxing more. Because it is such a technical and specific thing.

Interviewer What particular aspect of beatboxing?

G. Savage Just performance elements I think. There are certain things when you have battles. You get judged about how musical you are, how technically you are, and so on. How well you perform, because loads of beatboxers just come on stage, look at the ground, and they do their things and the audience it's really boring. So, I tend to look at the audience eye contact looking at other people, and I think it makes a huge difference to performance.

Interviewer When you rehearse to perform better, to have a good presence on stage, do you perform particular gestures?

G. Savage Sometimes, but I think, most of the time it's involuntary. You don't realise it, you are beatboxing and before you know it your hands are doing weird things and most of the time I do the just to keep myself in time. I know other beatboxers that beatbox with his hand behind his back and sort of waves like that. And occasionally, I would do that. There are some quite rigid beatboxers, a guy called Alem who won the World Championship last year, he is like [Grace Savage assumes a rigid body posture, and start to move her hand forward and backwards with the palm parallel to the body and straight, rigid fingers]. He is really stuffing and proud around the stage. It is just really interesting how each person has his personality and manifests itself.

Interviewer Are you aware of your body movements when performing?

G. Savage It depends because when you have got a microphone, you have got just one hand free. When you are doing it without you, have got both.

Interviewer Does the microphone influence the performance?

G. Savage I think so. Yeah. It came to my attention that beatboxers did these things when I was in a group called the Vocal Orchestra with Shlomo. He made a very good point. There were eight singers and beatboxers together, and if you would have performed with the microphone, you can't see people's facial expressions. Instead, when people are singing, you can see what they are doing, and you can see their emotions, but if half of your face is covered. Thus, the only indicators that the audience has about the sound you are doing are the hands. Let's say eight people and just one is doing the base, and one person is doing the hi-hat. The only indicator that the audience has that I'm making that sound is that my mouth matches the sound. He [Shlomo] made us describe the sound with our hands in performance. It was from that moment on I started doing it more on stage. I realised that the audience would connect to my sounds in that way.

Interviewer Do you try to match what you play with your mouth with gestures to try to express the musical concept?

G. Savage I don't say I do consciously now, but I think that a few years ago, I did and it became a natural thing. I don't think I think about it ever. I do know when beatbox bass sounds I move the hand from a low position to a higher position, in respect to the chest, with palm upwards and moving the fingers. Sometimes this movement of fingers, like if you want to imitate the loose leaps sound. Most beatboxers try to emulate the movement of the leaps with their fingers. There's another guy who I call Hobbit who does that quite a lot, which is like "woo sister".

Interviewer Do gestures help you perform?

G. Savage Performing without moving, it's not as fun. It's like you are not enjoying yourself. I mean you can do it, but even I feel my hands are doing this [G moves its hands pointing with the index finger following a rhythmical pattern].

Interviewer You performed on the street, theatre, festivals. loads of different venues. Does the venue influence performance?

G. Savage Yeah. Performing in an intimate venue where I can see people's faces and eyes, I get more nervous. If I was performing on this tour *Life in Progress* where the venue was between two and four thousands people each night, which is the most I have ever performed in front of, but because it's not my show. Because I'm behind a thing and I'm doing the same 15 minutes piece and like never get nervous. It doesn't really affect me, I don't really get into it. It just becomes routine. Even when I see four thousand people. If I was to perform just me in a room here with twenty people watching me with no lights, no equipment I'll be really nervous. Because there's intimacy and they can just focus on me, feel my breath, and I can see their faces, that yeah.

Interviewer Does the physical size of the venue affect your gestures?

G. Savage Yeah. I guess I would make my gestures bigger, with a bigger audience. Because you know, there are people on different levels, so, you have to perform even for the people up there as well. If there are few people in a small room, you can make smaller gestures that they can notice.

Interviewer Before you said something really cool, which is that a movement is something which happens in space. Do you see your voice as a sound source in space?

G. Savage Yeah, totally. A lot of it comes down to having a good sound guy as well. I have done some performances outside the street and there are no walls to enclose the sound. If you have got like a little shit amp you can beatbox with all your power, but the sound that comes out, it's just a horrible fart sound, and the audience is like: "oh, you're not very good!" and they walk away. But you can be in a room with, and the sound guy knows what he is doing. You can make the tiniest sound with the microphone and reverberate it. It makes a huge difference.

Interviewer Would you like to control the spatial location of the sound during a performance, like with panning effects?

Interviewer Yeah.

Interviewer Would it influence your way of performing?

G. Savage I don't know, because if it's not something I'm not experienced. I wouldn't be able to tell. But that happens anyway with sound, doesn't it!? In certain rooms or in particular spaces the audience misses a lot of the bottom-end. Some people get the full sound others don't. It depends on what position.

Interviewer I heard your piece *Diamonds on the skin*, and there is loads of environment, of reverberation. But then, I saw a live performance of that song. And of course, a live performance cannot always be as the studio version, no one does that. In your case is it for technical reasons or you choose to do it in that way, for example, to distinguish the studio version from the live version. Particularly, I have seen you a performance of yours with a loop-station, and all the environmental sounds that are on the studio version they weren't there. Did you do it that way so you couldn't reproduce the whole thing live, or you wanted to be different?

G. Savage A bit of both. Normally I've a full band. I've got a DJ, a drummer and a bass player. So when we perform live, we perform the track version, as it is. But for that gig, I was booked as a solo performer. So that was my sort of my loop-station remix song I wouldn't normally present that song in that way. It's my drum and bass remix version of that, as a creative sort of interpretation.

Interviewer When you were working on all those environmental effects and reverbs. Is it something you wanted to do on your voice or is something that another guy advised you to do so?

G. Savage I'm working with a producer called Dee Adam, a female producer. She is a brilliant producer, and because my voice isn't like Diva Jassy J kind of voice, so it requires I don't know, smooth kind of tone. So she effected it by putting a lot of space reverb and stuff on it. So that's all down to her. She is really good in vocal production.

Interviewer Yes. She is, the audio processing on your voice is really enjoyable. You use a loop-station as well, is it something you feel comfortable to use.

G. Savage I've used the RC-300, which was a board for guitarists, the pedals are really heavy, so I've upgraded to this one which is a lot softer [she refers to the RC-505] to touch. It's like a toy, it's really lite. I would say 60% comfortable with it.

Interviewer Which is the other 40%?

G. Savage Only because there is so much of it to be explored. There are loads of ways you can use it. It takes a long time. It is like learning Logic or something like that. And you just have to be a geek for a long time.

Interviewer So it is not really that you are struggling with the interaction with the hardware. Like for example, I'm beatboxing, and I'm so bored to press a button there because I want to move my hands or something like that. Are you annoyed to be stuck in a fixed point?

G. Savage I do find that annoying. And watch an audience member who is a loop-station artist, and they are standing on stage, just pressing buttons. and clicking things on the laptop. I think that's really boring. Especially if you are not a musician or a person who does not understand that what they are doing is so impressive. I think they need to have a balance of being able to do it for a little bit but then coming away from it and performing to the audience. That's what they are paid for; to see you that engage with emotion and naturally perform. That is my personal preference. I don't like people pressing buttons. But also, it is very engaging to see someone beatboxing, recording and adding layers on the top, and that's like a fun thing. But if you two long minutes building a track per every single song, I just personally find it a bit boring. So, I always make sure that I'm building a track it's only for a minute-long before I actually get into the performance. But it can be quite limiting because I give myself fewer options to bring back in.

Interviewer Did you ever get to the point where, you are performing this and then should press that button, but you don't want to, but you have to.

G. Savage Yeah. Do you know Shlomo, he is a world loop-station champion, and he does a lot of improvised loop-station. He does loads of improvisation stuff. And he has been developing a microphone, where the loop-stations' buttons were on the microphone. So he can walk around the stage while he is looping. So he is not stacked at the machine. Something like that would be really cool.

G.1.2 Participant 2

Interviewer For how long have you been beatboxing?

Interviewer Five years.

Interviewer Are you right or left-handed?

Participant 2 I'm right-handed.

Interviewer With which hand do you hold the microphone?

Participant 2 Usually with the left hand. Because what a beatboxer would tend to do is that they are moving their hands like crazy. But, I just keep on time and all those kinds of stuff in my head, when I use my hand more. That's why I use my left hand. Because my right hand, my right-hand helps me to keep on track.

Interviewer When using the microphone, does the microphone affect your gestures?

Participant 2 Yes, it affects my gestures, because if I'm holding the microphone so I could be using two hands which would obviously drive to change. Every person is different, and it's totally just something that keeps us on time, sort of conductor. It's different when you have two hands to kind conduct yourself. When you have only one, you have to do different movements to keep yourself on time. I mean, for me personally, none of the movements are planned for the most part. They are not always the same. They are not specific movements.

Interviewer Do you use gestures more to conduct yourself only or also to describe the timbre of the sound?

Participant 2 Is not really so much to do with the sound. It could be like a hardcore, or the most old school hip-hop, and I'll be still doing the same movements. Mostly is to keep on time for sure. Sometimes, when it is specifically a performance, I'll switch from keeping on time to you know getting the audience hyper or something like that. Still, for most of the time, I'd say to keep on time. I may not be hitting at a specific tempo like I may be hitting just the off-beat, I may be hitting like the second and the fourth beat. I may be hitting one specific beat like the third. It really depends on what the beat will be and the time signature of the beat. If you talk to beatboxers if they are rehearsing by themselves, you'll never find us doing it with each other. It is more the performance aspect of it. Like to get into the audience to go crazy, or when you drop crazy beats. A lot of people go down a little bit when they are doing like a base sound, a lot of that it is like adding visual elements to the performance. Because people can only listen to music, right? So to add some sort of visual aspect to it.

Interviewer Do gestures facilitate expressing your musical idea?

Participant 2 Yeah. It is for sure because even if you see the reverse of it, if you watch some of the videos on YouTube, even for experienced beatboxers it is hard to tell who is doing which part of it, especially if there are good beatboxers that could be doing any part of it. So, unless they are making a sound that everyone knows, specifically their sounds then anyone could be doing their part interchangeably. And a lot of the times, in battles and stuff, that's even done where you have two beatboxers. One would be doing basses, for example, and one would be doing the hi-hat. Then switch to show kinda that they can do all parts and that they can snap parts, just to show how good they are basically.

But the point is that they make each other sounds. Usually, the better they are, the more identical it is. But like so, it is the exact same thing, because they can do exactly the same thing. It is hard, especially when they have the microphone in front of their mouth. When you don't, it becomes a little bit easier.

Interviewer Are you aware of the gestures that you use during a performance?

Participant 2 Yeah, I would say that I use a limited range of gestures, just because at the end of the day, the gestures aren't the focal point of the show. Like a lot of it is gonna be kinda reflective to the sound, like what you said, people using hand like that [Participant miming few gestures]. So like if I'm making an inwards kick sound, I'll do something like this [Participant moving the hand backwards]. If I do a forward then a backward sound I'll move my hand forward and backward like I'm doing with my mouth. for my performances. Yeah about general motion that I do and they usually will appertain to the sound. They do not necessarily have to be. It is usually not planned. Gestures are going with the feeling of the music, with the feeling of the sound, it needs to change, it can't be static unless you really rehearse it.

Interviewer Do you use technology as well?

Participant 2 I don't yet. That's something I'm kind deciding and figuring out at the moment because I love using loop-station as the ultimate toy of the beatboxer. Right now I don't have one, so I only used it a couple of times at a friend of mine, so I don't actively use them now. Loop-station changes the performance, to be honest. It is more like cheating, that to go and learn to make two sounds at once, rather than learning using the machine to layer the two sounds. You know, so it is just that when you start going toward that kind of stuff like you have much more layers, and you are building a song that kind of thing. Whereas, when you are beatboxing, it is much more impressive the sounds that you can do with your mouth. Like a loop-station, to be a very good loop-station performer, you don't need to have many sounds, they do not need to have the same capacity as beatboxers. But when someone it's just beatboxing, they really need to be a very high-level beatboxer, to make it as high quality of the show. It is much easier to use a loop-station than it is to figure out how to make three sounds on top of each other beatboxing. It's just different. Nobody would find a loop-station VS a normal beatboxer in a battle. Like it is just not the same thing. It just depends what you wanna do, where you wanna go, how you want to deliver the performance. Like authentic beatboxing, electric song cover, which case it's much easier using a loop-station.

Interviewer Would you consider a loop-station like a constraint to your performance?

Participant 2 It is just something you have to deal with. It comes very much down to the person. You'll find some people even whether they have got a loop-station or not. Some would move around some won't. Like you know, the best beatboxers have understood how to perform as well. They move around from one side to the other side, move to the centre for the crazy part. You have to learn to do those stuff, and the more advanced you become as a performer, not as a beatboxer. As a performer in front of people, that's when you start changing all those stuff up. Like to do all those kinds of stuff so, most people won't. Some may not be able to do that stuff, whether it's because they have a loop-station or not. Personally, I'm always pretty good moving around the stage and stuff like that too. Being able to move during a performance, get the audience engaged, add that visual moving around, keep them engaged. So being able to do that while I'm moving not only why I'm beatboxing, helps the performance even more. You do not have to beatbox a little bit and then go to the one spot of the loop-station, and then start moving again afterwards. Or even during the loop-station, your record certain parts, and then you walk around and then come back and continue. So, being able to not be tight to the one-spot at all would be helpful, but it's only helpful if you are a performer that moves around.

Interviewer Would you use gestural technology during a performance?

Participant 2 I may not go for it right away. Let's say accidentally. I make that gesture, because like my gestures aren't planned I just do it. Like if you watch someone free-styling and their hands are going all over the place and the more they get into it, the more, the less they realise you know. So, especially for novel performers, they have that sense that you go on stage and like when you are speaking in front of people, you get nervous and stuff like that. And usually, it doesn't go away, but you can just become more aware of it and like more control over that. So if you performed a bunch of times already then adding that wouldn't be so bad, because you kinda have control of your gestures. But if you are a novice performer trying to control your gestures would be very hard.

Interviewer Talking about music production. Would you use spatial effects?

Participant 2 No, like I wouldn't do it at least because it ruins a part of what beatboxers are doing, the authenticity of all of this that I'm doing with my mouth. So, the more you add to it, to change instead then clean it up it makes it like if you are not doing it with your mouth any more. So If I did something like that, you would see like they have the loop-station and then they have like an effects machine. If you are doing that you are obviously using an effects machine, which is fine, but it's just not organic. It doesn't give that feeling that I'm trying to express. Beatboxers who start to change their video with effects, but then you meet them in person, they cannot beatbox like that you know. Reverb specifically is a pretty soft thing. Reverb it is not that big a deal. But it is more just like when they are affecting the sound in a way that they couldn't do. So like making the sound unnatural. For example, with distortion, you could do a little bit of distortion by yourself, with your own voice, but not the same one the works on the computer. So when you distort the computer that's kinda a false thing. But distorting the beatbox a little bit that's not so bad. That's your own thing. But you use a computer to do it in a way that you can't do it, it's different. It's kinda the same way you use a microphone to reach the back of the room that's not affecting your performance, it's just because the room is too big. So using reverbs to make the room better and the bass lower because of effects the room better that's ok. But turning up the bass so that you are doing it at this pitch, like that the beat is going out ten times lower than you can actually do. That's when it becomes false.

Interviewer You were talking about the sound in a room to hear better. Would you also move the sound around the room using spatialisation effects?

Participant 2 Yeah that would be cool. Forwards to backwards. Frankly, I would have a very little bit of interest, just because at the end of the day people have ears at the left and right of the head, and so they wouldn't hear it as well. But left to right is something I'll be interested in. But I want a layer of complexity, like I wouldn't just one stream of sound, like a stream of sound go left and right. I wanna be able to put a bass sound on the left and 'k' sound on the right because in that way I could effect in different ways. So it's kind of a layered complexity instead of just direct a sound. Would be cool to have separate streams and for example being able to turn them down, by sound instead than the whole thing. Maybe, I would like to move a bass note like a constant [the interviewee beatboxing a kick sound with a regular tempo] across the audience and adding another sound to move to the other side. So it is like to control them as individual streams instead of, the one straight recording kind of thing.

Interviewer Does the venue affect your performance?

Participant 2 It does effects for a lot of reasons, even one being the playback issue. So if you have got one speaker at the side of you, what you are hearing is a lot more what's more in your head. But when you are like on a stage environment, you have speakers that face you and playback, so you can hear what you are doing. So it happens a lot because if I can't hear what I'm doing, I can't change what I'm doing. In a stage environment, so because of the speakers facing the audience, the speakers on the sides, and the speakers facing me, and I can hear the playback. So when I can't hear myself and let's say, whatever set the beat for, say 5 bars for example, and I mess up one of the sounds a little bit of the first bar, then I can change it for the next bar. But when I'm on the street, I do not have any playback, and I can't hear what it sound like, it is even more organic. Just like doing the beat, hoping that they are the best possible, there are also some of the effects that you do as a beatboxer to the audience. You cut out all the sudden the bar or rest, or with no sounds. When you have that in a performance venue, then it's much harder hiding, because it's the only sound that's happening, besides people talking in the audience. So when you go quiet, everything goes quiet. But when you are on the street performance, it's less effective because it affects only people that are watching or listening to you. If you stop it goes quiet, but it goes to the street level quiet because everyone is still talking, walking around whatever, street cars driving by, whatever the case is. So, I wouldn't do long pauses if I was on the street

VS when I'm performing like in stage setting, you know. There are all little things like that that it does change a little bit, and like I was saying before, I gotta catch people's attention much quicker. Then when I'm on the street than when I'm on the stage. I'll be doing a lot more and a lot of more various genres, or more complicated beats without as much introduction. Whereas on the stage, I can spend 5 minutes on one bit only, to build it up to take it down, to do bass parts, you know those kinds of things.

Interviewer Since you started beatboxing, did you feel your needs different during the performance like, towards technology.

Participant 2 To be honest, technology is something I'm still deciding to have or not. It's a hard decision for me because, on the one hand, I'm a beatboxer, so what I do is to make the sounds come from me. If I use a loop-station or something like that, it will almost change the stuff that I did. So, if I wouldn't have started to do the loop-station stuff, I feel I wouldn't have challenged myself as much as a beatboxer. You do not have to do as much as a beatboxer, you do not have to challenge to put sounds on top of each other, because you just loop them. I feel like I may get a loop-station and I may do loop-station performances, but I don't think that they would ever overtake the beatboxing itself. So you are doing an equal amount of beatboxing shows by themselves as beatboxing with technology. I don't want the technology to take half of my shows. I want stuff that enhances the performance and doesn't change the performance. But it changes and enhances as one thing.

Interviewer If you were to use gestural technology, what gestures were to use for what sound?

Participant 2 It is very difficult. When you start beatboxing, you start to move your arm as well, but when you start to speed up, maybe your arms won't follow your mind any more. It may happen when you make two sounds at once if you are doing just the base or just the kick that's two different sounds. But if you are doing the bass the kick and then a second sound, which has two sounds together then you need a third motion, so I'd be worried to run out of motions to do quickly, to switch quickly. I like the idea, but it has to be done very, very elegantly. If you do something in between, doing it without sound equalisation and with a sound engineer. Something in between because then the average performance may increase without hiring a sound engineer. Because if you are doing a significant amount of shows, you can't hire a sound engineer. It doesn't make sense to practice with someone once a week to only have two performances a year. It just doesn't work. No sound engineer is going to volunteer his time like that unless you pay him. It is too much time commitment and less money involved and passion involved. Let's say because I'm studying or I'm working I don't do a ridiculous amount of shows per year. Even if I do a couple of shows per month, it doesn't make sense to hire a sound engineer.

G.1.3 Participant 3

Interviewer Are you Right handed or left handed?

Participant 3 Left handed.

Interviewer With which hand do you hold the microphone?

Participant 3 I hold it with my right hand.

Interviewer Do you mostly perform using the microphone or do you perform a cappella?

Participant 3 Well, because I was part of an a cappella group, so when I'm with the group a cappella but when I'm solo with.

Interviewer For how long have you been beatboxing?

Participant 3 Just over two years.

Interviewer Do you also perform with other beatboxers, in a team too?

Participant 3 Yeah, with a singing group.

Interviewer How many of you?

Participant 3 Twelve.

Interviewer Is it a mix between singers and beatboxers?

Participant 3 Yes it is.

Interviewer While performing do you move your body or do you play gestures?

Participant 3 Yes definitely, I do move to keep time or to visualise what you are hearing. Gestures help to coordinate what you are doing. Let's say you have a kick drum here, the hi-hat here and the snare here. If you can do a rhythm with gestures, you can do it with your voice. I mean if you can visualise it ,you can do it with your mouth. For example, if you cannot do a very quick hi-hat but then you do play virtually, you just start to do it subconsciously. Also, you can put your hand low for a bass sound and backward for a reverse sound. So if I was doing a hi-hat sound in reverse, I would the hi-hat gesture in reverse.

Interviewer Which part of your body do you most move?

Participant 3 Generally my left hand.

Interviewer Is it because you hold your mic with the right one?

Participant 3 Yeah

Interviewer Are you aware of the gestures you perform?

Participant 3 Yea, they tend to be fairly similar. I can perform something for a kick drum, it keeps the time better and you may put your hands like on a drum-set, so you do [playing drum beats and gesturing as playing a drum], so you play gestures as you would play the drum, and I kinda try to do that mostly.

Interviewer Do you choreograph your performances or you move as you feel on the spot?

Participant 3 Yeah it is something that just happens. If you do gestures, you cannot do them in the same way. If you are performing them straight away or practice them you look like an idiot anyway. So when you are beatboxing you move your body naturally.

Interviewer Do you use technology in your performances?

Participant 3 I do use loop-stations and things like that.

Interviewer Do you like a particular one, or do you think that they are pretty much all similar?

Participant 3 I tend the BOSS's RC505, it is the one with buttons and sliders, and buttons help to keep yourself coordinated.

Interviewer Does the loop-station constrain your performance?

Participant 3 Well I think that the loop-station demonstrates a kind of aspect of it, when you perform without loop-station you are kinda showing off how much you can do at once, without any effects, when you are using the loop-station you are showing off a continuous thought process, because you have to think everything you are going to do. You are showing off the accomplishment of that rather than each single beat you can do. If you do too much in a loop-station, you can ruin it, you are really to break it down and build it back up.

Interviewer Is the physical interaction with the loop-station a limit to the performance?

Participant 3 It didn't bother me so much really, it's just something that is there, it doesn't really matter what it is. With it you are able to do more things, so all you really need as a beatboxer is something that you can use to combine your strength. So you can combine more that you can do with a solo performance. So, if I would have another physical controller it wouldn't make a big difference. It would give back the same kind of effect.

Interviewer Have you ever recorded a performance or a piece of music?

Participant 3 I did have one.

Interviewer Have you ever processed your voice for your recordings if not would you do it?

Participant 3 I do like to layer and stuff but I try to keep everything as clean as possible. I try not to manipulate the recording as much.

Interviewer Is it more because of a stylistic choice or it is because you do not like technology itself?

Participant 3 Yeah, it is more stylistic. I mean if you rely on manipulating the sound afterwards the sound during recording would not be as good.

Interviewer Have you used or would you use sound spatialisation?

Participant 3 I did what you say, I also did a module at school, a module called acousmatic music, so all those kind of thing spatialising everything in those cool environment

Interviewer Have you ever considered using it during live?

Participant 3 Yes, with my choir, we perform a kind of choreography and we are aware of the fact that we bring the music somewhere. Sometimes we do a mash-up or a medley, and so we sing one song in over here and then another over there, and the beatboxing in the middle to combine the two.

Interviewer When you do that are you aware also about the soundscape you are generating?

Participant 3 Yeah, you have to be aware of where the sound is coming from. I'm working on spatial sounds when the people are more spread out.

Interviewer You talked about visualisation before. Do you also apply it to sound spatialisation?

Participant 3 Yes, it is important that everyone can see what you are doing. It is a way of interacting. After a while that you perform, performers would do even more complicated movements, kinda to show off the complexity of the beat they are doing, it is kind of a trash topic really actually.

Interviewer Does the venue in which you perform affect your performance?

Participant 3 In a concert venue you know that if you are doing a beatbox solo, everyone would be watching you whereas if you are on the street you are kind of everybody is there but nobody is listening. And you just go with that, it feels like if you are practising in public really. In a concert venue, or concert you feel much more under pressure. Instead in a public space you feel comfortable because everybody is there but nobody is really watching.

Interviewer Does the PA affect the way you perform?

Participant 3 If you are using one speaker only, you feel like you have to do more with your voice, whereas when you are in concert environment you have got two speakers so the sound is not gonna be an issue and you are more engaging and being more physical on the show, rather than being focused on the audio only which is the public basket setting.

G.1.4 Participant 4

Interviewer For how long have you beatboxing?

Participant 4 I started 10 years ago, I started when I was quite young. I just watched an advert on tv, and I tried to replicate the music, and he said that it would sound better if he could have done another sound rather than the current ones. And then as you know there is a very big community online and stuff like that. Humanbeatboxing.com was one of the first online forums for beatboxing, and what's happened there is that loads of people connected on the online forums. That's how everything got started. The very first YouTube tutorials started getting made, and more people started connecting, and we moved on from that now. So we do not really use the online forums. We are now at the point that very beatboxers know each other and if they meet somebody else, it is really easy to bring them on to the scene. So that's the kind of aspect lost. There is not much of the documented history. There are few beatbox articles that I can refer to. There is also a website called "wetalkbeatbox". It's a fairly new blog, and it explains what happened with America and stuff like that. It may be definitely relevant to you.

Interviewer Are you right-handed or left-handed?

Participant 4 Right handed.

Interviewer With which hand do you hold the microphone?

Participant 4 That's a difficult one. I do hold it with my right hand, sometimes with beatboxing, you want to make the gestures, and it doesn't feel right. So sometimes if I know that I'm going to do something, I'll transfer the microphone to my left hand, and so I can use my right hand for the gestures.

Interviewer Do you mostly prefer to perform using the microphone or to perform a cappella?

Participant 4 I prefer the way my sounds sound a cappella, we always get the opportunity without a microphone, and you can have a sort of private setting. And I do enjoy that more. I do not like people who use the microphone just to sound louder. If I'm just at home with my hands-free, you know gestures and move freely, I feel better. That's how I prefer a cappella.

Interviewer Do you use gestures in your performance?

Participant 4 Yes. Some of them could be purely just to keep track of the tempo in your own mind. A lot of people use gestures, like me, use it as almost as a representation of the sound that you are trying to make. So if for example, I was like to do [Participant beatboxing] and stuff like that you know. Yeah, to represent the sound you are making (see video recording no. 1 in Appendix F.3).

Interviewer Do you gesture the tempo only, or also other features of the sound?

Participant 4 There are a lot of different gestures. It also depends on the settings. For example, I'm doing an acoustic set or something like that, you are free, fluent and you are generally doing it harmonically. But then in settings like a battle, for example, you can accentuate the sounds more, to keep the tempo. I would say that the harmonic is what's more used these days, and for keeping the tempo a lot.

Interviewer So would you say that it is more to describe more the musical idea of the beat or to describe more the sound itself?

Participant 4 I think it is more to describe the actual sound itself in certain circumstances. Say for example a thing most beatboxers do in battling, during a battle, they'll do a beat and then a silly sound. [Participant beatboxing], and stuff like that; and, it can be used actually to represent one sound or the musicality, the musical idea of the whole beat itself (see video recording no. 2 in Appendix F.3).

Interviewer when performing. Do you move your whole body or just a particular part of it?

Participant 4 I use my whole body. Some beatboxers would do a dance to keep a certain beat or in relation to the musical genre. I use my whole body.

Interviewer Would you use different parts of your body to describe different qualities of the sound?

Participant 4 Yeah, it happens a lot. So if I would do a specific beat, and then if I introduce something else into the beat and then the hand would inevitably come up and start keeping the timing as well.

Interviewer Do you choreograph your gestures or improvise them?

Participant 4 There are someones like for the hi-hat. Which people always do for hi-hat. They do mostly something like that for the kick [he moving his fist forward and backwards]. Even the snare [like the hi-hat as the hi-hat but the finger moves back-wards]. People usually do like that because the sound is inwards. But pretty much everybody actually does use [gestures]. Some people plan their gestures a lot more. I know a Malaysian guy, who has his hand behind his back, and until and so he is doing [Participant beatboxing with his arm behind his back and then putting out his hand when the beat changes suddenly]. Then he changes the position of his hand when the beat changes a little bit (see video recording no. 3 in Appendix F.3). I generally go with it, I do not have many pre-planned gestures, but you always find yourself doing the same sort of things with your arms. We also use stage spaces, like in a battle situation, the way you move above the stage obviously affects the stage presence. That can be used, as well as the same set of gestures in beatboxing, and loads of people go down, like to bend their knees. They are close to the floor for bass noises and stuff like that.

Interviewer Does the venue affect the way you perform?

Participant 4 Oh yea, if you are on a bigger stage. I suppose that it depends on the sound system you have as well. A lot of people get more confident with a louder sound system and a bigger stage. You saw people come out of their shell, move and gesture more. The thing is that it also depends on the crowd. If you are not performing to a beatboxing crowd, they won't understand what you are doing. And so you want to sort of try to help them to understand what you are trying to do, like gesturing you can see a sort of moving and jumping around on stage, and they get into it as well. On the street, you have to be much more focused on showing off and giving the main musical idea quicker. It is because it's got to be interactive and you have to interact with the others it does change a lot. Because if you do not understand beatbox, you do not understand the music in it. So yeah.

Interviewer Is the PA affecting the performance?

Participant 4 Let's say you are on the street, and you have one speaker only. You move more with liberty, if you are on a big stage, for example, and you have got a bit system, for example, you are less connected to the performance. It thinks it happens with everybody. Let's say if I was on a bigger stage with a louder sound system, it would perfectly change the performance. People would try bassier tones and stuff like that. That's a weird one.

Interviewer Do you see gestures helping the interaction between the performer and the audience?

Participant 4 Oh yea, definitely. They are almost as important as the sounds you are doing. When it comes to portraying your music and what you are beatboxing in a performance.

Interviewer Would use a PA on the street?

Participant 4 I prefer just to turn up. See, I do not like performing on the street too much because I find it a little be intruding. You are kinda forcing people to listen to you, I know many people who love it. You know, bringing your own amplifier plugging it in, you know. And the street is their stage, and I'm personally more comfortable setting something up and going with it.

Interviewer Do you involve technology in your performance?

Participant 4 No. Too complicated to manage. If I had wrist band that could distort the sound like that [Participant beatboxing while performing a fist and while rolling his arm clockwise he performs], and distorting the sound, that would be amazing (see video recording no. 4 in Appendix F.3).

Interviewer Is using technology cheating in beatbox?

Participant 4 This is linked with the beatbox battle culture, where they are not really interested in the actual music, but everything is focused on the actual beatboxing skills. But after a long time, it can be an instrument, and it can be music; it can be whatever, so a lot of people would consider it is cheating. But a lot of people do not care about that because it helps you to make music more efficiently, to improve upon what you can do naturally. I think you won't have any problem at all. I think maybe it can be integrated into the beatbox battle concept, at the same way loop-station started as well, so it is just something that adds a new element, to your musicality to your technique and stuff like that.

Interviewer Before you dig the wrist movement while simulating a distortion. Is it because you would prefer to use distortion, or was something that was just in your mind at that moment? Would you use something different as well, like panning and sound spatialisation which does not really affect the sound itself timbrically, but it has more influence on creating a soundscape?

Participant 4 I don't know if it does make sense for you, but beatbox is almost bi-dimensional. So you really struggle to put 3D ideas in your head. Through your mouth and if you can introduce some sound spatialisation into it, and controlling it through gestures, it would introduce a whole new element to the actual performance.

Interviewer And when you move on stage, do you consider your voice a point in space. Are you aware of the fact that you are moving the voice with yourself while moving onstage?

Participant 4 Yeah, you definitely do. So say if you are doing a beat, and you go quieter and lead towards the crowd stuff like that. And then maybe face the audience and make it louder. It's something quite hard to explain because it's something that we'd be able to do without the technology and definitely.

Interviewer If you are performing while spatialising your voice around you from one point to the other one, would it be annoying for the performance?

Participant 4 You are lucky with the beatbox community. There are many talented musicians and performers, so I think your idea will work quite well as they embraced the idea of loop-station, and they got used to the idea. There is a beatboxer called Reeps One, a few years ago he made something just amazing. It was the story of the box, his beat his sound was in his hands and stuff like that [Participant beatboxing while making a throwing gesture]. And that would work amazingly with the technology (see video recording no. 5 in Appendix F.3).

Interviewer Do you consider gestures as a way to portray the sound?

Participant 4 One million per cent, especially to a non-beatboxing crowd. They would be more involved in what you are doing, and you process it differently. There are many different ways that I could mean that, so I could do [Participant beatboxing while moving his neck and shoulder sideways following the tempo of the beatboxing], or sort of [Participant weaving his hand with the palm facing sideways following the tempo of the beatboxing]. So you have really to see them, and see their gestures, in the way that they move to understand what they are doing (see video recording no. 6 in Appendix F.3).

G.1.5 Participant 5

Interviewer Are you left-handed or right-handed?

Participant 5 I'm right-handed.

Interviewer With which hand do you hold the microphone?

Participant 5 Left hand.

Interviewer Do you perform mostly a cappella or using the microphone?

Participant 5 I do use the microphone mostly.

Interviewer For how long have you been beatboxing?

Participant 5 Since I was a little kid, I started to play noises with my mouth.

Interviewer Do you perform any movement or gesture during a musical performance?

Participant 5 Yea, I'm kinda figuring out a lot of stage presence. I rap and beatbox. When I rap, I'm more moving around and some hand gestures, but with beatboxing, I'm always moving my hands all the time.

Interviewer Do you perform a specific beat per each gesture?

Participant 5 Like when I do beatbox I can do like the metronome sometime and then I can do like to from top to bottom, and I make my hands go like [Participant simulating like if he was playing the drum] (see video recording no. 7 in Appendix F.3). When I make a Hi-hat sound, and I usually my hand in some sort of ways, and I can imagine like panning [Participant beatboxing] (see video recording no. 8 in Appendix F.3).

Interviewer What sound characteristics do you tend to describe through gestures?

Participant 5 It's kinda like the tom, so picture a drum to the right, it's kinda like the tone of that tom, it's going downward. Also, gestures keep your mind on like the tempo you know, [interviewee beatboxing while keeping the tempo with the index finger, and the like [performing again]. There is really a lot that you can do with your hand (see video recording no. 9 in Appendix F.3). There are quite a few people that I see doing like when you hit a beat, like just a kick I mean, I can see a lot of [Participant beatboxing the kick sound while moving his hand] (see video recording no. 10 in Appendix F.3).

Interviewer Do you play a specific range of gestures?

Participant 5 I can imagine like a [beatboxer playing a crescendo sound while moving his hand frontward and opening the hand's palm] crescendo, you know it's louder crescendo, maybe to open the hand could change the volume, or like panning something like you know two sides (see video recording no. 11 in Appendix F.3).

Interviewer	Do you use any technology in your performance?
Participant 5	I use a loop-station, and with the loop-station, I use a vocal effect processor. the loops-station has got flanger, super echoing things, and then you can go into it and change all of them based on the variable which you can affect.
Interviewer	Do you use a particular loop-station?
Participant 5	I really like the Boss R505. But right now I use a Digitech JamMan which is two pedals like in one channel basically.
Interviewer	Do you like it for some specific reasons?
Participant 5	I like the Boss R505 because I can change the song structure. I can make a chorus and then harmonise the chorus in one channel, and then with another channel, I can take that one off and go to the verse. So it's just the ability to have 5 different things going on.
Interviewer	Referring to what we were talking about before, does the loop-station affects the way you move on stage?
Participant 5	Yea, if I was a cappella, I would move a lot more. But when it comes to looping I have to pay attention to the loops, the timing and everything, I'm kinda preoccupied with that, instead than just the expression form. Then I can move around and be animated.
Interviewer	Do you find this annoying for the performance?
Participant 5	I do find it kinda restricting in a sense. Not as much annoying as restricting. It is kinda annoying that I can't jump around all over the place because it gives the gigs that emphasis.
Interviewer	Would you use something hands-free to overcome such a problem? Would it help the performance?
Participant 5	Yeah, if I had something like that connected to the computer, and I knew all the movement I needed to do, it would change a lot of the game really.
Interviewer	In addition to all effects, have you ever played using panning effects?
Participant 5	I've used a lot the flanger but not the panning really.
Interviewer	Does the type of PA affect your performance?
Participant 5	Yea, because in the concert venue, my voice is louder. Panning can happen, and that would be really cool you know, sending to the right or to the left, that would be really cool. Because a cappella you know, I can go like this or like that, but that's it. [talking with the hand on one side of his mouth and then the other one to drive the voice in a specific direction (see video recording no. 12 in Appendix F.3).
Interviewer	Would you use sound spatialisation? How do you integrate it to your show?

Participant 5 I see myself making a beat over here and then maybe the hi-hat over here, [the interviewee pointing in different directions and beatboxing]. People are kinda sense of confusion, because the ear the bass and then they ear the hi-hat over here, and would create a whole different way of feeling the music (see video recording no. 13 in Appendix F.3). I think sound spatialisation will change the whole game and will bring people to think music is a different way. It would make the person feel more different feelings. Because, I feel when people create a soundscape, it sounds like if it comes from the real-life, it changes your whole view for a second and it moves you from the present like just dripped off into the moody.

Interviewer How would you control sound spatialisation?

Participant 5 It could work with a sensor type of thing, and a loop-station and the panning is based on the movement wrist band or glove whatever, like an imaginary wall in front of you. When you hit it, it's starting to record.

Interviewer When you move on stage, do you feel like moving your voice with yourself? Or do you move without thinking about it?

Participant 5 When I'm moving without thinking, I'm usually rapping. I kinda do that [Participant rapping]. But when I'm beatboxing, I'm very aware of what I'm doing with all movement, so I could walk over here and then playing over there and then get lower (see video recording no. 14 in Appendix F.3).

Interviewer And are your gestures affected by the type of venue you perform in?

Participant 5 No, because you know because beatboxing is the thing. So on the street or in the venue is the same thing.

G.1.6 Participant 6

Interviewer Are you right-handed or left-handed?

Participant 6 I'm right-handed.

Interviewer With which hand do you hold the microphone?

Participant 6 It changes, when I perform, I use my right hand, and when I get tired, I feel to change. But I mostly use my right hand. This is because I've a heavy microphone.

Interviewer For how long have you been beatboxing?

Participant 6 19 years.

Interviewer Does the microphone affect how you move on stage?

Participant 6 Yes, when I've the microphone, I cannot move so much. When I'm performing without, I'm moving much more because I've more freedom and I've nothing to hold. So, yes, it affects my style.

Interviewer How would you define a gesture?

Participant 6 A gesture is a movement with meaning. For example, when I do hi-hat, I do [Participant moving the hands if as playing the drum]. I often show it with my hand like if I had a stick, so when I'm on stage, I do [Participant beatboxing]. So every time I make the cymbal sound I show it. That's learnt, it's trained. But sometimes, when I'm deeply inspired, I just do some kinda movements that I do not really think about (see video recording no. 15 in Appendix F.3). Yeah, that's a good question. For example, when I vocal scratching, I also show the movement. So I also do this movement with my hand [interviewee showing the scratching gesture], so the audience knows what I'm trying to do. So this is connected. But with other sounds, I do not do it to describe the sound pitch. But I always had this [cymbal sound] or bass drum like this with my hand going forward because it's really heavy (see video recording no. 16 in Appendix F.3). I also perform in a cappella group, and in that situation, I keep steel, not to move too much to do not disturb the scene too much.

Interviewer You won't move when performing a cappella because the other wouldn't move that much or it's because you won't take over the scenery? Is it influenced by the others or it comes from you?

Participant 6 It is actually because of the audience. When I'm the audience, and I see one only guy gesturing, it's annoying. And also to do not distinguish me from the group. It's not four people and beatboxer, but it's one unity.

Interviewer So would you play more gestures in a beatbox team?

Participant 6 I play loads of gestures when I'm doing a solo, a drum solo. But in a group, I always try that the music of the singers is the number one priority and not my beatboxing.

Interviewer Is it because it would be weird that you will be the only one moving between four singers?

Participant 6 Yes, when it's beatboxing only, it's ok because everybody is doing that. But when I'm with other singers, I cannot destroy the picture. So, with other beatboxers in a beatboxing setting, it's easier for me to move a lot because the other people are moving the same. The thing is that I'm also a trained singer, so I tend to separate the two worlds. The beatboxing and the more traditional vocal performance.

Interviewer Do you perform as a beatboxer or as a singer do you perform differently gesturally and musically?

Participant 6 Yes, I try to go by style. For example, I love jazz, so I use to emulate brass sounds. But once I had to beatbox on a Beethoven Piece and so I tried to beatbox in a Beethoven style. Not too heavy not too much movement, like if it would have been done in that era.

Interviewer Is the environment around yourself, affecting the performance as well? The venue or the audience?

Participant 6 Yes, very much. I'm a very emotional guy, so I react very strongly to the whole environment around me. If I'm performing for a very well educated audience I go for less "la-la-la", but in a 'cool' and jamming context then I'm different.

Interviewer Do you use technology for your performances?

Participant 6 Yeah, when I started, I used the microphone only. After 10 years, I started to use pedals, distortion pedals, loop-station, octaver pedals. I started to use stuff used by guitarist and instrumentalist, and then use the laptop. I do not use the laptop, but I've the hardware on it.

Interviewer How do you find the interaction with such hardware? Do you find the kind of interaction disturbing for the performance?

Participant 6 Yes, I do. One of my best wishes from years has been as small device, which I do not have to be bothered with process and routing I guess to answer your question in short. When I'm on stage, I've a very difficult job, so the audience expects me to sing, think about the musical structure and to beatbox, to be emotional to take them emotionally. For me, when using a loop-station and I cannot miss the beat, it's really taking my mind off. So I do not really like to use a loop-station on stage, and so that's why I do not really do that.

Interviewer You mentioned the musical structure. In some forums, beatboxers mentioned that gestures are also used to feel the beat and to think about the musical structure. Is that true for you as well?

Participant 6 I actually tried to disconnect myself from the hip-hop and more popular scene. So, I try to do not let the gestures affect my musical abilities. I try to be able to perform either with movements or without. It must be in my mind, and I prefer to do not depend on them.

Interviewer So is it something that you do when you feel to do it, and you do without thinking?

Participant 6 Yes, but even if I feel it, I can mostly control it. If I want to make a gesture, I do it, but the sound quality would be the same.

Interviewer Do you use any technology in your performance?

Participant 6 Yes, the effect which I use mostly is the octaver, because I'm a tenor. So I use it to have a lower voice, and then I also use reverb, and then a bit of granular synthesiser.

Interviewer Do you use the reverb just to make the sound a bit more smooth and more interesting or also to simulate a virtual soundscape?

Participant 6 I use the reverb to create an atmosphere, a mood. Definitely, it can be used to create a virtual room. But for me, the psychological effect on the people it's stronger.

Interviewer Do or would you make use of audio panning or sound spatialisation?

Participant 6 Well, that's something I would like to do, and once with my sound engineer. He was told that when I was performing certain beats, he needed to use the pan my voice accordingly with the direction of my hand.

Interviewer When you did that experiment with your sound engineer, did you enjoy it? Would you like to do it again?

Participant 6 The main point for me is the audience, and for some of them, it was just amazing. But some people were a little bit disturbed, in the sense that they think with the use of technology I wouldn't be able to do the same without. So there are people that are negative when they know that I use effects, they think it's cheating. I think that it is a psychological thing. It depends on if you care about what other people say and if you care about it. So I would use technology much more, but only if it makes life easier.

Interviewer Have you have ever used sound spatialisation during production or post-production?

Participant 6 Yes, I did work on a piece I realised with my band. I think that it could be used in a cappella music. In that sense, it is really moving forward to be produced in many different ways using many effects.

Interviewer Does the venue influence your performance? And the PA?

Participant 6 What I'm used to performing with is what I call a 'normal' sound system, which is composed of a main left and right and subwoofers. That's a thing I'm crazy about. When I perform for official gigs I do always require at least one subwoofer, I want to feel my voice, and without a subwoofer, you sound like a baby.

Interviewer So feeling the sound close to you helps you?

Participant 6 Yes, definitely, it makes me think harder if I cannot ear my voice as powerful I lack confidence.

Interviewer Does the venue affects your way of gesturing too?

Participant 6 I guess I'm always the same, but when I'm on the street then do not feel as good as in a show, and so not being comfortable on the street might make me move less.

G.1.7 Participant 7

Interviewer For how long have you been beatboxing

Participant 7 4 and a half years.

Interviewer Are you left-handed or right-handed?

Participant 7 I'm left-handed.

Interviewer With which hand do you hold the microphone?

Participant 7 With my left hand.

Interviewer Do you perform mostly using the microphone or a cappella?

Participant 7 I use to perform in both ways, but usually, live performances with a microphone. However, when the space is small, I tend to not to use the mic, like in a coffee shop. It's important to me to show people all I can do without any sort of amplification or technology. Even to show my mouth's movement.

Interviewer Do you enact gestures during live performances?

Participant 7 When I beatbox, I move my hands along with the music.

Interviewer Do you perform particular gestures per each beat.

Participant 7 Yeah. I usually go, you know, I try to say volume up and volume down. See, for example, when you go [Participant beatboxing] So with my hands I let my audience know that I'm about to turn up or turn down. So I go for example [Participant beatboxing while performing knob turning gesture with a thumb-index finger pinch hand pose] (see video recording no. 17 in Appendix F.3). And when I want to do a more electrical crunchy beatboxing, [participant beatboxing] (see video recording no. 18 in Appendix F.3).

Interviewer Do you work on those choreographies before the shows or is something that it comes up in your mind?

Participant 7 Yes, some of them are well prepared, and some of them instead is just performed as they come up in my mind. Sometimes things happen in the show, and I know how to react accordingly in the show.

Interviewer Do you tend to describe the quality of the sound or the tempo?

Participant 7 It is kinda both actually. If I'm going to make a bass sound I will usually go like this, you know something. For example if I'm doing something which is low [Participant beatboxing] (see video recording no. 19 in Appendix F.3).

Interviewer Do you describe the crunchy quality of the sound with the finger's movement or something else?

Participant 7 Yes, the crunchy sound and the low tone. The idea is that I want the audience to have a visual experience too. As for me, many beatboxers tend to move their hands along with the music. Instead, I do it deliberately. Every single gesture I do has a meaning in relationship with the music. I move in relation to the sound. For example, if I'm making a electronic kinda music, my gesture will be in relationship to the beat (tempo). For example. [Participant beatboxing]. And see I'm using my fingers here. Because I'm thinking the music in a different way, I send a different vibe, so I cannot do something like this

[Participant moving weaving the fingers upwards], I can do something like this using my finger if I'm doing something groovy and like party (see video recording no. 20 in Appendix F.3).

Interviewer Does the microphone affect the way you move?

Participant 7 Yes, it depends and in fact, mostly I will switch hand during the show because I also use to move my right hand. I usually do it when the audience is spread out, on your left and on your right and so you have to be fully open. For example [Participant beatboxing] (see video recording no. 21 in Appendix F.3). If the audience is on my right, it might be different. I hold the microphone with my left hand because in that way my body is opened to the audience, and that has a positive effect in how they receive you (see video recording no. 22 in Appendix F.3).

Interviewer Do you perform differently in different venues?

Participant 7 When I'm on the street usually, I've no have no mic available, and so you really have to know how to sound loud. When it comes to a coffee shop performance, it becomes much more intimate. Meaning that you do not have to be loud as on the street because the sound is contained within the space, and people who are listening to you are really close to you. Also, I make gestures at every concert. However, they depend on the situation I'm in.

Interviewer Does the PA affect the way you perform?

Participant 7 Yes, when performing acoustically, the beatboxer is in control of everything. Every sound which comes out from the mouth, it's created organically, without any aid or modification. Beatboxing using the microphone, the sound goes through the mic, through the cable and then through the speakers, so that creates a difference. It changes the sound completely. And what I usually do is to call somebody who can clean up the sound as I'm producing the music, and you need somebody to help you in that aspect. Another thing is that when you perform acoustically, people can see you. They can see your face, they can see the facial expression that goes along with the gestures and the music itself.

Interviewer Have you ever used loop-station, effects or anything like this?

Participant 7 No, though to do it, I haven't for my performances.

Interviewer Is it more for a stylistic choice. For example, some beatboxers argue that if you use a loop-station, it could be considered as cheating. What do you think about it?

Participant 7 I do not have any problem with it, with those people, I don't think it is cheating. It's just a different way of performing. Who says that looping it's cheating do not understand how hard it is to be a goo live looping artist.

Interviewer What part of looping would you define as the most difficult ones?

Participant 7 It is to put every beat and sound in the same time-frame. Before you had to use also your feet to loop, but now with the Boss R505, you can also use hands to interact with the loop-station. Something very difficult is to improvise. You create entire songs, but you can't improvise. There is one only Beardyman in the world.

Interviewer Keeping talking about the different ways of interactions with loop-stations, do you think that interaction required by the loop-station suits the human beatboxing musical performance?

Participant 7 I would say yes, but it depends on the beatboxer. Because you could loop for a minute and then improvise over it, you can sing, or rap over it and without recording any other. But there isn't room for a movement when it comes to looping, and this makes things harder for other performers who use to move around and me.

Interviewer And you personally would you deal with that technology or, would you look for something else which allows me to move on the stage?

Participant 7 I would choose to go on without for the beatboxing set and then to use looping for the second part of the show.

Interviewer In addition to loop-stations have you used any other technology?

Participant 7 No. I prefer that people would listen to what's coming out from my mouth, organically. I want people to understand what I'm capable of.

Interviewer Have you ever used spatialisation of panning effects?

Participant 7 Yes, with my sound engineer. It had to do with the position of the audience, and also when the music was going up and down the audio was going from one speaker to the other one.

Interviewer What do you mean by up and down? de volume or the pitch?

Participant 7 The volume.

Interviewer And when you perform a cappella do you realise that when you are moving you change the position of your voice in space too?

Participant 7 Oh yes, I definitely do it. If the crowd is spread out, I would move all over the place.

Interviewer Does the PA affect your performance?

Participant 7 Well, it really affect my feeling. I feel much more powerful.

G.1.8 Participant 8

Interviewer For how long have you been beatboxing?

Participant 8 10 years.

Interviewer Are you left-handed or right-handed?

Participant 8 Right handed.

Interviewer With which hand do you hold the microphone?

Participant 8 right hand.

Interviewer Do you perform gesture when beatboxing?

Participant 8 For a part of my shows, when I beatbox, I really move a lot. For the second part, I perform with the loop-station, so I'm not moving that much because I'm behind the table. Gestures usually come out by themselves, but I'm working on the way to portray my sounds using gestures.

Interviewer Do you perform a specific beat per each gesture?

Participant 8 It depends, for example, when I'm dropping some sounds you just follow the tempo, and it happens for sounds like the lip roll (see audio recording no. 23 in Appendix F.3). This is a sound which you can describe the tempo but then if I go like, [Participant performing] (see video recording no. 24 in Appendix F.3). I can show with the fingers like if I was playing like with a piano (see video recording no. 25 in Appendix F.3). When I play gestures, it's like I'm painting, it is not that I'm doing painting gestures, but I try to make them understand the structure of what I'm doing. That's why I use my hands.

Interviewer Does the microphone affect your gestures?

Participant 8 I use to hold the microphone with my right hand, but then sometimes I use to switch the mic because it looks better for the audience to perform gestures with my right hand.

Interviewer When you try to portray your gestures, do you have something visual in mind?

Participant 8 Yes, I think that every beatboxer has his own representation of the sound. For every sound I play, it is like if I was composing with Logic or Ableton, I see ever sound and picture the colour they may have.

Interviewer And so do you organise sounds like in layer and colours and not more particular pictures.

Participant 8 Yeah.

Interviewer Do you perform using a loop-station?

Participant 8 Yes.

Interviewer Do you feel like ti constrains your movements?

Participant 8 It's like playing the piano, I just play it, and I move in the way I've to move. I use the RC-505, which is incredibly like the best loop-station ever created for live shows. I tried all of them, and that is incredibly powerful. The only way you can make it better is making a better interface, the screen is pretty shit, it is pretty small. If I was working on other loop-station, obviously the design is better. For me, the best set up would be an iPad to control a good looper on your laptop. To use the iPad as an interface. This would be amazing.

Interviewer Have you ever used panning or sound spatialisation effects?

Participant 8 I never used them live, but when I recorded for my album, the sound engineer used panning on the track.

Interviewer How did your sound engineer use the panning?

Participant 8 So when you record the tracks live, you have one track only. So you cannot isolate sounds, but you can isolate frequencies, so he isolates the bass frequencies from the kick and snare, and after that he makes it move from left to right, for example in sounds like. And so, the beat is going from right to left in synch with the music.

Interviewer Would you use it also during live performances?

Participant 8 If I had a good sound guy with me all the time, yes.

Interviewer So you don't use it because of technical reasons and not because of other issues.

Participant 8 Yes, exactly when you work on a show with a sound engineer, you can do what you realise during post-production, but you cannot do it live.

Interviewer When you move while performing a cappella, have you ever realised that you also move the sound you are producing from one point to the other one?

Participant 8 Yeah for sure, it depends on what kind of audience you have. It is a small group. But you know performing using the mic and performing a cappella it's something completely different. There are sounds that you can do performing using the mic but that you can't do a cappella, especially bass sounds.

G.1.9 Participant 9

Interviewer Are you left-handed or right-handed?

Participant 9 Left-handed.

Interviewer With which hand do you hold the microphone?

Participant 9 Right hand.

Interviewer Do you play any gestures while performing?

Participant 9 I use my hand to keep in time.

Interviewer When you move your body, do you try to describe more the tempo of the beat or the quality of the sound, like the timbre?

Participant 9 Yeah. If for example, I do a scratching sound like on a deck I'll play the DJ gesture at the same time.

Interviewer Do you use technology in your performances?

Participant 9 I've seen others doing it, but I do not use it. I think I would benefit from it. I struggle to keep the breath, and so a loop-station would help me.

Interviewer Does the microphone affect your gestures?

Participant 9 I would say that I would gesture more if I would use the microphone because you can pick up every sound with it, even quieter ones.

Interviewer Have you ever used effects or any other sort of technology in post-production?

Participant 9 No, I haven't done it.

Interviewer And why? Was it a stylistic choice or for any stylistic reason?

Participant 9 Well, I had the possibility to afford it, I would use it, and it would be useful. So it's more a circumstantial reason.

Interviewer And what do you think about who says that to use loop-station is cheating?

Participant 9 I don't think it is cheating at all. It helps. For example, if you cannot make all these sounds, or sometimes it just does no come for how you feel, so with a loop-station, you can actually build up many layers. Then you can also make a song out of it, or sing. Instead, if you are doing all with your mouth, it would be almost impossible.

Interviewer So if you would have the chance to work with a sound engineer, would you take the opportunity to do it?

Participant 9 Yes!

Interviewer Have you used or would you explore panning or spatialisation effects?

Participant 9 I think that it is something I would explore, to get used to it and then to see how it goes, how I feel about it.

Interviewer Does the venue affect the way you perform?

Participant 9 Yes, and the audience really affects my performances. Once I was performing, and the audience gave me the strength to carry on.

Interviewer And the PA?

Participant 9 Well, I've always performed with one speaker only, on my right or left.

Interviewer If you were to use panning effects, what gestures would you use to control them?

Participant 9 I would say particular gestures, just to do something sure. I think that throwing it and receiving it to the left or to the right, I think that it goes well with the act or sort of visualisation to the movement.

Interviewer When you say visualisation? Do you picture yourself in manipulating it?

Participant 9 Yes, yes, for examples like a swing that you slap with your leg and you go to one side and then you go back. Like in a venue where you have speakers all over, you can move the speaker from the front to the back, middle, left or the right.

G.1.10 Participant 10

Interviewer For how long have you been beatboxing?

Participant 10 10 years now.

Interviewer Are you left-handed or right-handed?

Participant 10 Right handed.

Interviewer With which hand do you hold the microphone?

Participant 10 Right hand.

Interviewer Do you mostly perform using the microphone or a cappella?

Participant 10 If I'm performing on stage, I use the microphone.

Interviewer Do you do the same on the street?

Participant 10 I do not use to perform on the street, but if I was using the microphone would be the same.

Interviewer When you perform do move your body?

Participant 10 Yes, mostly my hands. A kind of motion that goes along with the music (see video recording no. 26 in Appendix F.3). My hand just goes up and down along with the beat [Participant weaving the hand up and down with the palm facing inwards] (see video recording no. 27 in Appendix F.3).

Interviewer Is then mostly to perform the tempo of the beat?

Participant 10 Yes one two, three four, kinda think, one, two, three, four [Participant weaving his hand up and down while counting] (see video recording no. 28 in Appendix F.3).

Interviewer When gesturing, do you also try to describe the quality of the sound, like if it is a low or high pitch sound?

Participant 10 Yes, sometimes when I'm on stage, and I do loads of bassy sounds, I tend to bring my body down when I'm on stage. [Participant bending his back downwards] (see video recording no. 29 in Appendix F.3). I'll play these kind of sounds [Participant beatboxing] and bring myself down [Participant beatboxing] (see video recording no. 30 in Appendix F.3).

Interviewer Do you perform any other particular gesture or movement?

Participant 10 Yes, my hands just go along with the music.

Interviewer Do you use an audio system to perform?

Participant 10 I'm at home, I tend to use one speaker only, instead if I'm on stage I will use a standard PA.

Interviewer Does the type of PA affect your performance?

Participant 10 When I use two speakers, the sound is a lot more clear, and I can hear myself much better. Also, with a PA, you can produce more bassy sounds.

Interviewer Does it affects the way you move on stage?

Participant 10 Yeah, it does.

Interviewer Is it because of the sound only or also because of the audience?

Participant 10 I think that it is because of both, and when you sound louder, it is bad to see you just still.

Interviewer When you perform without PA, do you realise that you also move your voice from one point to the other one of the performing space?

Participant 10 Since I'm without a mic, I would focus more on the sound, rather than the audience or my movements.

Interviewer Have you ever used a loop-station?

Participant 10 I would like to, but I don't at the moment.

Interviewer Is it because you do not like them or other reasons?

Participant 10 Well, it's because it's expensive.

Interviewer Have you ever processed your voice in alternative ways?

Participant 10 Yeah! Once in a chat with hundreds of beatboxers, we were playing at recording our voices and speeding it up

Interviewer How did you feel on that occasion? Was it ok to deal with technology?

Participant 10 It wasn't hard at all to deal with the phone. But it is not a real thing if you know what I mean. For example, there is the same on the app, but there it is not the same because you can't have everything quickly.

Interviewer Is it because you feel more secure, do you feel like the whole thing it's more stable?

Participant 10 The problem with all those apps and those interfaces is that you can't use them alone. And you really have to spend time to find out how they work.

Interviewer Would you use gestures to control sounds?

Participant 10 I would use it, mover, for a live kind of thing. If you say, I point all the sounds sort goes that away [Participant pointing at a direction, and spreading the fingers while extending the arm] (see video recording no. 31 in Appendix F.3).

Interviewer Some beatboxers consider the use of electronics like cheating, what do you think about it?

Participant 10 In my opinion with a loop-station, you are making real music, a real live, so your machine would perfectly work. I don't think it's cheating. It is a quite taboo sort of argument. It's like people who say that if you use your hands to produce sounds, [Participant beatboxing with his hand on the chin], that you are not really beatboxing that you use your hands to make the noise. (see video recording no. 32 in Appendix F.3). There are some people that consider it as cheating, who don't. There are a few sounds like [Participant beatboxing with his lips between two fingers with the palm towards is face], where you use your hands. So, are you beatboxing or not? (see video recording no. 33 in Appendix F.3).

G.2 Extracted Codes

1. Contaminations from other form of art
   (a) Dance
   (b) Visual Art
2. History of beatbox
3. Use of technology
   (a) Loop-station
   (b) Guitar pedal effects
   (c) Mixing
4. Performing
   (a) Solo
   (b) Duo
   (c) Trio
   (d) Group

Appendix H

Interviews to Harpists

H.1 Scripts

H.1.1 Interview with Jennifer Ellis

Interviewer What is your experience in using music technology in harp performance?

J. Ellis I've used some live electronics in performance in the context of improvisation or ensembles, where there is someone else in charge of the laptop. But my first experience into really being in charge of it myself, and creating it myself while performing, as well, was through a programme called The Atlantic Music Festival at Future Music Lab. And that's something run by Mari Kimura, and I'm not the only harpist that's come through that programme, there is also an Irish harpist who went through that programme and so the year before me. It's been a very good boom for us I would say. Mari Kimura is Julliard faculty, she is an amazing and renown violinist - in terms with her work with electronics and violin - and she runs this program that is really for performers to help them get the skills they need to be doing the live electronics side of their projects. So, it's, you know, really geared, at primarily performers not necessarily the tech folks and gives them the tech skills they need to get what they need to get that. And Mari, one of the things Mari did, she got access to the MMO motion sensors, which stands for Modular Musical Objects [J. Ellis meant the MO Musical Objects presented in Rasamimanana et al. (2010)]. So Mari has done a lot of performance with motion sensor, and she gave us access to this technology basically, rather than IRCAM did. We got to figure out how to use it on our instruments. So, for me, I ended up attaching it to one hand. I think what I ended up using was an extender, for like an iPhone if you wanted to put your iPhone like around your arm or something; and I kind of rigged up this funky thing so it could go around my hand roughly and so it became "so play" and all this. The motion sensor itself was so light; I was really surprised. I was very sceptical that it would work having it on my hand when I play, but it was so light, and it was so stable, I didn't go through off the balance of my hand or anything, so it was very doable.

Interviewer What do you mean by "I was sceptical"? Were you worried about something?

J. Ellis Yeah, I was worried about getting on the way of me playing, that it should wear it on my hand because, especially the method of harp that I play after each note, enclose fully into your hand, so I need to have room to be doing this sort of motion all the time. You don't necessarily want something with a ton of weight back here, because you don't wanna mess with your tendons too much, you don't wanna because of an injury or anything. And then the other thing is when you got the top of the harp, the space gets so condensed, right? Between the neck and the board, you don't have a lot of room to manoeuvre, and the idea of having, adding a sort of bulk that could get in the way is not ideal. So, I was appreciative of how small and lightweight this was, it really made it feel way more doable. Yeah, so, I ended up creating three movements, of a piece based on poetry by Edna St. Vincent Millay and that was really my first friend when using live electronics. Since then, I sort of been doing more, both running the works of other composers that use electronics, and play those as well as during my own. For a long time, I was hesitating to get that deeply involved in electronic acoustic performance because I didn't want to swap out harps, first of all. And, I didn't want to have to buy a giant expensive Camac blue harp as wonderful as those are. So I didn't want to have to deal with swapping harps, I didn't want to deal with buying a tone of expensive gear, but most importantly, I felt like there was, and I still feel so, there is so much unexplored territory on the acoustic harp. The scores that are written for us barely scrape the surface of what the harp can do, just acoustically. By start messing with electro-acoustics you're gonna start to lose, um, a sense of like what the harp can do acoustically that we don't want electronic to manipulate some of those sounds, a harp can produce them as it is. So, I've been holding back on getting involved, and I had some other musician colleagues encouraging me (who weren't harpists), encouraging me just to take the jump on it; and I did, and it's been a lot of fun. I still think there is a ton of room for development for the acoustic harp that's untapped potential, but I enjoy working in electro-acoustics settings as well.

Interviewer Before using the MO, have you ever use guitar pedals or other technologies?

J. Ellis No, so not before. That was in 2014, that I did the Atlantic Music Festival, so not before 2014. Pedals are an interesting thing with harp because, obviously, we already have so many pedals if you are working with a pedal harpist, which is what I am. There are two sides of that coin. On the one hand, we are really busy. Our feet are already really busy. So, unless you want to water down the chromatics of the piece, I'm not sure that pedals are the most successful approach for harpists, at least for pedal harpists. Just because our feet are already spoken for that, it means that you can only hit an electronic effect when you need to hit a chromatic, an accidental, that seems a little problematic. Not to mention that we have so many effects that come from the pressing the pedal and then holding it, like suspending in two notches, and you have to keep your foot there, and then your foot is like down for the count, you know, it can't go anywhere else. On the other hand, if a composer were to write a piece that is completely free of chromatics, of like, you know, of any chromatic changes during the piece, and then gave us a bunch of pedals. We are the one instrument, other than the organ that's spent a lot of time practising foot movements, fast foot movements. So we could probably make, induce some pretty virtuosic pedalling between a lot of different effects, and go really fastback and forward and that wouldn't face us at all. But it means that you have to sacrifice some other chromatics in the piece. I control everything from my laptop, and I find it easier to just, with one hand keep playing, keep playing when is a rest just, hit something real quick and come back. As much as possible I try to automate it all, through Max, but I haven't done it as much work with pedals.

Interviewer When working with the Mo, did you find the gestural interaction with the electronics difficult?

J. Ellis I find it really easy to integrate gesture with the electronics. Especially with the method of harp that I've been training. There is so much movement already, so it felt like a natural integration.

Interviewer Which movements did you explore?

J. Ellis First of all, I only use motion sensors for this one piece, because that's the only time that I had it on loan, so it's not the main thing that I do any more, it's just the first thing that I did.

Interviewer I saw on the video of your performance. You were taking off the hands from the instrument at some point and moving your hand. But I couldn't understand what movements were associated with each signal processing.

J. Ellis That was a little deliberate because I didn't want it to be like a false like "after I play I have to do something special to trigger it". I wanted the motions that I naturally did when I played to be what triggered the sound without having to really make any adjustments from how I would normally move on my instrument. And each movement of the piece "ears different premise like the third movement doesn't use the motion sensor at all, it's totally doable without the motion sensor it's just Max. The second movement uses the motion sensor more heavily, and it's applying different settings -there is no pre-recording sound in any of this movements, it's all sounds the harp is just producing, and it's just being modified. That one is more tight to the accelerometer aspect: so how fast or slowly was I raising. And sometimes, it seemed to me like maybe it wouldn't get the read that I was expecting as I would raise extra fast, and it wouldn't necessarily pick up on it, or I would raise slow and might not pick up on it. But I really routed in that way that it would welcome those hiccups, as an effective part of the piece. So, I wasn't that concerned with accuracy; like in terms of meeting the motions would be completely accurate, and I wasn't that concerned with like having the audience be able to determinate the exact acceleration of my arm. I was concerned with the idea of suddenly having to bring audience attention to the physicality of the harp because suddenly they are paying attention to my arms in a way that they wouldn't otherwise. In general, one of the things of that piece is playing with the idea of the invisible and the visible on the harp. So, really, all movements are in some way dealing with sounds that are typically hidden on the harp, were are supposed to be heard, that I'm then featuring like as highlights points on the piece. Since we are already playing with what it is and isn't invisible, and what is and isn't audible on the harp, you know the motions also feel right into that aesthetic.

Interviewer If you would, for example, be free and establish any relation between gesture and sounds, or you would have the perfect system to track the gestures and the best audio engine to realise the audio effects, what gestures would you associate with audio effects?

J. Ellis I think for me, what I would most like, that if it's in a program, like Max or Pure Data, that I would personally be able to assign in it the gesture to the effect. And that there would be that programming and flexibility. Because honestly, I think that different gestures work better for different pieces. I also think that different gestures are going to make more sense of different harps and for different body types. I think on the freedom as a performer, if the technology exists, that could really be modifiable, I think that would be the ideal.

Interviewer When you say that the gestures are related to which kind of harp, what do you mean?

J. Ellis Well, between harps of different size that necessarily changes your gesture, because there are certain places on the harp where you are just going to have to move differently because it's a different size. And if you are working with a level harp versus a pedal harp, the roll where your later feet can be involved with those gestures becomes very different. I ended up going in a different direction, but I really want to pick the motion sensors on my ankle, because I wanted to use those hidden pedal motions as a way to trigger the effects; so like feet become a different possibility on different harps. Different harpists, harp techniques, different methods of harp playing, have different harp gestures. Harpists have incredible different physicality on the harp. Part of that is like some real differences in teaching traditions, but the fact remains that there are just huge differences on how we approach gestures on the harp. So, the interaction with technology has to be as organic as possible, and also as achievable as possible for the harpist. You don't want someone that is used to move in a certain way to suddenly trigger a sound every time they do a movement that they see assemble to their technique and just routine. You want them to be able to control those movements, and so people who typically raise like the arm all the time after they play each note, that would be potentially a problem. I think the ability to customise, the ability to modify, it's very critical.

Interviewer When you talk about gesture, do you talk about simple nuances on the hand when after plucking the strings or you talk about big movements?

J. Ellis I've done acoustic pieces that are really based on gesture and movement where I'm moving the harp across the stage. It is part of the performance, and the harp is being laid down at different points of the performance and picks back up and set upright and lean over; so I have used more theatrical gestures. In an acoustic setting, primarily, my experience with motion sensors it's been quite limited because of access, so the only time it's been at the Atlanta Music Festival. But at that time, I was really interested in not having to change any of my gestures from what I normally would do on the harp. I wanted to be able to do exactly what I would normally do on the harp and have the changes I needed and wanted. And part of that, meant giving up control, I'm heavily an improviser, so that was great, as far as I was concerned. The second movement, I'm not going to be able to complete. I deliberately made the settings too sensitive for me to be able to control as a mere human and modulate. I wouldn't be able to exactly guess which sound I'm triggering by how fast I'm rising. I could probably guess which one of three, but it's more sensitive than what humans can perceive, and that was a deliberate choice because I didn't want to know necessarily what sound was going to come out. I wanted to have to react to it in real-time.

Interviewer Sometimes performers relay very much on technology when performing, meaning they expect that the technology will produce the sound they want, but they don't consider the fact that at the origin of the chain there is the harp sound. In these terms, have you explored any particular instrumental technique coupled with an electronic effect?

J. Ellis Primarily, in Weav-Weav-Weaving it's just me making harp effects. I use extended techniques or using manipulations of the disk of the top of the harp, using metallic "corde tone, using single string, glissandos", and that's the basis of the whole piece, using this acoustic effects. And then I'm using the electronics to enhance and highlight that. Some of them can be a little quieter than I would like in a performance, but you use electronics to help layer and boost and add reverb and all of this, and suddenly the audience can interact with that effect in a different way. Extended techniques are the lifeblood of what I do, and I think that electronics can be very helpful. Some of them sound really different when you are sitting at the harp that what actually travels to the audience, so it can be really great to help amplify, and get to share those with the audience in a different way than you would otherwise do acoustically. I had an acoustic piece of mine in which I did that. I used to send for applications and things. I found out after I was accepted somewhere that part of the reason I was accepted because they were so impressed with my electro-acoustic harp writing. But, that piece, the piece I signed in, I thought it was electro-acoustic, was not electro-acoustic, it was totally acoustic, it was just me using some extended techniques, that's all. And it was a fun piece with extended techniques. But I wasn't using any electronics. So, now I've gone a little obsessive with really making clear that my acoustic pieces are acoustic because I don't want people to assume that any cool sound that comes out of the harp is actually coming out of the computer; that's a problem.

Interviewer When you approach the harp to look for like new sounds, you try to move your hand differently and try new positions. In Using the MO, did you approach the technology the same way or you thought "is a piece of technology, so it must do what I write into to Max"? Have you considered the sensor as a musical instrument or as a piece of technology?

J. Ellis Well, I saw it more as a collaborator. You know when you think as the device being so separated from the computer because it wouldn't make any sound without the computer, is involved? And that as a unit, I saw it as something that is in conversation with art. It's the way I designed that piece. I wouldn't say that I feel that way about technology universally, but in that case, what seemed to work in the piece was the idea that the electronics and harp were generating material that the other had to respond to. It was about the interaction between the two constantly. What you were saying before, as far as when you first go and explore around, when I'm programming in Max or something, that feels pretty different than exploring on the harp. I wish it were more like exploring on the harp, but there is a lot of like "oh no! Why isn't this working? Let's figure out this problem". It's just a lot of troubleshooting. But when it comes actually to have motion sensors in my hand, and I was trying to figure out what the impact would be, that feels super exploratory, that feels like finding new sounds on my instrument; it feels very free, easy and fun.
I mean the buttons were fun too, but because it was my first time really working with Max, so there was just a lot to learn. That took up a slightly different part of my brain. I think, down the line, if I'd spend as much time with something like Max than I would spend on my harp, then you would have a really different situation. But, in the beginning, it has some catching up to do.

Interviewer When you were coding and using the gestures, what felt more musical? I mean, what felt more close to the composing process or more like performing process: exploring the software part, the coding part, the creation of the synthesisers or exploring the sounds but moving your hand, or mapping gestures?

J. Ellis It all did. It was like a really integrated process.

H.1.2 Interview with Úna Monaghan

Interviewer What is your background and your relationship with technology, how do you use technology and to do what, what kind of technology?

Ú. Monaghan I think with music, I started when I was seven. I played the harp for many years and did a degree in physics and astrophysics. So my music and academic career are kind of split when I was about eighteen, when I went to university. I began to play music just separately from my studies which were in physics. But after I did that degree, I wanted to play more music, so I did a master in sonic arts at the Sonic Arts Research Centre. At that point, my music and technology were still kind of separate because I trained as a traditional harpist, so I played some MIDI effects on the harp, and I worked as an electric-acoustical composer. When I was in my early twenties, I began to want to combine my harp playing and my music technology that I was working with. I also work as a sound engineer for other bands. In that work, I guess I have experience with sound processing a lot of the technology that might be used, so I'm comfortable with mixing desks, with microphones, with audio signals, and a lot of the audio equipment. It also means that in my work with other musicians I see a lot of the equipment available, so I guess this familiarity with the technology. Then I started a PhD in 2002 on new technology with traditional music, so because I trained as an art traditional music harp player, my interest was in combining that type of music with technology. And now, I work with different types of technology for live interaction when I perform.

Interviewer What technologies do you perform with?

Ú. Monaghan At the moment, I use a motion sensor and I wear it on my left hand, it's made by x-Io. I had earlier one of this, which was just a small circuit board, and I attach a switch and a battery and everything myself - that was a few years ago. And then last year, they brought a more advanced one which had a switch and a battery encased on the plastic already. I was able to upgrade to that rather than using the one that I had soldering together. That's what I use now for motion capture and depends on the piece what I do with it. I work with Max. I generally have a specific patch for each piece, and I also have some more general patches with different behaviours that I can access. But often, when I write a piece of music, I also have the main patch. As well as motion sensors, I have microphones attached to the harp. So I use that for pitch detection and the detection of silence. So that's really useful to decide how to stop the electronics whether or not the harp is actually being played.

Interviewer When you say motion capture, was it use "easy"? have you considered the device like a musical instrument? Have you used an exploratory approach?

Ú. Monaghan It wasn't straight forward really for me, I had help. I worked with other people who specialize in gesture capture. So on that device, you have nine data streams, so there are a gyroscope, accelerometer and magnetometer. I used the IRCAM gesture capture devices for the violin, and I was able to adapt some of those gestures recognition systems to the harp. But it's quite a specific one. So, I now use traditional music; the damping motion of the left hand is quite abrupt, and it's quite abrupt that it happens quite often in the course of cleans. I was able to use that because it's quite recognizable to the system. But some of my pieces that explore data from the magnetometer is much less easy to control, but I like to use that in the piece. So, in some other pieces, where I want to be able to predict what the gesture I do, I'm more likely to use more the damping motion. If I want the pick or is a bit more ambiguous and I have less control – which means that is a bit more improvisatory in terms of what or at least indeterminate to open the chance in what the electronics do – then, I'm able to use, or I would like to use someone's the data stream that I can. It depends on what I'm trying to do with the piece and how much control

I want to go with the electronics; whether I want them to be predictable or not.

Interviewer Sure. And you told me that you attach the sensors to the x-IO to your left hand. Was it for a specific reason, or it was just a random choice?

Ú. Monaghan Because this is the hand that moves to damp the strings. The strings that need to damp the most are the ones in the base because they ring the longest, so I use the left hand to damp those, and that's the one that has the most recognizable gesture.

Interviewer What specific gesture have you used?

Ú. Monaghan So the damping gesture. But it changes depending on the piece. So in some pieces, I use processing in Max to generate sound, and this damping motion can choose that. I often use signs that derives from the harp, so I will have in Max a buffer where the sound plays into, and then I use the damping motion to select places in that buffer and process them. I use objects that would take a snapshot of the buffer and play that drone. But I don't rely on the last very long, so it's not drone in length, but it's kind of like a not changing pitch for a short amount of time. But then I also can use them to trigger pre-recorded soundfiles. There are some other pieces online that you can hear on my website.

Interviewer If you compare any software with a gestural controller, which one of those would allow a more musical approach in manipulating the sound?

Ú. Monaghan Well, it depends, because every often when you control the parameters of the sound with a device, like a motion capture or with a microphone, it is in a way more musical because you are often using movements or gestures or cognitive processes that are related to your music-making. So in that way, it is more musical, but there comes to a point at the end of that skill where is not musical because it is very hard to arrive at a point where you can control it the same way you can control an instrument. I have to use both. I wouldn't say one is better than the other. It's very clear in software what you are controlling and what parameters; you can see that once you've moved to a parameter realm. It's not as clear anymore because the mapping isn't so good. I wouldn't say that a keyboard if I'm talking about a piano keyboard, is that helpful to me because I'm not a pianist. I also would say that a computer keyboard is helpful in terms of emotion and feeling because it is not the way we are trained to play our instruments. This is why I try to use gestures that are already existing in the harp repertoire. So that's sort of the thing that I do; with this motion sensor, I try to capture things that I already relate to the emotion and feeling that I put into the music. I don't like to use extended gestures that involve me weaving my hand after playing because I don't feel like I'm a dancer or a movement specialist.

Interviewer Do you rather prefer you use your own gestures and to adapt those to the technology?

Ú. Monaghan Sometimes, I mean, those are very useful because they are already related to my music-making, but I'm also comfortable working in software with a mouse. I also use the Korg Nano-Control. If you have a gesture that you know is what you use for emotion and musicality and you can capture it, then that's great. If you have a software parameter that you can control in a different way that's also great. So, it's a balance.

Interviewer If you would choose a controller, which o would you chose?

Ú. Monaghan ne Well, if I weren't suddenly allowed to keep one of my controllers, I would keep the gestural one. That's partly because the gestural one can take gesture, but it also has on/off buttons that can take like a button input as well. I think on the scheme of things that the one I use for movement is more useful to me than the Nano-Control. I could do the Nano-control stuff in the software, but nothing will capture my gesture like with the motion sensor.

Interviewer Have you ever used guitar pedals or any similar technology?

Ú. Monaghan I am familiar with it, but I didn't bother to use them. I found them from others' use very limited. I guess they provide an entryway into manipulating sound, but for me, I always knew that I wanted something that could let me exploit what I needed to do, to exploit the potentials of the software. If you have a guitar pedal, and for example the loop-stations, it would loop something and would then be tight to this particular rhythm. It also didn't really manipulate that, so once you play, it carries on producing the same thing; I knew I wanted something extra. And then the other pedals that we need in real-time to process while you are playing, I never liked the sound of them, I thought that they were developed for a particular instrument, for a start, but also for a particular genre of music most of the time. And then those sound where in my head quite tight to those guys performing those instruments and those genres of music. I wanted something extra, and I wanted something that I could control, not something that some else on the shop has decided what it sounded like and I would then use. I wanted more control, and I wanted something different, and I wanted something that I tailor to me. And for that reason, I thought that if got into guitar pedals it would be a waste of time and also expensive because these things cost like £50/60 at least. So, I didn't bother with them.

Interviewer During a live performance, would it be more reliable and accessible using the laptop, Max and soundcard, and whatever else it requires or a guitar pedal?

Ú. Monaghan Well I've never done live gigs with guitar pedals, but I know that they are a lot more reliable than the work that I do. And they are more accessible, so that is something that I have to deal with. When I work on live performance with this work, I need a longer soundcheck. I need to do different checks. I need to be able to predict possible problems and have plans for as they happen. It makes the whole performance a lot more stressful than if I was not working with this stuff, but for me, it's just a different step on the road. I'm not going to go backwards and stop using them because it's not what I consider to be my music, so I have to do it or don't. It does produce problems; it's stressful; it requires a long soundcheck, it requires a better engineer. All of these things mean that it's hard, for example, to perform in festivals or there is only a tiny chance. It restricts the contexts that I perform in, but it also expands them. It's the reality: it is not great in some ways, but it's the choice. I know that if I work in performance with guitar pedals, it could be more sure that it would work. They would also take a lot more of use – you know, you could pack all that into a bag, whether as with computer is expensive, is valuable, is what people make ticket; there are all these problems, but I play the harp, I'm used to problems.

Interviewer Do you also work with visual feedback? Lighting or video?

Ú. Monaghan I don't project anything with my performances. I find it the coding, and the process, and the interaction with the music and sound is so far enough for me. If I wanted to do visuals I would want to employ as best as a visual person; I'm not experimenting on visuals myself, right now. I think there is a lot to be done with sounds, and my focus is as a sound artist, so I haven't gone there yet. And I don't think I would necessarily. I just feel like I'm someone that likes to work through things: like if I'm already occupied with the sound, it would be kind of a specific reason for me to go and include visuals. I'm more likely to get a visual artist to do that.

Interviewer Do you design gestures considering the visual impact it would have on the audience or do you design them to facilitate the musical performance?

Ú. Monaghan Both of them. I work on all these things in parallel. If I want something to work with a gesture and that's going to happen, I'm aware of the visual feedback of the audience. Sometimes, I push the button on the motion sensor, and I want the audience to see it, sometimes I don't want to show that, so I don't make it obvious. But other times I want them to know what's going on. So, I make it more obvious. I'm conscious of these things all the time, and if they need to happen for the composition or for the sound, then they need to happen. If I also wanted them to be visible, I would leave them visible. It is always a parallel process between what is practical, what is compositionally desirable, what is desirable for the sound. I'm thinking of composition, performance and practicality of the software and coding in parallel. I work with these strands as important as the others. Because as you said with the performance: this work presents problems at a performance, and so it's not just composition and performance, so you need to think of is like whether it's going to be possible on a live stage. And not just possible, but also reliable. I know that all those things need to happen at once. I have certain pieces where I like the composition, but they are not going to work reliably. Then the reliability and dictations of the software and the coding with the performance of the lights would have an effect on the composition itself.

H.1.3 Interview with Audrey Harrer

Interviewer What technology do you use and how you use it?

A. Harrer I studied music composition in college, and I took the harp maybe five-six years ago. It was sort as a free thing to start over to create something, to kind of learn something from scratch, to have something fresh and new, to play with as a musician, because, you know, sometimes, you need that in your process. I started incorporating electronics mostly from an orchestration perspective because I would perform, and I would think "wow, I wish I had more sustain". So, I researched to find out what I could do to discover that so what I use for that it's either a delay pedal sometimes. I use a lot of guitar pedals. That part of the exploration first led me to the world of guitar pedals. There is this dealer who sells them, they call him "Boutique guitar pedal". He's a pedal dealers and guitar builders that work all around the country, and he knows a lot about this kind of things. You'll tell him "I would like to explore this kind of sounds" and he will come to your house, with a suitcase with this amazing pedals and you can try them with your instrument, which it's really hard because you cannot really take it to the store. I ended up exploring guitar pedals a lot first, that I picked up for the acoustic harp that I had, but now I have an electric harp. But I first started with delay. The freeze pedal is one of my favourite things, what it does is you put your foot on the pedal, and it will hold the sound, like freeze the sound, whenever you press the pedal; it acts sort like a suspend pedal for a piano. I played with distortion, I played with pitch shifters, so if I wanted a much lower active or a much higher active, or redouble. Especially, if I'm doing loops and I want something progressive, usually for a quick drum, I'll pitch the harp down an octave and then hit the harp. It will make it much deeper, "boomier" like drum noise. Then I used looping, once I explored a lot of the sounds, and fed a lot of the different colours I could use using pedals. I started to loop, to make compositions, vocal-harp compositions with other pedals. And then, I started to incorporate a small synthesiser to add a few more colours to things, and I also used a vocal processor that I vocal harmonies with.

Interviewer Do you approach electronic devices during a performance in the same way as a musical instrument?

A. Harrer To perform, everything combines to make in my mind sort of my set up for creating sound. Muscle memory is very important for all of that, having the table and my keyboard at the same height, knowing where my pedals are and having them in the exact same position. In that way, it is pretty much like performing it physically like any other instrument, because muscle memory is a very important part of how enables to execute all the layers and moves and things like that. In terms of sounds too, I just developed a relationship with what all my technology does. I understand that when I hit this button, I know exactly how the sound is going to react, you know what is going to do. So now, I very much like a traditional instrument, and I think that's why I prefer using hardware versus a laptop and software. Most of the times, because it is more gestural, it is more physical, rather than harping buried in a screen, and being less about movement and less about physicality.

Interviewer When using technology during a performance, does it disrupt your interaction with the harp?

A. Harrer It's like everything. It's like learning a passage at the piano, or it's like learning a line that is like challenging to sing. It's just about practice and getting into your body. Different types of complexity based on how technology is being used. If I'm with the pedal and I'm just playing the harp, a lot of times, the pedals are just about colour and sustain, with that freeze pedal I was telling you about, but that's much more sad at the beginning. If there is a part that I put on distortion it's pretty easy to know when it's coming, I have muscle memory for when I'm supposed to hit that, and that's about through practice. If I'm building a loop song though, it has to be very very rhythmically precise. Although it doesn't have any sort of glitches in it, so that part is a bit more challenging. Then, once I have those loops layers, my playing is a lot more improvisational based on how I make the song. I would make different choices each time, up to what pedal I decide to use or what sound I look for.

Interviewer Do you feel the need for new technologies?

A. Harrer I think for me it's been about getting the tools and learning how to use them, up until this point. It's knowing that I wanted a more aggressive grabbier sound and how do I achieve that; it's either through using a pedal or finding some piece of software that I can connect my set up to. I guess, I am at the point now where I've explored a lot of this. If I would have thought about things that would have made it easier, it might be triggers that are actually mounted on the harp. It might be adjustments to how the thing [the device] is done, maybe it's a pedal, maybe it's something that I do with my feet; or, maybe it's something that I do with my hands, or maybe it's something that I can trigger in a more gestural way. I've mainly found existing devices, and used them and learn how to use them. But as far as maybe an ideal way to be interacting with some in the type of performances I do, I think there is room for improvement and development of ways to act with those tools that might be physically more simple.

Interviewer Would you use a gestural controller for your performances? Would you adapt like kind of more your already learnt gestures that you used to play the harp also to play the electronic part? Or would you use different gestures, like pressing a button or removing a slider?

A. Harrer I could see room for both of those. For gestural things, if you are talking about like the glove thing, having something like that could be very interesting. Or if you play the string with a different part of your finger perhaps something else happens. Perhaps, if you move your hand at a certain distance from the harp, it creates a particular sound. I think the technology would have to be pretty tight to the harp itself for it to be very useful; otherwise, it is you leaning away from the instrument to interact with another tool. If it were to be more of a part of the instrument, that would be amazing. But I think that could be the same for anything: a guitar, a piano, any instrument. The closer the interactions can be to the actual instrument, the more detailed those interactions can be, because you are not moving all over the place.

Interviewer Have you ever used visual feedback during your performances, for example, lighting effects or projections?

A. Harrer Not that I've controlled myself. I mean, that would be really cool, but I haven't gotten there yet.

Interviewer Is learnability, affordability, availability, or any other issues of technology?

A. Harrer As far as the production of it goes, I have a pretty understanding of that technology. I work at Berklee College of Music, I'm the creative director here, so I do a lot with music and video. But in terms of including it onto my own performance, I think I just have so much going on already with the sound that adding a new layer of visuals it's kind of a lot to take on. I would probably consider doing more of that work if I had a larger group or if I had technical assistance during a performance. But most of what I do is me setting everything up and performing solo. I think a lot of it just has to do with the technology that I use; for the sound, making is so intense that an extra layer is just a little intimidating as a solo performer.

Interviewer You said to prefer pedals and loop-stations because you find them more physical and more gestural, and you don't like to use laptops because you don't like watching at people staring at a screen during the performance. Is it just because of this that you use pedals? Or you use pedals also because of any other reason?

A. Harrer I think it has to do with reliability; I've just observed laptops crashed during so many performances, and then everything shuts down. If I have any technical difficulties, it's only with one object, and I can tell exactly where that is, and it doesn't prevent or pause the performance at all. So, that would be another reason because I'm aware that you can use controllers to interact with your computer in a very similar way. You can put switches and all of that, but I think reliability has been the other issue. Because sometimes it crashes when I'm even trying to record something, I have good computers too, and the pieces of gear are very ragged because they're meant to be carried around and used.

Interviewer When you do looping, for example, do you use a pedalboard or others that you control with your hands?

A. Harrer I use a pedalboard, I use a vast loop-station that has three layers of loops (so I have three loops going at the same time), and you can also pre-record things into that. What I usually do for loops that require a lot of holds up, I'll do one performance where the audience can kind of see what it's happening so they can see the creative ingredients. But it can sometimes take a minute and a half or two minutes, so to prevent that from happening and to play more music during the curse of a set I'll programme some of those layers already, just to keep the times of the song from getting too long. I use a pedalboard and is a pretty advanced piece of equipment: like you can reverse the loops, you can change the timing of that, you can have up to three things going on at once, and you can layer on those three different loops as much as you want. You have to kind of cleverly figure out how to structure things, and that's the reason why it's so important to have access to a lot of or at least several layers of loops because to keep the performance interesting and to change things up in the course of a song you need that flexibility.

Interviewer What determines the choice of chose of sound signal processing? Is it because of the characteristics of the harp sound or your way of playing?

A. Harrer I think is because of the aesthetics of a lot of it. Both, aesthetic and stylistic choice, because my music kind of inhabits the world that I call it odd chamber and it appeals to people that like modern and classical music, but also people who like pop music. I think it helps to bring the harp into a space that people can understand and hear in a new way but also kind of understand what that sound is. I haven't dealt into using a lot of things like Max or Live or any sort of media proprietary product, just because I do try and keep my music very accessible. I do meditation music sometimes, and maybe for things that have more space or that might be less depending on structure and theme. I could imagine exploring those a bit more, but since melody and structure are an important part of my work, and I do like trying those connections with popular music, I think that's why those products seem to make sense with what I do.

Interviewer Have you used panning or spatialisation effects?

A. Harrer I used panning in some of the loop placement for that little background vocal part. I want that to be separate in distant from the lead vocal, so I'll make sure I use like a vocal preset that has a bit of a difference in the queue, and that's panned to a different space in the mix than the lead voice.

Interviewer On a pedal effect there is commonly a LED that indicates if it's on or not, and even though you cannot really hear it if you know that it's on, you can imagine the sound that the audience is listening at. But with sound spatialisation or panning is a bit more complicated. It can be hard to have spatialised audio on stage. Do you experience this issue?

A. Harrer I usually during soundcheck record a few things and then go and step out into the space and make any adjustments that I think is necessary, in terms of hearing it during a performance. If I have stereo monitors, which I often do, I usually have one speaker coming at me with the mix. I mean, I think that's the reality of a lot of performances, you don't have the audience perspective as a performer because it's almost impossible, I think.

Interviewer In that case would be more helpful to have a visual representation of the spatial position of the sound? For example, a display on the pedal?

A. Harrer Well, it depends. I think visual feedback would be most important for whatever is most important going on on a song, right? So if spatial orientation is an important part of what is needed to communicate what I'm trying to communicate, I would want the visual feedback, to make sure that I'm accurately doing that. But it could be overkilled because if it's not important to the song, I should be looking at the harp or whatever else it's that I'm playing or trying to execute or connecting with the audience. So visual feedback is very important by it coordinates directly to the importance of what's being played on a song.

H.1.4 Interview with Sofia Asunción Claro

Interviewer What is your experience in using music technology in harp performance?

S. A. Claro Beginning using music technology wasn't obvious because the repertoire for harp is quite small, that was one problem in my opinion. But also because it was also a contemporary expression, I wanted to know something about it. The first piece I asked was *Strings & Shadows* by Åke Parmerud, the Swedish composer. I was very happy because playing with electronics is like in a way being a soloist with the whole symphony orchestra, which is of course quite exciting. I mean, to play a solo harp piece it is a very intimate thing, but if you have the electronics then the expression you reach is the opposite, it is not so intimate, is much more open. That gave me another kind of interest in approaching the music, which was also very attractive. So that was the real reason why did I come into electronics. I was interested in what was going on now. It was a musical necessity of finding out how I could make the repertoire on harp broader, because it is so little.

Interviewer When you performed electronic music for harp, were you performing using tapes or also sensors and devices?

S. A. Claro I was using tapes and then interactive. Also using cameras, that would change the sound.

Interviewer Did you use any pedal effects, like for guitar, or any other device controlled through buttons or sliders?

S. A. Claro No, I didn't use any other.

Interviewer How did you find the performance with interactive technology like the cameras? Did you find it engaging, or did it kind of put away from the performance?

S. A. Claro I don't think so. I don't think that the effect of this technology was so strong in the expression of the music, I must say.

Interviewer Where do you see the power of the expression in the music? Do you see the expression in the composition itself or the effect of the technology more?

S. A. Claro Of course in the composition itself, but with the sound used in the electronics. And this matches together. It's very interesting because it is like playing along with the instruments, like in a symphony orchestra; for the soloist, for the harp or the chamber music, where the different sounds can be different instruments, so it's making the whole dialogue more attractive.

Interviewer Have you used technology by yourself or with the help of a composer or technician?

S. A. Claro No, I used it with a composer. I was just interpreting, only that.

Interviewer Is it for any particular reason?

S. A. Claro I was just interested in the music and what the electronics give but not as much to do it myself.

Interviewer Was the approach with technology difficult?

S. A. Claro No, it's ok, I don't mind at all, it's not complicated or leads to any negative experiences, it's fine.

Interviewer What do you think about using gestures to control the electronics?

S. A. Claro I think it is interesting to listen to the result of my own movement. It is because it's quite unexpected that because you and the sound will change. You are not playing the instrument, you are just making a movement. So I think that's interesting.

Interviewer Have you used lighting during a performance?

S. A. Claro No, I did not. Well, there was always light, a component of the piece, you know. It was for the necessity to see, no more than that.

Interviewer Have you experienced any technical problem that affected the performance?

S. A. Claro No, not at all. It was always a pleasure and a positive experience. It was always very pleasant. I never had such problems. When I practice, the practical things have to be always under control, and then I was very well prepared in that aspect. So I know exactly how things have to be done in order to make things function. I was only twice, in Grenoble (France) and Chile, that I had to stop. But it was not because of me, but of the person that was manipulating the technical part, the tape. If I were alone, it would be much better. It would never have happened. I was practising a lot with the tape, I knew how to manage very well.

Interviewer When rehearsing the piece that uses the camera, were you rehearsing by yourself or with a technical assistant?

S. A. Claro I think that someone was managing it [the technical part], because the camera was already in the right place where they could, you know, to look at my hands or my face, but then it works by itself. I had to do absolutely nothing.

Interviewer If you had to control the technical part as well would that represent an issue?

S. A. Claro No problem at all. You adjust to include it in the whole thing. You prepare everything, so it doesn't matter if I have to. In some pieces, I have played, for example, I had to stop, I had to stand up, and then walk around the harp and do things. I have practised all that. Doing other things like playing the percussions. All that it's a part of performing. It is a part of the whole thing. It has, of course, a different character than if you just play. It adds acting elements that belong to the piece.

Interviewer Do you want to add anything else about you or electronics?

S. A. Claro I'm very happy to have had the possibility, sometimes, to know the composer. That's the human part of making music. It was something completely out of nowhere. You have immediate feedback that helps you in fixing the mistakes in the way you perform. Electronic music gives you contact with other kinds of musicians and festivals; it opens your possibilities.

Interviewer About the interactive piece using the camera, where the gestures notated on the score?

S. A. Claro Yes, they were written on the score. And then, I can't say exactly how, but it was clearly written what I had to do. Just to move the hands, the arms or whatever. It was more or less explained clearly.

Interviewer Was also written the audio effect resulting from such movements?

S. A. Claro It was only written if I had to move fast or slow. No, the electronic wasn't written, I had to listen a lot to the electronics.

H.1.5 Interview with Arnaud Roy

Interviewer What is your experience in using music technology in harp performance?

A. Roy I work with a French harp manufacturer called Camac. Back in 2009, I was called by them to work on a new type of app and it's basically is a MIDI harp. Back in time was a prototype, which is now developed further by another team because the engineer who built the first MIDI harp did not continue the project. Now I'm working with IRCAM, in Paris, is a contemporary music centre, and I work with them as a beta tester to debug give feedback about the instrument. In live performance, I work with Ableton Live, and actually, the workflow is very simple because the harp acts like a keyboard and so I can control all the MIDI interfaces in Ableton Live. I also used a MIDI pedal because as a harpist, my hands are busy, and I cannot use it for the computer. I use a MIDI foot pedal to change the sound of the harp, I can play a synthesiser, a violin or whatever. So that was my first configuration, my first live set with MIDI harp. And now I'm working with another harp. It's simpler. It's just an electric harp. For this, I use a Leap Motion to modify in real-time the sound of the harp. It's another project, and I use to live set to perform.

Interviewer Do you think that interaction with the electronics through gestures is closer to the way you perform, or you were fine pressing buttons or the keyboard of the laptop or anything else?

A. Roy I think the Leap Motion is great because it adds signage for the audience. It is something special when I move my hands the sound changes. So I think this actually is very interesting for people, the Leap Motion. For me is like a texture. Like you know a pad, I usually transform the sound of the harp with a guitar pedal. I use it to make a very long sustain on the sound of the harp-like it transforms the sound of the harp-like in a pad. And then with the Leap Motion, I use guitars or something like that to transform it. I think it's a very interesting way to perform, to have good control of your sound, and it's very accurate. I think the Leap Motion is accurate so I can do whatever I want with it. Sometimes it's a little bit buggy, a little, but I feel comfortable with it.

Interviewer Where do you place the Leap Motion when you play the harp?

A. Roy On the laptop. I work as a video game music composer, so my live set, I perform some video music I wrote. I remix them and play them on the harp. In *FlyByNo* live, I used on this video a Leap Motion but is not a MIDI harp. It's just an electric harp so every time the sound it's a harp sound, and it's modified by plugins in Ableton Live.

Interviewer When you use the Leap Motion, you have to stop playing and then move the hand far away from the harp. Do you find it this gestures disrupting the performance?

A. Roy No, in fact, it's a new way to perform. In the introduction of its video you can see that with my right hand I play a cord on the harp and with my foot I block it, I freeze it, with a guitar pedal and transform it into a pad. And then, with the Leap Motion with my left hand, I can modify on the pad. And also, I have another plugging which is called Movement, and it transforms the pad into an electronic rhythm. It's kind electronic dance music.

Interviewer Would you look for something that enables you to control the sound while you play the harp without taking the hands off the harp?

A. Roy Yes, about the MIDI harp, they are building a new version of the MIDI harp. I asked the engineer to have control on the board of the harp. It would be much more convenient to control the MIDI and anything else. Why not a Leap Motion directly on the board, for example? I think this new instrument, and I hope they are going to implement this new kind of controller directly on the board of the harp. Maybe in a few years, hopefully, we can see this kind of instrument.

Interviewer You mentioned some of the signal processing techniques that you apply on the harp. For example, the sustain to make the sound longer or to freeze the sound. Do have you any specific reason for using these approaches?

A. Roy I think is just a question to have a different kind of sound, because the harp's ADSR envelope is like a plug. I'm very happy to play the harp, but sometimes I want to play like the violin or strings and a more pad-like sound. I think sometimes, I need this, but not always, so yeah, it's good for that.

Interviewer In reference to the MIDI harp, how does it feel to play a different instrument? Does it come naturally or something that you have to get used to it?

A. Roy No, it's not natural; it's very weird. You have to learn to muffle the string much more because the MIDI harp senses when the string is stuck. So you have to muffle the strings much more frequently than in a traditional way, so as a technique is a little bit different. When you play a monophonic synthesiser, for example, to play the bass part of the song, then you can play like a bass guitar on the harp, and you can forget about muffling the strings, so it's very enjoyable. Sometimes, it's more difficult, for example if you want to play the piano on the harp, it's very difficult because all the strings are together and you have to muffle, it's more difficult. But for example with monophonic instruments it's very enjoyable and very simple to play. So, it depends on the kind of instrument you play.

Interviewer Have you tried different gestural controllers than the Leap Motion?

A. Roy For now, I have little time to do live performance. I'm quite stuck in my studio, so I have this Leap Motion sync and I'm sticking to it for now. But in the future, I will work on another MIDI harp prototype and so when this new product will come to life and hopefully will become commercially available, then I'll make new live set with this new instrument. But for now, I stick with this because I don't have much time to experiment other hand gesture systems.

Interviewer Have you ever integrated lighting in your performances?

A. Roy Not lights, but video. For *HarpJamX* I worked a video guy as a due. We had two computers and there were link by MIDI and with the MIDI clock. He used Modulate to perform video. So it was very simple because we didn't have much time to do this live set. Basically, when I play on the swing, there is a video and some effects on this. My teammate he almost does nothing on stage, only to see if the video is fine. Every video is controlled by MIDI harp. So we worked like that. It is video sampling; it's very in fact very simple.

Interviewer Before to use the Leap Motion or the MIDI harp, were you using technology, like guitar pedals or other effects?

A. Roy Yes, I have a rock band and, I use a very simple guitar pedal effects like delay and distortion. When I was younger, my first set was with a looper, a Bose looper, so I played, recorded a loop, then I played percussions, then I can improvise on the loop.

Interviewer Do you think that interacting with sound processes through gestures and the MIDI harp has improved your workflow?

A. Roy It's a very good question. I work as a composer and also as a sound engineer, so I'm very passionate about technology. But sometimes, technology can push you away from being creative because as a result many stimulus, so many instruments, so many plugins. So, I have to constrain myself to learn another technology because my main goal is to create music and no to learn new tools. Sometimes, I have to tell myself "it's enough, you have too many, already too many stuff, and you can work with this". And yeah, it's the same with synthesisers, when you start to own a synthesiser you want a second one, and another one and so on. I was very happy to learn the MIDI harp because I was one of the firsts to do this. I was very proud of it, spent much time to learn the technology and learn how to play. But, in my live show, I played the MIDI harp maybe four or five times. So it's not much compared to the time I spend to create the pieces. It's a form of research, so you have to be patient.

Interviewer Did it take much also to learn how to combine the Leap Motion with the harp?

A. Roy No, because I use Geko, is a little software. It's very intuitive, so I can easily transform the Leap Motion data into MIDI data. With Ableton Live it's so easy to control plugins, so maybe I have to spend two hours understand how it works, but it's very easy.

Interviewer So in a way was faster to learn how to use combined the Leap Motion with the harp and to effect the harp with the Leap Motion than to play MIDI harp?

A. Roy Yes because, with the MIDI harp, you have to create your own sound-set. The MIDI harp is way much complicated because the structure of the performance: you have to adjust sounds, to make loops, to record on real-time. And with the Leap Motion is more simple, I have a playback, I improvise on the playback, and play with the Leap Motion so yes, is less complicated.

Interviewer Do you think that in a way, gestural technology might be a solution for keeping things simple as well to elaborate the sound of an instrument?

A. Roy I think it's really enjoyable because when I started to play with the Leap Motion, I was amazed by the way my hands were controlling sound. It's kind of magical, is like you are feeling the sound, like you are a sculpture, you know? And you can make a face or something like a landscape. It's really magical. I think it is a really enjoyable performance for the performer and also for the audience because the audiences feel that when you are moving your hand the sound changes. I think, maybe, another idea is to have lights. People love that.

H.2 Extracted Codes

1. Instrumental technique
    (a) hand gestures
    (b) foot movements

2. Sound processing
   (a) Dynamic sound processing
   (b) Reverb effect
   (c) Freeze effect
3. Use of technology
   (a) Learnability
   (b) Gestural technology (Leap Motion, IMU, cameras)
   (c) Gesture design
   (d) Guitar pedals
   (e) Video projections
4. Performing
   (a) Solo
   (b) Duo
5. Score notation

Appendix I

Score *The Wood in the Water*

with thanks to Balandino Di Donato
for making this musical adventure possible

Eleanor Turner

The Wood and the Water

for electric harp with gesture controlled sound spatialisation and effects,
gestures based on British Sign Language and spoken word.

The Wood and the Water

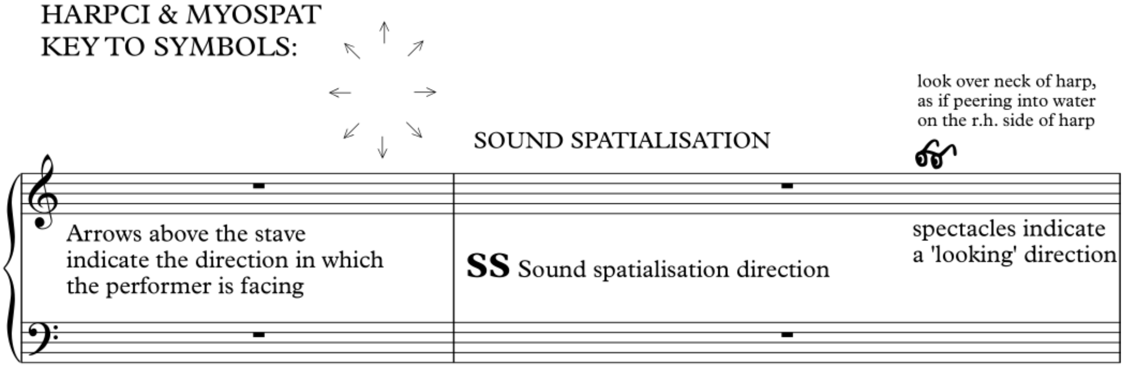
transcript of a sign language poem

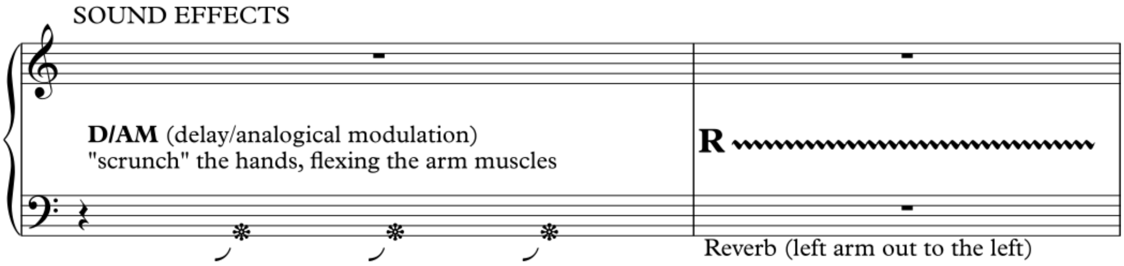
Make some marks on a piece of paper;
How much can I tell this page?
It's that strange, mysterious symmetry again,
A pain I can't bear to gauge.
I see my false reflection in the pond.
I dip a finger in and watch as
The ripples circle outwards and disappear at the brim.

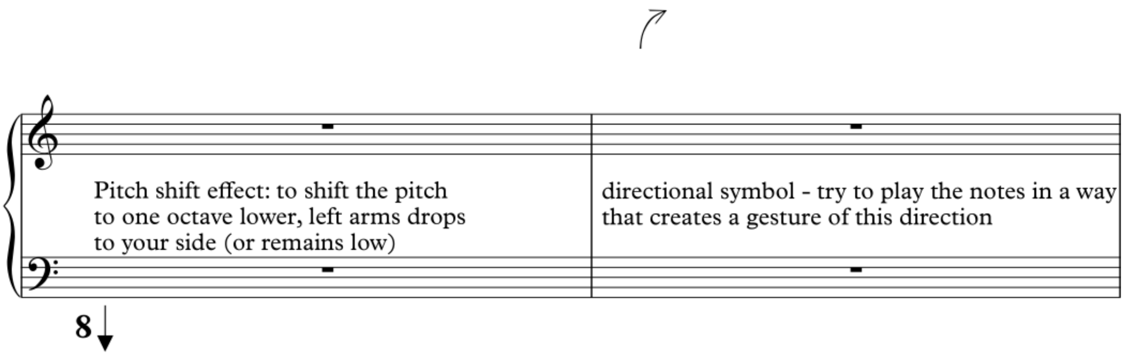
The forester approaches. He takes my two hands
And ties them behind my back. He says
"You're not what I want. You don't even know who you are!
The birds in my trees are real. The call of the wind in my branches: that's real.
But look down there in that pond. Who is she?"
I look and I look until the pond freezes over.
And I wonder if he saw me.

Eleanor Turner

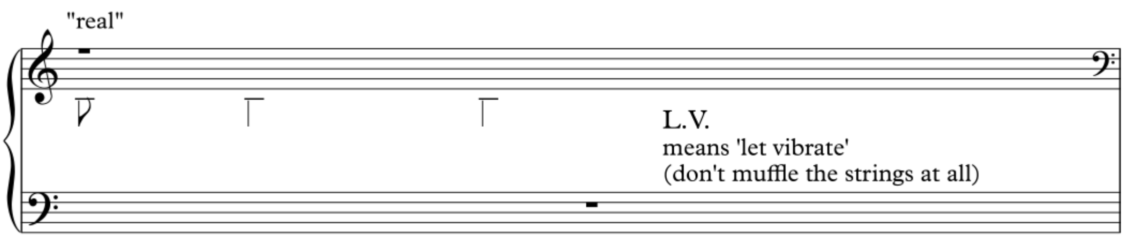
HARPCI & MYOSPAT
KEY TO SYMBOLS:





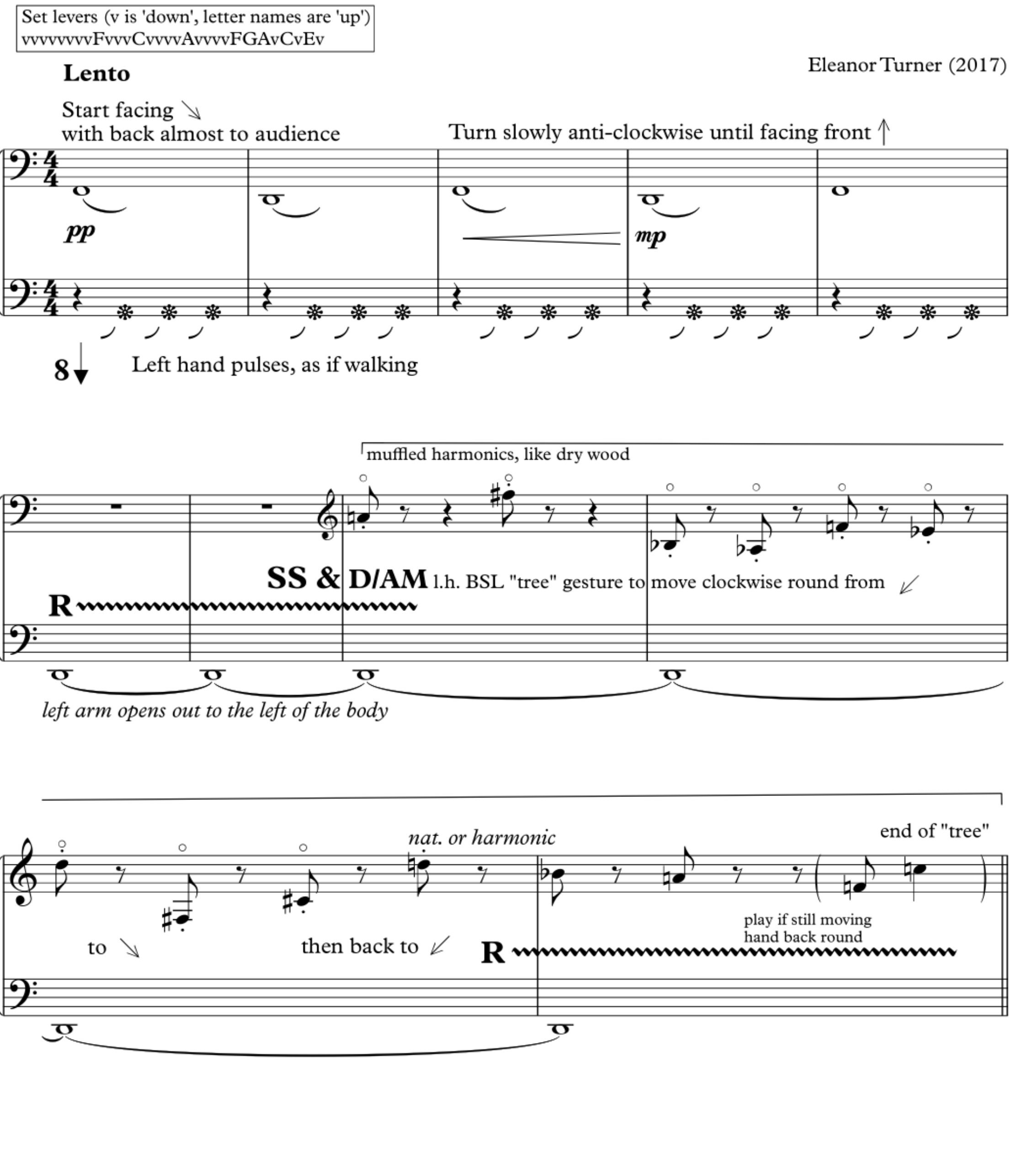


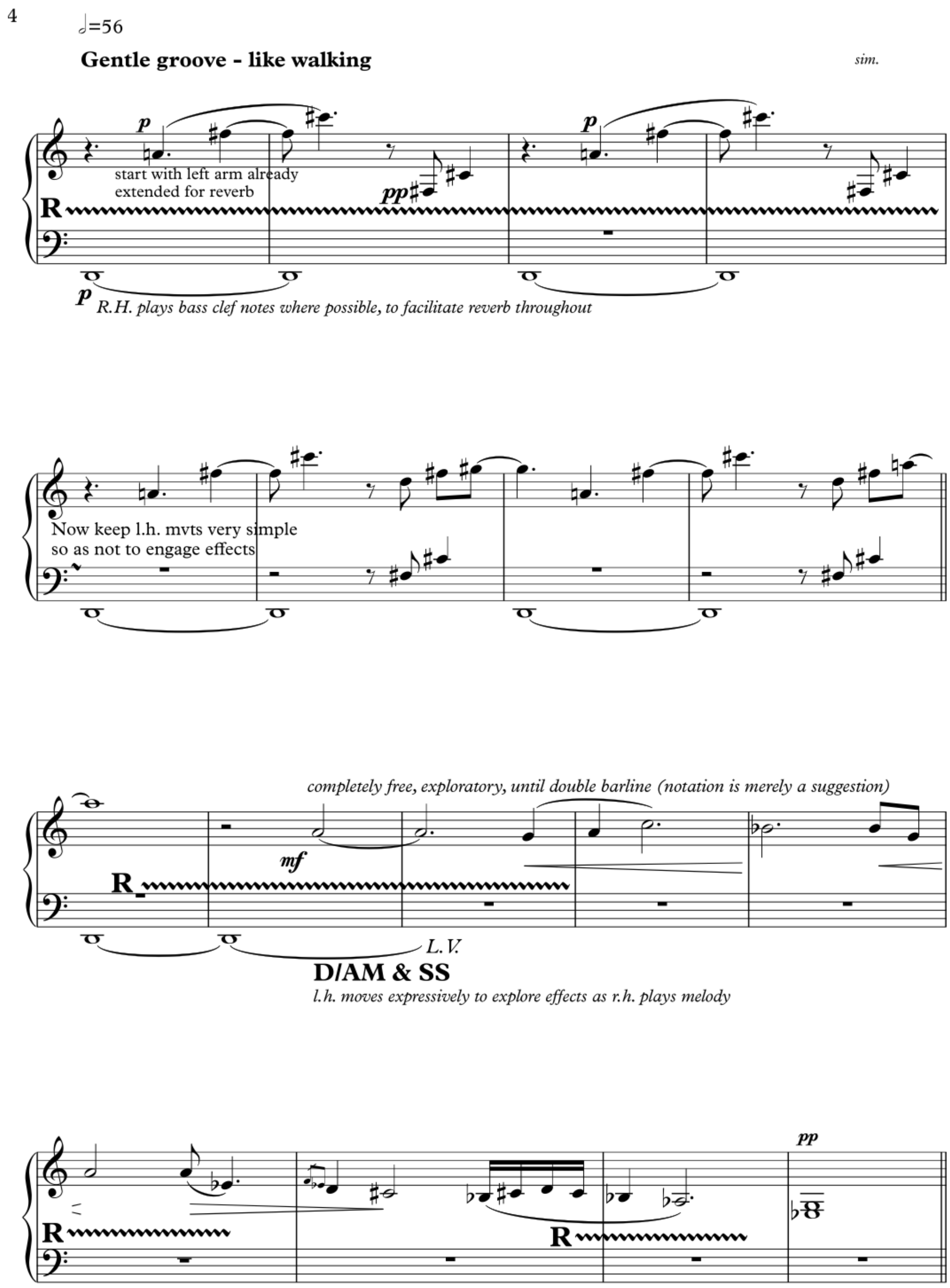
words in speech marks are BSL signs; sometimes these can be done whilst playing, sometimes they need their own time (after playing) and occasionally they are incorporated into the notes themselves (for example, shown below with notehead-less stems, the sign can be 'played' in the rhythm)



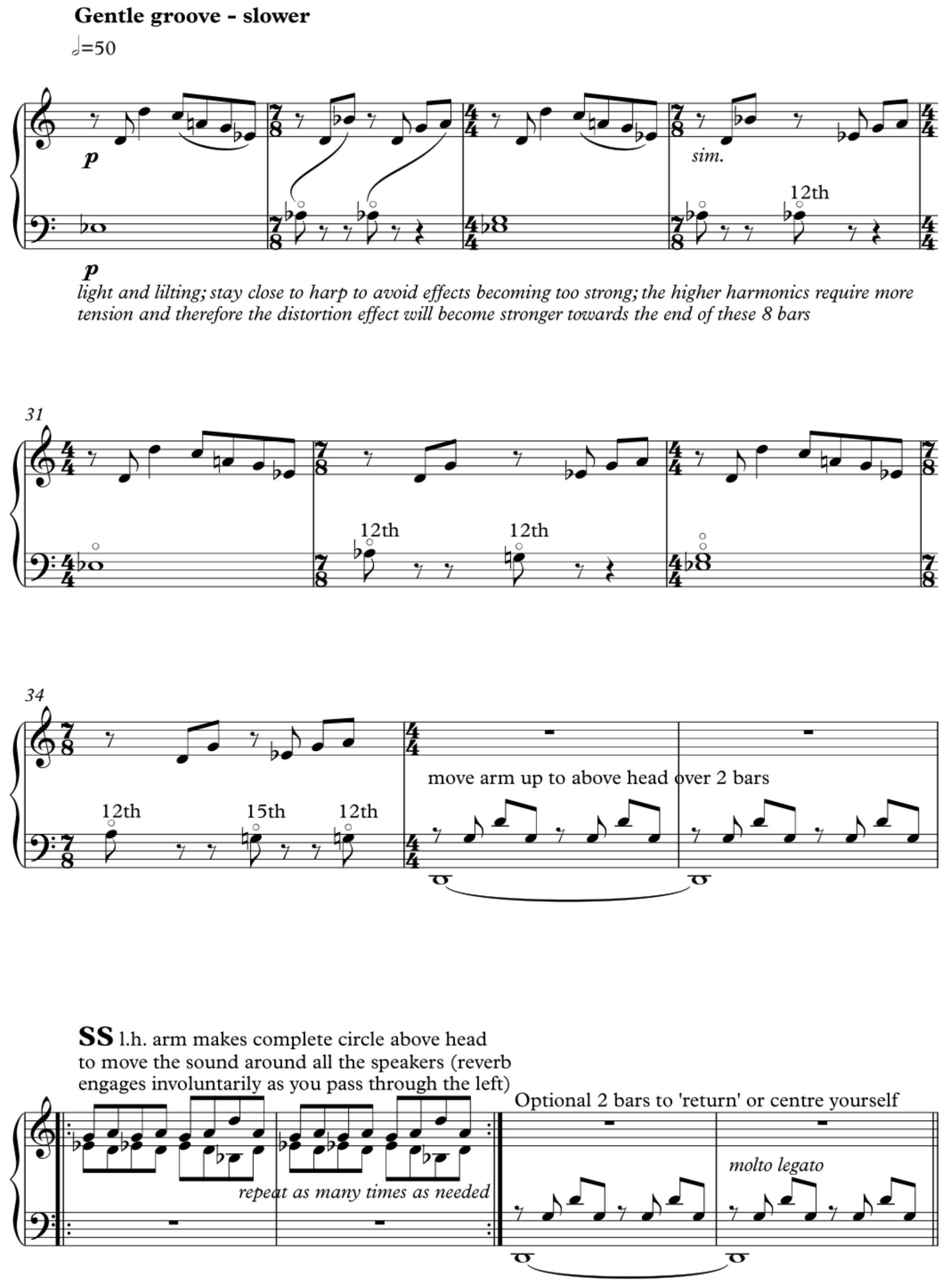
The Wood and the Water

Set the scene of walking into a forest, filled with wonder and starting to hear noises from all directions. The left hand 'scrunch' (making a tight fist) engages the delay/analogical modulation effect, giving a grainy pulsing effect.

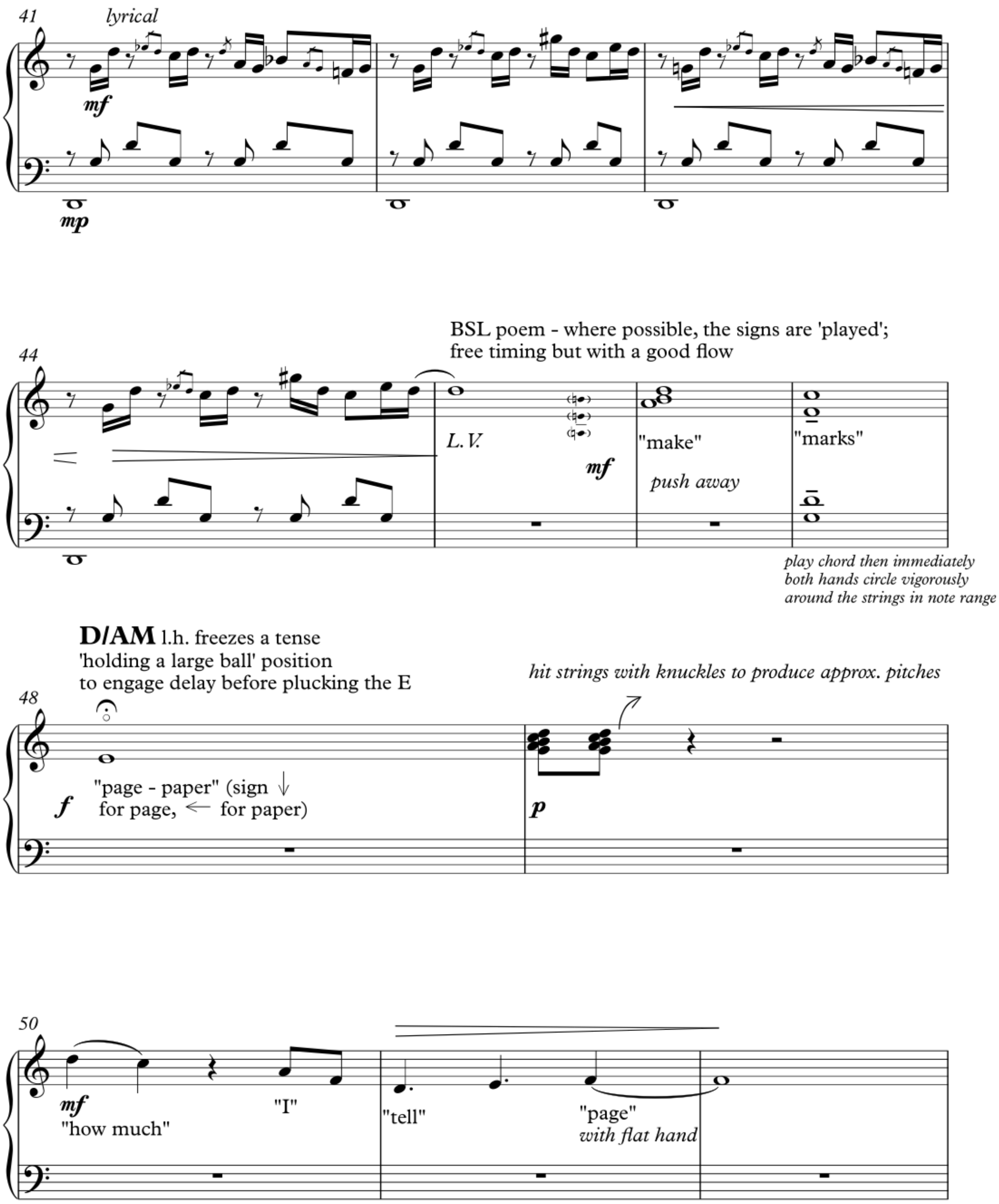



4
𝅗𝅥=56
Gentle groove - like walking
sim.
p
start with left arm already
extended for reverb
pp
R
p
R.H. plays bass clef notes where possible, to facilitate reverb throughout
Now keep l.h. mvts very simple
so as not to engage effects
completely free, exploratory, until double barline (notation is merely a suggestion)
R
mf
L.V.
D/AM & SS
l.h. moves expressively to explore effects as r.h. plays melody
R
R
pp

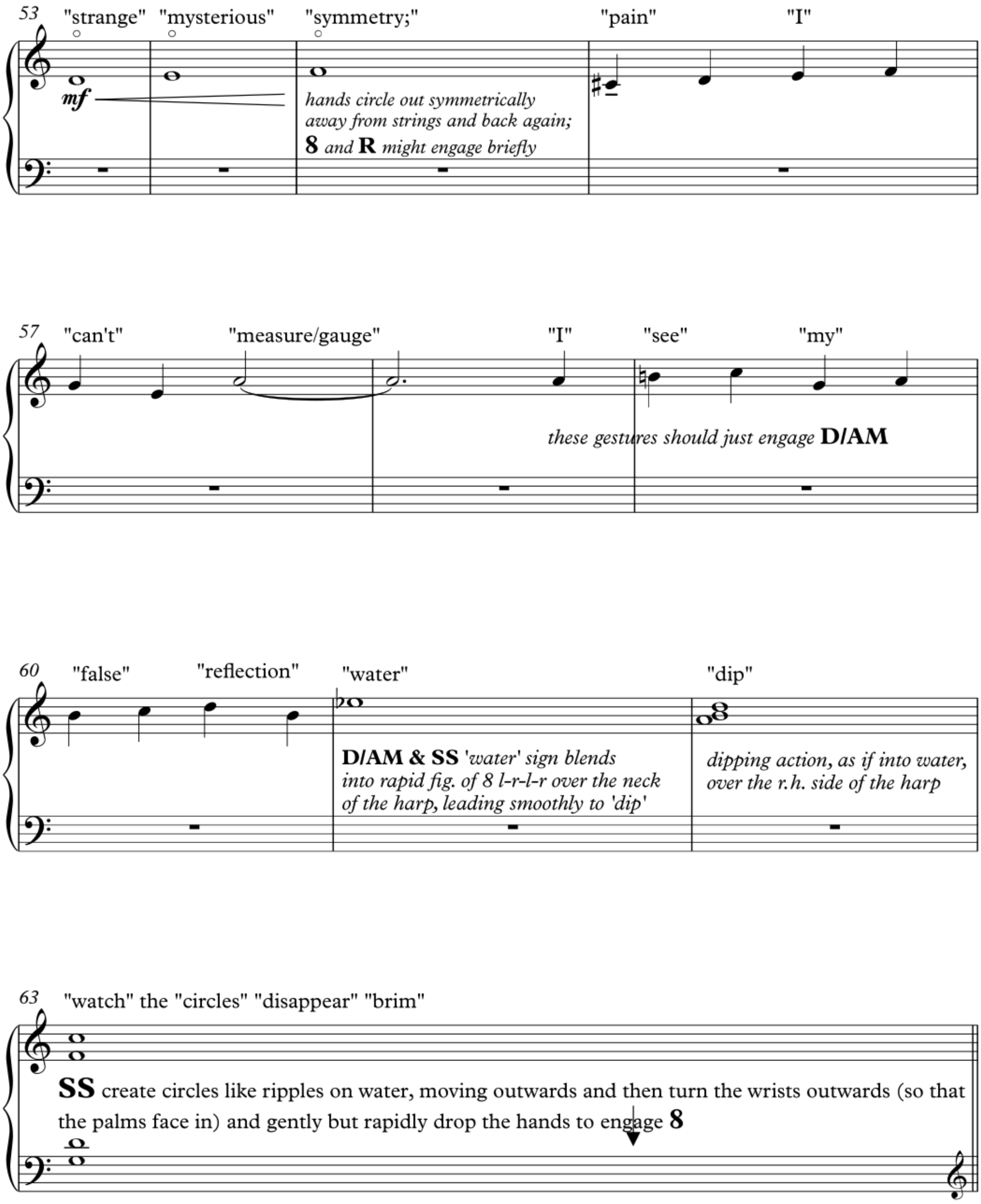
5


Gentle groove - slower
𝅗𝅥=50
p
sim.
12th
p
light and lilting; stay close to harp to avoid effects becoming too strong; the higher harmonics require more tension and therefore the distortion effect will become stronger towards the end of these 8 bars
31
12th
12th
34
12th
15th
12th
move arm up to above head over 2 bars
SS l.h. arm makes complete circle above head to move the sound around all the speakers (reverb engages involuntarily as you pass through the left)
Optional 2 bars to 'return' or centre yourself
repeat as many times as needed
molto legato

6


41
lyrical
mf
mp
BSL poem - where possible, the signs are 'played';
free timing but with a good flow
44
L.V.
mf
"make"
push away
"marks"
play chord then immediately
both hands circle vigorously
around the strings in note range
D/AM l.h. freezes a tense
'holding a large ball' position
to engage delay before plucking the E
hit strings with knuckles to produce approx. pitches
48
f
"page - paper" (sign ↓
for page, ← for paper)
p
50
mf
"how much"
"I"
"tell"
"page"
with flat hand

7


53
"strange" "mysterious" "symmetry;" "pain" "I"
mf
hands circle out symmetrically
away from strings and back again;
8 and R might engage briefly
57
"can't" "measure/gauge" "I" "see" "my"
these gestures should just engage D/AM
60
"false" "reflection" "water" "dip"
D/AM & SS 'water' sign blends
into rapid fig. of 8 l-r-l-r over the neck
of the harp, leading smoothly to 'dip'
dipping action, as if into water,
over the r.h. side of the harp
63
"watch" the "circles" "disappear" "brim"
SS create circles like ripples on water, moving outwards and then turn the wrists outwards (so that
the palms face in) and gently but rapidly drop the hands to engage 8

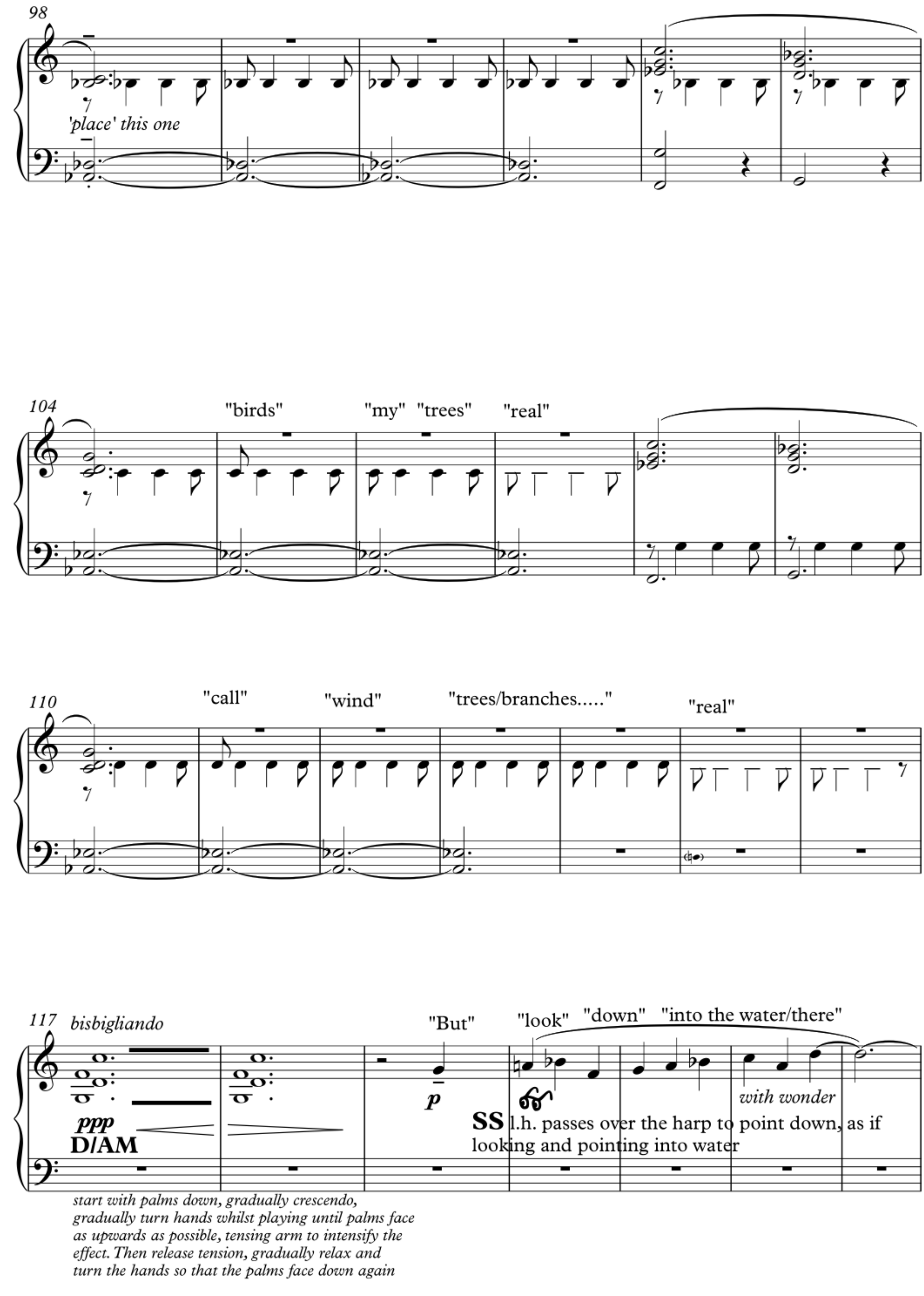
8


TRANSITIONAL SECTION
64
pp
'2 in a bar' feel; fluid
mysterious and reflective
p
67
70
this part of the poem will enage reverb, pitch shift effect and delay/AM in passing
"man"
"belonging"
"forest"
"comes"
L.V.
"he"
"takes"
"2 hands"
"ties"
"behind"
"back"
R
L.V.
8
74
Slow
8
D/AM
push l.h. up as if stretching to the sky
loco
R

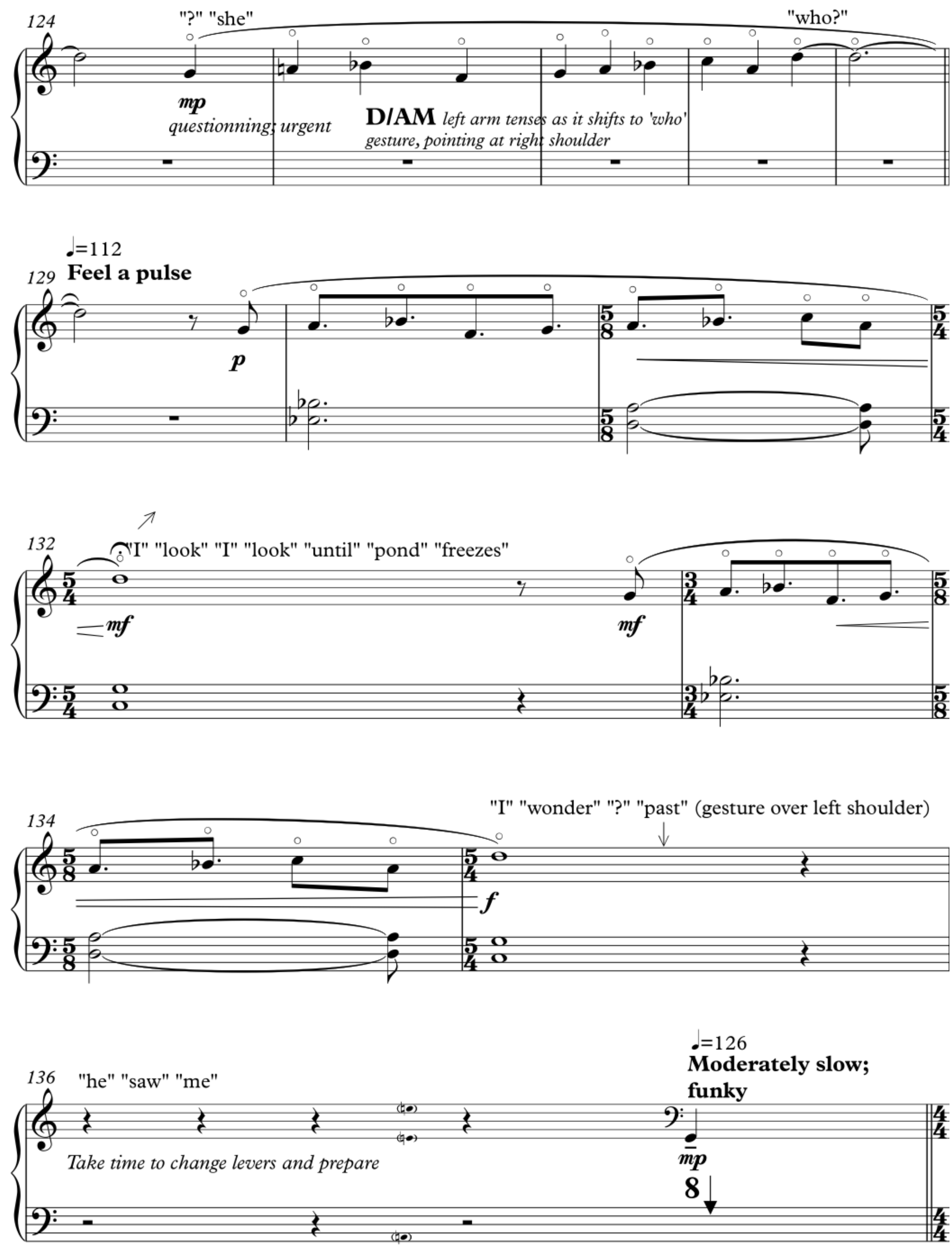
9


Optional 4 bars to explore more effects - very free (bb78-81)
78
p
82
Slow, steady pulse ♩=66
Optional 4 bars to regain pulse (bb 82-85)
85
Walking ♩=66
88
transition smoothly into 3
♩.=136-144
1 in a bar feel
ppp
(repeated note)
p
92
subito ppp
(melody Eb-G-C)

10


98
'place' this one
104
"birds"
"my"
"trees"
"real"
110
"call"
"wind"
"trees/branches....."
"real"
117
bisbigliando
ppp
D/AM
"But"
p
"look"
"down"
"into the water/there"
SS
l.h. passes over the harp to point down, as if looking and pointing into water
with wonder
start with palms down, gradually crescendo, gradually turn hands whilst playing until palms face as upwards as possible, tensing arm to intensify the effect. Then release tension, gradually relax and turn the hands so that the palms face down again

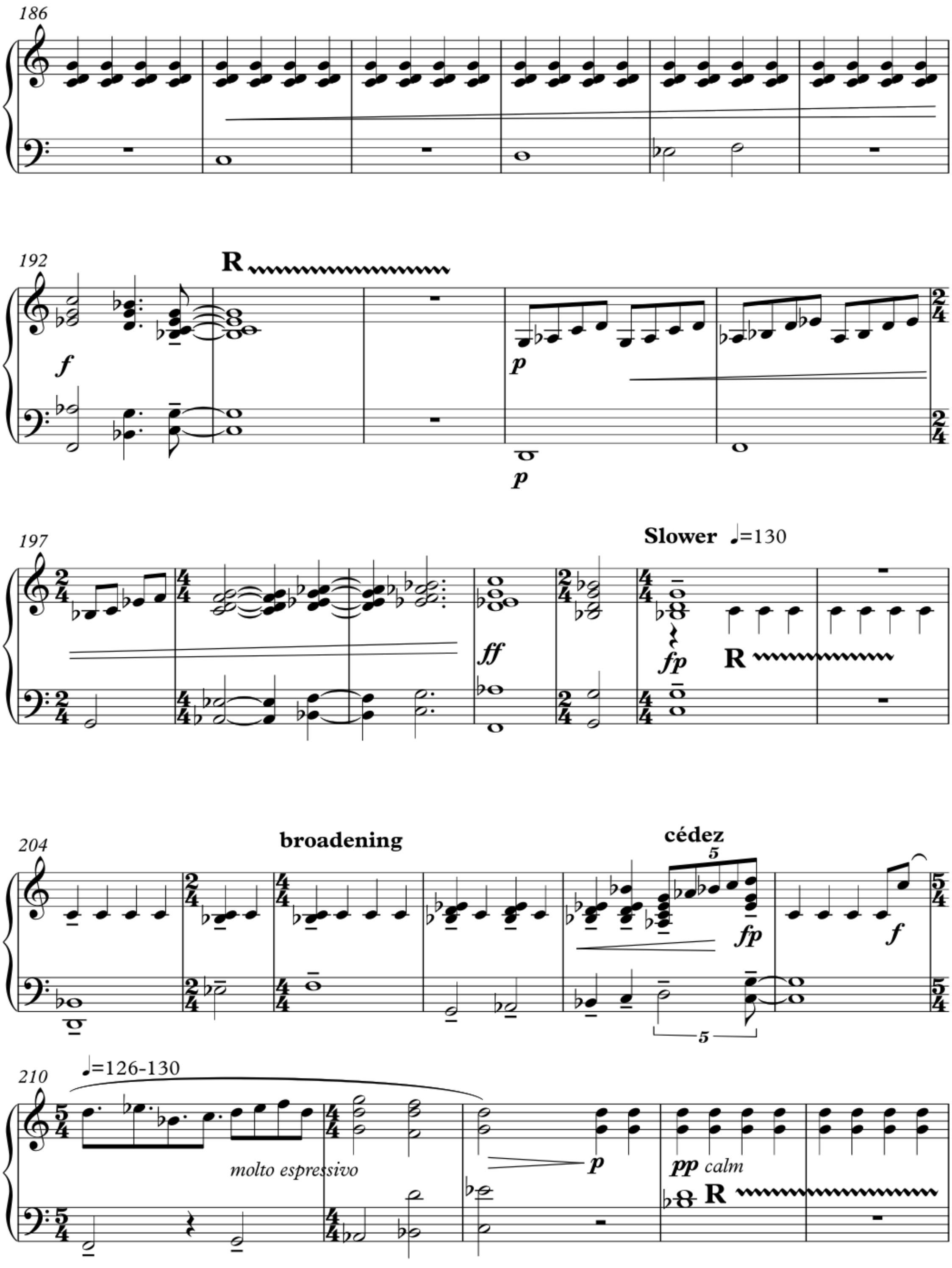
11


124
"?" "she"
"who?"
mp
questionning; urgent
D/AM left arm tenses as it shifts to 'who' gesture, pointing at right shoulder
♩=112
129
Feel a pulse
p
132
"I" "look" "I" "look" "until" "pond" "freezes"
mf
mf
134
"I" "wonder" "?" "past" (gesture over left shoulder)
f
136
"he" "saw" "me"
Take time to change levers and prepare
♩=126
Moderately slow; funky
mp
8


12
137
139
SS & R
loco
144
arm lowers again
mf
8
150
SS & R
loco
(loco)
f
155


160
f
166
mp
mp
p
170
sim.
mp
f
mp
f
175
♩=144
Determined and energetic
180
*insert spoken word (see end of score for block lyrics of my original rap
mp
sim. - played crisply and rhythmically

14


186
192
R
f
p
p
197
Slower ♩=130
ff
fp
R
204
broadening
cédez
5
fp
f
5
210
♩=126-130
molto espressivo
p
pp calm
R

215
220
mf
224
mp
228
p
sim.
Meno mosso
rit.
R
to nothing
8

16

WHISPERED WORDS
Rap for final section
(from bar 180)
This could be exchanged with any relevant spoken word: rap, poem or dialogue

So when I was nineteen years old
And I told you I had a new life growing inside me,
And I watched as your eyes flit from my face to my finger,
To see if a man had made claim to me.
Curiosity gets the better of me too, sometimes,
But you don't realise:
Accumulated effect of your whispered words
Made my body feel absurd. Have you heard?
"Too young to become a mum"
Nature's possibility overrides this inconsistency,
What's the name of the man who began this event of life inside me?
Unaccountable due to inavailability,
Blameless through his absency!
But only God has a name and an aim for me;
And there was never any sense in shaming or blaming me.
And so I rearranged the pain
And I turned it into the fire that fuelled the gain in me.
But how did I dig those whispered words into
My soil and my soul and grow from them?
Well, they became the deep down ocean rumble and jumble
Of waves and words that crashed and broke against me.
And now those whispered words are only ever heard from the
Caged birds who should be out soaring across the
World but instead they are sitting,
Alone, absurd.
In their cages.
Spitting feathers.

Eleanor Turner

Bibliography

Ableton (2017). Live. Available from: `https://www.ableton.com/en/live/`. Accessed: 20 June 2017.

Abras, C., Maloney-Krichmar, D., and Preece, J. (2004). User-Centered Design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications*, 37(4):445–456.

Abreu, J. G., Teixeira, J. M., Figueiredo, L. S., and Teichrieb, V. (2016). Evaluating Sign Language Recognition Using the Myo Armband. In *Proceedings of the 2016 XVIII Symposium on Virtual and Augmented Reality*, SVR, pages 64–70. IEEE.

Acon Digital (2014). Verberate Surround. Available from: `https://acondigital.com/products/verberate-surround/`. Accessed: 1 February 2016.

Activision (2017). Guitar Hero. Available from: `https://www.guitarhero.com/uk/en/game/controller`. Accessed: 24 November 2017.

Adesara, A. and Ashank, P. (2014). 3M - Myo Music Mixer. Available from: `http://aakashadesara.github.io/M3/`. Accessed: 14 December 2017.

Aimi, R. M. (2007). *Hybrid Percussion: Extending Physical Instruments Using Sampled Acoustics*. PhD thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.

Antao, L. P. S. (2017). Cooperative Human-Machine Interaction in Industrial Environments. Master's thesis, Universidade do Porto, Porto, Portugal.

Apfelstadt, H. (1988). What Makes Children Sing Well? *Update: Applications of Research in Music Education*, 7(1):27–32.

Arfib, D., Couturier, J.-M., and Kessous, L. (2005). Expressiveness and Digital Musical Instrument Design. *Journal of New Music Research*, 34(1):125–136.

Arganini, S. (2018a). Follow Me. Available from: `https://youtu.be/2c1lP4hveGQ`. Accessed: 5 March 2018.

Arganini, S. (2018b). Follow Me. Interviewed by B. Di Donato, 26 January.

Arief, Z., Sulistijono, I. A., and Ardiansyah, R. A. (2015). Comparison of five time series EMG features extractions using Myo Armband. In *Proceedings of the 2015 International Electronics Symposium*, IES, pages 11–14, Surabaya, Indonesia. IEEE.

Artiphon (2017). Artiphon. Available from: `https://artiphon.com/`. Accessed: 24 June 2017.

Atau Tanaka (2015). Myogram. Available from: `https://youtu.be/G6H1J2k--5I`. Accessed: 24 March 2016.

Baalman, M. A. J. (2010). Spatial composition techniques and sound spatialisation technologies. *Organised Sound*, 15(3):209–218.

Barbosa, J., Malloch, J., and Wanderley, M. M. (2015). What does "Evaluation" mean for the NIME community? In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '15, Baton Rouge, LA, USA.

Bates, E. (1997). *The Composition and Performance of Spatial Music*. PhD thesis, Trinity College Dublin, Dublin, Ireland.

Bates, E. (2009). *The Composition and Performance of Spatial Music*. PhD thesis, Trinity College Dublin, Dublin, Ireland.

Bau, O., Tanaka, A., and Mackay, W. E. (2008). The A20: Musical Metaphors for Interface Design. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '08, pages 91–96, Genoa, Italy.

Bayle, F. (1993). *Musiques acousmatiques, propositions... positions*. Institut national de l'audiovisuel, Paris, France, buchet-chastel edition.

Beatbox Battle (2015). Alem vs NaPoM - Final - 4th Beatbox Battle World Championship. Available from: `https://youtu.be/Nday2C4rBs0`. Accessed: 24 March 2016.

Beller, G. (2015). Babil-on V2. ISEA2015, AV Disruption, Vancouver BC, Canada, 17th of August 2015.

Beller, G. and Aperghis, G. (2011). Gestural Control of Real-Time Concatenative Synthesis in Luna Park. In *P3S, International Workshop on Performative Speech and Singing Synthesis*, pages 23–28, Vancouver, Canada.

Bellotti, V., Back, M., Edwards, W. K., Grinter, R. E., Henderson, A., and Lopes, C. (2002). Making Sense of Sensing Systems: Five Questions for Designers and Researchers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '02, pages 415–422, New York, New York, USA. ACM.

Benford, S. (2010). Performing musical interaction: Lessons from the study of extended theatrical performances. *Computer Music Journal*, 34(4):49–61.

Benford, S. and Giannachi, G. (2011). *Performing Mixed Reality*. The MIT Press, Cambridge, Massachusetts.

Benford, S., Schnädelbach, H., Koleva, B., Anastasi, R., Greenhalgh, C., Rodden, T., Green, J., Ghali, A., Pridmore, T., Gaver, B., Boucher, A., Walker, B., Pennington, S., Schmidt, A., Gellersen, H., and Steed, A. (2005). Expected, Sensed, and Desired: A Framework for Designing Sensing-based Interaction. *ACM Transition on Computer-Human Interaction*, 12(1):3–30.

Berberian, C. (1966). Stripsody. C.F Peters Corporation.

Berio, L. (1968). Sequenza III. Universal.

Berweck, S. (2012). *It worked yesterday: On (re-)performing electroacoustic music*. PhD thesis, University of Huddersfield, Huddersfield, United Kingdom.

Bevilacqua, F. (2012). New tools, new objects for gesture-based musical expression. Available from: `http://interlude.ircam.fr/?p=471`. Accessed: 12 August 2017.

Bevilacqua, F., Rasamimanana, N., Fléty, E., Lemouton, S., and Baschet, F. (2006). The augmented violin project: Research, composition and performance report. In *Proceedings of the 2006 Conference on New Interfaces for Musical Expression*, NIME '06, pages 402–406.

Birmingham City University (2014). Birmingham City University's Research Ethical Framework. Available from: `https://bcuassets.blob.core.windows.net/docs/BCU-%20Research_Ethical_Framework.23.11.10.pdf`. Accessed: 11 February 2020.

Bokowiec, M. A. (2011). V!OCT (Ritual): An Interactive Vocal Work for Bodycoder System and 8 Channel Spatialization. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '15, pages 40–43, Oslo, Norway.

Boyali, A. and Hashimoto, N. (2016). Spectral Collaborative Representation based Classification for hand gestures recognition on electromyography signals. *Biomedical Signal Processing and Control*, 24(C):11–18.

Boyatzis, C. J. and Watson, M. W. (1993). Preschool Children's Symbolic Representation of Objects through Gestures. *Child Development*, 64(3):729–735.

Boyer, E. O., Vandervoorde, L., Bevilacqua, F., and Hanneton, S. (2015). Touching sounds: audio virtual surfaces. In *2015 IEEE 2nd VR Workshop on Sonic Interactions for Virtual Environments*, SIVE, pages 43–47.

Braun, V. and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101.

Bühler, K. (1982). The deictic field of language and deictic words. In *Abridged translation of K. Bühler (1934): Sprachtheorie*, pages 9–30. John Wiley & Sons Ltd.

Bullock, J., Michailidis, T., and Poyade, M. (2016). Towards a Live Interface for Direct Manipulation of Spatial Audio. In *Proceedings of the International Conference on Live Interfaces*, ICLI, pages 134–141, Falmer, UK.

Bullock, J. and Momeni, A. (2015). ml.lib: Robust, Cross-platform, Open-source Machine Learning for Max and Pure Data. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '15, pages 265–270, Baton Rouge, Louisiana, USA.

Caramiaux, B. (2016). Myo-maxpd. Available from: `https://github.com/bcaramiaux/Myo-maxpd`. Accessed: 13 February 2015.

Caramiaux, B., Susini, P., Bianco, T., et al. (2011). Gestural Embodiment of Environmental Sounds: an Experimental Study. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '11, pages 144–148, Oslo, Norway.

Carter, T. (2002). *Monteverdi's Musical Theatre*. Yale University Press.

Chadefaux, D. (2013). *Musician/instrument interaction: the case of the concert harp*. PhD thesis, Université Pierre et Marie Curie, Paris, France.

Chadefaux, D., Le Carrou, J. L., and Fabre, B. (2013). A model of harp plucking. *The Journal of the Acoustical Society of America*, 133(4):2444–2455.

Chadefaux, D., Wanderley, M. M., Le Carrou, J. L., Fabre, B., and Daudet, L. (2012). Experimental study of the musician/instrument interaction in the case of the concert harp. In *Proceedings of the Acoustics 2012 Nantes Conference*, pages 1639–1644, Nantes, France.

Chaigne, A. and Kergomard, J. (2016). *Acoustics of Musical Instruments*. Modern Acoustics and Signal Processing. Springer New York, New York, New York, USA.

Chandler, D. W. and Grantham, D. W. (1992). Minimum audible movement angle in the horizontal plane as a function of stimulus frequency and bandwidth, source azimuth, and velocity. *The Journal of the Acoustical Society of America*, 91(3):1624–1636.

Chowning, J. M. (1971). The simulation of moving sound sources. *Journal of the Audio Engineering Society*, 19(1):2–6.

CICM, C. (2014). HoaLibrary - High Order Ambisonics Library. Available from: `http://hoalibrary.mshparisnord.fr/en`. Accessed: 12 January 2017.

Clozier, C. (2001). The Gmebaphone Concept and the CybernéPhone Instrument. *Computer Music Journal*, 25(4):81–90.

Connor, S. (2000). *Dumbstruck: A Cultural History of Ventriloquism*. Oxford University Press.

Cook, N. (2000). Demise of the Work Ethic: Jimi Hendrix's Improvisation as Performance Art. In *Royal Musical Association Conference Performance*, Southampton, United Kingdom. Page numbers not obtainable.

Cook, P. (2017). 2001: Principles for Designing Computer Music Controllers. In *A NIME Reader*, pages 1–13. Springer International Publishing, Cham.

Cook, P. R. (2001). Principles for Designing Computer Music Controllers. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '01, pages 3–6, Seattle, WA.

Cook, P. R. and Trueman, D. (1998). A Database of Measured Musical Instrument Body Radiation Impulse Responses, and Computer Applications for Exploring and Utilizing the Measured Filter Functions. In *Proceedings of the International Symposium on Musical Acoustics*, pages 303–308.

Cronin, D. (2014). *Usability of Micro- vs. Macro-Gestures in Camera-Based Gesture Interaction*. PhD thesis, California Polytechnic State University, San Luis Obispo, California, USA.

Cruz, I., Lopes, A. R., and Dias, J. (2015). CARDIMAX - Working prototype. Available from: `https://youtu.be/jCZlqKOcfnM`. Accessed: 8 May 2016.

Cycling '74 (2007). Max/MSP 5. Available from: `https://cycling74.com`. Accessed: 12 September 2014.

da Silva, H. P., Fred, A., and Martins, R. (2014). Biosignals for Everyone. *IEEE Pervasive Computing*, 13(4):64–71.

Davidson, J. W. (2001). The Role of the Body in the Production and Perception of Solo Vocal Performance: A Case Study of Annie Lennox. *Musicae Scientiae*, 5(2):235–256.

Davidson, J. W. (2007). Qualitative insights into the use of expressive body movement in solo piano performance: a case study approach. *Psychology of Music*, 35(3):381–401.

De Amicis, V. and Di Donato, B. (2015a). VoicErutseG. Frontiers Festival. Available from `https://youtu.be/xUo1TyymAQc` and `https://youtu.be/r7_NN8ppu3g`. Accessed 5 Jannuary 2021. Birmingham, United Kingdom, 17 March.

De Amicis, V. and Di Donato, B. (2015b). VoicErutseG. EmuFest. Available from `https://youtu.be/pQOY-YsKHPY` and `https://youtu.be/gdhnKbxUkUg`. Accessed 5 Jannuary 2021. Rome, Italy, 7 October.

De Amicis, V. and Di Donato, B. (2015c). VoicErutseG. ElectroAQustica. L'Aquila, Italy, 24 June.

Deldjoo, Y. (2009). *Wii Remote Based Head Tracking in 3D Audio Rendering*. PhD thesis, Chalmers University of Technology, Gothenburg, Sweden.

Dennett, J. I. (2015). Essential gesture. Master's thesis, The University of Melbourne, Melbourne, Australia.

Dertien, E. (2016). Use BITalino to Graph Your Biosignals and Play Pong! Available from: `https://makezine.com/projects/use-bitalino-graph-biosignals-play-pong/`. Accessed: 15 February 2020.

Desantos, S., Roads, C., and Bayle, F. (1997). Acousmatic Morphology: An Interview with Francois Bayle. *Computer Music Journal*, 21(3):11–19.

Desmet, F., Nijs, L., Demey, M., Lesaffre, M., Martens, J.-P., and Leman, M. (2012). Assessing a Clarinet Player's Performer Gestures in Relation to Locally Intended Musical Targets. *Journal of New Music Research*, 41(1):31–48.

Devaney, K. (2016). Star Cluster. Birmingham, United Kingdom, Available from: `https://youtu.be/mOig6ULoxNI`. Accessed: 20 September 2017.

Devaney, K., Turner, E., and Di Donato, B. (2017). Star Cluster. Birmingham, United Kingdom, Available from: `https://youtu.be/9ToP33Ki2SE`. Accessed: 20 September 2017.

Di Donato, B. (2016). Myo Mapper Wiki. Available from: `https://github.com/balandinodidonato/MyoMapper/wiki`. Accessed: 14 February 2017.

Dolscheid, S., Shayan, S., Majid, A., and Casasanto, D. (2013). The thickness of musical pitch: Psychophysical evidence for linguistic relativity. *Psychological Science*, 24(5):613–621.

Dooley, J. and Di Donato, B. (2016). Dancing with sound and light. Available from: `http://integra.io/projects/dancing-with-sound-and-light/`. Accessed: 14 December 2017.

Dourish, P. (2001). *Where The Action Is: The Foundations of Embodied Interaction*. MIT Press.

Ehmann, W. (1968). *Choral Directing*. Augsburg Fortress, Minneapolis, USA.

Eitan, Z. and Granot, R. Y. (2006). How Music Moves. *Music Perception: An Interdisciplinary Journal*, 23(3):221–248.

Ellis, J. (2014). Weav-Weav-Weaving I, II and III. Available from: `https://www.youtube.com/playlist?list=PL5GBxMmvpzqFdsANyZ3N1ONgB_QeYTyCj`. Accessed: 13 August 2017.

Engel, L. (2001). Body poetics of hip hop dance styles in Copenhagen. *Dance chronicle*, 24(3):351–372.

Enhancia (2018). Venice Beach. Available from: `https://youtu.be/MfnQb59qa0Y`. Accessed: 12 March 2018.

Erickson, R. (1975). *Sound Structure in Music*. University of California Press.

Expressive E (2017). Touché Factory Presets: Citykeys - "Burton". Available from: `https://youtu.be/X9TEzZUQlOQ`. Accessed: 24 November 2017.

Expressive E (2018). Touché. https://www.expressivee.com/buy-touche. Accessed: 24 November 2018.

Fiebrink, R. (2009). WekiInputHelper. Available from: `http://www.wekinator.org/input-helper/`. Accessed: 05 June 2015.

Fiebrink, R. and Cook, P. R. (2010). The Wekinator: a system for real-time, interactive machine learning in music. In *International Society for Music Information Retrieval Conference*, Utrecht, Netherlands. Available from: `https://www.researchgate.net/publication/228719072_The_Wekinator_A_System_for_Real-time_Interactive_Machine_Learning_in_Music`. Accessed: 26 March 2015.

Fiebrink, R., Cook, P. R., and Trueman, D. (2011). Human Model Evaluation in Interactive Supervised Learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 147–156, New York, New York, USA. ACM.

Fiebrink, R., Trueman, D., Britt, C., Nagai, M., Kaczmarek, K., Early, M., Daniel, M. R., Hege, A., and Cook, P. R. (2010). Toward Understanding Human-Computer Interaction In Composing The Instrument. In *Proceedings of the International Computer Music Conference*, ICMC, pages 135–142, New York, New York, USA.

Fishkin, K. P. (2004). A taxonomy for and analysis of tangible interfaces. *Personal and Ubiquitous Computing*, 8(5):347–358.

Flyotw (2014). REEPS ONE EXHIBITION // A.D.O. Available from: `https://youtu.be/z5VMmsqPRlI`. Accessed: 25 November 2017.

Françoise, J., Schnell, N., Borghesi, R., and Bevilacqua, F. (2014). Probabilistic Models for Designing Motion and Sound Relationships. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 287–292, London, United Kingdom.

Frassinetti, F., Bolognini, N., and Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research*, 147(3):332–343.

Garrett, G. E. and Goudeseune, C. (1999). Performance Factors in Control of High-Dimensional Space. In *Proceedings of the Interaction Computer Music Conference*, Beijing, China.

Gaver, W. W. (1991). Technology Affordances. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '91, pages 79–84, New York, NY, USA. ACM.

Geiger, C., Reckter, H., Paschke, D., Schulz, F., and Poepel, C. (2008). Towards Participatory Design and Evaluation of Theremin-based Musical Interfaces. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 303–306, Genoa, Italy.

Gentner, D. and Goldin-Meadow, S. (2003). *Language in Mind: Advances in the Study of Language and Thought*. MIT Press.

Giannakis, K. (2006). A comparative evaluation of auditory-visual mappings for sound visualisation. *Organised Sound*, 11(03):297–11.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Houghton Mifflin.

Gibson, J. J. (1977). The theory of affordances. *Hilldale, USA*, 1:2.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception.* Boston: Houghton Mifflin.

Gillian, N. and Paradiso, J. A. (2014). The Gesture Recognition Toolkit. *The Journal of Machine Learning Research*, 15(1):3483–3487.

Giordano, M. (1996). Modellizazione e simulazione del campo acustico in cavitá a geometria e caratteristiche acustiche variabili. Master's thesis, Università degli Studi Di Roma, La Sapienza, Rome, Italy.

Glaser, B. (1978). *Theoretical Sensitivity.* The Sociology Press.

Godøy, R. I. (2010). Gestural affordances of musical sound. In Godøy, R. I. and Leman, M., editors, *Musical Gestures: Sound, Movement, and Meaning*, pages 115–137. Routledge, New York, New York.

Godøy, R. I., Haga, E., and Jensenius, A. R. (2006). Exploring Music-Related Gestures by Sound-Tracing: A Preliminary Study. In *Proceedings of the 2nd International Symposium on Gesture Interfaces for Multimedia Systems*, GIMS2006, Leeds, United Kingdom.

GonzaloNV (2015). Myo-Ableton. Available from: `https://github.com/GonzaloNV/Myo-Ableton`. Accessed: 13 February 2015.

Graham, R. and Bridges, B. (2014). Gesture and Embodied Metaphor in Spatial Music Performance Systems Design. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '15, pages 581–584, London, United Kingdom.

Grandhi, S. A., Joue, G., and Mittelberg, I. (2011). Understanding naturalness and intuitiveness in gesture production. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI, pages 821–824, Vancouver, BC, Canada.

Grant, J. (1987). Improving Pitch and Intonation. *The Choral Journal*, 28(5):5–9.

Graugaard, L. (2008). Sound Shapes. Available from: `https://vimeo.com/16665607`. Accessed: 13 February 2015.

Gupta, N., Barreto, A., and Ordonez, C. (2002). Spectral modification of head-related transfer functions for improved virtual sound spatialization. In *Proceedings of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages II–1953–II–1956.

Hair, H. I. (1981). Verbal Identification of Music Concepts. *Journal of Research in Music Education*, 29(1):11–21.

Halfyard, J. K. (2000). Before night comes: narrative and gesture in Berio's *Sequenza III* (1966). *National Arts Education Archive Occasional Papers in the Arts and Education*, 8:79–93.

Hammond (2012). Leslie. Available from: `http://hammondorganco.com/products/leslie/`. Accessed: 25 October 2014.

Hammond, J. (2014). Air Scratching. Available from: `https://vimeo.com/107164256`. Accessed: 11 December 2014.

Hantrakul, L. and Kaczmarek, K. (2014). Implementations of the Leap Motion in sound synthesis, effects modulation and assistive performance tools. In *Proceedings of the International Computer Music Conference and Sound Sound and Computing Conference*, ICMC, SMC, pages 648–653, Athens, Grece.

Harrison, S., Sengers, P., and Tatar, D. (2007). The Three Paradigms of HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI 2007, pages 1–18, San Jose, USA. ACM.

Hewitt, D. (2011). Choreographic approaches to music composition for a new musical interface: The eMic. In *Proceedings of the International Symposium on Performance Science 2011*, pages 169–174, Toronto, Canada.

Hewitt, D. (2013). eMic: Developing Works for Vocal Performance using a Modified, Sensor Based Microphone Stand. In *Proceedings of CHI '13 Extended Abstracts on Human Factors in Computing Systems*, pages 2943–2946, Paris, France. ACM Press.

Hibbard, T. T. (1994). *The Use of Movement as an Instructional Technique in Choral Rehearsals*. University of Oregon, Oregon, USA.

Hinderks, A. (2015). UEQ - Online. Available from: `http://www.ueq-online.org/`. Accessed: 8 March 2015.

Hipke, K., Toomim, M., Fiebrink, R., and Fogarty, J. (2014). BeatBox: End-user Interactive Definition and Training of Recognizers for Percussive Vocalizations. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*, AVI '14, pages 121–124, Como, Italy.

Holland, S., Wilkie, K., Mulholland, P., and Seago, A. (2013). Music Interaction: Understanding Music and Human-Computer Interaction. In Holland, S., Wilkie, K., Mulholland, P., and Seago, A., editors, *Music and Human-Computer Interaction*, pages 1–28. Springer London, London.

Holman, T. (2008). *Surround Sound: Up and running*. Taylor & Francis.

Holzinger, A. (2008). *HCI and Usability for Education and Work*, volume 5298. Springer-Verlag Berlin Heidelberg, Graz, Austria.

Hsu, W. and Sosnick, M. (2009). Evaluating Interactive Music Systems: An HCI Approach. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 25–28, Pittsburgh, PA, United States.

Huang, J.-Y. A. (2011). Effective harp pedagogy: A Study of Techniques, Physical and Mental. Master's thesis, University of Canterbury, Canterbury, United Kingdom.

Hunt, A. and Wanderley, M. (2002). Mapping performer parameters to synthesis engines. *Organised Sound*, 7:97 – 108.

Hunt, A. D., Wanderley, M. M., and Paradis, M. (2002). The importance of Parameter Mapping in Electronic Instrument Design. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '02, pages 88–93, Dublin, Ireland.

Idiosynphonic (2012). Reeps One Interview for Idiosynphonic. Available from: `https://youtu.be/FrmnvkL6qVE`. Accessed: 25 September 2016.

InAVate (2015). K-Array's Owl is moving head loudspeaker. Available from: `https://youtu.be/8cj5EjdXTnU`. Accessed: 19 April 2015.

Integra Lab (2013). Integra Live. Available from: `http://integra.io/integralive/`. Accessed: 27 November 2014.

Integra Lab (2016a). PitchShifter.module. Available from: `https://github.com/BirminghamConservatoire/IntegraLive/blob/develop/modules/PitchShifter.module`. Accessed: 4 June 2016.

Integra Lab (2016b). StereoReverb.module. Available from: `https://github.com/BirminghamConservatoire/IntegraLive/blob/develop/modules/StereoReverb.modulee`. Accessed: 4 June 2016.

IRCAM (2012). Mubu. Available from `http://forumnet.ircam.fr/product/mubu-en/`. Accessed: 13 February 2019.

Ishijima, R., Ogawa, K., Higuchi, M., and Komuro, T. (2014). Real-time Typing Action Detection in a 3D Pointing Gesture Interface. In *Proceedings of the 5th Augmented Human International Conference*, AH '14, pages 20:1–20:2, Kobe, Japan.

Jensenius, A. R., Wanderley, M. M., Godøy, R. I., and Leman, M. (2009). Musical Gestures Concepts and Methods in Research. In Godøy, R. I. and Leman, M., editors, *Musical Gestures: Sound, Movement, and Meaning*, pages 12–35. Taylor & Francis, New York, New York, USA.

Jessop, E. (2009). The Vocal Augmentation and Manipulation Prosthesis (VAMP): A Conducting-Based Gestural Controller for Vocal Performance. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '09, pages 256–259, Pittsburgh, PA.

Jing, P. and Ye-peng, G. (2013). Human-Computer Interaction using Pointing Gesture based on an Adaptive Virtual Touch Screen. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 6(4):81–91.

Johnson, B., Norris, M., and Kapur, A. (2014). tactile.motion: An iPad Based Performance Interface For Increased Expressivity In Diffusion Performance. In *Proceedings of the International Computer Music Conference and Sound and Music Computing Conference*, ICMC-SMC 2016, pages 798–802, Athens, Greece.

Johnson, B., Norris, M., and Kapur, A. (2016). speaker.motion: A Mechatronic Loudspeaker System for Live Spatialisation. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '16, pages 41–45, Brisbane, Australia.

Johnson, D. and Tzanetakis, G. (2017). VRMin: Using Mixed Reality to Augment the Theremin for Musical Tutoring. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 151–156, Copenhagen, Denmark. Aalborg University Copenhagen.

Jordà, S., Geiger, G., Alonso, M., and Kaltenbrunner, M. (2007). The reacTable: Exploring the Synergy Between Live Music Performance and Tabletop Tangible Interfaces. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction*, TEI '07, pages 139–146, Baton Rouge, Louisiana.

Jot, J.-M. (1999). Real-time spatial processing of sounds for music, multimedia and interactive human-computer interfaces. *Multimedia Systems*, 7(1):55–69.

Jules, F. (2016). myo-for-max. Available from: `https://github.com/JulesFrancoise/myo-for-max`. Accessed: 1 February 2017.

K-array (2015a). KW8. Available from: `https://www.k-array.com/en/concert/kw8.html`. Accessed: 27 May 2015.

K-array (2015b). Owl Manager. Available from: `https://www.k-array.com/en/software/owl-manager.html`. Accessed: 27 May 2015.

Kaltenbrunner, M. (2013). reacTIVision. Available from: `https://github.com/mkalten/reacTIVision`. Accessed: 25 October 2014.

Kaltenbrunner, M. and Reas, C. (2003-2020). Processing. Processing Foundation. Available from: `https://processing.org/`. Accessed: 25 October 2014.

Kamkar, S. (2015). MyoOSC. Available from: `https://github.com/samyk/myo-osc`. Accessed: 13 February 2015.

Kapur, A., Lazier, A. J., Davidson, P. L., Wilson, S., and Cook, P. R. (2004). The Electronic Sitar Controller. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '04, pages 7–12, Hamamatsu, Japan.

Karjalainen, M., Valimaki, V., and Janosy, Z. (1993). Towards High-Quality Sound Synthesis of the Guitar and String Instruments. In *Proceedings of the International Computer Music Conference*, ICMC, pages 1–8, Tokyo, Japan.

Karplus, K. and Strong, A. (1983). Digital Synthesis of Plucked-String and Drum Timbres. *Computer Music Journal*, 7(2):43–55.

Kendon, A. (1994). Human Gestures. In *Tools, Language and Cognition in Human Evolution*, pages 43–62. Cambridge University Press.

Kendon, A. (2004). *Gesture: Visible Action as Utterance.* Cambridge University Press, Cambridge, United Kingdom.

Kew, C. (2015). Toward a Beatboxology. Available from: `https://www.humanbeatbox.com/articles/toward-a-beatboxology/`. Accessed: 27 November 2017.

Kiefer, C., Collins, N., and Fitzpatrick, G. (2008). HCI Methodology For Evaluating Musical Controllers: A Case Study. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 87–90, Genoa, Italy.

Knapp, R. B. and Lusted, H. S. (1990). A Bioelectric Controller for Computer Music Applications. *Computer Music Journal*, 14(1):42.

Kondonassis, Y. (2006). *On Playing the Harp.* Carl Fischer LLC.

Kronlachner, M. (2012). übersetzen - vertimas.

Kronlachner, M. (2013). The Kinect distance sensor as human-machine-interface in audio-visual art projects. Master's thesis, University of Music and Performing Arts Graz, Graz, Austria.

Kuhns, C. (2014). *Beatboxing and the Flute: Its History, Repertoire, and Pedagogical Importance.* PhD thesis, Florida State University Libraries, Florida, USA.

Kulesza, T., Amershi, S., Caruana, R., Fisher, D., and Charles, D. (2014). Structured labelling for facilitating concept evolution in machine learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3075–3084.

Kuperberg, B. (2016). MyoOSC. Available from: `https://github.com/benkuper/MyOSC`. Accessed: 13 February 2015.

Kurosawa, K. and Davidson, J. W. (2005). Nonverbal behaviours in popular music performance: A case study of The Corrs. *Musicae Scientiae*, 9(1):111–136.

Lagacè, F. c. K. and Dagher, A. R. (2017). X/Centris Delirium Machine: Les bruits de l'esprit. Available from: `https://youtu.be/R90_UCJ0VcQ`. Accessed: 14 December 2017.

Lazar, J., Hochheiser, H., and Feng, J. H. (2010). *Research Methods in Human-Computer Interaction.* Wiley Publishing.

Le Carrou, J. L., Wahlen, E., Brasseur, E., and Gilbert, J. (2008). Two dimensional finger-string interaction in the concert harp. *The Journal of the Acoustical Society of America*, 123(5):3122–3122.

Leap Motion (2014). Leap Motion. Available from: https://www.leapmotion.com/. Accessed: 19 December 2014.

Lech, M. and Kostek, B. (2013). Testing A Novel Gesture-Based Mixing Interface. *Journal of Audio Engineering Society*, 61(5):301–313.

Lee, M. and Wessel, D. (1992). Connectionist Models for Real-Time Control of Synthesis and COmpositional Algorithms. In *Proceedings of the International Computer Music Conference*, ICMC, pages 277–280, California, USA.

Leman, M. (2007). *Embodied Music Cognition and Mediation Technology*. The MIT Press.

Leman, M. (2009). Music, gesture, and the formation of embodied meaning. In Godøy, R. I. and Leman, M., editors, *Musical gestures: sound, movement, and meaning*, pages 126–153. Routledge.

Leman, M. and Maes, P.-J. (2015). The role of embodiment in the perception of music. *Empirical Musicology Review*, 9(3-4):236–246.

Liao, M.-Y. (2002). *The Effects of Gesture and Movement Training on the Intonation and Tone Quality of Children's Choral Singing*. PhD thesis, The University of Sheffield, Sheffield, United Kingdom.

Liao, M.-Y. and Davidson, J. W. (2007). The use of gesture techniques in children's singing. *International Journal of Music Education*, 25(1):82–94.

Lossius, T., Baltazar, P., and de la Hogue, T. (2009). DBAP - Distance-Based Amplitude Panning. In *Proceedings of the International Computer Music Conference*, ICMC, Montreal, Canada. Available from: http://hdl.handle.net/2027/spo.bbp2372.2009.111. Accessed: 29 September 2014.

Lunches, G. (2013). Adam Matta: Vocal Performance Artist. Available from: http://www.fahrenheit-212.com/adam-matta-vocal-performance-artist/. Accessed: 1 March 2014.

Lyon, E. (2014). The Future of Spatial Computer Music. In *Proceedings of the International Computer Music Conference and Sound and Music Computing Conference*, ICMC-SMC 2016, pages 850–854, Athens, Greece.

Lyon, E., Knapp, R. B., and Ouzounian, G. (2014). Compositional and Performance Mapping in Computer Chamber Music: A Case Study. *Computer Music Journal*, 38(3):64–75.

Mackay, W. E. (1990). *Users and customizable software: A co-adaptive phenomenon*. PhD thesis, Citeseer.

Mackay, W. E. (1991). Beyond Iterative Design: User Innovation in Co-adaptive Systems. Technical Report EPC-1991-130, Rank Xerox EuroPARC, Cambridge, United Kingdom.

Mackay, W. E., Fayard, A.-L., Frobert, L., and Médini, L. (1998). Reinventing the Familiar: Exploring an Augmented Reality Design Space for Air Traffic Control. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '98, pages 558–565, Los Angeles, California, USA. ACM Press/Addison-Wesley Publishing Co.

Magnusson, T. (2010). Designing constraints: Composing and performing with digital musical systems. *Computer Music Journal*, 34(4):62–73.

Mainsbridge, M. M. (2016). *Body as instrument: an exploration of gestural interface design*. PhD thesis, University of Technology Sydney, Sydney, Australia.

Marentakis, G., Malloch, J., Peters, N., Marshall, M., Wanderlay, M., and McAdams, S. (2008). Influence of performance gestures on the identification of spatial sound trajectories in a concert hall. In *Proceedings of the International Conference on Auditory Display*, ICAD, pages 1–8, Paris, France.

Marrin Nakra, T. T. A. (2000). *Inside the conductor's jacket: analysis, interpretation and musical synthesis of expressive gesture*. PhD thesis, Massachusetts Institute of Technology.

Marshall, M. T., Malloch, J., and Wanderley, M. M. (2009). Gesture control of sound spatialization for live musical performance. In Sales Dias, M., Gibet, S., Wanderley, M. M., and Bastos, R., editors, *Gesture-Based Human-Computer Interaction and Simulation: 7th International Gesture Workshop, GW 2007, Lisbon, Portugal, May 23-25, 2007, Revised Selected Papers*, pages 227–238. Springer Berlin Heidelberg, Berlin, Heidelberg.

Marshall, M. T., Wanderley, M. M., and Malloch, J. (2007). Non-conscious Gesture Control of Sound Spatialization. In *Proceedings of the 4th International Conference on Enactive Interfaces*, ENACTIVE/07, Grenoble, France. Available from: `https://www.researchgate.net/publication/228754329_Non-Conscious_Control_of_Sound_Spatialization`. Accessed: 3 October 2014.

McCormick, E. (2015). Design Media Arts student integrates beatboxing into visual artwork. Available from: `http://dailybruin.com/2015/06/08/design-media-arts-student-integrates-beatboxing-into-visual-artwork/`. Accessed: 16 November 2015.

McNeill, D. (2005a). Gesture: a psycholinguistic approach. In *The Encyclopedia of Language and Linguistics, Psycholinguistics Section*. Elsevier Science. Available from: `http://mcneilllab.uchicago.edu/pdfs/gesture.a_psycholinguistic_approach.cambridge.encyclop.pdf`. Accessed: 5 April 2015.

McNeill, D. (2005b). *Gesture and Thought*. University of Chicago Press.

Meehan, K. (2011). *Not Just a Prey Voice: Cathy Berberian as Collaborator, Composer and Creator*. PhD thesis, Washington University, St. Louis, Missouri, USA.

Merchel, S. and Groth, S. (2010). Adaptively Adjusting the Stereophonic Sweet Spot to the Listener's Position. *Journal of the Audio Engineering Society*, 58(10):809–817.

Merleau-Ponty, M. (1945). *Phénomènologie de la perception*. Gallimard, Paris.

Michailidis, T. (2017). Non-verbal Interaction through vibrotactile feedback. Available from: `https://youtu.be/n1x0fVHA2iw`. Accessed: 14 December 2017.

Michailidis, T. and Di Donato, B. (2016). Vibrotactile feedback for interaction in dance performances. In *Abstracts of InDialogue International Symposium*, Nottingham, United Kingdom. Available from: `https://www.researchgate.net/publication/309812834_Vibrotactile_feedback_for_interaction_in_dance_performances`. Accessed: 15 December 2016.

Milgram, P., Takemura, H., Utsumi, A., and Kishino, F. (1995). Augmented reality: a class of displays on the reality-virtuality continuum. In Das, H., editor, *Proceedings SPIE 2351, Telemanipulator and Telepresence Technologies*, volume 2351, pages 282–292. SPIE.

Mi.Mu (2016). mi.mu. Available from: `http://mimugloves.com/`. Accessed: 15 April 2017.

Mittelberg, I. and Waugh, L. R. (2009). Metonymy first, metaphor second: A cognitivesemiotic approach to multimodal figures of thought in co-speech gesture. In Forceville, C. J. and Urios-Aparisi, E., editors, *Multimodal Metaphor*, pages 329–358. Mouton de Gruyter, Berlin, New York.

Monk, A. (2007). Lightweight techniques to encourage innovative user interface design. In Wood, L. E., editor, *User Interface Design: Bridging the Gap from User Requirements to Design*, pages 109–129. CRC Press.

Moog (2014). Theremini. Available from: `https://www.moogmusic.com/products/etherwave-theremins/theremini`. Accessed: 12 December 2014.

Mooney, J. R. (2005). *Sound Diffusion Systems for the Live Performance of Electroacoustic Music*. PhD thesis, The University of Sheffield, Sheffield, United Kingdom.

Mulder, A. G. E., Fels, S. S., and Mase, K. (1997). Empty-handed Gesture Analysis in Max/FTS. In *In Proceedings of Kansei - The Technology of Emotion, AIMI International Workshop*, pages 87–91, Genoa, Italy.

Müller, M. (2007). *Information Retrieval for Music and Motion*. Springer, Berlin, Heidelberg.

Nagashima, Y. (2003). Bio-Sensing Systems and Bio-Feedback Systems for Interactive Media Arts. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 48–53, Montreal, Canada.

Nicolls, S. (2010). Seeking Out the Spaces Between: Using Improvisation in Collaborative Composition with Interactive Technology. *Leonardo Music Journal*, 20:47–55.

Nicolls, S. (2011). Brief Thoughts on using Live Electronics: The Inside-out Piano and touring. *eContact*, 13(2). Available from: `http://econtact.ca/13_2/nicolls_electronics.html`. Accessed: 24 November 2017.

Nielsen, J. (1993). Iterative User-Interface Design. *Computer*, 26(11):32–41.

Norman, D. (2002). Emotion & Design: Attractive Things Work Better. *interactions*, 9(4):36–42.

Norman, D. A. and Draper, S. W. (1986). *User Centered System Design: New Perspectives on Human-computer Interaction.* Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Norman, D. A. et al. (1988). *The psychology of everyday things*, volume 5. Basic books New York.

Normandeau, R. (2009). Timbre Spatialisation: The Medium is the Space. *Organised Sound*, 14(3):277–285.

Nymoen, K., Haugen, M. R., and Jensenius, A. R. (2015). MuMYO - Evaluating and Exploring the MYO Armband for Musical Interaction. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '15, pages 215–218, Baton Rouge, Louisiana, USA.

O'Modhrain, S. (2011). A Framework for the Evaluation of Digital Musical Instruments. *Computer Music Journal*, 35(1):28–42.

Orio, N., Schnell, N., and Wanderley, M. M. (2001). Input Devices for Musical Expression: Borrowing Tools from HCI. In *Proceedings of the CHI'01 Workshop on New Interfaces for Musical Expression*, NIME '01, pages 15–18, Seattle, WA.

Overholt, D. (2005). The Overtone Violin: a new computer music instrument. In *Proceedings of the International Computer Music Conference*, ICMC, Barcelona, Spain.

Oviatt, S. (2003). Multimodal Interfaces. In Jacko, J. A. and Sears, A., editors, *The Human-Computer Interaction Handbook*, pages 286–304. L. Erlbaum Associates Inc, Hillsdale, NJ, USA.

Owens, P. (2005). *Cognitive load theory and music instruction.* PhD thesis, University of NSW, Sydney, Australia.

Pajares, M., Ayala, P., Fajardo, I., Vicente, D., and Grana, M. (2004). Usability analysis of a pointing gesture interface. In *Proceedings of the International Conference on Systems, Man and Cybernetics*, volume 3, pages 2652–2657.

Park, Y., Heo, H., and Lee, K. (2012). VOICON: An Interactive Gestural Microphone For Vocal Performance. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '12, Ann Arbor, Michigan. Available from: `http://www.nime.org/proceedings/2012/nime2012_199.pdf`. Accessed: 10 October 2014.

Pauletto, S. and Hunt, A. (2006). The sonification of EMG data. In *Proceedings of the International Conference on Auditory Display*, ICAD, pages 152–157, London, United Kingdom.

Peterson, L. (2009). K-nearest neighbor. *Scholarpedia*, 4(2):1883.

Picart, B., Brognaux, S., and Dupont, S. (2015). Analysis and automatic recognition of Human BeatBox sounds: A comparative study. In *ICASSP 2015 - 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4255–4259. IEEE.

Pinto, A. G., Dias, G., Felizardo, V., Pombo, N., Silva, H., Fazendeiro, P., Crisostomo, R., and Garcia, N. (2016). Electrocardiography, electromyography, and accelerometry signals collected with BITalino while swimming: Device assembly and preliminary results. In *Proceedings of the 12th International Conference on Intelligent Computer Communication and Processing*, ICCP, pages 37–41. IEEE.

Plux (2014a). Bitalino. Available from: `http://bitalino.com/en/`. Accessed: 10 January 2015.

Plux (2014b). BITalino's APIs. Available from: `http://bitalino.com/en/development/apis`. Accessed: 10 January 2015.

Poliniak, S. (2013). Incorporating beatboxing into a cappella performance. *Teaching Music*, 21(2):46.

Porres, A. T. and Manzolli, J. (2007). Adaptive Tuning Using Theremin as Gestural Controller. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 363–366, New York City, NY, United States.

Precision Music Technology (2015). Leviathan. Available from: `http://precisionmusic.technology/`. Accessed: 13 February 2015.

Proctor, M., Bresch, E., Byrd, D., Nayak, K., and Narayanan, S. (2013). Paralinguistic mechanisms of production in human 'beatboxing': A real-time magnetic resonance imaging study. *The Journal of the Acoustical Society of America*, 133(2):1043–1054.

Protopixel (2014). The Glitch Chamber. Available from: `https://vimeo.com/98075534`. Accessed: 24 January 2015.

Ramakrishnan, C., Goßmann, J., and Brummer, L. (2006). The ZKM Klangdom. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '06, pages 140–143, Paris, France.

Rasamimanana, N., Bevilacqua, F., Schnell, N., Guedy, F., Flety, E., Maestracci, C., Zamborlin, B., Frechin, J.-L., and Petrevski, U. (2010). Modular Musical Objects towards Embodied Control of Digital Music. In *Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '11, page 9–12, New York, NY, USA. Association for Computing Machinery.

Ratcliffe, J. (2014). Hand Motion-Controlled Audio Mixing Interface. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '14, pages 136–139, London, United Kingdom.

Rauschenberger, M., Schrepp, M., Perez-Cota, M., Olschner, S., and Thomaschewski, J. (2013). Efficient Measurement of the User Experience of Interactive Products. How to use the User Experience Questionnaire (UEQ). Example: Spanish Language Version. *International Journal of Interactive Multimedia and Artificial Intelligence*, 2(1):39–45.

Reactable (2016). Rotor. Available from: `http://reactable.com/rotor/`. Accessed: 19 June 2016.

Reeves, S., Benford, S., O'Malley, C., and Fraser, M. (2005). Designing the Spectator Experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, pages 741–750, Portland, Oregon, USA.

Refaeilzadeh, P., Tang, L., and Liu, H. (2009). Cross-Validation. In Liu, L. and Özsu, M. T., editors, *Encyclopedia of Database Systems*, pages 532–538. Springer US, Boston, MA.

Roads, C. (2012). From grains to forms. In *Proceedings of the International Symposium Xenakis. La musique électroacoustique Xenakis*, Paris, France.

Rocchesso, D., Mauro, D. A., and Monache, S. D. (2016). miMic: The Microphone As a Pencil. In *Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '16, pages 357–364, Eindhoven, Netherlands.

Roland, O. L. (2014). *The Seaboard: discreteness and continuity in musical interface design*. PhD thesis, Royal College of Art, London, United Kingdom.

Roli (2016). JUCE. Available from: `https://juce.com/`. Accessed: 8 June 2016.

Roomi, S. M. M., Priya, R. J., and Jayalakshmi, H. (2010). Hand Gesture Recognition for Human-Computer Interaction. *Journal of Computer Science*, 9(6):1002–1007.

Rosenbaum, D. A., Slotta, J. D., Vaughan, J., and Plamondon, R. (1991). Optimal movement selection. *Psychological Science*, 2(2):86–91.

Roussou, M. (2004). Learning by Doing and Learning Through Play: An Exploration of Interactivity in Virtual Environments for Children. *ACM Computers in Entertainment*, 2(1):10–10.

Roy, A. and Bouchet, M. (2017). FlyByNo Live - The Families (Lumeris Theme). Available from: `https://youtu.be/gJOFqnLiyoY`. Accessed: 8 June 2017.

Rusconi, E., Kwan, B., Giordano, B. L., Umiltà, C., and Butterworth, B. (2006). Spatial representation of pitch height: the SMARC effect. *Cognition*, 99(2):113–129.

Saberi, K., Dostal, L., Sadralodabai, T., and Perrott, D. R. (1991). Minimum audible angles for horizontal, vertical, and oblique orientations: Lateral and dorsal planes. *Acta Acustica united with Acustica*, 75(1):57–61.

Sabri, S. (2017). Nu body. ACE Maker Monday Commissioning Project. Accessed: 14 December 2017.

Sad, A. T. N. (2016). ReactiFI model changing. Available from: `https://youtu.be/U9YLrc5mVeM`. Accessed: 29 July 2016.

Salvati, D., Canazza, S., and Roda, A. (2011). Sound spatialisation control by means of acoustic source localization system. In *Sound and Music Computing Conference*, SMC 2011, pages 284–289, Padova, Italy.

Salzedo, C. (1952). Modern Harp Technique: Gestures Have a Vital Part in Playing the Harp. *Etude, The Music Magazine*, 70(1):9.

Sapthavee, A., Yi, P., and Sims, H. S. (2014). Functional Endoscopic Analysis of Beatbox Performers. *Journal of Voice*, 28(3):328–331.

Savage, G. and Di Donato, B. (2016). Music Gesture Beatbox. Music Tech Fest. Berlin, Germany, 29 May.

Schaeffer, P. (1966). *Traité des objets musicaux*. Seuil.

Scheider, L., Grassmann, S., Kaminski, J., and Tomasello, M. (2011). Domestic dogs use contextual information and tone of voice when following a human pointing gesture. *PLOS ONE*, 6(7):e21676–6.

Schlessinger, D. M. (2012). Concept Tahoe: Microphone Midi Control. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '12, Ann Arbor, USA. Available from: `http://www.nime.org/proceedings/2012/nime2012_202.pdf`. Accessed 18 April 2016.

Schnell, N., Röbel, A., Schwarz, D., Peeters, G., Borghesi, R., et al. (2009). MuBu and friends–assembling tools for content based real-time interactive audio processing in Max/MSP. In *Proceedings of the International Computer Music Conference*, ICMC, Montreal, Canada.

School, T. C. (2017). CIS Ontario Student Visual Art Festival 2017 — Joaquin Manay. Available from: `https://www.cisstudentvisualartsfest.com/joaquin-manay`. Accessed: 24 November 2017.

Schrepp, M., Hinderks, A., and Thomaschewski, J. (2014). Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios. In Marcus, A., editor, *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience: Third International Conference, DUXU 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part I*, pages 383–392. Springer, Cham, Germany.

Sengers, P. and Gaver, B. (2006). Staying Open to Interpretation: Engaging Multiple Meanings in Design and Evaluation. In *Proceedings of the 6th ACM Conference on Designing Interactive Systems - DIS '06*, page 99, University Park, PA, USA. ACM Press.

Sengpiel, A. (2014). Calculation of the Speed of Sound in Air and the effective Temperature. Available from: `http://www.sengpielaudio.com/calculator-speedsound.htm`. Accessed: 8 February 2015.

Serafin, S., Trento, S., Grani, F., Perner-Wilson, H., Madgwick, S., and Mitchell, T. (2014). Controlling Physically Based Virtual Musical Instruments Using The Gloves. In *Proceedings of the International Computer Music Conference*, NIME '14, pages 521–524, London, UK.

Shan, G. and Visentin, P. (2012). EMG Applications in Studies of Arts. In *Applications of EMG in Clinical and Sports Medicine*, pages 201–208. InTech.

Smola, A. J., Williamson, R. C., and Bartlett, P. L. (2000). New Support Vector Algorithms. *Neural Computation*, 12:1207–1245.

Steeves, C. (1984). *The effect of Curwen-Kodaly hand signs on pitch and interval discrimination within a Kodaly curricular framework*. PhD thesis, Music, University of Calgary, Alberta, Canada.

Stoffregen, P. (2015). DmxSimple. Available from: `https://github.com/PaulStoffregen/DmxSimple`. Accessed: 24 November 2015.

Stowell, D. (2010). *Making music through real-time voice timbre analysis: machine learning and timbral control*. PhD thesis, Queen Mary University of London, London, United Kingdom.

Stowell, D. and Plumbley, M. D. (2010). Delayed Decision-making in Real-time Beatbox Percussion Classification. *Journal of New Music Research*, 39(3):203–213.

Streeck, J., Goodwin, C., and LeBaron, C. (2011). Embodied Interaction in the Material World: An Introduction. In Streeck, J., Goodwin, C., and LeBaron, C., editors, *Embodied Interaction, Language and Body in the Material World*, pages 1–10. Cambridge University Press, Cambridge, United Kingdom.

Svanæs, D. (2000). *Understanding interactivity: steps to a phenomenology of human-computer interaction*. Norges teknisk-naturvitenskapelige universitet.

Szanto, G. (2014). Guest Post: The Myo™ Armband + DJ Player. Available from: `http://developerblog.myo.com/guest-post-myo-armband-dj-player/`. Accessed: 24 November 2015.

TaikaBox (2016). TaikaBox DigiDance Workshop. Available from: `https://youtu.be/EuQZSNm6Ut4`. Accessed: 8 December 2016.

TaikaBox (2017). MachineMovement. Available from: `https://youtu.be/Un2rP4ZyYNM`. Accessed: 14 December 2017.

TaikaBox (2018). TaikaBox and MyoMapper. Interviewed by B. Di Donato, 27 January 2018.

Tanaka, A. (1993). Musical technical issues in using interactive instrument technology with application to the BioMuse. In *Proceedings of International Computer Music Conference*, ICMC, Tokyo, Japan.

Tanaka, A. (2012a). BioMuse to Bondage: Corporeal Interaction in Performance and Exhibition. In *Intimacy Across Visceral and Digital Performance*, pages 159–169. Palgrave Macmillan UK, London, United Kingdom.

Tanaka, A. (2012b). Sensor-Based Musical Instruments and Interactive Music. In Dean, R. T., editor, *The Oxford Handbook of Computer Music*. Oxford University Press.

Tanaka, A. (2015). Myogram. Available from: `http://research.gold.ac.uk/17508/`. Accessed: 29 May 2015.

Tanaka, A., Altavilla, A., and Spowage, N. (2012a). Gestural Musical Affordances. In *Proceedings of the 9th International Conference on Sound and Music Computing*, Copenhagen, Denmark.

Tanaka, A. and Knapp, R. B. (2002). Multimodal Interaction in Music Using the Electromyogram and Relative Position Sensing. In *Proceedings of the 2002 Conference on New Interfaces for Musical Expression*, NIME '02, pages 1–6, Dublin, Ireland.

Tanaka, A. and Nicolls, S. (2017). Suspension. In Meta Gesture Music: Embodied Interaction, New Instruments, and Sonic Experience. Goldsmiths Press.

Tanaka, A., Parkinson, A., Settel, Z., and Tahiroglu, K. (2012b). A Survey and Thematic Analysis Approach as Input to the Design of Mobile Music GUIs. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '12, Ann Arbor, Michigan. University of Michigan.

TechCrunch (2014). Roli Seaboard Grand Piano Demo. Available from: `https://youtu.be/7mKJ8CJ2OuA`. Accessed: 24 November 2017.

Thalmic Labs (2015a). Myo SDK. Available from: `https://developer.thalmic.com/`. Accessed: 8 January 2016.

Thalmic Labs (2015b). Myo SDK Getting Started. Available from: https://developer.thalmic.com/docs/api_reference/platform/getting-started.html. Accessed: 8 January 2016.

The Beatbox Collective (2012). Pass the Sound. Available from: https://youtu.be/a4wBj8eC3Mg. Accessed: 25 November 2017.

The MathWorks (2014). Curve Fitting Toolbox. Available from: https://www.mathworks.com/products/curvefitting.html. Accessed: 8 February 2015.

Theremin, L. S. and Petrishev, O. (1996). The Design of a Musical Instrument Based on Cathode Relays. *Leonardo Music Journal*, 6(1):49–50.

Thum, T. (2013). The orchestra in my mouth. Available from: https://www.ted.com/talks/tom_thum_the_orchestra_in_my_mouth. Accessed: 8 May 2015.

Tiburon, E. (2015). Rory McIlroy PGA Tour. Available from: https://www.easports.com/pga-tour/rory. Accessed: 24 November 2017.

Trimble, O. C. (1934). Localization of Sound in the Anterior-posterior Vertical Dimensions of 'Auditory' Space. *British Journal of Psychology. General Section*, 24(3):320–334.

TroikaTronix (2016). Isadora. Available from: https://troikatronix.com/. Accessed: 14 December 2016.

Tsandilas, T., Letondal, C., and Mackay, W. E. (2009). Musink: Composing Music Through Augmented Drawing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 819–828, Boston, MA, USA. ACM.

Tubb, R. and Dixon, S. (2015). An Evaluation of Multidimensional Controllers for Sound Design Tasks. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 47–56, Seoul, Republic of Korea.

Turchet, L. (2016a). The Hyper-Hurdy-Gurdy. In *Proceedings of the Sound and Music Computing Conference*, SMC 2016, pages 491–497, Hamburg, Germany.

Turchet, L. (2016b). The Hyper-Zampogna. In *Proceedings of the Sound and Music Computing Conference*, SMC 2016, pages 485–490, Hamburg, Germany.

Turner, E. and Di Donato, B. (2016). The Wood and the Water. Available from: https://youtu.be/gYu4Za-1E48. Accessed: 16 January 2017.

Turner, E. and Di Donato, B. (2017a). The Wood and the Water. Music Program of the Audio Mostly 2017 Conference. Available from: https://audiomostly.com/2017/program/music-program/ and https://youtu.be/3n_2y3OerVQ. Accessed 5 January 2021. London, United Kingdom, 25 August.

Turner, E. and Di Donato, B. (2017b). The Wood and the Water. Electronic Music Week 2017. Available from: https://youtu.be/p0_VR9N4x8w. Accessed 2 Jannuary 2021. Shanghai, China, 20 October.

Turner, G. and Kenny, D. T. (2012). Restraint of body movement potentially reduces peak SPL in western contemporary popular singing. *Musicae Scientiae*, 16(3):357–371.

UKF (2014). UKF Meets - Reeps One. Available from: https://youtu.be/z6pc4kM2eqA. Accessed: 18 March 2014.

Valbom, L. and Marcos, A. (2005). WAVE: Sound and music in an immersive environment. *Computers & Graphics*, 29(6):871–881.

van Buuren, A. (2014). Behind the scenes – Armin van Buuren controls lights and stage effects live using MYO Armband. Available from: https://youtu.be/Wrc1c8g2FPk. Accessed: 1 November 2014.

Van Raalte, C. (1998). Old dog, new tricks: Biofeedback as assistive technology. In *California State University, Northridge Center On Disabilities 1998 Conference*.

Varèse, E. (1932-1934). Ecuatorial.

Varga, B. A. (2013). *From Boulanger to Stockhausen: interviews and a memoir*. University of Rochester Press, Rochester, New York.

Verillon, P. and Rabardel, P. (1995). Cognition and artifacts: A contribution to the study of though in relation to instrumented activity. *European Journal of Psychology of Education*, 10(1):77.

Vigliensoni, G. and Wanderley, M. M. (2012). A Quantitative Comparison of Position Trackers for the Development of a Touch-less Musical Interface. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '12, Ann Arbor, USA. Available from: http://www.nime.org/proceedings/2012/nime2012_155.pdf. Accessed: 8 February 2015.

Villegas Curulla, G. (2016). Fantasie pour violoncelle. Available from: https://youtu.be/EioMZD9LbF0. Accessed: 14 December 2017.

Villegas Curulla, G. (2018). Fantasie pour violoncelle. Interviewed by B. Di Donato, 25 January.

Visi, F. and Coorevits, E. (2015). Kineslimina. Gala Concert of the 11th International Symposium on Computer Music Multidisciplinary Research (CMMR), Available from: https://youtu.be/yZna5W4KZnA. Accessed: 24 March 2016.

Visi, F., Coorevits, E., Schramm, R., and Miranda, E. R. (2017). Musical Instruments, Body Movement, Space, and Motion Data: Music as an Emergent Multimodal Choreography. *Human Technology*, 13(1):58–81.

Visi, F., Schramm, R., and Miranda, E. (2014). Use of Body Motion to Enhance Traditional Musical Instruments. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '14, pages 601–604, London, United Kingdom.

Wakefield, J., Dewey, C., and Gale, W. (2017). LAMI: A Gesturally Controlled Three-Dimensional Stage Leap (Motion-Based) Audio Mixing Interface. In *Audio Engineering Society Convention*, Berlin, Germany. Audio Engineering Society. Available from: `http://www.aes.org/e-lib/browse.cfm?elib=18661`. Accessed: 18 July 2017.

Ware, C. (1997). *Basics of vocal pedagogy: the foundations and process of singing.* McGraw-Hill.

Waves (2015). Wave 360° Surround Tools. Available from: `https://www.waves.com/bundles/360-surround-tools`. Accessed: 12 June 2015.

Wehner, M. (2012). Man to machine, applications in electromyography. In *EMG Methods for Evaluating Muscle and Nerve Function*, pages 427–453. BoD, Books on Demand.

Wessel, D. and Wright, M. (2002). Problems and Prospects for Intimate Musical Control of Computers. *Computer Music Journal*, 26(3):11–22.

Wilkie, K., Holland, S., and Mulholland, P. (2009). Evaluating musical software using conceptual metaphors. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology*, BCS-HCI '09, pages 232–237, Cambridge, United Kingdom.

Wilkie, K., Holland, S., and Mulholland, P. (2010). What Can the Language of Musicians Tell Us about Music Interaction Design? *Computer Music Journal*, 34(4):34–48.

Wilkie, K., Holland, S., and Mulholland, P. (2013). Towards a Participatory Approach for Interaction Design Based on Conceptual Metaphor Theory: A Case Study from Music Interaction. In Holland, S., Wilkie, K., Mulholland, P., and Seago, A., editors, *Music and Human-Computer Interaction*, pages 259–270. Springer London, London.

Wilson, S. and Harrison, J. (2010). Rethinking the Beast: Recent Developments in Multichannel Composition at Birmingham Electroacoustic Sound Theatre. *Organised Sound*, 15(3):239–250.

Witten, I., Frank, E., Hall, M. A., and Pal, C. J. (2014). Weka 3. Available from: `https://www.cs.waikato.ac.nz/ml/weka/`. Accessed: 24 November 2015.

Witten, I. H. and Frank, E. (2005). *Data mining: practical machine learning tools and techniques.* Elsevier, San Francisco, CA.

Wozniewski, M., Bouillot, N., Settel, Z., and Cooperstock, J. R. (2008). Large-Scale Mobile Audio Environments for Collaborative Musical Interaction. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '08, pages 13–18, Genova, Italy.

Wu, J. C., Yeh, Y. H., Michon, R., Weitzner, N., Abel, J., and Wright, M. (2015). Tibetan Singing Prayer Wheel: A Hybrid Musical-Spiritual Instrument Using Gestural Control. In *New Interfaces for Musical Expression*, NIME '15, pages 91–94, Baton Rouge, Los Angeles, USA.

x-io Technologies (2016). x-IMU. Available from: `http://x-io.co.uk/x-imu/`. Accessed: 1 August 2015.

Xambó, A., Laney, R., Dobbyn, C., and Jordà, S. (2013). Video Analysis for Evaluating Music Interaction: Musical Tabletops. In Holland, S., Wilkie, K., Mulholland, P., and Seago, A., editors, *Music and Human-Computer Interaction*, pages 241–258. Springer London, London.

Zaunschirm, M., Frank, M., and Zotter, F. (2016). An Interactive Virtual Icosahedral Loudspeaker Array. In *Proceedings of 42nd Annual Conference on Acoustics*, DAGA 2016, Aachen, Germany. Accessible from: `https://iaem.at/Members/zotter/2016_zaunschirm_virtualico_daga.pdf`. Accessed: 19 December 2016.

Zbikowski, L. M. (1997a). Conceptual Models and Cross-Domain Mapping: New Perspectives on Theories of Music and Hierarchy. *Journal of Music Theory*, 41(2):193–225.

Zbikowski, L. M. (1997b). Des Herzraums Abschied: Mark Johnson's Theory of Embodied Knowledge and Music Theory. *Theory and Practice*, 22:1–16.

Zivix (2017). Jamstik. Available from: `https://jamstik.com/`. Accessed: 24 November 2017.

Zotter, F. and Sontacchi, A. (2007). Icosahedral Loudspeaker Array. Technical Report 39/07, Institut für Elektronische Musik und Akustik, Graz, Austria.