



HAL
open science

Exploration in Movement-Sound Interactive Systems

Victor Paredes

► **To cite this version:**

Victor Paredes. Exploration in Movement-Sound Interactive Systems. Sound [cs.SD]. Sorbonne Université, 2024. English. NNT : 2024SORUS425 . tel-04901689

HAL Id: tel-04901689

<https://theses.hal.science/tel-04901689v1>

Submitted on 20 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Exploration in Movement-Sound Interactive Systems

Victor PAREDES

Thèse de doctorat de Sorbonne Université

ED130 EDITE - Ecole Doctorale Informatique Télécommunications Electronique

UMR 9912 STMS - Sciences et Technologies de la Musique et du Son
IRCAM · CNRS · Sorbonne Université

Equipe ISMM - Interaction Son Musique Mouvement

Composition du jury

Florent BERTHAUT <i>Université de Lille</i>	Rapporteur
Alexander Refsum JENSENIUS <i>University of Oslo</i>	Rapporteur
Stefania SERAFIN <i>Aalborg University</i>	Examinatrice
Indira THOUVENIN <i>UTC (Compiègne)</i>	Examinatrice, Présidente du Jury
Frederic BEVILACQUA <i>IRCAM</i>	Directeur
Jules FRANÇOISE <i>Université Paris-Saclay</i>	Co-Encadrant

Abstract

Movement-sound interactive systems enable interactions with computer-generated sounds through sensing technology, typically wearable motion sensors or cameras. A large body of research has been reported on such systems by communities around New Interfaces for Musical Expressions (NIME), Sound and Music Computing, and Movement & Computing. These works have various applications in music, dance, sports, and rehabilitation. Although this research field has matured in several aspects, the different learning processes implied in designing and performing with such systems remain elusive as they cannot be described or formalized from a unique point of view. This PhD dissertation aims to investigate, through the lens of *exploration*, the learning and appropriation of movement-sound interactive systems.

This work is an interdisciplinary effort combining different methodological approaches from Engineering, Human-Computer Interaction, and Research-Creation. First, interviews with artists, composers, and educators using movement-sound technologies have been carried out to ground our research in artistic and pedagogical contexts. Second, an interactive system named *mosX* was designed to be useable and adaptable in different research and artistic contexts through various cycles of experiment and technological developments. *mosX* allows users to explore sound samples through body movements measured by wireless sensors. It uses *polyspring*, a Python package we developed to manipulate sound database 2-dimensional representations. *Polyspring* was motivated by a preliminary experiment of movement-sound exploration. Third, a controlled experiment using *mosX* was conducted, where participants explored a set of sounds mapped to the posture of their dominant arm. Quantitative metrics were created to analyze the exploration, supplemented by qualitative data. The analysis proved relevant for quantifying the space explored by participants and eliciting their experiences and different early learning processes. Fourth, we present, from a first-person perspective, performance-led research utilizing *mosX* along with other gesture sonification strategies.

This research work can thus be seen as a loop between user-centered and participatory design, technology development, and performance-led practices. We found that interactive technologies are rarely the starting point of artistic practice but rather a tool embedded in the socio-cultural context in which artists evolve. It is their experiential qualities that are found to bring something meaningful to the practice, both as a reflector of the action and as a tool to experiment and create beyond norms. Both the interviews and the experiments show that open-ended exploration is a necessary step to understanding the possibilities and limitations of those tools and is an actual part of learning. As interactive systems are at the crossroads of several domains, learning how to use such systems was found to englobe multiple intertwined meanings: understanding the technology that supports them, getting better at crafting meaningful interactions, exploiting the resulting interaction, and ultimately, being able to transmit the related knowledge to others.

Résumé

Les systèmes interactifs mouvement-son permettent d'interagir avec des sons générés par ordinateur grâce à des capteurs de mouvement portés à même le corps ou des caméras. De nombreuses recherches ont été menées sur ces systèmes par les communautés autour des Nouvelles interfaces pour les expressions musicales, de l'informatique sonore et musicale, et de l'informatique pour le mouvement. Ces travaux ont diverses applications dans les domaines de la musique, de la danse, du sport et de la rééducation. Bien que ce domaine de recherche ait atteint sa maturité sur plusieurs aspects, les différents processus d'apprentissage impliqués dans la conception et la performance avec de tels systèmes restent difficile à saisir car ils ne peuvent être décrits ou formalisés de manière unique. Cette thèse de doctorat vise à étudier, sous l'angle de l'*exploration*, l'apprentissage et l'appropriation des systèmes interactifs mouvement-son.

Ce travail est un effort interdisciplinaire combinant différentes approches méthodologiques de l'ingénierie, de l'interaction humain-machine et de la recherche-crédation. Tout d'abord, des entretiens avec des artistes, des compositeurs et des éducateurs utilisant des technologies de mouvement et de son ont été réalisés pour ancrer notre recherche dans des contextes artistiques et pédagogiques. Ensuite, un système interactif appelé *mosX* a été conçu pour être utilisable et adaptable dans différents contextes artistiques et de recherche à travers divers cycles d'expérimentation et de développements technologiques. *mosX* permet aux utilisateurs d'explorer des échantillons sonores grâce à des mouvements corporels mesurés par des capteurs sans fil. Il utilise *polyspring*, un package Python que nous avons développé pour manipuler les représentations bidimensionnelles des bases de données sonores. *Polyspring* a été motivé par une expérience préliminaire d'exploration mouvement-son. Troisièmement, une expérience contrôlée utilisant *mosX* a été menée, au cours de laquelle les participants ont exploré un ensemble de sons correspondant à la posture de leur bras dominant. Des mesures quantitatives ont été créées pour analyser l'exploration, complétées par des données qualitatives. L'analyse s'est avérée pertinente pour quantifier l'espace exploré par les participants et pour obtenir leurs expériences et les différents processus mise en place lors de l'apprentissage initial. Quatrièmement, nous présentons, à la première personne, une recherche sur des performances artistique utilisant *mosX* ainsi que d'autres stratégies de sonification gestuelle.

Ce travail de recherche peut donc être considéré comme une boucle entre la conception centrée sur l'utilisateur et participative, le développement technologique et les pratiques basées sur la performance artistique. Nous avons constaté que les technologies interactives sont rarement le point de départ de la pratique artistique, mais plutôt un outil intégré dans le contexte socioculturel dans lequel les artistes évoluent. Ce sont leurs qualités expérientielles qui apportent quelque chose de significatif à la pratique, à la fois en tant que reflet de l'action et en tant qu'outil permettant d'expérimenter et de créer au-delà des normes. Les entretiens et les expériences montrent que l'exploration ouverte est une étape nécessaire pour comprendre les possibilités et les limites de ces outils et qu'elle fait partie intégrante de l'apprentissage. Les systèmes interactifs étant à la croisée de plusieurs domaines, il s'est avéré que l'apprentissage de l'utilisation de ces systèmes englobe de multiples significations entrelacées : comprendre la technologie qui les soutient, s'améliorer dans l'élaboration d'interactions significatives, exploiter l'interaction qui en résulte et, enfin, être capable de transmettre les connaissances correspondantes à d'autres.

Remerciements

Je souhaite commencer par remercier Frederic et Jules, sans qui je n'aurais pu aller au bout de ce travail. Ils ont été pour moi un soutien permanent autant sur le plan académique que humain. Ils m'ont enseigné comment être chercheur tout en me laissant beaucoup de liberté dans mes actions. J'ai toujours ressenti de leur part une bienveillance qui m'a permis de m'exprimer et de travailler sans crainte de jugement. De plus, ils sont restés ouverts à tout ce qui ne touchait pas directement à ma recherche : musique, création, enseignement, sport, et m'ont toujours encouragé à les poursuivre. Pour tout cela je vous remercie, vous êtes pour moi les exemples d'une recherche saine, humaine et captivante.

Cette thèse représente l'aboutissement d'un parcours scolaire et d'un parcours musical qui n'ont cessé de se croiser. Tout au long de ces enseignements j'ai croisé de nombreuses personnes qui m'ont aidé, orienté et motivé dans la poursuite de mes études. Tout d'abord, je remercie mes enseignants du Conservatoire d'Hérouville-Saint-Claire et en particulier Laurent Letassey, qui m'a donné goût à la musique et à l'écoute attentive du son, ainsi que Norbert Genvrin, qui a toujours soutenu mon projet de travailler dans la technique musicale. Je remercie aussi mes enseignants du secondaire du Lycée Salvador Allende ainsi que les enseignants de la prépa PTSI-PT du Lycée Dumont d'Urville. Ils sont en grande partie responsables, de par leurs enseignements et conseils, de mon orientation académique. Enfin je remercie les enseignants du département de génie mécanique de l'ENS Cachan qui m'ont transmis le goût et les compétences de l'enseignement et qui m'ont orienté vers l'IRCAM.

Je tiens aussi à remercier de nombreuses personnes à l'Ircam avec qui j'ai vécu ces quatre dernières années et qui m'ont énormément apporté. J'ai partagé beaucoup avec mes collègues de bureau et de l'équipe ISMM. Je remercie aussi tous mes amis rencontrés au laboratoire : Aliénor, Ninon, Thomas, Mathilde, Vincent, Apolline, David, Sarah, Lenny, Giovanni et beaucoup d'autres. Bien sûr, j'ai aussi été soutenu par de nombreuses personnes en dehors de l'IRCAM et du monde de la recherche. Je remercie mes parents à qui je dois tellement. Ils m'ont donné accès à la musique et à la possibilité de mener des études longues. Ils m'ont appris la patience et la persévérance mais aussi à l'importance de prendre soin de soi autant au travail qu'en dehors. Je remercie mes amis d'enfance et d'adolescence : Maël, Lola, Zoé, Noé et Camille, avec qui j'ai partagé de nombreux moments de vie depuis bientôt 15 ans et qui ont été à mes côtés de près ou de loin pendant ces années de doctorat. Je remercie aussi mes amis rencontrés à l'ENS Cachan : Mathilde, Guillaume, Julien, Antoine et Samuel qui m'ont apportés soutien et conseil ces quatre dernières années. Je tiens à remercier particulièrement Marina, avec qui j'ai eu le plaisir de partager ces trois dernières années de doctorat, et qui a été une source de motivation et de joie tout au long de ce projet.

Tout ce travail n'a pas été simple, souvent compliqué même. Mais j'ai passé un moment formidable et je garderai longtemps les souvenirs du temps passé avec vous tou-te-s.

Contents

Abstract	i
Résumé	ii
Remerciements	iii
1 Introduction	1
1.1 Context	1
1.2 Motivations and Aims	2
1.3 Methods	2
1.4 Contributions	4
1.5 Dissertation Overview	5
2 Background	7
2.1 Movement-Based Interaction with Sound	7
2.2 Learning Movement-Sound Interactive Systems	13
2.3 Chapter Summary	18
3 Interviews with Practitioners: Learning and Appropriating Movement-Sound Interactive Systems	19
3.1 Introduction	19
3.2 Method	20
3.3 Theme 1: Technology Mediates Practice	21
3.4 Theme 2: Navigating Usability, Normativity and Obsolescence	23
3.5 Theme 3: Learning and the Development of a Personal Practice	24
3.6 Discussions	27
3.7 Chapter Summary	29
4 Shaping 2-D Sound Spaces for Movement-Sound Exploration	31
4.1 Introduction	31
4.2 Exploring 2-D Sound Spaces with Movements: First Prototype and Experiment	33
4.3 Manipulating a 2-D Point Set Distribution: Theoretical Background	39
4.4 Toolbox Description	44
4.5 Examples and Perspectives	48
4.6 Chapter Summary	49
5 Exploration in Movement-Sound Interactive Systems	51
5.1 Introduction	51
5.2 <i>mosX</i> : movement sound explorer	52
5.3 Exploration Data Analysis	59
5.4 Experimental protocol	63
5.5 Results	66
5.6 Discussions	76

5.7 Chapter Summary	80
6 Artistic Applications of Movement-Sound Interactive Systems	81
6.1 Introduction	81
6.2 Methods	82
6.3 <i>Circus Lab</i> : Collaborative Interaction Design with Shared Sensor Data	82
6.4 <i>Sympoiesis</i> : Movement Sound Exploration	85
6.5 Spectators' Experience Design	96
6.6 Discussions	100
6.7 Chapter Summary	101
7 Discussions	103
7.1 Learning and Appropriation are Intertwined with the Development of Practice	103
7.2 Exploration Supports Design, Learning, and Appropriation	105
7.3 Designing for Exploration	109
7.4 Chapter Summary	113
8 Conclusion and Perspectives	115
A Appendix	119
A.1 Chapter 3	119
A.2 Chapter 4	121
A.3 Chapter 5	121
List of Figures	122
List of Tables	124
List of Acronyms	125
Bibliography	126

CHAPTER 1

Introduction

1.1 Context

Embodied interaction has been a key concept in the design of human-machine interactions for more than two decades. Dourish formulated it in his 2001 seminal work as the reunion of two concepts: tangible computing and social computing, together with phenomenology as a conceptual layer. Social Computing acknowledges the sociological context surrounding and influencing interactions with any information system, while tangible computing is an attempt at “*unifying the physical and electronic worlds to create a blend which is more closely matched to our daily experience and abilities*” (Dourish, 2001, p.190).

Around the same time, gestural control of digital music instruments (DMIs) became a formalized object of study, which gave birth to the New Interfaces for Musical Expression (NIME) community (Wanderley, 2001). It brings together researchers, musicians, and designers working on this topic. Free from physical and acoustic constraints, DMIs allow for a wide variety of sounds, many impossible to produce by pure acoustic means, but with the cost of losing the physicality inherent to acoustic instruments. As Dobrian and Koppelman stated, “*Computer interfaces can dissociate gesture from result to varying degrees by the way that software intermediates the relationship between gesture and the resulting sound*” (Dobrian and Koppelman, 2006). Practitioners and researchers in music and performing arts investigated both conceptual and practical aspects of the interaction with music, which led to the notion of embodied musical cognition (Leman, 2007) and embodied musical interactions (Lesaffre et al., 2017). Embodiment refers to human action as expressed through corporeal actions and body movements. Embodied musical interactions consider the interaction with music as lived by the body in relation to its musical environment. An environment that encompasses sounds, people, places, and, more broadly, all the context around music creation. As such, these interactions are both physical and social.

1.2 Motivations and Aims

Our general aim in this research is to design embodied interactions with sounds that enrich the movement experience and foster kinesthetic awareness across various practices, including music, dance, sports, or physical rehabilitation. In particular, we focus on full-body interaction through cameras or wearable sensors such as inertial measurement units (IMUs) that can measure accelerations and rotations of body parts. Throughout this dissertation, we refer to movement-sound interaction technologies, which encompass a set of technological objects and technical skills that enable users or performers to produce sounds through their movements using sound synthesizers.

Using such movement interfaces involves learning processes to reach a certain level of expertise. Many new interfaces are used only by their creator or a single performer, thus eluding issues in skills acquisition by a wider group of users. Recent reflections in the NIME community expressed challenges on the adoption and longevity of the systems created (Morreale, 2017; Marquez-Borbon and Martinez-Avila, 2018; Masu et al., 2021). Moreover, beyond NIME, there is no coercive community of practice around DMIs like there is around traditional musical instruments, but rather heterogeneous communities with different practices. We can then wonder whether it is simply a question of time and opportunity to reach a broader range of users or whether we should reconsider, at least partially, how we design and learn these so-called “*new interfaces*”?

This PhD project was initially motivated by the idea that embodied learning and exploration processes of movement-sound interactive systems should be fully investigated to promote widespread adoption and practice. This PhD was initiated in synergy with a collaborative project on the notion of *Learnability in Movement Interaction*¹. This research can be rephrased through this general following question that guided our thinking:

What are the relationships between exploration, learning, and appropriation of movement-sound interaction technologies?

In particular, what types of exploration exist in such a context? Can we find a measure of movement exploration? What is the role of movement exploration when learning and appropriating movement-sound interactions? How can we design to foster exploration? These are ambitious general questions for longer research prospects beyond the scope of a single PhD. For this PhD research, we established the following specific goals:

1. study of the learning and appropriation of these systems in artistic contexts;
2. development of tools to study exploration and learning of movement-sound spaces;
3. controlled experiments in movement-sound space exploration to investigate different processes of exploration and early learning;
4. practice-based research-creation in collaboration with artists.

1.3 Methods

Our research questions on exploration, learning and appropriation are central to NIME research, and relates to several scientific disciplines, including sonic interaction design and sensorimotor learning. Such a question cannot solely be answered either in a laboratory setting or through artistic creation. It calls for a broad *interdisciplinary* approach “*where*

¹Element project, <https://element-project.ircam.fr/>

integration of the contributions of several disciplines to a problem or issue is required,” as described by Stember (1991). Such integration is far from trivial, and several methodological frameworks were created to frame and support interdisciplinary endeavors. Two frameworks, in particular, propose entangling research methods drawn from different disciplines related to design and performing arts. Method triangulation (Mackay and Fayard, 1997) and performance-led research in the wild (Benford et al., 2013). Their structure and relationship between activities are schematized in figure 1.1.

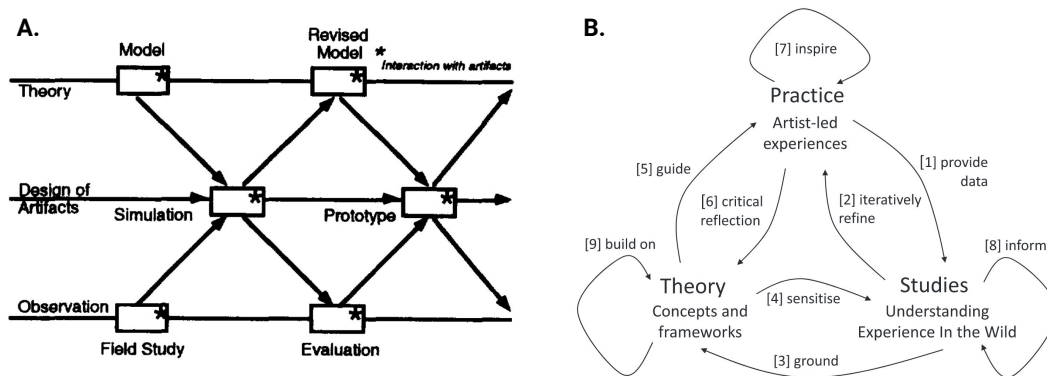


Figure 1.1: A. Method triangulation in Design, from (Mackay and Fayard, 1997); B. Performance-led research in the wild framework, from (Benford et al., 2013)

Mackay and Fayard advocate “[the use of] more than one research approach to address the same question,” particularly between scientific and design disciplines. They propose a framework alternating between theory and empirical observations of people interacting with “artificially created” artifacts, as shown in figure 1.1A.

Performance-led research in the wild is dedicated to searching how interactive technologies can enable artistic installations and performances. Central to performance-led research is the involvement of artists who shape the research questions and design processes. This ensures that the technological interventions are situated in the artistic context and address real needs and desires within performance practice. In this framework, research combines three activities: practice, theory, and studies. The framework proposed by Benford et al. introduces nine key relationships between and inside those activities that define how they interact with each other.

Figure 1.2 reflect the methods I followed in my research. It borrows from the triangulation framework the central aspect of the design activity. We indicate relations between activities using two of the nine key relationships introduced by Benford et al.: informing and guiding. It also integrates performance-led research in the wild as axes, as described below.

- *Observation* refers to field studies and experiments, including interviews with practitioners and controlled studies.
- *Design* refers to activities led by the researcher in the laboratory to enable specific interactions in performance practices or support lab studies.
- a *Practice* axis is added to the framework. *Practice* represents an opportunity to test and refine *Designed* artifacts. *Practice* may involve design, but it is led, like the research, in collaboration with artists. It is situated in a performance context constrained by a strict deadline: the premiere (no hoping for any extension here), and requires a working

and reliable system.

Finally, the *Theory* axis is not represented in figure 1.2 as we do not focus our triangulation on theory as an axis for research activity leading to contributions. Nonetheless, we performed a systematic review of the state-of-the-art, informing and sensitizing the various studies, designs, and practices.

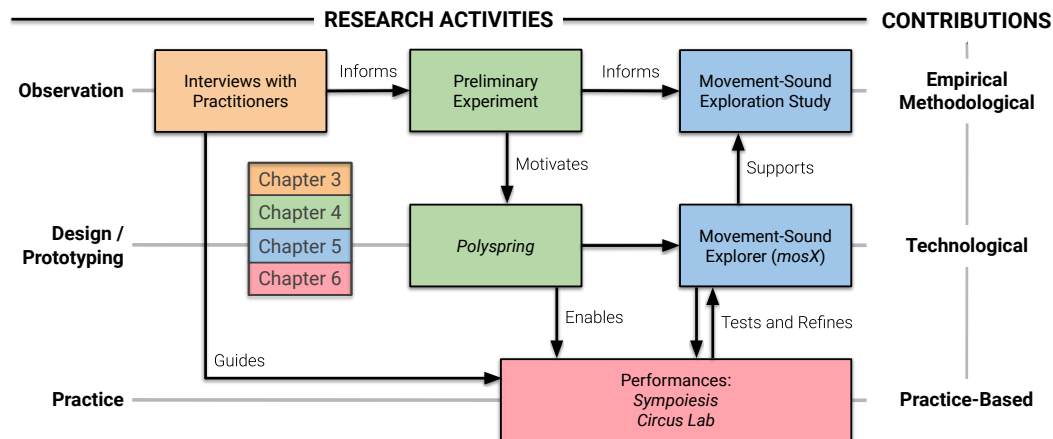


Figure 1.2: Research activities are organized on three axes that relate to each other. Four types of contributions follow from those activities.

1.4 Contributions

This dissertation presents contributions of four types: empirical, methodological, technological, and art-based.

1.4.1 Empirical Contributions

The empirical contributions brought by field study and controlled experiments are:

- Interviews with expert artists using movement-sound interaction highlighting the intertwined nature of learning and appropriating technologies and the central role of exploration in these processes.
- Experiments demonstrating early learning through exploration of movement-sound interactive systems and identifying different modes of exploration when learning movement-sound interactions.

1.4.2 Methodological Contributions

Methodological contributions demonstrated in controlled experiments are:

- A method to visualize and quantify the exploration of a movement-sound interaction space based on Kernel Density Estimation techniques.
- A drawing-based method to elicit mental representations of the audiomotor map that participants acquired when exploring movement-sound interaction spaces.

1.4.3 Technological Contributions

Technological contributions are:

- *polyspring*: a Python package to design sound spaces for interaction by manipulating 2-dimensional sound corpus representations.
- *mosX*: an interactive system that enables the embodied exploration of a sound space.
- Max4Live objects that enable multiple musicians to receive signals from movement sensors and process them directly in the musical software Ableton Live.

1.4.4 Practice-Based Contributions

Performance-led research yielded the following contributions:

- *Sympoiesis*, an art-science collaboration with dancers where exploration-based movement-sound interaction was employed as a spectator experience design tool.
- *Circus Lab*, a collaboration between musicians and circus artists where movement-sound interactions amplified performers' actions.

1.5 Dissertation Overview

Chapter 2 introduces broad concepts and recent results on digital musical instruments, sonic embodied interaction, sensorimotor learning, and exploratory learning.

Chapter 3 describes interviews with six artists involved in movement-sound interactions. The themes that emerged from the thematic analysis gave us insights to guide the integration of interactive technologies in performances as well as inform a preliminary experiment in sensorimotor learning described in Chapter 4

Chapter 4 reports on a preliminary experiment using a prototype for movement-sound interaction that enables exploring a 2-dimension sound space from either a 2D planar surface or from the arm movement. The results motivated the development of the Python package *polyspring*, presented in a second part, which offers a greater capacity to manipulate the sound space.

Chapter 5 presents a movement-sound exploration experiment that uses a novel system called *mosX*, based on *polyspring* (presented in follows in Chapter 4). A quantitative method to analyze exploration data is introduced. The results underscore the effectiveness of exploration-based learning in early audiomotor map acquisition.

Chapter 6 presents *Sympoiesis* and *Circus Lab*, two artistic collaborations with dancers and circus artists. Both used movement-sound interactions with different spectator experience design objectives. The creation processes are detailed to highlight the discussions and tensions that appear in such collaborations, with an emphasis on the necessary sharing that hybrid practices entail.

Chapter 7 discusses the key results of this dissertation and proposes considering pedagogical design as part of the design process of movement-sound interactions.

CHAPTER 2

Background

The main question addressed in this dissertation concerns the link between exploring and learning movement-sound systems for the performing arts. This chapter introduces the core concepts and related studies that this work refers to. First, digital musical instruments focus mainly on movement-sound embodied interactions; second, sensorimotor learning and exploratory learning. The following chapters will introduce more specific related works when necessary.

2.1 Movement-Based Interaction with Sound

This thesis's central objects of study are systems that enable the interaction between human movements and sounds generated by a computer. Studies of such systems lie at the interface of sound processing, embedded electronics, sensorimotor learning, computer sciences, and human-computer interactions (HCI) in particular.

2.1.1 Digital Musical Instruments

At the end of the 20th century, electronic music is already several decades old. In 1948, Pierre Schaeffer broadcasted the first concert of *musique concrète* made entirely of recorded materials from the environment or instruments. As composer Elliot Schwartz pointed out in 1978, this moment was especially significant for music:

The product of his work was a finished sound object meant to be heard directly over loudspeakers, a one-to-one communication between composer and listener that bypassed the need for realization by an interpreting performer. (Schwartz, 1978, p.37)

The absence of performers in some of the early electronic musical creations was surely the most important break with acoustic music at the time. It presaged a shift in what it meant to compose and play music. This momentum was confirmed when Max Mathews and his team performed the first music pieces played with a computer. He observed in 1963 that “*the computer has proved to be of more use to the composer who wishes to produce new compositions with new sounds than to the performer who wishes to duplicate existing music.*” (Mathews, 1963, p.28)

Nonetheless, electronic musical instruments such as the Theremin or the Ondes Marthenot were designed decades before and are still played nowadays. Although revolutionary for their time concerning gestural control and sound synthesis techniques, these new instruments were designed to extend the acoustic instrumental tradition regarding musical content and found their place in the modern repertoire. Arguably, the shift in how electronic music was to be performed emerged when Buchla and Moog built their first synthesizers in 1963 and 1964, respectively. They responded to demands from various composers to make electronic music creation more accessible and usable in concert settings. These first models ended up using knobs, faders, and variations of keyboards as inputs. This would remain the norm for interaction with sound synthesis for several years, with the addition of string, percussive, and wind form factors directly inspired by acoustic instruments as displayed in figure 2.1.



Figure 2.1: Roland A-80, Nyle Steiner EVI, SynthAxe, and Roland SPD-8 are MIDI controllers that afford interactions transferred from acoustic physical instruments.

As computers grew increasingly more affordable and efficient, it became possible to prototype electronic music programs for real-time applications on consumer-grade systems using programming languages such as PureData, Max/MSP, or SuperCollider. This gave birth to an exponential development of digital sound synthesis techniques. Parallely, sensors have become more and more affordable and widespread. At first, sensor data were processed in a computer and sent to synthesizers for real-time interaction with electronic sound synthesis or directly controlled sounds produced by a computer program. These interfaces detached from the “*instrument-like*” controllers of the time. The union of digital sound synthesis and sensor-based interfaces gave birth to a new type of instrument called Digital Musical Instruments (DMIs). By the early 21st century, the movement had grown to the point where this practice became an object of study that gathered researchers, designers, and artists under the New Interfaces for Musical Expression (NIME) community (Jensenius and Lyons, 2017).

Control and Expressivity

The main objective of DMI design is to create a computer-based instrument that can be played by a performer and seen and heard by an audience (O’Modhrain, 2011). Wessel and Wright introduced a conceptual framework that models interactions between a performer and a digital musical instrument (Wessel and Wright, 2002), represented in figure 2.2. The performer has musical intentions that are transformed into physical action by the motor program. The sensors from the gestural interface transform performers’ movements into control signals that are then processed to output control parameters. A controller maps those signals to the synthesizer parameters. Finally, the synthesizer generates a sound signal from control parameters that is amplified and converted to mechanical sound waves by headphones, speakers, or less usual electromechanical transducers.

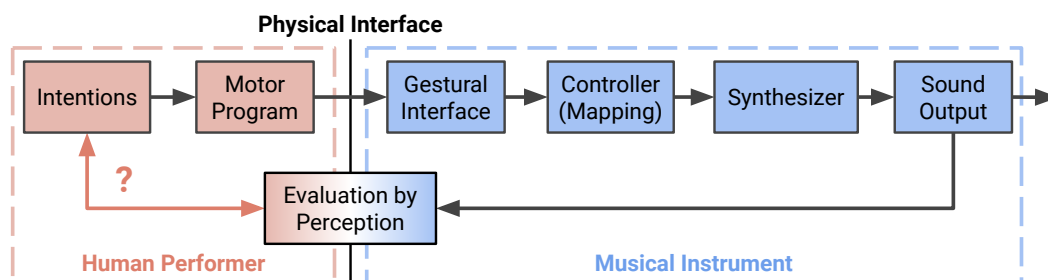


Figure 2.2: Framework for interaction with a computer-based musical instrument originally proposed by Wessel and Wright (2002). The author mentions the question of how intentions evolve when exploring the interaction.

Wessel and Wright noted that “*this diagram is schematic and incomplete. One aspect that is not well captured by it is the way in which performers’ intentions are elaborated upon by discovery of new possibilities afforded by the instrument. Experimental and otherwise exploratory intentions are certainly dear to the authors*” (Wessel and Wright, 2002). They outlined the difficulty of modeling the interaction between the perception and the intentions when exploring and experimenting with such systems.

Expressivity is naturally central to the NIME community and DMI design in general. Dobrian and Koppelman defines it as “*the gestural nuances that a live performer adds to a composed material.*” They positioned the control of a DMI as a property of an instrument that enables expressive use by a performer. A DMI must propose a mapping that is concurrently repeatable, learnable, and “*sufficiently refined*” to allow for an intimate and complex control of the instrument (Dobrian and Koppelman, 2006). Fels et al. uses the term “*transparent mappings*” to describe its capacity to enable expressivity for the performer and appreciation of this expressivity by the audience (Fels et al., 2002). O’Modhrain investigated the evaluation of DMIs by the various stakeholders accompanying their design and use. Questions of expressivity were raised when evaluating the instrument’s playability (O’Modhrain, 2011). He found that the first approach was to consider specific tasks that transfer expressive characteristics borrowed from traditional instruments to DMI playing. Such evaluation of the ability to support expressive playing supports the objective of transparency proposed by Fels et al..

However, Gurevich and Treviño nuances the need for transparency of the performer’s expressive intents to the audience (Gurevich and Treviño, 2007). This representation of a unique artistic intent to be perceived identically by all audience members has been questioned by visual arts for decades. Making artistic intent as transparent as possible should not become

the sole purpose of DMI design and practice, as it is not necessarily the artist's responsibility to make sense of the proposed artistic performance.

The question of spectators' experience when attending a DMI performance has been raised by several researchers. Reeves et al. (2005) proposed a framework for spectator's experience design that applies to any interactive setup where interaction either hides or amplifies manipulations and effects. Berthaut et al. proposed to display a 3-D representation of the internal mechanism of a DMI to amplify the manipulations of the performer on the instrument for the audience (Berthaut et al., 2013). Screen-based amplification is also investigated to connect the audience with immersive virtual musical instruments (Berthaut et al., 2014).

Mapping Strategies

Historically, the interaction in DMIs has been formalized through the notion of the so-called *mapping*. Once the choice of gestural interface and sound synthesis has been made, mapping the control parameters to the synthesis parameters is the final step in enabling control (Hunt et al., 2000). Over the years, many approaches have been proposed to design complex multidimensional mappings, using diverse methods such as physical models (Momeni and Henry, 2006; Johnston et al., 2008), machine learning (Fiebrink et al., 2011) or geometrical approaches (Van Nort et al., 2014). In the latter approach, discrete points serve as control structures for generating continuous mappings, where several techniques can be used to define and transform the geometry of the intermediate mapping layer.

Zbyszynski et al. (Zbyszyński et al., 2021) proposed a regression method (Neural Network Regression Model) to set the mapping between gestural data to the audio space. Françoise and Bevilacqua proposed a technique to map movements to sounds "*through interaction*" (Françoise and Bevilacqua, 2018). The user performs gestures over different target sounds, providing a finite set of continuous corporeal demonstrations of desired embodied associations. From those few examples, real-time movement recognition to synthesis parameter generation was made possible.

Other machine learning approaches include unsupervised mapping learning. Fasciani and Wyse proposed an approach to automatically map control parameters of general-purpose gestural interfaces to synthesis parameters using dimensionality reduction techniques and statistical analysis of perceptual features of the synthesized sound (Fasciani and Wyse, 2012). Murray-Browne and Tigas adopts an intermediary approach by using variational autoencoders to generate a latent space from unlabelled gestural examples (Murray-Browne and Tigas, 2021).

West et al. led a mapping design experiment with nine NIME users who had to create a mapping between a given interface and synthesizer in an hour. They found that participants alternated between short phases of exploration and experimentation when building mapping between control and sound synthesis parameters (West et al., 2020). A second study highlighted the difficulty of evaluating mapping design and the importance of context and subjectivity in doing so (West et al., 2021). On the other end of the spectrum, Fiebrink and Sonami reflected on an eight-year collaboration between the authors. Laetitia Sonami insisted on the time to invest in order to create a meaningful interaction with the instrument (Fiebrink and Sonami, 2020).

2.1.2 Embodied Sonic Interactions

The arrival of personal computers in workplaces and homes gave rise to a new field of study focusing on how users interact with computers and how to design these interactions, namely human-computer interaction (HCI). Three stages, or so-called waves of HCI design, are identified. This section accounts for the origin of HCI and its recent evolutions, particularly the shift toward embodied interactions and the emergence of sonic interaction design.

Human-Computer Interaction

The first study on human interaction with computers emerged from the field of *human factor engineering*, or *ergonomics*, in Europe. At the beginning of the 20th century, the focus was on dividing tasks between humans and machines. “*Humans*” were then considered a necessary factor with skills and limitations that delimited which functions they could perform (Cott and Kinkade, 1972). When computers were introduced in the workplace, they were not regarded as different from machines operated by humans, with programmers mediating between users and systems. At the end of the seventies, most computer users were not scientists or programmers but workers with daily computer use whose jobs did not focus on computing. At the time, programming was necessary for operating a personal computer, which few users were willing to do.

HCI emerged partly from these changing conditions. Analog to human factors, HCI studies added that designing for interaction between humans and machines required thinking about the cognitive coupling between them, which pushed further than simply designing better displays (Carroll, 1987). This first wave of HCI studies was led mainly in laboratories to study users’ behaviors regarding design choices such as command naming conventions (Grudin and Barnard, 1984). In 1995, Bannon advocated shifting “*from human factors to human actors*” (Bannon, 1995) and, in particular, “*from the laboratory to the workplace*”. The emphasis was then on the context in which the interaction took place and became the central concept of the second wave, pointed out Bødker (2006).

The third wave of HCI challenged the common values of technology favored in the second wave, such as efficiency, to focus on experience and meaning-making (Bødker, 2015). Harrison et al. explained that the third paradigm defined the interaction as a “*form of meaning making in which the artifact and its context are mutually defining and subject to multiple interpretations*” (Harrison et al., 2007).

Embodied Interactions

Dourish formulation of embodied interaction in HCI practice participated certainly to the shift toward the third wave. He argues that embodiment, referring to the integration of physical and social experiences in interaction, should be a fundamental consideration in the design and analysis of interactive systems (Dourish, 1999). Traditional approaches to HCI have often disregarded the importance of embodiment, emphasizing abstract models of cognition and information processing instead. Embodied interactions offer a more comprehensive understanding of how individuals engage with technology by encompassing cognitive processes, social contexts, and material environments. Importantly, he formulated several principles for embodied interaction design. The third and fourth ones are: *users, not designers, create and communicate meaning*, and *users, not designers, manage coupling*. We introduce only those two as they directly invite designers not to overstep their roles. In particular, coupling refers to the link between the internal representation of an artifact and the context in which it is

used.

Sonic Interaction Design (SID)

The new modes of interaction that emerged from the NIME community are, for the most part, thought of as musical instruments meant to be played in a concert setting. As such, these interactions could be associated with the so-called second wave of HCI, where the concert hall is analogous to the workplace contexts of the second wave Bødker (2015). Partially built from this core of research, sonic interaction design emerged as an *”active medium that can enable novel phenomenological and social experiences with and through interactive technology“* (Franinović and Serafin, 2013, p.vii). Recognizing that sonic interactions possess experiential qualities that can take music beyond the status of cultural products, sonic interaction design pulled interactive music systems out of studios and stages to infuse non-musical interactions with sound in everyday life contexts, fostering what could be considered *”third-wave”* sound interactions Tanaka (2019). This multi-sensory approach is grounded in ecological psychology, which declares that perception and action are intertwined within the environment (Wicker, 1984). SID builds on sounds’ expressive and emotional qualities to create interactive experiences that go beyond the limitations of traditional screen-based interfaces.

2.1.3 Fingertip and whole-body interactions with sounds

Physical Interfaces: From Tablet To Motion Sensors

Several Digital Music Instruments, using various sound synthesis techniques, use visual feedback and 2D interfaces. The Reactable is one of the landmarks that allow for complex synthesis techniques and collaborative playing (Jordà et al., 2007). More recently, new multi-touch interfaces with pressure sensors were used to develop new musical instruments. Schwarz et al. conducted a survey about the use of 2D touch interfaces for musical expression (Schwarz et al., 2020). Commercial products targeting musical applications, such as the Lightpad Block by Roli or the Erae Touch by Embodme, have multi-touch and pressure sensors with embedded low-resolution visual displays. Another multi-touch sensor with pressure, the *Sensel Morph*, can be used with swappable interface overlays. All these interfaces include, by design, both visual and audio interaction modalities.

Touch-based interaction on a surface using digital means relies on different action-sound coupling than the ones found in acoustic instruments and, more broadly, in any physical sound-producing object (Jensenius, 2022). Jensenius distinguishes *”acoustic”* instruments, where the action-sound coupling is given by physical properties, and *”electro-acoustic”* instruments, where the action-sound coupling is designed. In particular, it is possible to associate sound with any action using motion-sensing technology. Various features of body movement can be tracked using for example cameras Sentürk et al. (2012); Yoo et al. (2011); Han and Gold (2014); Dahlstedt and Skånberg (2019), 3D scanners Bernardo et al. (2017), or Inertial Measurement Units (IMUs) that embed accelerometers, gyroscopes or magnetometers Aylward and Paradiso (2006); Mitchell et al. (2012); Madgwick and Mitchell (2013); Medeiros and Wanderley (2014); Françoise and Bevilacqua (2018); Brown et al. (2018). All these approaches are based on capturing the dynamic of the moving body. Some interfaces, such as the *Myo* (Nymoen et al., 2015) or the *MindRove Armband*, rely on electromyography to measure muscle activity. Such sensors enable the detection of efforts and micro-movements (Françoise et al., 2017). In both cases, there is generally no visual feedback, and the user/musician must rely on their own proprioception. Haptic feedback (passive or active) can enhance or

facilitate gestural performance (Magalhães et al., 2020). Different levels of control can be designed for the movement-sound interaction, from a deterministic relationship such as found in acoustic-like control or with some indeterminacy compatible with more exploratory or improvisational control.

Movement-Sound Interactions in Dance

Movement-sound interactive systems do not limit their artistic use to DMIs. Enabling expressive movement sonification necessarily interests other movement-based performing arts. Here, we present a selection of works focusing on exploration, collaboration, and skill acquisition. Dahlstedt and Skånberg created an art piece in which performers explore a sound space using a depth *Kinect* camera (Dahlstedt and Skånberg, 2019). Françoise et al. proposed a dispositif called *CO/DA* in which one or two dancers improvise with sensors strapped to their bodies. The interaction is live-coded by a third performer, which creates a unique, constantly evolving interaction (Françoise et al., 2022). Fdili Alaoui reports on creating an interactive dance piece, revealing tensions that appeared when integrating technology (Fdili Alaoui, 2019). Similarly, Sullivan et al. reports on a case study of designing interaction for a dance performance. They highlighted the lack of standardization of design methods in such a context and the fact that it is not a deficiency but rather a reality that can serve as an example to other HCI contexts “*that are also resistant to standard methods and may call for more flexible design practices*” (Sullivan et al., 2023). Interestingly, Giomi highlighted the shift of focus from musical interactions to the somatic knowledge that movement sonification brings (Giomi, 2020). Such observation is essential in identifying the nature of the coupling that occurs between movement-sound interactive systems and dance performance creation. In movement-oriented practices such as dance, movers also use sound interaction to create installations for people “*to discover new interpretation of their dancing body*” (Bermudez et al., 2011) or foster kinesthetic awareness (Candau et al., 2017). Interactive systems can also help music education through embodied interaction of concepts such as conductors’ gestures (Bevilacqua et al., 2007).

2.2 Learning Movement-Sound Interactive Systems

The notion of learning in movement-sound interactive systems can be viewed from different angles. This section presents research works on the subject from various perspectives. First, considering the link these systems create between movement and sound, learning to use them requires the acquisition of sensorimotor skills. In particular, the notion of audiomotor map acquisition is presented. Second, the notion of exploratory learning in HCI is introduced.

2.2.1 Audiomotor Maps Acquisition

Learning how to use a movement-oriented musical interface is close to learning how to play an acoustic instrument. It comes down to learning which movement produces a particular sound: a sensorimotor map learning (van Vugt and Ostry, 2019). In this case, the user does not necessarily learn how to perform a new movement but rather to memorize the link between a movement and the sound he or she wants to produce. This application can be referred to as a sound-oriented task (Bevilacqua et al., 2016). Many other applications have been considered for movement sonification, mostly in sport and rehabilitation (Schaffert et al., 2019; Bevilacqua et al., 2017). These cases aim to improve motor performances, focusing on movement-oriented tasks.

Multiple components must be learned to enable virtuosity in sensorimotor acquisition, and different learning processes apply to those components (Wolpert et al., 2011). To optimize performance, the body relies on several control strategies. One of their goals is to compensate for any external or internal perturbations. In particular, the neuromotor system is subjected to noise, which produces motor variability as an output. This motor variability has task-relevant components that affect the effort in the targeted direction and task-irrelevant components that do not. It is this first component that needs to be compensated for.

However, motor variability can not be reduced only as a consequence of noise. It is now accepted that, during the early stage of practice, the motor system creates high-dimensional spaces with several solutions and explores those solutions through variability. This behavior is commonly labeled as “exploration” (Sternad, 2018), and it is the base mechanism of reinforcement learning. However, it has come to light that noise in the neuromotor system indeed impairs reinforcement learning (Therrien et al., 2018).

Supporting the idea that motor variability promotes learning, Wu et al. have shown that, during practice, the temporal structure of motor variability is reshaped by the neuromotor system to synchronize it with the trained task to improve learning (Wu et al., 2014). Those results are in keeping with reinforcement learning theory.

In the context of musical instrument playing, there are still many unknowns about the actual mechanisms that drive and facilitate sensory learning in this context. This is an object of recent research (Caramiaux et al., 2018). In particular, Van Vugt and colleagues proposed very intriguing motor control research in recent years that is particularly relevant to our studies. Their research is motivated by the will to understand the processes of learning musical instruments. This implies that the user learns a novel sensorimotor map that establishes the correspondence between sound properties and gestures without any visual feedback. Such sensorimotor maps are called *audiomotor maps*. They argue that such learning is different from motor learning experiments that employ visual feedback since they rely on adapting from previously known sensorimotor maps by the participants. On the contrary, learning audiomotor maps necessitates establishing ‘from scratch’ an association between specific movements and sounds. In this case, the very first phase of being exposed to such an audiomotor map is called *early learning*. Below, we summarize key findings from three papers.

First, they established that learning an audiomotor map can be effective even without feedback on an error. They found that two possible maps can be learned: one “*connecting sensory targets with motor commands*” and “*an error map linking sensory errors to motor corrections*” (van Vugt and Ostry, 2018b). This clearly shows that learning an audiomotor map can be performed initially by exposing the participant to a series of audiomotor instances without providing error-based feedback.

A second article showed that participants do not seem to “*interpolate*” linearly between previously known locations in the audiomotor map van Vugt and Ostry (2018a). Instead, it appears that they could reach an intermediate position on average as if the movement was following probability distributions approximately centered on the learned anchors.

In their third paper of this series, van Vugt and Ostry proposed to test several mechanisms that are known to be key components of learning in other paradigms, with the assumption that all were sufficient but not necessary to display learning (van Vugt and Ostry, 2019). One mechanism that has been shown to drive sensorimotor learning efficiently is error-based learning (Wolpert et al., 2011). The neuromotor system is given a target to reach and feedback

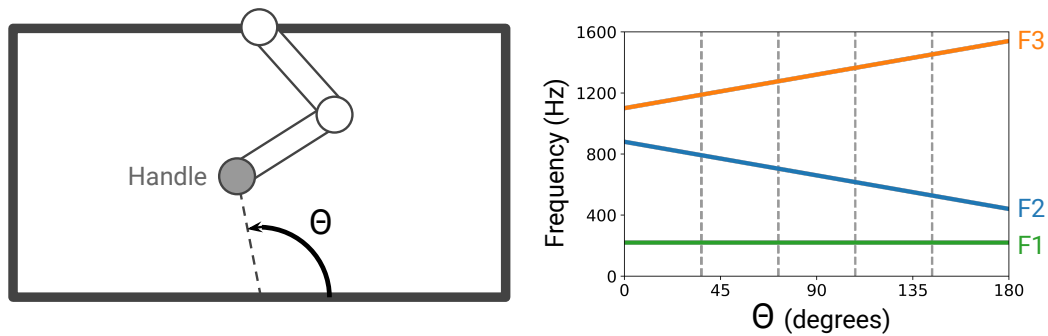


Figure 2.3: Experimental setup from (van Vugt and Ostry, 2019). The angular orientation of the handle is mapped to the frequency content of three sinusoidal oscillators.

after the movement is performed. Therefore, it can estimate the error by comparing the expected outcome with the output and trying to compensate for it. Van Vugt and Ostry have investigated, among other things, if learning is possible without target error. They used a simple one-to-one mapping, linking the angular position of a planar handler (in cylindrical coordinates) to the frequency content of three pure tones. The figure 2.3 presents the experimental setup. The objective of the participants was to learn which position produced which sound. The training consisted in placing the handler in space and listening to the corresponding sound output played for half a second. A group trained with sound targets, as described before, and another trained without targets, only with feedback sounds. Both groups appeared to have reduced the mean angular error after training, showing learning. It is consequently possible to learn a simple mapping with free exploration.

In addition, Van Vugt and Ostry tested two other hypotheses: if learning is possible for participants who passively undergo the exploration (assisted by a robotic arm) or for participants who explore a permuted mapping. The permuted mapping does not preserve distances between movements and sounds: a slight movement can result in a large variation of the frequency content and vice versa. It appeared that both paradigms induced learning. Learning achieved with permuted mapping is essential for more complex mapping learning that includes discontinuous sound feature variations.

In summary, these contributions show that early learning of audiomotor space could be possible with the gestural exploration of an audio space without any visual feedback. Moreover, their results indicate that a partially discontinuous audiomotor space, possibly found in corpus-based sound synthesis, does not prevent learning. It is thus interesting to use and confront these results for the design of sonic embodied exploration and interaction. We present in the Chapters 4 and 5 experiments that build on these results to study the early learning of more complex movement-sound interactions without visual feedback.

2.2.2 Exploration and Appropriation

The previous section considered exploration from a motor learning perspective as a mechanism through which the body finds solutions to achieve sensorimotor tasks. Van Vugt and Ostry's study put forward a more active form of exploration to acquire audiomotor maps. Such an approach is relevant to our research as active exploration is central to appropriation (Zappi et al., 2018)

Questions of appropriation and longevity in DMIs

Technology appropriation is a societal challenge that is not limited to interaction technologies. The question of how people appropriate technologies in an ever-evolving world is crucial to understanding and addressing inequalities in technology usage (Gradoz and Hoibian, 2019). In interaction design, appropriation considerations are essential in creating interactions that “*the users understand and are comfortable enough with to use it in their own way*” as described by Dix (2007). In DMI design, appropriation encompasses as much technological appropriation as the personal creativity that an individual will express through the use of the instrument. Zappi et al. (2018) proposed two studies of DMI appropriations. The first study showed that constrained interactions can promote the discovery of idiosyncratic techniques and that tighter constraints actually lead to deeper exploration. The second study proposed a “*hackable*” instrument that the performer could rewire. The range of proposed modifications highlighted that it is the performers who made sense of the instrument design by leveraging openness and modularity.

Moreover, appropriation should not be reduced to an individual action led alone. Communities of practice benefit the development of individual artistic practices (Marquez-Borbon and Martinez-Avila, 2018; Marquez-Borbon, 2020). Through the multiplicity of individual perspectives, they foster deeper comprehension and appropriation, favoring the longevity of instruments and practice.

What is Exploration?

The word *exploration* encompasses different processes that are not formalized identically between practices. The french national center for textual and lexical resources proposes several meanings to *exploration*¹. Two in particular resonate with our research:

1. systematic study of a little-known or little-studied reality or object;
2. the action of apprehending something through the senses in order to know its qualities.

Exploration, therefore, encompasses actions aimed at discovering the unknown. The former definition refers to systematic actions to uncover objective knowledge about a little-known object. As such, it is abstracted from the individual, and the unknown lies in the holes in the current state of shared human knowledge about this object. In the latter, it is not objective knowledge that is sought through exploration but rather subjective appreciation of the perceived qualities of an object. In such a case, exploration is necessarily embodied, as it relies on the body’s perception to uncover qualities.

Exploration in HCI: Exploratory Learning and Navigation

The term *exploration* covers a variety of concepts in Human-Computer Interaction. This section focuses on a few works that are interested in studying exploratory behaviors in data exploration, software learning, and navigation.

When the first personal computers were introduced in the office in the early 1980s, gaining expertise in software such as the word processor involved several hours of classroom training, and thus, after every upgrade. As computers have become more and more commonplace in the office, the way workers learn to use them has changed fundamentally. When multiple software were used, each filled with many features with only a few used by a given user

¹<https://www.cnrtl.fr/definition/exploration>, visited on July 18th, 2024

Nielsen et al. (1986), classroom training was gradually replaced by other means of learning: interaction with other users, reference to the manual for specific tasks and trial and error (Rieman (1996)). The latter falls into what is called exploratory learning. Several studies in the nineties highlighted the value of exploratory learning and the evidence for its use in workplace contexts (Rieman (1996); Wiedenbeck and Zila (1997); Kitajima and Polson (1996); De Mul and Van Oostendorp (1996); Howes and Payne (1990); Trudel and Payne (1995)). Rieman defines exploratory learning as investigating the possibilities offered by a system without following “*precisely sequenced training materials*” (Rieman (1996)). Learning occurs without training or following a manual. This does not imply that the exploration is not task-driven but instead that users “*direct their own learning*” (Wiedenbeck and Zila (1997)).

Rieman differentiates the initial exploration to obtain a first overview of the software and the “*just-in-time*” task-driven exploration of specific functionalities, which he found to be favored by users. This assumes the existence of a target from which the error is calculated. This target corresponds to the resolution of a given problem. For instance, finding in a word processing software how to change a list numbering style without resorting to a manual or a web search is a form of exploratory learning that often involves navigating menus. In these workplace settings, “*exploratory learning for pleasure is rare*”. Other studies proposed similar distinctions, such as structured/unstructured exploration (Trudel and Payne, 1995) or guided/unguided exploration (De Mul and Van Oostendorp, 1996). As the primary motivation for exploration is productivity, and open-ended exploration is often found counterproductive, other works focused on the workplace to optimize the exploratory learning process in computer-human interactions (Trudel and Payne (1995)). In this context, exploration mainly refers to a trial-and-error process where the user decides what is tried, and the error is calculated as the distance from the result to the given goal. More recent exploratory learning literature focuses on exploratory learning environments, which are open-ended pedagogical software that allows students to solve problems through trial and error to acquire skills in a given field (Roll et al. (2010)).

The notion of navigation in HCI is closely related to exploration. Navigating an environment involves creating a cognitive map that will lead to decision-making when choosing a direction (Tolman, 1948). The environment can be real, such as a city, in which case navigation is called wayfinding. But it can also be virtual, as in an information system. In this second case, one of the navigation challenges is to help users find what they’re searching for, usually in extensive collections of data (Vainio (2010)). There are many ways to navigate in an environment. These include goal-directed navigation, where the movement in space is directed towards a target (for instance, a city monument or an online document), and exploratory navigation, where the movement is not directed towards a predefined target. Interestingly, Darken and Sibert showed that “*principles extracted from real-world navigation aids can be seen to apply in virtual environments*” (Darken and Sibert (1993)). Virtual environments designate both information systems and 3-dimensional virtual environments. Exploring the parameters of a sound synthesizer is an example of exploratory navigation in virtual environments. Several modern digital synthesizers comprise hundreds of parameters, making exploring outside pre-recorded presets slow and challenging to learn from. Scurto proposed to support this navigation through collaboration with an autonomous agent that simultaneously adapts the mapping on multiple parameters using reinforcement learning directed by user’s feedback (Scurto, 2019).

2.3 Chapter Summary

This Chapter introduces broad concepts and recent results on digital musical instruments (DMIs), embodied sonic interaction, sensorimotor learning, and exploratory learning. DMI design aims to enable expressive control of computer-generated sounds through various gestural interfaces. Building partly on those works, embodied sonic interactions generalize human actions' sonification to contexts outside traditional musical endeavors. This design practice acknowledges the experiential potential of sound that is often disregarded in interactive system design. In particular, movement sensors such as cameras or inertial measurement units enable whole-body interactions with artificially created sounds. Such interactions require the acquisition of audiomotor maps linking movements to sounds. Recent studies showed that learning movement-sound interactions through exploration was possible. As several DMI design research showed, exploration is one of the key processes of appropriation. However, exploration encompasses different processes and understanding in software learning, navigation, and sensorimotor learning. These observations call for specific studies with exploration as the central theme to understand how it relates to the learning and appropriation of movement-sound interactive systems.

CHAPTER 3

Interviews with Practitioners: Learning and Appropriating Movement-Sound Interactive Systems

This Chapter contains written material published in (Paredes et al., 2022)¹. The paper was a collaborative effort with Jules Françoise and Frederic Bevilacqua.

3.1 Introduction

The study of movement-sound interactive systems, that are designed and developed at the intersection of several disciplines, is one of the important areas of NIME research (Jensenius, 2022). First, these systems are conceptualized through embodied interaction and phenomenology, where mind and body are not separated (Leman, 2007). Second, the development of movement-sound interactive systems is also directly driven by technological advances. Precisely, the wide availability of motion-based technologies has opened new possibilities to either study musical gestures (Godøy and Leman, 2010), develop Digital Music Instruments (DMIs) (Jensenius, 2022), or experiment with gestures, movement and bodily awareness in dance (Françoise et al., 2017) Finally, these movement-sound interactive systems are also driven by artistic endeavors in music, dance, installations, which in turn directly influences how these systems are used and modified.

Interestingly, pioneers such as Michel Waisvisz, Laetitia Sonami, Atau Tanaka, and David

¹Paredes V., Françoise J., and Bevilacqua F.: Entangling Practice with Artistic and Educational Aims: Interviews on Technology-based Movement Sound Interactions. In *New Interfaces for Musical Expression (NIME)*, Auckland, New Zealand, June 2022.

Wessel, to name a few, developed and practiced over the years their own instruments/interfaces, even before the NIME conferences started (Jensenius and Lyons, 2017). Since then, many new approaches and interfaces have been proposed and evaluated at NIME conferences (Jorda, 2004; Rasamimanana et al., 2011). Moreover, as NIME is largely driven by academic research, it often focuses on “*new*” digital musical instruments rather than on singular practices over time. Understanding how artists build “*mastery*” of their digital instruments requires analyzing a complex network of questions, including skill learning, programming, improvisation and composition, and transmission to others. Several artists who first developed their once novel instruments pursued to develop their own expert practices. While their systems might have been built in collaboration with NIME researchers, the following years of practicing and learning, and sometimes restarting with different systems, might not have always been reported.

For this reason, we wanted to raise these questions and confront them with artists who have practiced with movement-sound interactive systems for several years. The goal of this research is to document through interviews how professionals from different backgrounds (including performers, composers, dancers, and educators) question their practices. Inspired by recent studies based on thematic analysis from semi-directed interviews (West et al., 2021; Robson et al., 2022), we contacted six different artists to reflect on the following questions: (1) which technological systems do they use and why; (2) how do they learn, experiment, and practice these systems; (3) how do they transmit such systems. Transmission is meant here as the act of disseminating and/or teaching knowledge and systems to peers or students.

The structure of the chapter is as follows. First, we recall some key related works, and we describe the method. The major part of the article consists in the description of three major themes that emerged through the interviews. The results are then discussed with regards to the design of movement-based interactions.

While user studies have been reported on notions such as usability, appropriation (Zappi and McPherson, 2014), or sensorimotor learning (van Vugt and Ostry, 2019), we wanted to confront these topics with the view of different practitioners using movement-sound interactions in creation. Reflexive Thematic Analysis, as described by Braun and Clarke (Braun and Clarke, 2006; Braun et al., 2019), is a relevant tool for theme extraction from qualitative data such as interview transcripts.

3.2 Method

Following the approach of Robson et al. (Robson et al., 2022) and West et al. (West et al., 2021), we conducted semi-directed interviews and thematic analysis to understand better how experts relate to questions of learning, practice, and relation to technology.

3.2.1 Participants

We contacted 6 artists who used and/or designed movement-sound interactive systems in their practice. All artists collaborated or were in contact with the last two authors, but not the first author who led the interviews. All artists accepted to participate in the study and gave their consent to the publication of the article. We led three interviews at the beginning of 2021 and the rest at the end of the year. We ensured gender balance and covered various practices: we interviewed 2 music performers (Mari Kimura and Philippe Spiesser), 2 dancers (Bertha Bermudez and Yves Candau), 1 teacher (Fabrice Guédy) and 1 composer (Michelle Agnes Magalhaes). Participants’ background are detailed in appendix A.1.

3.2.2 Study Design

This study aimed to investigate how practitioners engage in learning and appropriating technologies that mediate movement and sound in artistic or pedagogical contexts. The interviews were semi-structured around five topics with a set of questions that could be adapted to the practice of the interviewee: (1) background and motivation in using movement sensors; (2) the process of learning movement and/or sound; (3) the transmission of such systems; (4) technology appropriation (by performers and public); (5) limitations and perspectives. A set of general questions is given in Appendix A.1

Two interviews were conducted in person, and the rest were conducted over video conference. The interviews lasted between 90 and 120 minutes. They were automatically transcribed first using Microsoft Speech to Text service and then manually corrected using the audio recording.

3.2.3 Analysis

We decided to do a reflexive Thematic Analysis (TA) on the transcripts, as described by Braun and Clarke (Braun and Clarke, 2006; Braun et al., 2019) TA is a qualitative data analysis method, well established in the social sciences and psychology, that enables researchers to extract patterns of meaning called *themes* across a dataset. After familiarization with the data, it consists of a first phase of coding and a second phase of theme development. Reflexive TA differs from other TA techniques in that it acknowledges the subjectivity that lies in qualitative data interpretation. The theme development emerges from the interpretative work of the researcher made *from* the codes. The codes do not come from a codebook but are *generated* by the researcher when analyzing the data. The themes should not be defined prior to the coding and must emerge as patterns from those codes. Moreover, reflexive TA is “*theoretically flexible*” and is not bound to certain methodologies that were initially meant for psychotherapy research and psychology. It applies to a broad range of domains and more generally suits the exploration of participants’ subjective experiences.

All three contributors individually coded the transcripts and regrouped them under potential themes. We then converged together toward three main themes: (1) Practitioners consider technology as mediating practice, (2) they navigate usability and normativity, and (3) their learning process is entangled with the development of a personal practice.

3.3 Theme 1: Technology Mediates Practice

During the interview we explicitly asked the artists to position their work and approaches with respect to the use of *technologies*. We did not provide any definition of what technologies could refer to, and let them describe how they situate themselves with this large topic. We can describe their discourses by *what* technology is, *why* using it and *how*.

3.3.1 What is Technology, Why Use it and How?

First, they either use the term “*technology*” in a very general sense (which could be induced by the fact we use it in questions), or they refer to the specific tools they have been using in their work (such as the Kinect, smartphones, etc). They do not refer to specific categories as used in engineering fields such as motion capture, machine learning or AI, to cite examples of categories currently widely used. Interestingly, YC even considers “*A somewhat extended*

notion of technology. As one could say that language is a technology, that writing is a technology” (YC).

Importantly, technology is neither considered as a “*theme*” or “*goal*” in their practice, but rather as something embedded in a socio-cultural context, e.g. as “[*musical*] *instruments with today’s technology*” (PS). They all consider that integrating technology in a personal practice is a complex process that demands significant engagement to become a meaningful part of the practice as opposed to development made “*for the sake of technology*” (BB). By considering technologies an integral part of the practice, thus as a mean of expression, such an approach contrasts with other fields involving research and development which might prioritize novel technologies within a solutionist perspective.

Technology is therefore never thought of as a “*solution*” to a given problem, nor as a means to realize an artistic vision. It is rather seen as a beneficial “*perturbation*”, enabling them to “*move out of their comfort zone*”(PS), or something they have to deal with. In all cases, it should be grounded in a methodological approach that might need to be invented. The possibility of looking for systems failures or finding inspiration in bugs is also explicitly described.

I like when there are bugs, because it’s in these interstices that there are often opportunities to find things that you wouldn’t have found if everything had been done beforehand (FG).

3.3.2 Technology Can Act as a Mediator for Transmitting Concepts and Ideas

Most artists also considered interactive technologies as a support for transmission. The most prominent example is FG who uses, among other things, movement technologies to help teach music theory to young students. Using those technologies is a way to create bridges between tradition in music education and the world of computer science, and “*it is very clear that the de-compartmentalization that it produces in students is something that works.*” (FG). BB has never experienced movement interaction technologies as “*a tool for the process of creation to be enhanced. But for [her], it has been a tool really for transmission.*” A tool that she uses to initiate and teach non dancers movement practices. For BB, working with interactive systems requires producing a synthetic model of the practice, which benefits its transmission:

It helped me to understand [...] how to make it accessible. Because reducing, summarizing, also helped us change the transmission, develop a clearer terminology, be more conscious, really, of what we make. (BB)

3.3.3 The Mediating Role of Technology Might be Unclear for the Audience

Performances with movement technologies are still new to the general audience. YC, PS and FG reported that there was an effort to be made towards the understanding and/or acceptance of this interaction by part of the public. PS and YC pointed out that

There is this attitude, as a first response to a new system, of wanting to understand what is going on. (YC)

The technology becomes central in the performance, overshadowing the artistic proposition. For MK, however: “*if you have the motion sensor that is moving, people can actually see and relate to the electronic sounds better.*” In this perspective, the interaction actually links the audience to the produced sound. But further work is needed for spectators to adapt to

the interaction, particularly with non tangible interfaces, where the underlying link between movement and sound is blurred :

That is, the smallest gesture you make becomes an identifiable and interpretable musical gesture for an audience. (PS)

3.4 Theme 2: Navigating Usability, Normativity and Obsolescence

As interviews revolved around technology, interviewees offered insights regarding the usability of interactive systems. In particular, the potential and ease of use of some technologies and feedback modalities over others was discussed. Critically, the question of the normative power of technologies over their users was a recurring theme of the interviews, as well as issues of obsolescence.

3.4.1 Technology Requires Commitment and Friction

Most participants have a history of practice that involves the use of wearable sensors embedding IMUs. They are appreciated because they are easy to set up, are minimally invasive and have a low latency. In particular, mobile phones greatly facilitate setup. While usability is appreciated, interviewees still emphasized that working with technology takes time and requires dedication, engagement, and commitment. Reflecting on a decade of experience, MK reported:

I've practiced well enough using just one thing. [...] But in order to refine your movement and try to incorporate that into a tool that's useful for musical expression, you really need to spend time with it. (MK)

Issues such as calibration and parameterization have a learning curve but are critical for both performance and transmission. All interviewees emphasized a desire to escape a triggering paradigm found limited and often associated with a “*gadget*”. Yet, apprehending more complex forms of interaction can be difficult. For PS, as complexity increases, “*you can get overwhelmed very quickly because you become a human synthesizer and so all you do is sound.*” Apprehending such increasing complexity is not easy, and interpreters need to learn to play with the subtleties of movement. Yet, interviewees often valued the friction that occurs over the development of a practice:

Because the tool resists, the user will persist a little bit in doing it and that's where he'll find things that I find interesting. Typically, in the gesture tracking, [...] it wouldn't have worked if it had worked right away. It's a paradoxical way of putting it, but I think that's what can make you think bigger. (FG)

3.4.2 Normative Technologies should be Avoided

Practitioners value friction — to some extent — as a mechanism going against potentially normative effects of technology. In creative practices, technologies should not be designed to solve problems and facilitate use to the extent of “*infantilization*”. What motivates FG is “*this interweaving*”:

We are not going to use ready-made technology. [...] What is interesting is to associate it from the start, to think about several disciplines at the same time, which will then be mixed together. You're not going to be interested in

transmitting a technological brick, but really more the process of reflection behind the creation. (FG)

Several interviewees reflected on the history of technology in art, drawing comparisons between early experiments with technologies that aimed to broaden the scope of possibilities and some contemporary technologies that have become normative. Most interviewees make efforts to foster creativity and personality in their transmission of the technologies rather than a form of imitation and replication:

We do our best to reproduce the best performance but as soon as we think like that, everything is screwed up because we will put ourselves in a mechanical mode. [...] And as soon as we are there, focused on the result, we are no longer in the present. (MAM)

In creation and pedagogy, technology therefore aims to emancipate students and support the development of a personal practice rather than help them reproduce an “*ideal*” result.

3.4.3 Obsolescence

The notion of obsolescence consistently emerged throughout discussions with all practitioners, even though the researchers did not introduce it as part of the interview structure. With the rapid pace of technological innovation, interactive systems can rapidly become outdated and unusable. Such obsolescence creates tensions among artists with regards to the preservation of creations.

Ligeti didn't want to go into interactive because he thought that the technology will be obsolete, and he's correct, that if you write in a certain way and if that technology dies, your piece dies. (MK)

For MK, such obsolescence is not problematic because her works, her creations will survive. Yet, obsolescence can make transmission hard, as platform updates might prevent students from using the software. For BB, what matters is the persistence of methods and knowledge over time rather than the technology itself:

The installation, maybe if we put it back it doesn't work anymore. It's not a maybe, it is the reality. [...] That's a problem, you could say, it's really a problem because the tool itself is obsolete. But not the creation process. What did we do during the creation? That's not obsolete, because we're still here. So we have to talk about that, we have to leave a trace because I think it's important, I think it's quite unique. (BB)

3.5 Theme 3: Learning and the Development of a Personal Practice

Learning was one of the key themes introduced in the interviews. Our analysis revealed key learning processes and insights into how technologies can be designed to structure movement learning. Importantly, interviewees emphasized that learning is highly intertwined with the development of a personal practice.

3.5.1 Discovering, Experimenting, and Challenging Limits

Approaching a new form of interaction involves several phases of discovery, exploration, and integration. When experimenting with a new interactive system, people tend to rely on their existing skills:

So I think in a learning system, you always start with the things you know first and try to apply them to a new environment. I tried to look for those gestures that I knew. For example, in this virtual instrument system, I imagined hitting a skin, I imagined hitting a block, shaking the maracas, doing a deathstroke. And so I tried to reproduce these gestures, I started from there. (PS)

In this early stage, the quality of attention is essential: the mover must remain open to the feedback provided by the system to get acquainted with a new action-perception loop. This involves moving while listening to the system's responses, as emphasized by MK: *"You know hand movements or body movements, but then use their movement to incorporate the sensors and the first thing I usually do with the students is that I have them wear it right now [...] and see if there's any relevant thing that you can use as an expression"*. Often, as previously mentioned, this exploration is initially driven by the desire to understand the interaction at play.

Naturally, all interviewees emphasized the need for experimentation and exploration. FG, describing learning in a pedagogical context: *"At the workshop, we try without having a clear goal, but we know that it is still interesting to try. [...] In fact, the desire itself often refers to something that we want to see, like a child who, even if he doesn't know what the toy is for, will try to throw it"*. Whether practicing for a choreographed or improvised outcome, many participants mentioned improvisation as a key marker in the learning process. Improvisation can be seen both as a means to explore the capabilities of an artifact and a form of challenge to oneself, which assesses one's ability to interact and express oneself with the system. Such experimentation involves exploration and exploitation, and repetition is key in acquiring embodied skills: *"Through the repetition of the same gesture, you can learn, and that's exactly what we do. So you can learn from repetition, that's what you do in a dance studio."* (BB). Experimentation is considered essential to explore the instrument but also as a process that extends beyond the use of the technology itself:

It's the very activity of coding that is interesting, to think in a certain way. Afterwards, we go back to the piano, to the violin, but we have understood things differently, and we have explored a whole space. That's what it allows you to do, to experiment and to try, but not with the aim of necessarily making an application. (FG)

Once the initial phase of discovery leads to an understanding of the interaction, performers often challenge the system, exploring its limits to understand where it breaks.

One interesting thing is that we always try to do two things: one, we understand the system, so we do it, we are obedient, [...] and then we try to break the system, and see when it doesn't work. [...] Almost all the dancers did the same. (BB)

Finding limits involves exploring the potential and shortcomings of the technology, but it is also seen as a process of one's exploration of their own abilities and habits. Technology *"helps me to understand the material better and to get out of a habitus that normally I would have in the studio"* (BB).

3.5.2 Technologies can be Designed to Facilitate Learning

The way interactive systems are designed can affect their learnability in the context of open-ended learning processes that integrate with a personal practice. Several factors that could inform the design of learnable interactions emerged through the interviews. First, as highlighted above, constraints provide valuable stimulation for learning, for they drive attentional processes and challenge the performer's habits. Systems can be designed to stimulate movement through metaphors and images or more implicitly so that movement emerges through interaction.

The identification of sound-movement [...] will be learned by the body, but I don't need to verbalize it. He will understand if he turns it like this that it won't work because it's not the right sound. But I don't need to write a detailed instruction for that because the own object, the device itself will indicate that it's not in the right direction. It's also about trusting the body. The body is not stupid. (MAM)

For PS, interactive systems can be designed for learning by relying on play: *"[a child] will learn gestures very quickly, by himself. So if we can frame it, find an interaction, it's a bit like video games. You go through different stages."* However, several interviewees mentioned the need for adaptability so that *"you make your own path"* (PS). It is important to find a balance between *"a technology that adapts to the interpreter"*, and the need for interpreters to adapt to, learn from and appropriate the technology.

It's not a tool that I could pass on to anyone. It's a technology, it's a way of seeing things. After that, it's up to you to adapt it, to imagine an interesting use. (FG)

3.5.3 Learning is Intertwined with the Development of Personal Practice

Reflecting on the details of one's learning process seems to be challenging, because, contrary to other fields involving movement learning, such as sports, that are task-oriented, expression is often the primary purpose. Learning appears to be inseparable from the development of a personal practice: quoting BB, *"I think that at the level of movement acquisition or understanding it was not about the form or the execution, but rather about the personal practice, how you work."* Depending on the focus of the interviewee, technology can mediate with the theoretical underpinnings of the practice, can foster movement practice and kinaesthetic awareness, or can support creativity. When technology is used to support an embodied practice such as dance, its reflective qualities can help mediate embodied experiences, helping the mover become aware of their body, of particular movements, of felt experiences:

It helps to be conscious of what you do [...] because you have a response, the movement becomes almost tactile. It becomes tangible. The movement has a trace, either in sound or in images, that reflects what you just did. (BB)

As such, the learning process is not merely about gaining pure control over the instrument, but about developing skills within a personal practice by drawing upon the shifts in the lived experience induced by an external feedback. Sound feedback brings attention to the movement from another modality, therefore structuring experience.

Technology can structure learning but can also structure experience. There's something a little more marked, with fewer dimensions, and that allows you to channel the experience. (YC)

Technology can therefore shape movement, but it affects perception more broadly, quoting MAM: “*what I really want is for musicians to come out of an experience like that transformed. That they can act on their field of perception*”. As a result, most interviewees view their interactions with technology beyond the pure control of an instrument as a reflective and conversational process. Technology is considered beyond a tool or an instrument, as “*a partner, who helps me to think about what I am doing*” (BB).

What characterizes the entanglement of learning and practice development is the search for a compromise between freedom and constraints. The use of extrinsic feedback on movement brings constraints that are found stimulating for learning, as a way to structure attention because “*There is a big difference when it’s a personal choice and when it’s a choice that comes from outside. And also the personal choices, often, happen more automatically. So there’s really a gain in having choices that are made from the outside*” (YC). For YC, changes and ruptures of the interaction over time were found particularly stimulating:

I would say that the things that are very strong in the experience are moments of surprise, for example where there is something established that I can interact with, and I develop some sense of what can be done, and then all of a sudden there is something else that manifests itself, that appears in response to something that I am doing, but that was not there before. (YC)

Such external constraints drive performers out of common patterns and habits. Yet, artists use technology to support an expressive practice, and all interviewees expressed the need for a certain degree of freedom: “*it was important to find myself as an interpreter. I trigger things, but I also have the possibility to be free between some triggers, to feel breaths*” (PS). Such freedom is necessary so that interaction lets personality and idiosyncrasies emerge among different interpreters: “*Obviously the piece is totally different with the same parameters. That’s what they can influence, they can put their personality without having to change the system exactly*” (PS).

3.6 Discussions

3.6.1 Expressivity in Question

An important question raised by the NIME community, at least at its start, concerns the notion of “expressivity” — which is included in the NIME acronym (Dobrian and Koppelman, 2006; Gurevich and Treviño, 2007). The use of embodied interaction involving movements and the body for music performance certainly questions how musical expression can be conveyed to the public. Interestingly, even when explicitly questioned, the notion of expressivity was regarded as somehow ill-defined or not considered an important topic to discuss frontally. Therefore, it would seem that the use of gestures is not meant to enhance or facilitate expressivity, as one could naively believe.

The notion of “*expressivity*” appeared indirectly when the interviewees, especially the ones concerned with music performance, did mention spontaneously that the public reception could be problematic. The “*expressive intention*” behind the movements and gestures might not always be perceived by the audience, especially if they are unfamiliar with such approaches of technology. “*How should spectators experience a performer’s interaction with a computer?*” (Reeves et al., 2005), Reeves et al. questioned nearly two decades ago. The reactions of part of the audience raise a more general question: who are sound-movement interaction systems designed for? It seems that they grow from the need to push certain

boundaries in one's practice and therefore help interact differently with sound or gesture, the finality remaining the same: nurturing creativity. So, technology should not be the center of attention when transitioning from the studio to the stage. The use of new technologies on stage requires thinking about the perception of the performance by the audience.

To face this challenge, Capra et al. proposed to implement spectator experience augmentation techniques (SEAT), that help improve the audience experience, as shown by Capra et al. (Capra et al., 2020). Similarly, Berthaut et al. proposed a visualization of the internal mechanism of the DMI with 3-D shapes to amplify musicians gestures to the audience. As noted by Robson et al., "*the word 'audience' suggests a passive role, whereby you are rewarded with culture simply by virtue of turning up*" (Robson et al., 2022). Robson et al. have interviewed sound artists who are engaged in situated sonic practices. One of the interviewees noted that "[Visitors] need to do something to bring the whole installation together, move around and fetch different parts." ². Similar to sound installations where receivers must actively make sense of the piece by changing their spatial relationship to it, it would be beneficial in the process of "*restoring trust with the spectators*" (Capra et al., 2020) to help get them from being part of a passive audience to active receivers, not of a technological demonstration, but of an artistic proposition.

3.6.2 Questioning the role of technology

In all interviews, the view of technology is ambivalent, navigating between normativity and emancipation. While it is considered central to current artistic practices, as part of a global socioeconomic system, it is also seen as a "*tool*" to experiment and create beyond norms. This creates an interesting paradox, since the need for experimentation might go against the establishment of standards that are necessary for community building and transmission. Moreover, obsolescence is unavoidable, based on evidence of using technologies for several years, and yet the general concepts that are developed are intended to resist technological evolution. This brings important questions about transmission and pedagogy, some of which have been brought up by several researchers (Zayas-Garin et al., 2021; Pigrem and McPherson, 2018). We believe that it opens a large and interesting debate, since this necessitates to establish what in our practices is truly independent from technological artifacts.

3.6.3 Supporting Learning

It appears that pedagogy inherited from standard music practices does not encompass the variety of musical practices that exists today. Particularly in improvisation practices, as shown by Hayes (2019), the notion of graduating from novices to expert with a predefined educational path does not apply, as it puts the skillful musical instrument expertise before the value of "*the instantiation of multiple sensitivities of the person as a whole*" (Hayes, 2019). The role of interactive systems is fundamentally perceived as elements that should be sufficiently modular, and assembled in different manners. As such, interviews highlighted that systems can be designed to facilitate learning, through metaphors, images, play, and interactive feedback supporting movement execution. This can be linked to the use of different types of affordances as suggested by Altavilla et al. (2013). Because such systems alter the action-perception loop, they bring external constraints that structure the lived movement experience and stimulate attention and kinesthetic awareness. Yet, designing systems in a way that eases the learning process involves several challenges. Crafting the right level

²Quote from the interview of Roswitha von den Driesch and Jens-Uwe Dyffort, Sound artists from Berlin, Germany.

of detail for movement-sound interactions can hardly be optimized for all learners because the learning process is open-ended and intertwined with the user's own personal practice. As such, interviews show that performers often co-evolve with the system, progressively uncovering subtleties in movement expression. Designing such adaptive systems seems promising but requires delineating what should change and what should remain consistent within a learning process that is always personal.

3.7 Chapter Summary

This chapter investigates how practitioners of movement-sound interactive systems relate to questions of learning and appropriation with technological systems. We conducted interviews with six artists: one performer, one performer/composer, one composer, one teacher/composer, one dancer/teacher, and one dancer. The interviews we conducted were fruitful in bringing a different light to NIME research about movement-sound interactive systems. As discussed by other researchers in this field, it is intrinsically difficult, or maybe fundamentally impossible, to translate global methods borrowed from traditional music practices in composition and pedagogy to practices based on interactive technology. Nevertheless, the insights given by the artists still reveal how important it is to privilege strong concepts over detailed implementation, the importance of openness and modularity of the system, and how constraints and perturbations can be fruitful in artistic contexts. Finally, they expressed that experimentation and exploration through improvisation are essential to appropriating interactive technologies.

CHAPTER 4

Shaping 2-D Sound Spaces for Movement-Sound Exploration

This Chapter (from section 4.3) contains written material published in (Paredes et al., 2023)¹. Jules Françoise and Frederic Bevilacqua collaborated on the paper.

4.1 Introduction

Interviews highlighted the importance of exploration and experimentation in learning movement-sound interactive systems. Studies in audiomotor map acquisition by Van Vugt and Ostri (van Vugt and Ostry, 2018b, 2019) showed the possibility of learning movement-sound spaces through exploration. We are interested in bridging these controlled studies with our team's ecosystem of artistic practices and the one described in Chapter 3.

This Chapter proposes a prototype for movement-sound interaction that enables the exploration of a 2-D sound space created with the concatenation-based sound synthesis technique *CataRT*. By *2-dimensional sound space*, we refer to a sound synthesizer with two parameters. In the case of a space created with *CataRT*, the two dimensions correspond to the sound descriptors used to project the sound grains on a 2-D plane. Sound spaces produced with these techniques are traditionally explored with fingertip interaction on 2-D touch surfaces (Schwarz et al., 2020). We propose to control this synthesis with movement sensors strapped to the arm, enabling whole-body interaction with a corpus of sounds. In this Chapter, we investigate the early learning of this system through exploration and identify

¹Paredes V., Bevilacqua F., and Françoise J.; Polyspring: A Python toolbox to manipulate 2-D sound database representations. In 20th Sound and Music Computing Conference (SMC 2023), Stockholm, Sweden, June 2023. Zenodo. doi: 10.5281/zenodo.8399016

that the inhomogeneity of the sound space hinders its exploration. We then propose a tool for manipulating the distribution of sound grains over the 2-D sound spaces.

We presented an empirical study focused on movement-sound exploration. Following the preliminary experiment reported in Chapter 4, we were interested in quantifying and qualifying the exploration of movement-sound spaces in exploratory learning contexts. Our first contribution is to propose a movement-sound interactive system called *mosX* that uses an abstract audio space controlled with wearable sensors. We further established an analysis method using kernel density estimation (KDE) to visualize exploration data. We proposed an empirical method to evaluate the participants' internal representation of audiomotor maps based on their drawings or sketches. This experiment revealed significant insights into how individuals learn, represent, and navigate complex sound spaces through embodied interaction. The findings underscore the effectiveness of exploration-based learning in early audiomotor map acquisition. In particular, our KDE-based method produced quantitative data and visualizations that enabled the identification of different exploratory behaviors between participants.

With data recording capabilities, the prototype enables us to conduct experiments on audiomotor map learning in different conditions. First, we report on a preliminary experiment with six expert users, which gives us initial insight into the possibilities of learning such a space by exploring it in just a few minutes. In particular, we investigate **(1)** to what extent exploration of a movement-sound interaction can lead to early learning of the specific association between the gesture and the sound, using either a 2D flat interface or more complex body movements, and **(2)** how exploration is expressed by expert participants using both words and drawing. This experiment shows that learning the mapping between movement and sound with whole-body interactions was particularly difficult for participants, partly due to the inhomogeneity of sound distribution over the 2-D sound space.

Following these results, we propose a method that enables the control of the points distribution, from homogenization to custom density function, and the setting of adaptable border geometries. The method presented here is based on the *unispring* algorithm introduced by Lallemand and Schwarz (2011), with improvements at both theoretical and practical levels. First, we present several theoretical contributions that simplify the use of the base physical model and facilitate the real-time manipulation of the sound map structure. In particular, we propose to estimate a key parameter from the dataset to provide users with few function arguments, allowing them to use this tool without knowledge of the underlying physical models. Additionally, we formalized a simpler density manipulation technique using Gaussian mixtures as inputs, which allows users to manipulate the distribution of the sound clouds in real time. Second, we provide the community with an open-source Python package called *polyspring*. It enables the control of a sound dataset distribution starting from any type of 2D projection. It can distribute the dataset uniformly while preserving neighborhood relations between points. A user-defined density function can further shape the final distribution of sounds. This distribution can also be restricted to custom boundaries. A Python script and a Max patch communicating over OSC are also provided to manipulate a sound dataset in the Max package *Mubu*.²

²<https://forum.ircam.fr/projects/detail/mubu/>

4.2 Exploring 2-D Sound Spaces with Movements: First Prototype and Experiment

4.2.1 First Prototype

The system designed for this study is based on corpus-based concatenative synthesis techniques. Those methods rely on analyzing and recombining databases of recorded sounds. Such methods involve the segmentation of audio recordings, their parametrization (for example, with audio descriptors), and their resynthesis using selection methods. Various implementations have been proposed, such as the real-time corpus-based concatenative synthesis system CataRT (Schwarz et al., 2006), used by several research groups and musicians (Einbond et al., 2016; Zbyszyński et al., 2021; Schwarz, 2007), the tools developed in the Fluid Corpus Manipulation project (FluCoMa) (Tremblay et al., 2019, 2021), earGram (Bernardes, 2021) or the AudioStellar system (Garber et al., 2020). 2D representation of the sound parametrization greatly facilitates its direct manipulation by the performer through projection or dimensionality reduction from the possibly high dimension of the sound description. The relationship between the user gestural input and the selection of such 2-D reduced space can be formalized as a special case of the so-called gesture-sound mapping (Hunt et al., 2000). This synthesis technique (and its variants) is typically used with a 2D projection of the multi-dimensional sound descriptor space on a screen. It allows for easily selecting sound segments with a mouse or a touchpad, as shown in figure 4.1. Other types of feedback have been proposed, such as haptic feedback (Magalhães et al., 2020).

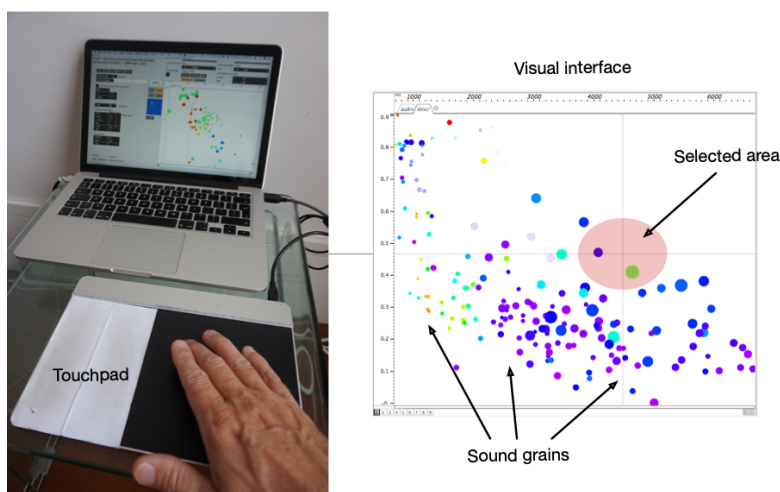


Figure 4.1: 2-dimensional touch interaction with a sound space created with CataRT.

Sound Synthesis

Sound spaces were designed and implemented with *CataRT*,³ a corpus-based concatenative synthesis system Schwarz et al. (2006), freely available for Max (Cycling'74) with the MuBu package⁴ Schnell et al. (2009). It allows for the exploration of a large corpus of recorded sounds. Recordings are segmented in so-called *sound grains* (typically 500ms) and analyzed using *audio descriptors* Peeters (2004). In our case, for each grain, the average value of two descriptors is computed:

³<https://forum.ircam.fr/projects/detail/CataRT-mubu>

⁴MuBu package: <https://forum.ircam.fr/projects/detail/mubu/>

- **x**: *spectral centroid* barycenter of the frequency spectrum, related to sound perception with low-high brilliance.
- **y**: *periodicity*, the ratio between the energy of the harmonic content (harmonic partials) to the total energy. Periodicity is typically high for sounds from musical instruments (with defined pitch), while it is generally low for “noisy” sounds. It is calculated using the YIN algorithm (De Cheveigné and Kawahara, 2002).

We chose the combination of those two descriptors because they are generally weakly correlated. Consequently, 2-D spaces created with these descriptors tend to display better spreading of sound grains.

Sound Corpora

Two different sound spaces were specifically designed for this experiment. Each includes six to eight clusters of sounds using recordings of a typical object–surface interactions, such as rubbing, hitting, shaking, or rolling different metallic, ceramic, glass, or wooden objects. The segmentation and analysis of the sounds from a cluster produce a spatially cohesive group of sound grains in the audio space. Some outlier grains were removed to facilitate discrimination between the different clusters. Corpora are represented on the right-hand side of the figure 4.2.

Interfaces

Two different interfaces could be used interchangeably with the sound synthesis system:

2D Tablet : The first interface considered for this experiment is a touch surface, the Sensel Morph. Although it allows for multi-touch control and measures touch force and area, we only used a single-touch position and scaled its coordinates to the 2-D sound space. The surface is $138 \times 240mm$, but the actual playable area is a centered $130 \times 130mm$ square directly mapped to the square containing the grains in the sound space, as displayed in figure 4.2. The rest of the surface is not used. At the center of the tablet, a circular sticker helps users identify the center of the space.

Whole-body interaction using IMU sensors : the second interface consists of two inertial measurement units (IMUs) attached to the dominant arm. One to the forearm, just before the wrist, and one to the arm, just before the elbow. They measure the angle between the vertical axis and the arm and forearm, respectively. The arm angle is mapped to the x-axis of the sound space, and the forearm angle is mapped to the y-axis as shown in figure 4.2. The data is sent from the IMUs to the computer via Open Sound Control (OSC) streaming over WiFi.

In both cases, the clusters are distributed within a square in the audio space (see figure 4.2). At the center of every corpus is a circular area where no grain is played. Instead, the user hears a reference sound. This sound is a continuous synthetic sound that contrasts with the rest of the corpus. Its purpose is to help identify the center of the space.

Notably, by design, the whole-body interaction case cannot be reduced to the planar case in 3D. The space defined by the two angles in the whole-body interaction implies large movements of both the upper and lower arms. Even if the sensors were solely placed on one arm, the term “*whole-body interaction*” was chosen here since this mode implies large arm movements that impact the whole body. It requires participants to adjust their balance and indirectly implies movements of other body parts (this is in contrast with the tablet where the

participants are seated).

Blindfold Interaction

The experimental set-up is designed to be explored blindfolded. CataRT is traditionally used with a screen (see figure 4.1) so that sound grains can be explored by pointing. We are interested in gestural learning and, in particular, sensorimotor learning through proprioception. The facilitation of gestural learning through visual feedback has already been addressed (Jeanne et al., 2017). In our case, sound feedback is not an aid but one of the modalities that must be learned in conjunction with movement. Blindfold exploration supports this objective.

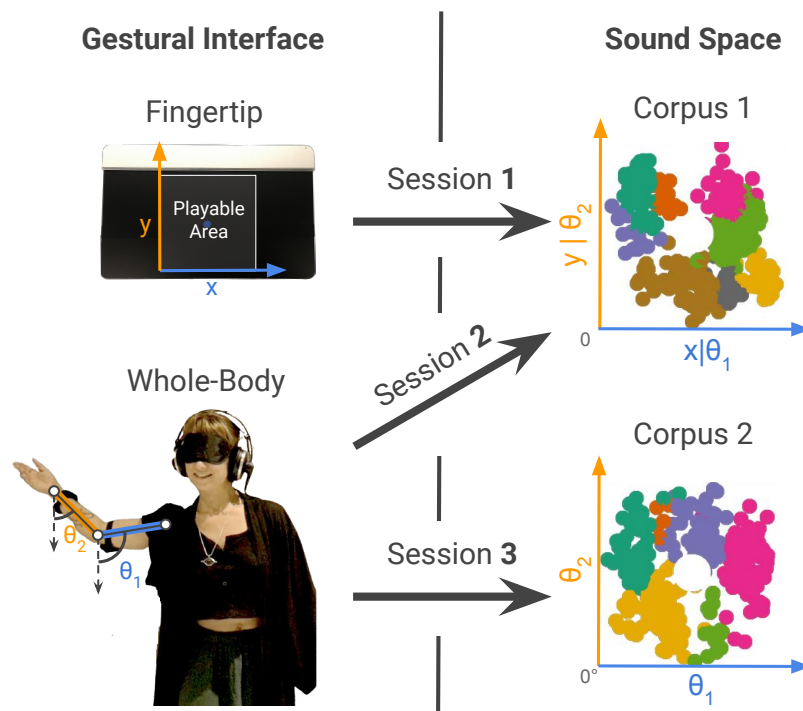


Figure 4.2: Schematic overview of the experimental setup. The parameters from movement interfaces (left) are mapped to two dimensions of a sound space (right) that uses corpus-based concatenative synthesis. The user position determines which grain is played. The first interface is a touch surface, the Sensel Morph. The second interface consists of two IMUs placed on the arm and the forearm. They measure the angle between the vertical axis and the arm and forearm, respectively. The positions received from the tablet or the angles from the IMUs are mapped to the two dimensions of the sound space.

4.2.2 Experiments

Participants

We recruited 6 participants (3 females and 3 males) with expertise in one field related to the study: either dance, electronic music, or sound design. The two dancers (coded DA1/2) were more likely to explore beyond everyday movements and give more informed feedback on the movement space. The 2 sound designers (coded SD1/2) had expertise in sound perception and classification and concatenative sound synthesis. The 2 musicians (coded MU1/2) were trained in electronic practices and had experience with granular synthesis. Participants had

no prior knowledge of the experimental setup. They all signed a consent form to participate in the research experiment.

Method

We purposefully designed the system without visual feedback, and participants were blindfolded. In particular, they could not see their arm and, therefore, could not ensure pose reproduction by looking at their body. Consequently, they could rely only on the sound feedback and on their proprioception. We hypothesize that users create a mental representation during exploration to help memorize sound-movement association. We elicited sketches or drawings to complement interviews to get insights into such participants' cognitive representation of the sound corpora and the interaction. After exploration, they were given a squared 18cm wide paper to draw their representation of the audiomotor space using a pen or pencil.

Procedure

The experiment consists of three consecutive learning and trial sessions with 2 interfaces (tablet or IMUs) and 2 corpora: **1.** corpus 1 with tablet, **2.** corpus 1 with IMUs, and **3.** corpus 2 with IMUs, as shown in figure 4.2.

Each session was separated in three different phases, explained below:

Exploration: Participants explore one corpus freely for 5 minutes with either the touch surface or the IMUs to learn the sound/position relationship;

Target Task: The system plays A pre-recorded movement trajectory twice, and the participants hear the corresponding sounds. They must then reproduce the sounds played as best as possible in one attempt. Each task consists of 6 different trajectories. A trajectory must travel through 2 clusters;

Drawing and Interviews: Participants explain their understanding of the relationship between sounds and movements using sketches/drawings and talking aloud.

Analysis

We recorded the trajectories in the descriptors plane during the exploration and the target task. These were examined along with the participants' drawings. Interviews were transcribed from video recordings and analyzed in relation to their sketches/drawings and exploration.

4.2.3 Results and Discussion

Quantitative Analysis

We computed several metrics on the target task results to assess their learning of the movement-sound space: the duration of trajectories performed by participants and their length (distance traveled in the audio space). Second, we computed a score measuring whether participants found the target sound clusters during the tasks.

The trajectory length in the audio space is higher in sessions 2 and 3 compared to session 1 (1.m=0.85 std=0.11, 2.m=1.77 std=0.25, 3.m=1.48 std=0.21). This means that with whole-body interaction, the trajectory between target clusters is less direct in the audio space compared to the tablet, which was expected. Nevertheless, we observed no significant difference across sessions regarding average trajectory duration.

A score was computed to evaluate if the performed trajectories reached the targeted sound clusters. It is averaged across participants and trajectories and reported in figure 4.3. As expected, participants obtained the best scores with the tablet ($m=1.55$, $std=0.26$). Nevertheless, the score for body interaction shows that, generally, at least one cluster was reached. Interestingly, the average score progressively increases between sessions 2 ($m=1.1$, $std=0.23$) and 3 ($m=1.27$, $std=0.26$), showing that participants seem to improve over the session, even if the corpus changed between these sessions. This trend, to be confirmed by further experiments, is encouraging since it shows that participants can learn or adapt to the system.

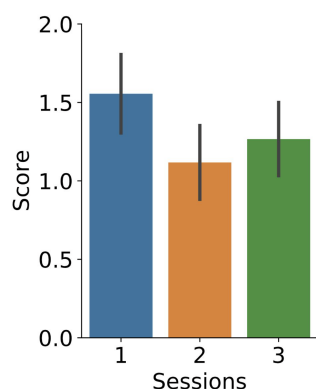


Figure 4.3: Average score measuring the number of target sound clusters reached for one trajectory

Qualitative Analysis

Touch Surface Control During the first session (corpus 1, touch surface), all participants succeeded in associating most parts of the touch surface with sound textures. Since they knew, before the first exploration, that the target trajectories all started from the center point, they explored the space in star-shaped trajectories, going back and forth between the center and the extremities: *“I think it’s very starry as a distribution, you have the center, and you have the little branches, that’s how I visualized it.”* (MU1). Examples of such exploration trajectories are shown in figure 4.4. Participants’ drawings are consistent with the patterns of explorations. The number of clusters participants identified is consistent with the corpus (8 sound clusters and 1 silent area). However, the various groups of sound are represented with approximately the same size. Sometimes, they cannot remember which sound is associated with a cluster, even though they recognize its existence and position.

Transferring from a planar to a joint angle representation With this task, we wanted to investigate how participants could learn a 2-D movement-sound space mapped to a joint angle representation to examine how they manage an interaction change while keeping the same sounds. Participants explored the same corpus in the first and second sessions. In this case, the sound space’s x-axis (respectively y-axis) was mapped to the angle between their dominant arm (respectively forearm) and the vertical. As expected, all participants found the transition from tactile to body interface difficult, even if they already knew that sound space. Interestingly, it seems that the audiomotor space needed to be relearned. Participants explained they could not transfer the relationships between clusters they identified in the first session. Although the frontiers were the same, the best they could do was associate a position with a sound: *“I had specific places that had names. But I did not have enough time to get the ‘in between’, to get how they float into each other”* (DA1).

a specific methodology to the exploration: “*I lost a little bit of the nuances. I was thinking ‘what sound do I want to make? How do I make it?’ and I tested the big areas to see if I could repeat them.*” (DA1). After a second exploration with the body interaction, they could better explain exploration processes using such interfaces: “*I repeated the same gesture several times, but it required such precision that I could not always do the same movement twice. And, sometimes, I discovered new sounds that way, and I could not tell if I discovered them with my hand, with my elbow, or with the angle of my arm.*” (DA2).

In this case, we found that participants used idiosyncratic movements to explore the whole audiomotor space. This led to certain parts of the audio space not being visited during exploration, as illustrated by trajectories in figure 4.5. One of the participants commented: “*With the tablet, after a minute or two I had identified the zones and so I could spend time on the sounds, but now I don’t have the feeling that I have finished exploring. I don’t know the limits of my body in space, I don’t know if I have explored all the gestures.*” (MU2). Importantly, the appropriation of such interaction modality depends strongly on the somatic expertise of the users. We observed that the two dancers implied significantly more of their whole body during the arm-movement interaction compared to the other sound experts. Even if moving the arm could enable one to reach all sound clusters, we believe such interaction might be more fruitful if the body is involved, as discussed by Candau et al. Candau et al. (2017).

4.2.4 Limitations and Area for Improvement

We were interested in evaluating early learning by exploration of our movement-sound interaction prototype using either a touch-based or a movement-based interface. Overall, this first study allowed us to validate the setup from a technical point of view. We found that the transfer from the tablet to the body space is limited and might not be investigated further. Nevertheless, repeated trials with whole-body interaction led to improvement, but several limitations seemed to hinder exploration. The corpora used in this experiment were built to facilitate exploration: as little empty space as possible with clearly delimited clusters of approximately the same size. Even with these objectives in mind, the resulting corpora were not entirely satisfying. Although each cluster had the same number of grains, they did not occupy the same area. In addition, some zones in corpora presented few to no sound grains. With *CataRT*, it is possible to move sound grains manually, but it is difficult to fill empty spaces without breaking cluster spatial coherence. Moreover, some participants did not reach all parts of the sound space because of the unusual body configurations required. One solution would be to adapt the sound space to the shape of the space they explored.

These observations led us to design a tool for manipulating the sound grain distribution using a spring-mass physical model called *polyspring*. The following sections present the tool’s theoretical background, a description of the toolbox features, and some examples of use.

4.3 Manipulating a 2-D Point Set Distribution: Theoretical Background

The preliminary experiment presented above highlighted the need for an easy way to manipulate the distribution of sound grains in *CataRT*. The problem is formulated as follows: given a set of points randomly distributed in a 2-D space with closed boundaries, find a transformation to those points that preserve their initial neighboring relationship while spreading them evenly

in the space. An illustration of the initial problem with the expected results of the solution is given in figure 4.6.

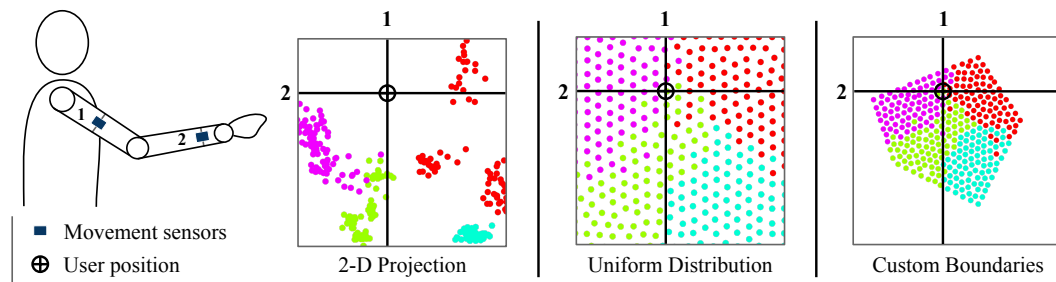


Figure 4.6: Exploration of a sound database using motion sensors mapped to the control space. The 2-D audio space might present empty space and tight clusters compared to the uniformly distributed space. The control space can be adapted to the user’s motion range using a region with custom boundaries.

The toolbox we designed to answer this problem, *polyspring*, builds upon several algorithms and methods to manipulate a set of points in a 2-dimensional plane. After introducing the base *unispring* algorithm initially proposed by Lallemand and Schwarz (Lallemand and Schwarz, 2011), we present several modifications and additions we made to this algorithm. Then, we describe a cost-effective interactive method to shape the density of the distribution.

4.3.1 Related Works

Gerard Roma and colleagues made several contributions concerning the visualization and mapping of sound collections. In Roma et al. (2019), they present a novel framework for sound space creation for interaction. Features are extracted from the sound corpus using a neural encoder that produces high-dimensionality features. A segmentation algorithm chops initial files into sound fragments based on a novelty algorithm. Eventually, each fragment is projected in a low dimensional space (2-D or 3-D) using a dimensionality reduction technique. The playability highly depends on the visual characteristics of the 2-D representation obtained from this last step. For Isomap and the Fruchterman-Reingold algorithm, the authors highlight a trade-off between perceptually meaningful clusters and good spatial occupation when dealing with the number of neighbors. They also experimented with Self-organizing Maps (SOMs), which is a mapping technique between high-dimensional data and a 2-D grid layout. Still, SOMs can assign multiple data points to the same output location, which is not a feature we want in our case.

More recently, Roma et al. produced a more general framework for visualization Roma et al. (2021) where they evaluated several dimensionality reduction algorithms for music creation applications. They found the uniform manifold approximation and projection (UMAP) algorithm to be more suited for this use case. One of UMAP parameters is the number of neighbors for projected points. This parameter impacts the clustered aspect of the 2-D projection, therefore allowing for control of the spread of the points over the space. They presented a solution to de-cluster points by mapping the output of a dimensionality reduction technique to a grid using the Hungarian algorithm. By oversampling the output grid, the initial distribution of points can be preserved while avoiding overlapping points. While helpful in getting a uniform distribution, this technique requires defining a grid that matches the desired distribution and the correct number of points, which can be difficult to generate

for general cases (typically arrangements not initially following a grid).

4.3.2 Unispring Algorithm for Density Manipulation

The *unispring* algorithm (Lallemand and Schwarz, 2011) manipulates a 2-dimensional set of points to distribute it across a given region uniformly or according to a user-defined distribution. The *unispring* algorithm comprises two steps explained below: the pre-uniformization and uniformization, as illustrated in figure 4.7.

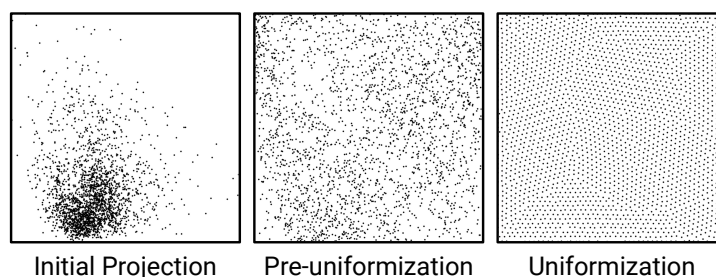


Figure 4.7: The *unispring* algorithm takes the 2-D projection of a dataset and distributes it uniformly in the space using a mechanical analogy between a triangle mesh and a spring network. A first pre-uniformization step is used to spread the points in the space to make the simulation converge faster.

Pre-uniformization The first step of the algorithm consists in scattering the initial distribution by performing a “*pre-uniformization*” step. It replaces each point’s x (resp. y) coordinate with the index of this coordinate in the sorted list of all x (resp. y) coordinates in the set. Note that these intermediate coordinates are scaled to preserve the initial range of the x -axis (resp. y -axis). This intermediate distribution fills the space while maintaining the order between points from the initial projection. This first step will help the physical algorithm to converge faster.

Uniformization The second step is based on the *distmesh* algorithm (Persson and Strang, 2004), introduced by Persson and Strang, that generates meshes using a simple mechanical analogy between a triangular mesh and a 2D spring network. The triangle mesh is created using a Delaunay triangulation: each point is linked with its neighbors, forming triangles whose circumcircles do not contain any point. The physical model is created by considering each connection between vertices as a spring. The obtained model is a network of joints linked by springs. Only the springs’ repulsive actions are considered, bringing the modeled behavior closer to that of a gas. Consequently, a point is subjected to a force whose direction and amplitude depend on the distance from connected neighbors. Each connected point applies a force in the direction of the connection with an amplitude proportional to the length of the connection:

$$f = \begin{cases} k(l_0 - l) & \text{if } l < l_0 \\ 0 & \text{if } l \geq l_0 \end{cases} \quad (4.1)$$

with l the current spring length, l_0 the rest length, and k the spring stiffness.

The position at a given step is solved using a forward Euler method. To end the simulation, a stop criterion is checked at every step: if, for all points, the displacement at a step is under a certain threshold, then the simulation stops. Further details on the algorithm can be found in the original *distmesh* article (Persson and Strang, 2004).

The region where the points are distributed consists of boundaries that cannot be crossed. If a point moves past a boundary, it is brought back to the closest point on the boundary. The user defines the region as a *signed distance function* that returns the distance to the nearest point on the boundary, positive if inside the region or negative if outside. We shall see in Section 4.4.2, that in our implementation, we modified this definition to simplify user interaction.

The target density is defined by the *element size function* $h(x, y)$. For a uniform target density, all springs have the same rest length so that the final distance between points is the same everywhere. We can modify the density in certain parts of the region by specifying different rest lengths. It is not required to specify the exact rest length of the springs, $h(x, y)$ only gives the relative distribution over the domain. The actual rest length of a spring whose center is at a (x, y) coordinate is then calculated using the following equation:

$$l_0(x, y) = \sqrt{\frac{\sum l_i^2}{\sum h(x_i, y_i)^2}} * h(x, y). \quad (4.2)$$

4.3.3 Additions and Modifications to the Unispring Algorithm

The aim of the *polyspring* toolbox is to allow users to manipulate a 2-D point distribution easily without having to tweak the parameters related to the underlying physical model. For this purpose, we made the following additions and modifications to the Unispring algorithm.

Calculation of the Springs Rest Length for a Uniform Distribution

The results of the physical model used to move the points highly depend on the choice of the springs' rest length since the final distance between points will be approximately the rest length. So, it must be chosen according to the target distribution, the number of points, and the region dimensions. We deduce an estimate of the target distance l_0^{uni} between points for a uniform distribution from the region area and the number of points. Let a uniform target distribution of density $d = \frac{N}{A}$ with N points contained in a region of area A . Since all pairs of points linked by the Delaunay triangulation must be at the same distance, they form equilateral triangles with their neighbors, as shown in figure 4.8. Since each point is the vertex of 6 different triangles, we can write the density of points in a single triangle as:

$$d = \frac{3}{6A_{tri}} = \frac{N}{A} \quad (4.3)$$

with $A_{tri} = (l_0^{uni})^2 \frac{\sqrt{3}}{4}$, the area of a triangle. Resulting in:

$$l_0^{uni} = \sqrt{\frac{2}{\sqrt{3} \frac{N}{A}}}. \quad (4.4)$$

This calculation does not consider the points on the borders of the region: a point on the border is the vertex of fewer than 6 triangles. Therefore, the density in Eq. (4.3) is underestimated in these cases, resulting in an overestimation of the rest length. However, as stated in (Persson and Strang, 2004), "*it is important that most of the bars give repulsive forces to help the points spread out across the whole geometry, [...] which can be achieved by choosing l_0 slightly larger than the length we actually desire.*" This slight overestimation

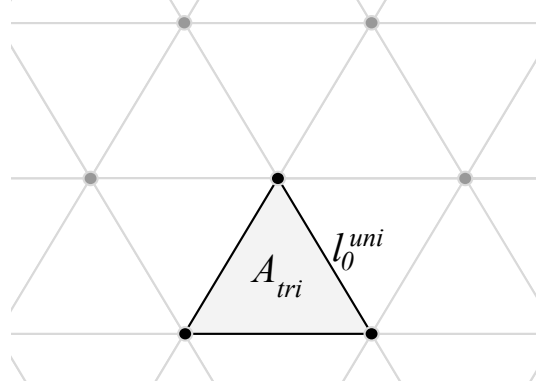


Figure 4.8: For a uniform distribution, points should be arranged so that they form equilateral triangles. Therefore, the spring rest length l_0^{uni} must be chosen to verify the target density.

of l_0 favors the spreading of points over the region.

User-defined Distribution: Non-uniform Rest Length

The *distmesh* algorithm introduces an *element size function* $h(x, y)$ that is used to specify the relative density in various parts of the region through a non-uniform rest length of the springs in the physical model. This method is implemented in the toolbox, albeit with some differences in the user-specified function. Instead of defining an *element size function*, we introduce a *density function* ρ , such that $h(x, y) = \frac{1}{\rho(x, y)}$, which specifies the relative density of points in different parts of the region. We made this change to focus the interaction on the density of points at specific coordinates rather than the distance between points. The *element size function* derives directly from the analogy of the spring network, the elements referring to the connections made by the Delauney Triangulation. Like h , this function is relative and does not need to sum to any particular value across the region—a location where $\rho = 2$ will be twice as dense as if $\rho = 1$.

The rest length of a spring whose center is at the coordinates (x, y) is calculated using the uniform rest length l_0^{uni} that we obtain using Eq. (4.4) and the ρ function that the user provides:

$$l_0(x, y) = l_0^{uni} \cdot \sqrt{\frac{N_c}{\sum_{i=1}^{N_c} \left(\frac{1}{\rho(x_i, y_i)}\right)^2} \cdot \frac{1}{\rho(x, y)}} \quad (4.5)$$

Where N_c is the number of connections and (x_i, y_i) is the center of the connection i . Note that in Eq. (4.5), the sum of l_i^2 from Eq. (4.2) is replaced by the average rest length l_0^{uni} from Eq. (4.4) which is a better estimate of the average rest length when there are large empty spaces in the initial distribution.

The user can specify any density function $\rho(x, y)$ as long as it verifies that $\forall(x, y) \in \mathbb{R}^2, \rho(x, y) > 0$. For instance, in figure 4.11B, $\rho(x, y) = x + y + 1$. This means that, since $\rho(0, 0) = 1$, $\rho(1, 0) = \rho(0, 1) = 2$ and $\rho(1, 1) = 3$, the distribution is three times denser at $(1, 1)$ and two times denser at $(1, 0)$ or $(0, 1)$, than at $(0, 0)$. The resulting distribution is displayed in figure 4.11B. Note that coordinates have been scaled to $[0, 1]$ in this case.

Gaussian Attractors: Manipulating the Distribution during Interaction

Originally, *unispring* was designed to obtain a “static” representation of the database, as stated by the authors: “we mean that the interface is determined once by the user and does not require further adjustments in the course of the interaction process” (Lallemand and Schwarz, 2011). Our goal in designing *polyspring* was to let users manipulate and interact with the distribution concurrently. While it is possible to interact with the sound corpus during the mass-spring simulation process, the speed of the simulation highly depends on the number of points and cannot be anticipated easily.

We created a method to interact with the point distribution using an analogy with a force field that can move points away from their stable position. Users can specify this force field through a mixture of Gaussians. After uniformization using the spring network method, the points are pushed away from their initial positions. The gradient of the force field directs the displacement vector of a point, and its norm is equal to the field’s value at the point coordinates. The field is scaled between 0 and J_0^{uni} so that the point displacements are bounded. An example of a distribution obtained with a mixture of three Gaussians is given in figure 4.12. While this model allows for inputs other than Gaussian mixtures, it is difficult to anticipate its behavior without understanding the process behind creating the displacement vector. So, we restricted the input possibilities to make it more accessible while allowing for complex final distributions.

4.4 Toolbox Description

The *polyspring* toolbox implements a modified *unispring* algorithm with its extensions described above. This toolbox consists of a Python package that provides classes to manipulate the distribution of a set of points inside a given region. It also includes a Cycling’ 74 Max package that enables its use with the MuBu⁵ library. A screenshot of the demonstration patch is displayed in figure 4.9. All the scripts and patches are provided in a repository⁶ and a video tutorial of the Max patch is available online⁷ and described in Appendix A.2.1.

4.4.1 Features

We now detail the main features of *polyspring*, illustrated with an example built using the Max implementation for *CataRT-MuBu*. In the screenshot of the Max patcher in figure 4.9, the green and yellow parts (“Initialize Corpus” and “Choose Representation”) are dedicated to the creation of the corpus and the initial projection. The blue part (“Shape Distribution”) uses *polyspring* to manipulate the sound fragment distribution. The visualizer on the right side displays the current distribution in a square. Note that it is automatically scaled to the initial distribution range on both axes and does not change its limits unless the descriptors are modified.

Initialization and uniformization The corpus is created in our example using *CataRT*’s automated segmentation and analysis process. The user imports one or more sound files that are first segmented (using, for example, a fixed segment length or onsets) and then analyzed with a series of N audio descriptors. Then, the user can choose two audio descriptors used to visualize all the sound segments as points in a 2D map, which corresponds technically to

⁵<https://forum.ircam.fr/projects/detail/mubu>

⁶<https://github.com/ircam-ismm/polyspring>

⁷https://drive.google.com/drive/folders/1aJmROHvKKQV1LZy3_BwteLnyeAt2aePq?

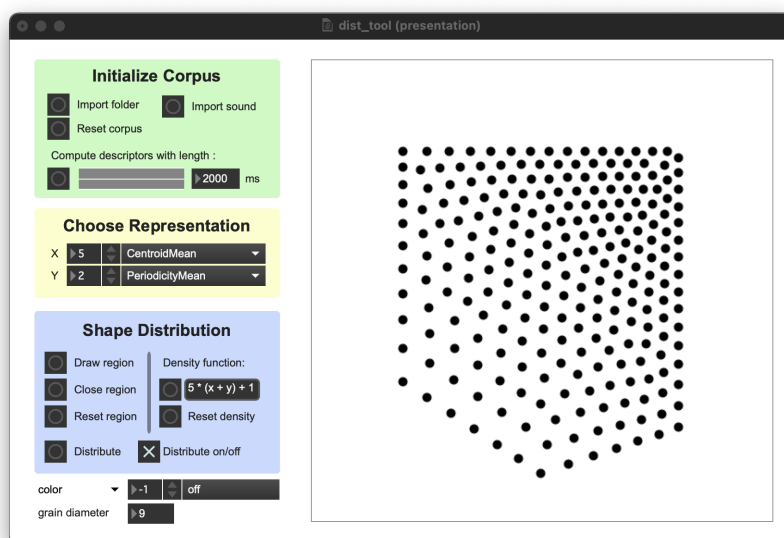


Figure 4.9: Screenshot of a Max patch demonstrating *polyspring* capabilities. The user can draw custom regions directly on the visualizer and type any density function of x and y , such as $\rho(x, y) = 5(x + y) + 1$.

the projection of the 7-D audio descriptors space to a 2-D plane. All of this is implemented using the externals provided in the *MuBu* package.

The interest of *polyspring* lies in the possibility of modifying the distribution of points representing each sound segment. By default, the region is set as the non-oriented bounding box of the set of points, and the density function is set to 1. When pressing the “*Distribute*” button, the points get uniformly distributed across this region. Since the visualizer is scaled to the initial range in both directions, it is equivalent to the default region, and the points fill the whole visualization space. The projection of a sound corpus in a descriptor space (spectral centroid and periodicity), before and after uniformization, is represented in figure 4.10.

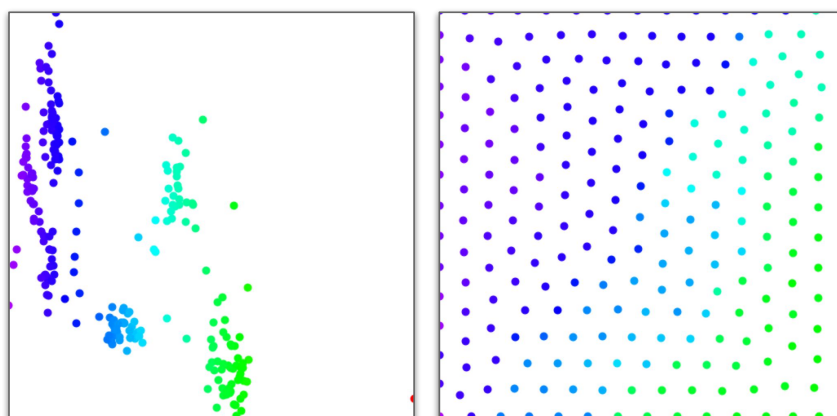


Figure 4.10: Example of a basic uniformization of a sound corpus. The left scatterplot represents the projection of the segmented corpus over two descriptors: the spectral centroid and the periodicity. The scatterplot on the right depicts the representation resulting from a uniformization of the control space.

Custom region Users can specify a custom region by drawing a polygon on the visualization space. After clicking the “*Draw region*” button, they can click on the visualizer to specify each vertex of the polygon sequentially. Clicking the “*Close region*” button will close the polygon by linking the last defined vertex with the first one. The uniformization algorithm then distributes points inside this region, as illustrated in figure 4.11A. Since the region definition relies on the package *shapely*, any polygon that can be created in *shapely* can be used to shape the distribution.

Custom density The user can define the density function ρ as a function of x and y (See section 4.3.3). For instance, in figure 4.11B, $\rho(x, y) = x + y + 1$. The user can type any function existing in the *numpy* package by specifying “*np.*” before the function name (eg. *np.exp*, *np.sqrt*...).

The Python class accepts any density function that takes two coordinates as arguments and returns a float indicating the relative density at this position. The example uses a formal expression definition, but we can also input a function obtained from a Kernel Density Estimation (KDE).

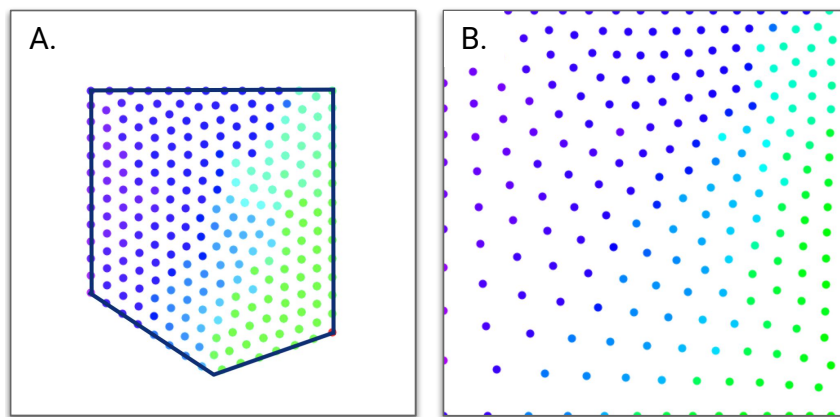


Figure 4.11: A. Example of the uniformization of a sound corpus over a user-defined polygonal region. B. Example of a distribution obtained with a user-defined density function $\rho = x + y + 1$.

Real-time interactions with the distribution, based on Gaussian attractors, are presented in a separate window as three configurable 2-D Gaussian functions exposing the Gaussian parameters: the center (μ_x, μ_y) , the spread (σ_x, σ_y) , and the angle θ . Figure 4.12 shows the Max window with the chosen parameters and the resulting distribution. Since the points are attracted from their original position by different Gaussians spread around the space, it can create empty spaces in the map. This distribution is not equivalent to what would be obtained by providing the Gaussian mixture as the density function ρ in the spring network model, which is significantly slower to compute – and, therefore, not appropriate for real-time manipulation. For example, this method takes less than $30ms$ compared to several seconds with the spring network. However, the resulting distribution does not guarantee the complete coverage of the space and is more difficult to anticipate. This complements the density definition using the ρ function, which is slower but affords more control over the final distribution.

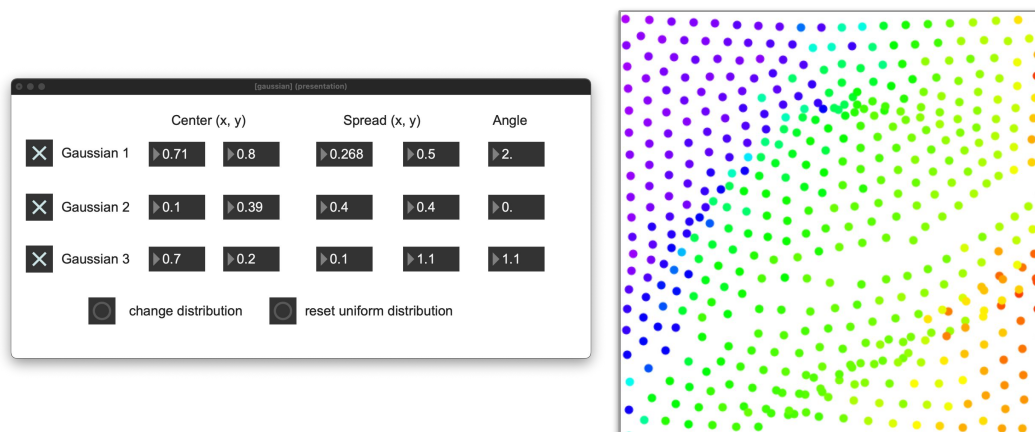


Figure 4.12: A distribution made using 3 gaussian attractors. The user defines the parameters for each Gaussian.

4.4.2 Implementation

Polyspring Implementation in Python The models described in section 4.3 are implemented with Python using standard packages for scientific computing, except for the region definition. All Matrix operations use *numpy*, and we use the Delauney triangulation from the *scipy.spatial* submodule. The Gaussian attractors use *numpy.gradient* to compute the gradient of the force field over a grid and the *scipy.interpolate.grid* data to interpolate the gradient at each datapoint from the gradient grid.

In the base *distmesh* algorithm, the main difficulty for region management user-wise is to create the *signed distance function* that returns the distance from a point to the closest boundary point (positive if inside, negative if outside). While it allows for any region to be defined, it isn't very easy to create in most cases.

We decided to use the *shapely* package⁸, a set of classes aimed at set-theoretic analyses and manipulation of planar features that uses functions from *GEOS*, a C/C++ library for computational geometry. *Shapely* provides a polygon class with a highly efficient point-in-polygon (PIP) test and closest points calculation. Using this package, any region can be easily defined as a simple polygon or more complex regions with holes using set operations. In our implementation, the user provides the region as a shapely *Polygon* or *MultiPolygon* object.

Using Polyspring in Max We created a Max patch and a Python script to interface *polyspring* with a MuBu container, which we used to make the above example. The Max object can be used to distribute any set of points contained in a MuBu track. The Max patcher and the Python script communicate using OSC (Open Sound Control) over UDP (User Datagram Protocol). Data are sent and received on the Max side using the base Max UDP implementation (*udpsend* and *udpreceive* objects). On the Python side, the communication is managed by the *python-osc* package⁹ that allows for an OSC address \leftrightarrow callback matching system. Thus, the communication consists of packets sent over UDP from one process to another. If both processes run on the same machine, the packets are sent to the local host IP address (127.0.0.1), and each process uses different receive ports (8011 for Max and 8012 for Python by default).

⁸<https://shapely.readthedocs.io/>

⁹<https://pypi.org/project/python-osc/>

4.5 Examples and Perspectives

Polyspring is structural to the design of the *movement sound explorer (mosX)* system described in the next chapter, which enables the exploration of a 2-Dimensional *CataRT* space with movement sensors. It is also used in the art-science performance *Sympoiesis* presented in chapter 6 where dancers explore a sound space during the performance.

The second part of the *Sympoiesis* performance was centered around slime molds, a unicellular organism whose growing behaviors have been thoroughly studied (Bonner, 2015). We wanted to use slime mold growth video recordings to control sound synthesis for this performance. A dataset of several hours of field recording in a laboratory has been scattered in a *CataRT* space. Since the videos consisted of upper views of a Petri dish, it was interesting to restrict the sound grains to a circle using *polyspring* to be able to superimpose them, as shown in figure 4.13. The uniform distribution also ensured the sound could be triggered everywhere in the circular region.

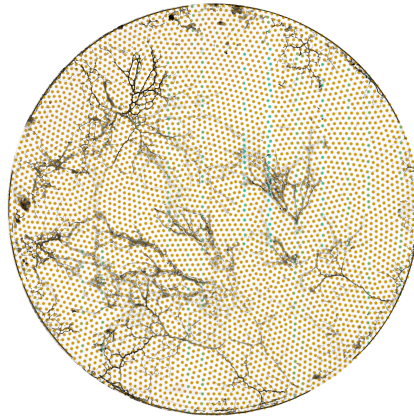


Figure 4.13: Picture of Slime Molds after several days of growth in a petri dish superimposed on a *CataRT* space of 4796 points uniformly distributed in a circular region. *Slime molds picture credit: Laurane Le Goff*

Other use cases might take advantage of the toolkit, such as the one proposed by Jensenius et al., where the position of a musician on stage, tracked by a camera, is used to control a *CataRT* space (Jensenius and Johnson, 2010). With a simple one-to-one mapping between the location on the floor and the 2D sound parameters (centroid and periodicity), the performer triggered segmented violin sounds. We can imagine an extension of such a paradigm with the possibility offered by *polyspring*: other sound parameters could be easily explored while assuring that all the sound cloud points are well distributed in a performance space of any geometrical form.

Polyspring is sufficiently generic to be used in a variety of applications. However, our main use cases always used the MuBu package. Communicating between Max and Python through OSC is not optimal in terms of performance. To improve this, we are implementing this toolbox as a Pipo¹⁰ process that will run in the Max process and should be substantially faster than the Python implementation.

Finally, the *unispring* algorithm should be generalizable to 3-D or higher-dimensional spaces. This allows for mapping between several movement descriptors and sound space

¹⁰<https://ismm.ircam.fr/pipo/>

dimensions in the case of gesture control, for instance. Uniformization over N dimensions would guarantee the presence of sound points in all directions, which is particularly interesting when exploring high-dimensional spaces that are difficult, if not impossible, to visualize. Delaunay triangulation can be performed in n -Dimensions, already implemented in the `scipy.spatial.Delaunay` function and the repulsive action between two points could, therefore, be calculated using the n -D Euclidean distance. However, ensuring the convergence of the generalized model is not trivial; the region borders would be more complex to define, and keeping the simulation stable with a forward Euler method would require a lot of experimenting with the parameters. An alternative solution could be to work with the Voronoi diagram of the point set, the dual of the Delaunay triangulation, and use the Voronoi relaxation method to obtain an evenly spaced set of points. A general algorithm for n dimensions exists. However, the custom boundary and custom density constraints require modifications to the base algorithm.

4.6 Chapter Summary

This Chapter presents a prototype of movement-sound interaction based on corpus-based concatenative synthesis and controlled with a touch-based or a movement-based interface. Overall, this first study allowed us to validate our interest in studying, in controlled experiments, whole-body exploration of a sound space. In particular, we found that such embodied interaction, even when manipulating only two sound dimensions, cannot generally be described as a simple extension of fingertip planar exploration. Nevertheless, we found that repeated trials with whole-body interaction led to improvement, but several technical limitations seemed to hinder exploration. In particular, the inhomogeneity of the grain distribution in the sound space made exploration difficult. In response, we presented the theoretical basis and the Python implementation of the *polyspring* toolbox. It allows for the design of sound corpora where 2-dimensional sound clouds are distributed over a specific space while maintaining local neighborhood relationships between sound parameters. This opens new possibilities by enabling more control, for example, on the *CataRT* 2-dimensional representation.

CHAPTER 5

Exploration in Movement-Sound Interactive Systems

The content of this chapter is currently being worked on for submission to ToCHI in August 2024.

5.1 Introduction

This chapter follows the preliminary experiment presented in Chapter 4 to further study early learning of movement-sound interactions. The prototype introduced previously allows for the exploration of a sound corpus projected on a 2-D plane using *CataRT*. Participants explored the same sound space with planar fingertip interaction and full-body interaction. The preliminary experiment has shown that whole-body exploration differs significantly from fingertip exploration on a tactile surface. Participants found most sound clusters with the touch-based interface and drew a good approximation of the sound space on paper. With the same sound space, whole-body exploration through arm movements disturbed their planar representation, proposing a radically different interaction with the sound corpus. Still, such an exploration using the arm's articulations appears to represent an interesting embodied sonic experience. In addition, participants showed a variety of behaviors in their exploration of space with this interface, particularly in the way sound space is covered. However, the inhomogeneity of the distribution of sound grains complicated the exploration and learning of space through whole-body interaction in a short time.

Following on from these initial observations, our goal is to investigate further early learning when exploring a movement-sound space. Importantly, we aim to investigate how early learning of audio-motor map and exploration can occur *without* visual feedback, focusing

on proprioception coupled to auditory feedback. Such an approach calls for different questions compared to the case of learning with visual feedback which has already been studied (see (Jeanne et al., 2017) for example).

In this context, we aim to describe and understand users' exploration experiences rather than assess task performance. Our primary research questions in this chapter are: **(1)** Can we characterize exploration through quantitative measures? **(2)** How do exploration patterns vary across individuals and tasks? **(3)** How do people build mental representations of movement-sound spaces through exploration?

To address **(1)**, we developed a method for characterizing exploration trajectories based on Kernel Density Estimation techniques (KDE). We calculate quantitative indicators adapted to the parameter space of the *mosX* system to generate visualizations of participants' exploration patterns. We show that this method enables us to identify exploration patterns and characterize them through the temporal evolution of these visualizations.

To answer **(2)** and **(3)**, we conducted a study where participants explored a movement-sound space. We adapted the initial prototype presented in Chapter 4 using *polyspring* to create a novel movement-sound exploration system called *mosX*. It is dedicated to sound corpus exploration using body postures measured by wearable sensors (wireless Inertial Measurements Units). Using the system, we designed a movement-sound space that links arm articulations to ambient sound textures. Informed by the preliminary experiment, this space should be simple enough to be learned quickly. We designed an experiment with 12 participants who explored the movement-sound space created with the *mosX* system during four 5-minute sessions, alternating task-oriented and open-ended exploration. Analyzing the outcomes using our KDE-based method, we observed that participants could learn most of the structure of the movement-sound space in only 20 minutes of exploration. Quantitative results revealed that participants shared similar exploration patterns, particularly in task-oriented sessions, but presented significant differences in their exploration and points of interest during open-ended exploration. Along with these measures, we gather feedback from participants with interviews as well as sketches of their perception of the movement-sound space. The sketches proved interesting for eliciting each participant's representations of space and showed that some drawings shared similarities in how they represent space. Finally, we discuss this result in regard to cognitive representation literature and navigation literature.

The following section presents a selection of research works related to exploration and learning in interactive systems. The third section describes the *mosX* system. The fourth section proposes a method for analyzing exploration data based on KDE. The fifth and sixth sections report on the experiment led with *mosX* and the seventh section discuss the results.

5.2 *mosX*: movement sound explorer

5.2.1 Preliminaries: Studying Sensorimotor Learning in the Performing Arts

Our initial motivation for the prototype presented in Chapter 4 was to draw inspiration from studies in audiomotor map acquisition by Van Vugt and Ostri (van Vugt and Ostry, 2018a,b, 2019). Their proposed system was simple regarding movement and sound, allowing for a controllable experiment and accurate measurement of the learning rate. Our aim is to propose an interaction with a correspondence between movement and sound that can be learned in the time of a simple laboratory experiment, while offering richer sounds and movements than those traditionally used in sensorimotor learning experiments. In particular, we want to

engage the user in a full-body interaction with sound, closer to works with dance or music practitioners (Françoise et al., 2022; Fernandez et al., 2017; Nabi et al., 2024). Importantly, the interaction is designed as a continuous mapping between postures and sound textures that must be embodied, and we deliberately chose a different interaction paradigm from a pointing task where the user would navigate positionally using a 3D tracker to trigger sounds. A video demonstration of the system is provided in the supplementary materials.

This first prototype we designed consisted of controlling the *CataRT* corpus-based synthesis with arm movements by directly “*plugging*” the output of the orientation of IMU sensors strapped to the arm to the axis of the 2-D descriptor plane. This setup provided participants with intriguing experiences. However, it showed several flaws due to the difficulty of manipulating the distribution of sound grains over the 2-D plane, making the movement-sound space particularly hard to learn. We created the *polyspring* package to address these issues.

The revised prototype, *mosX*, for *movement-sound explorer*, is a system that enables the exploration of a sound corpus through body configurations based on *CataRT* and *polyspring*. Through *polyspring*, the sound space can be adapted to different scenarios from the point of view of movement and user expertise. It was used here to study movement-sound exploration in a more controlled sonic environment than in the preliminary experiment. Below, we present the movement-sound space, with particular attention to the system’s design specifications regarding the type of interaction we were interested in studying.

As a reminder, we define an *N-dimensional sound space* as a sound synthesizer with *N* “*human controllable*” parameters. In the case of a space created with *CataRT*, 2 sound descriptors are used to project the sound grains on a 2-D plane, producing a 2-D sound space. The *movement space* refers to the joint space that interacts with the gestural interface. Finally, the *movement-sound interaction space* or *movement-sound space* corresponds to the interaction space that results from the mapping enabled by the interactive system.

5.2.2 Design Specifications

Movement-Sound Interactive Systems for the performing arts relate to various practices due to their multimodal nature. For this system, we defined specifications inspired by constraints typically found in traditional musical instruments.

Continuous movement-sound mapping The system proposes continuous movement-sound relationships. Interviews highlighted the limitations and “*gadget*” aspect of triggering paradigms often found in electronic music controllers. Traditional instruments are not all continuous in pitch. Some, such as the guitar or the clarinet, have discrete notes that can be played. However, they all present continuous properties such as nuances, articulations, or note lengths, to name a few. It is these continuous properties that enable musical expressivity.

Reactivity and repeatability The system must enable learning of movement-sound relationships. A specific action in the movement space must produce the same sound outcome every time it is performed, and the time between the execution of the movement and the perception of the corresponding sound must be short enough to facilitate the perception of causality and foster a sense of agency over the system.

Surjective Mapping Traditional instruments are often “*surjective*,” meaning several movement or action possibilities can create similar sounds. Many instruments can play the same

note with different performers' movements. For example, a given note could be performed through different fingerings in string instruments, or different fingerings are possible for a given melody when playing a piano. While this might appear as a difficulty at first, it is generally not considered a problem but rather an “*opportunity*” to develop a variety of playing techniques. For this reason, we propose to consider what we call *surjective mapping*. While the whole sound space is covered by the movement space, which means that each point in the sound space is reachable, different points in the movement space can lead to the same point in the sound space, as illustrated in figure 5.1. No assumption is made regarding the nature of the points in the movement and sound space. It just refers to a specific state in each space.

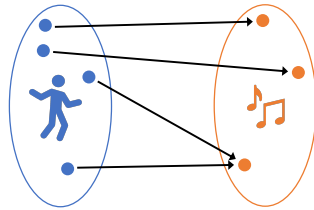


Figure 5.1: A surjective mapping ensures that every element of the codomain (here, the sound space) can be reached by at least one element of the domain (here, the movement space). Therefore, several elements of the domain can be mapped to the same element of the codomain.

5.2.3 Sound Space

Corpus-Based concatenative Sound Synthesis

The system designed for this study is based on corpus-based concatenative synthesis techniques. Those methods rely on analyzing and recombining databases of recorded sounds. Such methods involve the segmentation of audio recordings, their parametrization (for example, with audio descriptors), and their resynthesis using selection methods. Various implementations have been proposed, such as the real-time corpus-based concatenative synthesis system *CataRT* (Schwarz et al., 2006), used by several research groups and musicians (Einbond et al., 2016; Zbyszyński et al., 2021; Schwarz, 2007), the tools developed in the Fluid Corpus Manipulation project (*FluCoMa*) (Tremblay et al., 2019, 2021), *earGram* (Bernardes, 2021) or the *AudioStellar* system (Garber et al., 2020). 2D representation of the sound parameterization greatly facilitates its direct manipulation by the performer through projection or dimensionality reduction from the possibly high dimension of the sound description. The relationship between the user gestural input and the selection of such 2-D reduced space can be formalized as a special case of the so-called gesture-sound mapping (Hunt et al., 2000). This synthesis technique (and its variants) is typically used with a 2D projection of the multi-dimensional sound descriptor space on a screen. It allows for easily selecting sound segments with a mouse or a touchpad. Other types of feedback have been proposed, such as haptic feedback (Magalhães et al., 2020).

In *mosX*, sound spaces are designed and implemented with *CataRT*. It allows for the exploration of a large corpus of recorded sounds. They are segmented into so-called sound grains and analyzed using audio descriptors, which can then be used to select a single segment with the desired audio characteristics for playback. This sound synthesis technique enables the control of a rich set of sounds that can be blended by creating layers of (possibly overlapping) sound grains, thus controllable with a limited set of parameters.

Sound Grain Distribution

The projections created with CataRT can present inhomogeneity of the grain distribution in the sound space as shown in the figure 5.2. The empty areas in the space make exploration difficult without visual feedback. In response, we created the *polyspring* toolbox for CataRT described in Chapter 4. It allows for the design of sound corpora where 2-dimensional sound clouds are distributed over a specific space while maintaining local neighborhood relationships between sound parameters as shown in the figure 5.2.

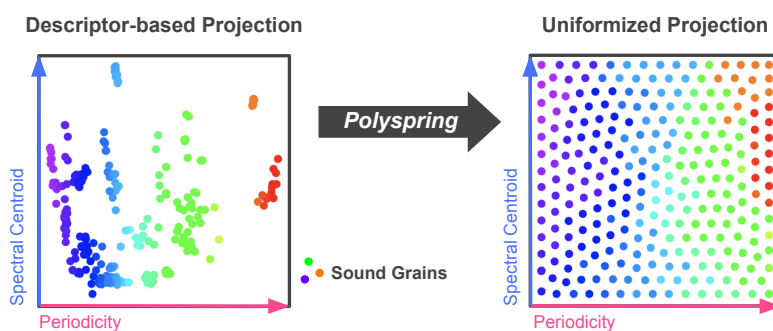


Figure 5.2: Example of basic uniformization of a sound corpus. The left scatterplot represents the projection of the segmented corpus over two descriptors: the periodicity (x-axis) and the spectral centroid (y-axis). Each dot represent a sound sample. The scatterplot on the right depicts the representation resulting from a uniformization of the control space.

Designed Sound Space

Sound sources The sound space we created using mosX kept the same number of clusters, but sounds were selected to create a corpus of 8 identifiable and discriminable sound classes organized into 4 categories:

- I Market - (a) crowd in a market, (b) sounds of grains shaken;
- II Fire - (a) fireplace crackling, (b) fireworks;
- III Nature - (a) water streams, (b) rainforest;
- IV Machine - (a) motorcycles, (b) electric arcs.

Each sound class comprises 4:00 minutes of different sound samples originating from similar environments. We chose the various sound labels according to the actual content of the sound samples, but the categories' names only indicate the intent behind the sound design. We did not expect the participants to identify those names but to be able to distinguish them ideally without ambiguity (which will be tested). We will now refer to the categories with their Roman numbers.

Each sound sample was cropped in 500ms sound grains, which produced approximately 480 sound grains per sound class. For each class, the grains were projected on a 2-D space using the spectral centroid and periodicity descriptors, after which the obtained distribution was uniformized with *polyspring*. The four obtained spaces were then projected on a larger plane. The organization of sound classes and categories on this plane are presented in figure 5.3. The four categories are located in a quadrant of the space. Inside each category, the corresponding sound classes were superimposed. Note that each sound class was mirrored so that their respective origin met at the center of the space at coordinate (0.5,0.5), and

descriptors increased outward of the center on the x and y-axis.

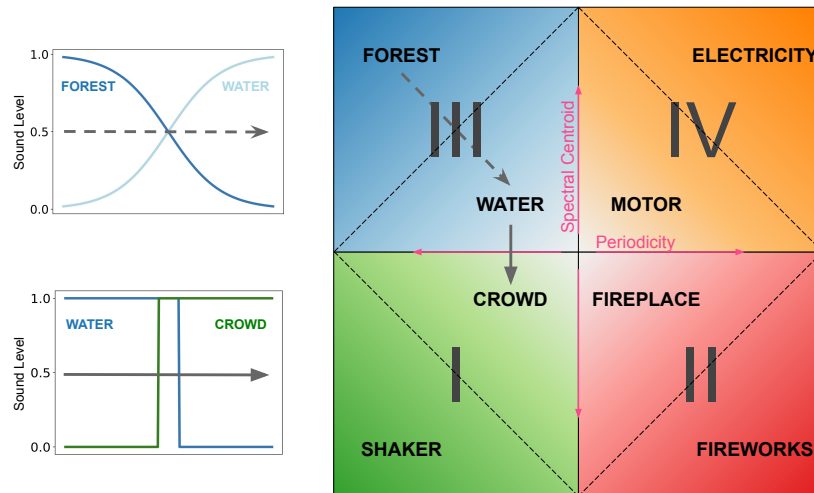


Figure 5.3: Organization of the sound space. The frontiers between categories are abrupt; the sound transitions quickly from one sound class to another. Transitions between sound classes within categories are seamless, allowing them to blend into each other.

Transitions The frontiers between categories (in solid lines) are clear. The sound transitions quickly from one class to the other between categories (e.g., from water to crowd) with a short overlap of both classes. However, inside each category, a crossfade allows for a smooth transition (in dashed lines) from one sound class to another (e.g., forest to water for class I). The resulting space is made of 4 separate categories, inside which sound classes can blend.

Sound Grain Triggering The inputs taken by *CataRT* are the coordinates of a point in the sound space. We note these coordinates (x, y) in our 2-D case (see figure 5.3). The simplest triggering mode consists of playing the closest grain to this input point, either once or repeatedly, as long as it remains the closest to the input point. In the first case, the input must vary to trigger new points, whereas in the second case, the input can remain stationary.

We decided to favor the second case to simplify the exploration, but we added the following behavior to make the stationary sound more interesting. Each time an (x, y) couple is received, all grains within the radius $r = 0.08$ to the input are retrieved (see figure 5.4), the whole region being a 1×1 square. If the closest grain differs from the one calculated with the previous input, then the new grain is played. If it is identical to the previous closest grain, then another grain within the radius is randomly chosen and played. This method allows for generating slightly changing sound textures even if the user is not moving. Since, in a given sound class, points close to each other share similar descriptors, the resulting texture sounds coherent. When the input is stationary, the sampling period is approximately $60ms$. When the inputs move, each new closest grain is played. Therefore, the dynamic of the input affects how many grains are played per second, which directly translates to a variation in loudness.

5.2.4 Movement Space

To make the movement space accessible and easy to explore in a short time, we decided to focus on the postures of the dominant arm. We will not follow strict anatomical naming and

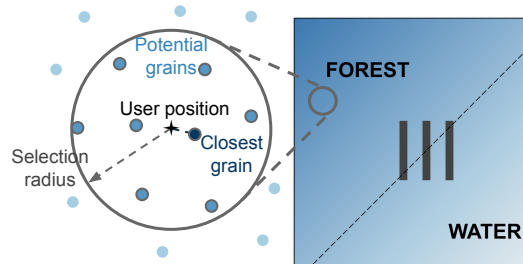


Figure 5.4: The user input determines which grain is played. First, we identify all the grains lying in a circle of radius $r = 0.08$, centered on the user position. The closest grain to the center is played unless it is the last played grain. In this case, any other grain inside the circle is randomly chosen.

refer to “*arm*” as the limbs going from the shoulder to the hand (also called upper limbs), to “*upper arm*” as the limbs between the shoulder and the elbow, and to “*lower arm*” as the limbs from the elbow to the wrist. The participants were standing up and could move their arms freely in every direction. To measure the orientation of the arm and the forearm, we strapped two inertial measurement units to the arm: one to the upper arm and one to the forearm, just before the wrist (see figure 5.5). The resulting space from the user’s perspective is a joint space with 5 Degrees of Freedom (DoF):

- 3 DoF from the shoulder (azimuth, elevation, and humeral rotation);
- 1 DoF from the elbow (flexion);
- 1 DoF from the forearm (pronosupination).

Only the forearm rotation around its axis (pronosupination) is considered here. Abduction and wrist flexion are not considered since the sensor cannot measure them.

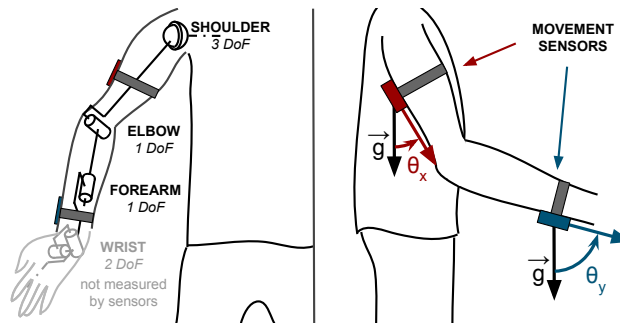


Figure 5.5: Kinematic diagram of a simplified arm with 7 degrees of freedom (DoF). Movement sensors are strapped to the upper arm and the forearm. Since the wrist’s movement does not impact the sensors’ measurements, only 5 DoF are measured. The orientation of each sensor relative to the vertical is extracted and used as input for the movement-sound mapping.

Movement Sensors Fusion The movement sensors are two inertial measurement units. They include a 3-axis accelerometer, gyroscope, and magnetometer. Using a sensor fusion algorithm on accelerometer and gyroscope data, we calculate the orientation of each sensor. Since we do not use magnetometer data, we obtain only the pitch and tilt of the sensors and not the azimuth. From these data, we calculate for each sensor the angle between the first

axis of the sensor and the gravitational force direction (θ_x, θ_y) (see figure 5.5). This forms the control parameters.

5.2.5 Resulting Movement-Sound Space

The mapping concludes with the connection between the control parameters (θ_x, θ_y) obtained by the sensors and the synthesizer parameters (x, y) with a factor of 180° .

$$\begin{cases} x = \frac{\theta_x}{180^\circ} \\ y = \frac{\theta_y}{180^\circ} \end{cases} \quad (5.1)$$

To summarize, the movement-sound space links the posture of the dominant arm (figure 5.5) to a 2-dimensional sound space (figure 5.3) created by projecting sound grains into a 2-D plane and uniformizing the obtained distribution. The complete transformation can be written as:

$$(q_1, q_2, q_3, q_4, q_5) \xrightarrow{\text{Sensors}} \begin{pmatrix} acc_{xi} & acc_{yi} \\ gyr_{xi} & gyr_{yi} \end{pmatrix} \xrightarrow{\text{Fusion}} (\theta_x, \theta_y) \xrightarrow{\text{Scaling}} (x, y) \xrightarrow{\text{Polyspring}} \text{Sound Grains} \quad (5.2)$$

With:

- $\{q_i\}_{i \in [1,5]}$ the parameters of the articulation space corresponding to the five degrees of freedom presented in figure 5.5 ;
- acc_{xi} and gyr_{xi} respectively the accelerometers and gyroscopes data of sensors x (or y) in direction $i \in [1, 3]$;
- (θ_x, θ_y) the angles between the first axis of each sensor and the gravitational force direction as displayed in figure 5.5;
- (x, y) the sound space parameters.

5.2.6 Resulting Interaction

The obtained interaction follows the design specifications given above.

Since the sensor latency over Wi-Fi is kept at around $5ms$ with no obstacles between the antenna and the user, and the processing time of sensor data is negligible (around 1 ms), the system feels reactive to movement input. Even with the random process used when the input is stationary in the sound space, the resulting sound texture at a given position only evolves slightly. Therefore, for the same input, the system will deliver the same texture, allowing for a repeatable experience.

The system allows for a continuous interaction between movement and sound. The space's granularity (approximately 4000 sound grains over the whole space) allows for varying sound textures, even with slight input variations.

Finally, the mapping is surjective. We give some examples of arm configuration to sound outcome relations. To point at the outer corner of each category, participants must have their arms in the following configurations:

I bottom-left (origin) arm extended, at rest along the body;

II bottom-right arm bent with the elbow pointing upward;

III top-left arm bent with the elbow pointing downward;

IV top-right arm extended, raised towards the ceiling.

These positions are the only ones that enable the corners of the sound space to be pointed. On the contrary, any movement that keeps sensors at the same angle related to the direction of gravity force produces no variation in the sound space. Therefore, the mapping is invariant by azimuthal rotation around the vertical axis. For instance, one position that points to the center of the space is to have the arm fully extended horizontally. But, starting from this position, bending the elbow while keeping the forearm horizontal will not move the position in the sound space. This rotational invariance means that various intermediary positions between the center and the extremes can be used to point inside a category. The resulting mapping is interesting for exploration because the number of movements to move between sounds in the sound space is potentially infinite.

5.3 Exploration Data Analysis

We present in this section a method to analyze exploration data based on Kernel Density Estimation (KDE) techniques. Given a set of points \mathbf{X}_i corresponding to consecutive positions in the explored space we want to visualize the explored space as well as determined what proportion of the space has been explored in total, as well as in subparts of the space. In section 4.2.3, we directly analyzed scatter plots of all \mathbf{X}_i to obtain information about visited parts of space as shown in figures 4.4 and 4.5. Superimposing points in this manner give little information about the distribution of those points in the space. At best, we identify the unvisited part of the space. To quantify the surface area that has been explored, further computation is required.

Using KDE for data visualization and analysis has been proposed by Tukey in 1977 for univariate cases (Tukey, 1977). David W. Scott made an overview of multivariate density estimation techniques in (Scott, 2015), which extends toward trivariate and quadrivariate data analysis. Multivariate KDE has been used in various contexts such as economics (Jin and Kawczak, 2003), geovisualization (Maciejewski et al., 2010; Chai et al., 2024), or eye-tracking data analysis (Jansson et al., 2013; Blascheck et al., 2017). In earlier work, he presented methods focused on bivariate density estimation and evaluation, starting with histograms and moving to kernel-based estimation methods (Scott, 2004). The later are particularly interesting for 2-dimensional visualization of bivariate distributions. Our method adapts bivariate Kernel estimators as presented by Scott to the case of exploration data analysis. We focus here on the case of 2-D spaces.

Let be a set of 2-dimensional points $\mathbf{X}_i = (x_i, y_i)$. To analyze it we propose to estimate the density function \hat{f} of its underlying distribution. The figure 5.6 shows an initial set of exploration points and the map we compute using KDE. From the density function estimated we can then compute local and global coverage estimations. This section first presents this method principle and illustrates it with data from a 5-minute exploration with the *mosX* system.

5.3.1 Kernel Density Estimation

Following are descriptions of visual and quantitative analysis based on the data distribution. Both rely on estimating the probability density function (PDF) of the distribution using the kernel density estimation (KDE) technique.

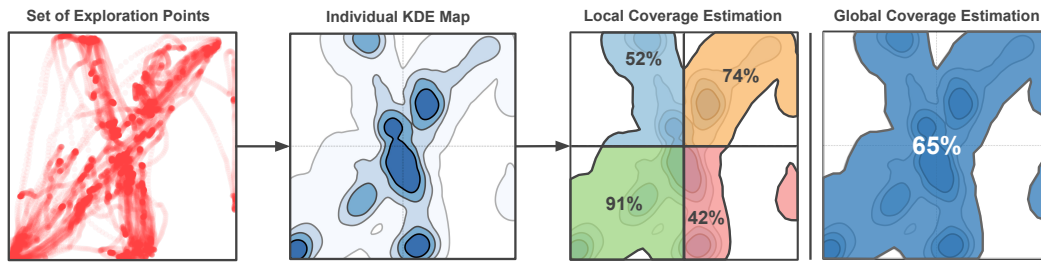


Figure 5.6: Overview of our analysis method in 2-Dimensions. A set of exploration points is a series of 2-Dimensional positions. From exploration points we compute a map representing visited parts of the space using kernel density estimation. We can estimate from these maps the proportion of the space that has been covered, locally, in subparts of the space, as well as globally, over the whole space.

Basics of Bivariate KDE

KDE is a nonparametric method used to obtain an estimate of the probability density function (PDF) of an unknown distribution. As such, no assumption is made about the distribution itself, and it can be applied to any set of points, but some parameters must be adjusted. We describe here the basic principle of the KDE method and how the various parameters were chosen in our case.

Given a finite set of points $\mathbf{X}_i = (x_i, y_i)$, the density at the point \mathbf{X} is calculated using the density estimator \hat{f} :

$$\hat{f}(\mathbf{X}) = \frac{1}{n} \sum_{i=1}^n K_{\mathbf{H}}(\mathbf{X} - \mathbf{X}_i) \quad (5.3)$$

with n the number of observations and $K_{\mathbf{H}}$ the so-called scaled kernel. $K_{\mathbf{H}}$ is a bi-variate function corresponding to a radially symmetric kernel K with varying bandwidth. Each observation contributes to the global density through the kernel K , centered on the measured point and scaled and rotated with the 2×2 symmetric positive bandwidth matrix \mathbf{H} . The adjusted Kernel is expressed as

$$K_{\mathbf{H}}(\mathbf{X}) = |\mathbf{H}|^{-1/2} K(\mathbf{H}^{-1/2} \mathbf{X}) \quad (5.4)$$

In the general case. The general form of the bandwidth matrix is:

$$\mathbf{H} = \begin{bmatrix} h_1^2 & h_{12} \\ h_{12} & h_2^2 \end{bmatrix} \quad (5.5)$$

Where h_1 , h_2 and h_{12} are the smoothing parameters. h_1 and h_2 change the scale of the kernel in directions 1 and 2, while h_{12} impacts how much the kernel is rotated. Note that if $h_1 = h_2$, the Kernel remains radially symmetrical and cannot be oriented in any direction, making the parameter h_{12} useless. On top of the three smoothing parameters, we must choose the kernel K . It is commonly accepted that the choice of smoothing parameters is the most important one as it will decide how many distribution modes (local maximums of the density) this technique will reveal. If the bandwidth is too large, separate distribution modes could blend and look like a single, wider distribution mode. On the contrary, a bandwidth that is too narrow will introduce artifacts, creating parasite distribution modes inside of a single mode. The impact of the smoothing parameters is described in length by Wand and Jones (1993). The choice of the Kernel itself marginally impacts the estimated distribution (Jones, 1994).

Choice of KDE Parameters

Multiple methods exist to estimate the smoothing parameters from the data automatically. Scott's or Silverman's method gives a general rule of thumb, but they do not yield good results with multimodal data (Wand and Jones, 1993). As we do not expect them to concentrate only on one point of the space, multiple distribution modes should appear in the distribution. We processed and analyzed the data in Python using the KDEpy package. This package includes the FFTKDE class that uses a fast Fourier transform to compute the estimator \hat{f} . This class computes the density function ten times faster than the *gaussian_kde* class from the *scipy* package. However, for bivariate data, this solution only proposes to estimate the bandwidth using Silverman's or Scott's rule of thumb. So, we had to choose the smoothing parameters ourselves. Moreover, we could only specify the bandwidth as a scalar, corresponding to the case where $h_1 = h_2 = h$ and $h_{12} = 0$.

We chose the bandwidth parameter according to the sound space's granularity level. At a given point \mathbf{X} , the participant can hear sound grains positioned as far as $r = 0.08$ from \mathbf{X} (with the space being a 1×1 square); see figure 5.4. Therefore, the participants are not subjected to a per-grain representation of the sound space but rather to a local zoom of all the sound grains around a given location. Consequently, we chose a finite support kernel (the Epanechnikov kernel) scaled with a radius of $r = 0.08$. By using the same scaled Kernel for each exploration data, we compute a distribution representation that reflects the sound space as perceived by the participants rather than the distribution of the trajectories considered independently of the sound interaction.

5.3.2 Visual Analysis of the KDE

We first plotted each control parameter point of exploration sessions on the 2-D plane corresponding to the sound space. For example, the scatter plot for a 5-minute exploration session is shown in figure 5.7A. This scatter plot comprises the 30,000 points recorded during the exploration (300 seconds at 100Hz). By reducing the opacity of the dots, the areas where points are the most concentrated and superimposed stand out in solid red.

With this view, it is possible to see various areas with concentrated points corresponding to zones where the participant stayed and returned multiple times. For instance, on the lower boundary line of the figure, we observe two clusters of points, one at the origin of the space and one in the middle of this boundary. Here, with an opacity of 5%, we reach full opacity when 20 points are superimposed. This visualization gives us a rough idea of how much space has been explored and which areas were the most visited. However, such a scatter plot presents too much information at a fine granularity. By estimating the density distribution over the space, we abstract all those points to a continuous function that will enable us to draw contour plots. The whole process is presented in figure 5.7.

The figure 5.7B displays a 3-dimensional view of the density estimated with the KDE technique. The x and y axis correspond to the scatter plot axis in figure 5.7A, and the z -axis is the density value estimated for each (x, y) coordinate. The 2-D plots on the "walls" for the figure correspond to the distribution of points for each variable. From the 3-D KDE plot, we can see the distribution's local maximums (distribution modes). Here, we identify two local maximums on the lower y border previously identified on the scatter plot. As highlighted by Scott in (Scott, 2004), the visualization of the density for each variable independently results in hidden distribution modes. The y variable plot displays a high mode at $y = 0$ while we know another mode lies behind it. Similarly, the x variable plot shows the highest central mode at the center, which hides the lower mode behind it. The 3-D plot also has its flaws

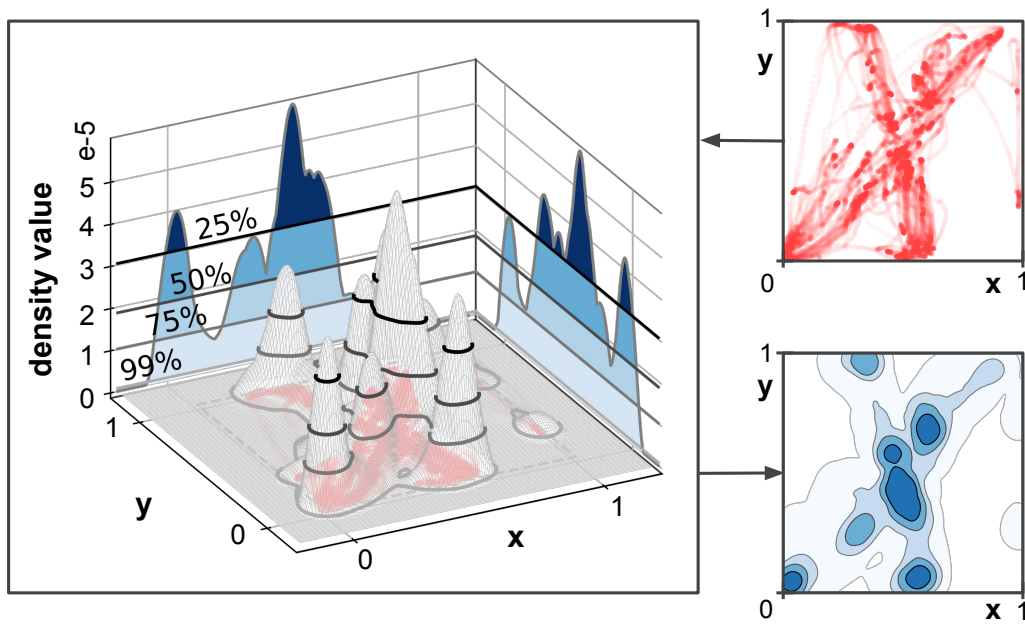


Figure 5.7: Overview of KDE computation for visually interpreting exploration traces. An exploration trace (A) is a time series of 2D positions. To extract the distribution contours from the observations, we estimate a probability density function using the kernel density estimation (KDE) technique from this set of points (B). Then, we calculate the contours corresponding to upper percentages of the highest density regions (C).

because masking necessarily follows from the choice of point of view. Note that, to create this figure, the estimated PDF \hat{f} is evaluated on a regularly spaced grid of 200×200 . The density values are scaled so that the density over the evaluated grid sums to 1.

To visualize both the covered area of the exploration as well as the various distribution modes, all in one plot, we use contour plots. We draw iso-density lines at particular levels to highlight regions with densities higher than the given level. This projection allows the visualization of the distribution modes and the covered area, as shown in figure 5.7C. However, the choice of density levels is essential to seeing these features. Using the 3-D visualization of the PDF to identify relevant levels and craft contour plots on a case-by-case basis could be possible, but it is neither objective nor consistent.

A more relevant approach is to select levels representing iso-proportions of the density. For any given percentage, there is a unique level above which lies this proportion of the probability mass. For instance, above the 25% iso-proportion level lies 25% of the probability mass. Moreover, this level defines the smallest region that contains this probability mass. Put otherwise, a point drawn from this distribution has a 25% chance of being inside the contour drawn at 25% probability. Therefore, iso-proportion level contours indicate the densest regions of the distribution. This contours are analogous to univariate quantiles but are not strictly bivariate quantiles, see (Siddiqui, 1960).

To calculate an approximation of the level γ_α corresponding to the proportion α , $0 < \alpha \leq 1$ of the finite grid of evaluated densities $(d_{ij})_{i,j \in \llbracket 1, N \rrbracket^2}$ we use the following method:

1. flatten the grid $(d_{ij})_{i,j \in \llbracket 1, N \rrbracket^2} \rightarrow (f_i)_{i=1}^{N^2}$

2. sort the list in ascending order $(f_i)_{i=1}^{N^2} \rightarrow (s_i)_{i=1}^{N^2}$;
3. find the smallest i_{lvl} such that $\sum_{i=1}^{i_{lvl}} s_i > 1 - \alpha$;
4. the approximate level $\hat{\gamma}_\alpha$ is given by $\hat{\gamma}_\alpha = s_{i_{lvl}}$.

With this technique, the levels at which the contours are drawn are distribution-dependent. The levels must be recalculated for each distribution. When comparing participants or sessions, the contours will not be drawn at the same levels but will display the same information regarding proportions. We chose 25%, 50%, and 75% for visualization.

5.3.3 Coverage Estimation

From the KDE estimation, we can evaluate what portion of the space has been explored, in its entirety as well as for individual categories. To do so, we need to define a criterion to decide whether or not a point in space has been visited. We can use the iso-proportion levels computed before to craft an exploration indicator. We consider a point in space as explored by a participant for a given session if the probability that it comes from the distribution of this session is above a threshold ϵ . This method is used by Jansson et al. to define the outlier region of eye-tracking data (Jansson et al., 2013). ϵ designates the probability that a point outside the outlier region comes from this distribution. We used $\epsilon = 0.01$ to plot the contours corresponding to the explored region figure 5.7. This contour does not contain all the points for this session. In doing so, a location in space is considered explored when a minimum density of points has been reached, and outliers are excluded from the 99% contour.

We define the total coverage as the area of the 99% contour divided by the total area of the parameter space. Since the total area of the parameter space is 1, the area inside the 99% contour directly gives us an estimate of coverage. With this method, we find a coverage of approximately 73% for the example displayed in figure 5.7. We also define partial coverage as the coverage of one of the categories area, resulting in four partial coverage per session.

5.4 Experimental protocol

We designed an experimental protocol to study individual behaviors when discovering, exploring, and learning the movement-sound space. We formulated the following hypotheses regarding movement-sound space exploration and learning: (*hypothesis 1*) it is possible to learn the movement-sound space by exploration; (*hypothesis 2*) learning by exploration involves phases of exploration of the space and exploitation of acquired knowledge; (*hypothesis 3*) movement-sound exploration can be prone to ‘blind spots,’ i.e. unexplored region; (*hypothesis 4*) the embodied representation by the participants of the acquired audiomotor map can differ from the sound space representation.

We created a protocol with several exploration sessions. We begin and end with free exploration sessions (S1 and S4) separated by two task-orientated sessions presented below (S2 and S3). Notably, the protocol is not meant to compare the effect of each session independently of each other, which would have required randomizing the order, but rather to quantify the evolution of the exploration through a specific session order identical for all participants. S2 and S3 will serve to verify hypotheses 1 and 2. S2 focuses, in particular, on single-position learning, or landmark knowledge if we refer to navigation terminology, and S3 on trajectory learning, or procedural knowledge. The protocol described lasts approximately an hour and involves one participant at a time. The Ethics Committee of Sorbonne Université has validated the protocol (reference CER-2021-061).

5.4.1 Participants

We recruited 12 participants (6 Female, 6 Male, 3 left-handed, $M=26.4$ yo, $SD=4.1$, age range=21-34) with no previous familiarity of movement-sound interaction technologies. Participants were recruited through the author's contacts. Before starting the experiment, participants gave informed consent to participate and be recorded (video and sound).

5.4.2 Phase 1 - Familiarization with the Sound Space

This experiment's first phase aims to familiarize the participants with the sound space and ensure they can discriminate between the four categories.

Sound Labelization

The first experiment reported in chapter 5 showed that each participant perceived differently each sound class from a semantic perspective, and therefore imposing a name or label made memorizing sounds more challenging. Moreover, discovering and memorizing sounds during exploration added a cognitive task that participants struggled with. Consequently, the participants are invited to “*give a noun or adjective to describe what each sound evokes*” while the conductor plays each sound class manually using the mouse. They write the names of the sound classes in a table organized by categories (four columns) and then choose a category name for each pair of sound classes. They can refer back to this table at any moment during the experiment. The presentation of the sounds is organized by categories, which they also have to name. The sounds are then played again to validate their choices. These labels are used throughout the experiment to designate the various sound classes.

Identification Test

To ensure that participants can differentiate the 4 categories of the study, they then follow a sound identification test. Each category (figure 5.3) is divided into 4 subsquares, which creates a grid of $4*4$ subsquares. 16 test sounds are generated from it using each subsquare's center as (x, y) input coordinates. Therefore, the test sound can either be made of a single sound class or a combination of two classes of the same category. They listen to a test sound and have to press a button to designate the category they think it belongs to. The test consists of several cycles randomly going through the 16 test sounds (urn). The test is considered successful when there is at most 1 wrong answer per cycle after a minimum of two cycles.

The various sound classes proved sufficiently discriminable for all participants to pass the sound identification test with only the two minimum cycles required.

5.4.3 Phase 2 - Movement-Sound Space Exploration Sessions

Once the identification test is successful, participants strap the sensors to their dominant arm to begin the exploration phase. This phase is divided into four 5-minute sessions of unconstrained use of the system. Before the first session, only the surjective property of the mapping is explained: “*you will now explore the sound space you just listened to through the postures of your arm. Each arm posture produces a sound, and different postures can produce the same sound*”.

S1 - Discovery Session

During the first sessions, participants have five minutes to explore the movement-sound space for the first time. They do not have an explicit task to achieve in this session. However, as they know they are exploring the previously discovered sound space, we expect participants to try to find all the sounds with their arms, thus exploring the space with an implicit task in mind that follows from the previous phase.

S2 and S3 - Explicitly task-oriented Sessions

The second and third sessions are explicitly task-oriented. Before each session, participants are presented with a task they will have to achieve at the end of the session. After the explanations, they explore the space again for 5 minutes, preparing for the task, and then immediately carry out the exercise after the five minutes. A sound signal is given to them after 4 minutes. The first tasks were set out as follows:

Task A: Category pointing task. “A category is given to you. You will have to put your arm in a position corresponding to the given category without sound feedback. When you want to validate the position, say ”OK,“ and you will hear the sound that corresponds to this position for 3 seconds. This is repeated 16 times, 4 times for each category.” This task requires the participants to memorize at least one arm configuration for each category (4 categories in total) and be able to recall it without sound feedback. For instance, this ability is analogous to a violinist’s ability to play the right note directly without adjusting the position of their fingers on the neck of the instrument.

Task B: Trajectory task. “You can switch seamlessly from one sound class to the other in a category without going through the other categories. You will be given a category and have to go back and forth from one sound class of the category to the other. You will cycle through each category two times. You will move without sound feedback during the first series and with sound feedback during the second one.” This task, more complicated than the previous one, requires that the participants find each sound class individually and a trajectory inside each category, with smooth transitions, as shown in figure 5.3.

S4 - Non task-oriented Session

For the last session, no instructions were given to the participants. They were invited to use the system for five minutes as they wanted. This open session aims to observe the exploratory behavior of participants once they are accustomed to the space and have acquired a good part of the audiomotor map. In this session, we expect participants to manifest the most diverging behaviors. Will they still explore for new sounds inside of sound classes or for possibilities that the system could offer? Or will they concentrate their attention on a particular zone of the space?

Interview

After each session (tasks included), participants are interviewed about the session and their understanding of the movement-sound space. For the discovery session, the discussion concentrated on the sounds they heard the most, the sounds they did not hear, and the strategy that guided their exploration. For the task-oriented sessions, the discussions focus on the perceived difficulty of the task and the strategy they followed to train for it. For the last session, they were asked what they were focusing on during the session and how their understanding of the system changed since the first session.

5.4.4 Phase 3 - Sketching the Mental Representation

The experiment's last phase aims to elicit the mental representation of the participants' audiomotor map that they built during the exploration sessions. This representation should display the proximity relationships between sounds. A perfect representation of the sound space would be the figure 5.3, but we expect the mapping of the sound to the posture of the arm will distort the space. Participants are given the following instructions to obtain this representation: "*Propose a 2D representation of the different categories and sounds*". They have to draw inside a frame on the same page as the table containing the names of the sound classes and categories they filled in during the familiarization phase. The instructions are purposefully left open to let them decide how to organize their representation.

5.5 Results

We analyzed the exploration sessions, the task outcomes, the interviews, and the sketches. We aim to characterize the exploration sessions and find if similar patterns stand out between participants or if the exploration is user-specific. First, we present the participants' performance on the tasks proposed after sessions 2 and 3. Then, we report on global exploration measures of coverage and presence that give insights into exploration patterns at the group level throughout the study. Finally, we present in-depth analyses of individual maps to gain insights into individual differences and discuss participants' visual representations of the sound space.

5.5.1 Task Performance

Task A: Pointing to Sound Categories

In task A, participants were asked to adopt a position that would result in a point inside a specific category in a randomized order sixteen times, four times per category. Figure 5.8 displays the outcomes of task A for each participant. The results showed that participants could place the categories correctly in relation to each other, but the absolute positions of several categories were shifted (Quantitative results will be shown in Section 5.5.2). Interestingly, the most reproducible positions were found for category II, while categories I and IV outcomes were more spread out. This is an interesting result because these categories could be reached by simply lowering the hand downward or raising it upward, corresponding to the corners of the sound space, as demonstrated by P2 and P10. In several instances, the main diagonal $x = y$ emerges clearly with several points falling on it. It is particularly clear with P1, P4, P6, and P11 results. Interestingly, the third category displays the most shift or spreading, with most points ending up around the central area while being the most covered location during exploration, as shown in figure 5.12.

Task B: Navigating in Sound Categories

The second task required participants to move back and forth between sound classes of the same category. The first trial happened without sound feedback and the second one with sound feedback. The participants' trajectories for each trial are plotted in figure 5.9. Similarly to task A, participants' trajectories are coherently positioned in relation to each other. A perfect result for this task would be four diagonal lines forming an X, with each line going from one corner to the other without crossing a border.

Most participants misplaced at least one sound class, stopped their trajectory somewhere

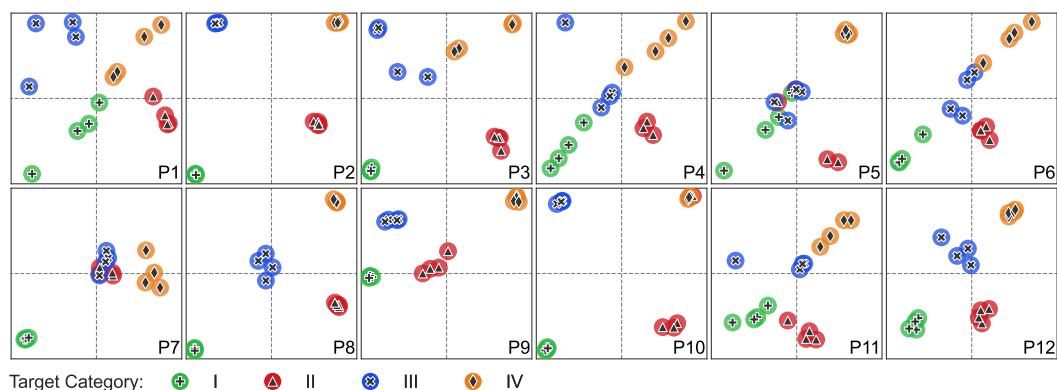


Figure 5.8: First task trials for each participant. They were asked to point to each category in a randomized order sixteen times, four times per category.

between one sound and the other, or overshot a border. For instance, participants tended to move along the y -axis when reproducing the category III trajectory, misplacing at least one of the isolated sound classes. Most participants have overshot one of the inner sounds. For instance, when moving inside category I, P12 overshot and entered category III. When training for the task, participants focused on finding the borders, P3 explained that “*at first I made sure I found a path and then I tried to go a little further. How far does the sound actually go? Does this path go further?*”; and also made sure not to cross them and leap into another category: “*I had to put my arm out to the side, but not too close to 90° because then I would transition to ‘waves’.*” (P4). This behavior made some participant undershoot their target. P8 explained that they felt “*there’s a problem of stability in certain positions. You can hear the sound, but as soon as you move a little, it’s no longer the sound, or there are things added.*” Interestingly, the addition of sound feedback did not appear to improve performance on the task, which supports the idea that they have learned the trajectory and can anticipate the sound feedback rather than being dependent on it.

5.5.2 Quantitative Task Results

To evaluate participants’ success at performing each task, we summarized the outcomes in table 5.1. For task A, a trial was considered successful if the corresponding point was inside its category. For example, P11 placed 1 point wrong for category II and 3 points wrong for category III. For task B, a trial was considered successful if the corresponding trajectory was inside its category, crossed the diagonal separating each sound class, and did not cross one of the borders.

Task A has a success rate of 84%, with ten participants obtaining 12 correct trials or more. Overall, the task was successful, with difficulties encountered mainly in categories II and III. Most errors in category III seem to come from overshooting. As the central area had been identified as the arm extended, precisely stopping the movement before crossing the border was an advanced task relying on good proprioception. With no sound feedback, participants must rely on proprioception only, which is particularly hard when the arm is extended. Task B presents a lower success rate with a high standard deviation. Some participants succeeded very well, while others had more difficulty. The criteria we applied for task B are restrictive. Several failed trials are due to overshooting, even though participants identified a trajectory correctly.

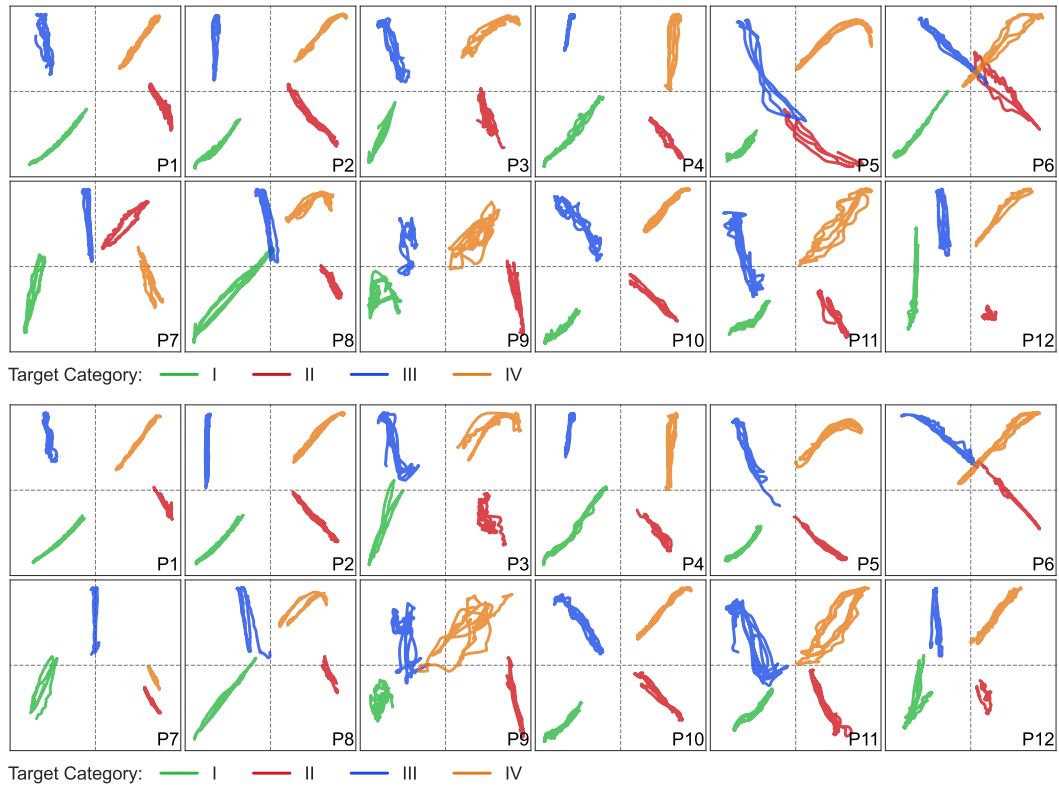


Figure 5.9: Second task trials for each participant. They were asked to move back and forth between sounds of the same category. (Top figure) results without sound feedback; (bottom figure) results with sound feedback.

Overall, the pointing task A was globally successful (success rate of 84%), while the trajectory task B appears to be more difficult (success rate of 63%). Importantly, these results describe the ability to reproduce the absolute position/orientation that participants had to learn through proprioception, and most errors are due to rotational or translational shifts, which was expected for persons with no particular somatic training. Nevertheless, the four different trajectories are distinct for all participants, demonstrating they associated a specific trajectory to each sound border.

5.5.3 Measures of Exploration Behavior

This subsection presents a quantitative coverage analysis calculated using the KDE technique.

Total Coverage analysis

The coverage estimation indicates the percentage of the parameter space participants have visited during one session.

First, we plot the temporal evolution of the coverage in figure 5.10. Logically, we find the result of figure 5.10 at the fifth minute: Session 4 averages a higher coverage than sessions 1, 2, and 3. These plots present an exponential evolution that seems to reach a threshold around the fourth minute. A five-minute exploration time appears to be an appropriate duration for exploring such movement space.

The boxplot of the coverage per session is shown in figure 5.11A. Interestingly, the

Table 5.1: Summary of task results per participants and categories.

Task Category	A				B				A $\Sigma/16$	B $\Sigma/8$
	I	II	III	IV	I	II	III	IV		
P1	3	3	4	4	2	0	2	2	14	6
P2	4	4	4	4	2	1	2	2	16	7
P3	4	4	4	4	1	2	2	2	16	7
P4	4	4	3	4	2	2	2	2	15	8
P5	3	2	1	4	2	1	0	2	10	5
P6	4	4	1	4	1	0	0	0	13	1
P7	4	2	0	2	0	0	1	0	8	1
P8	4	4	2	4	0	1	1	2	14	4
P9	4	0	4	4	2	0	0	1	12	3
P10	4	4	4	4	2	2	2	2	16	8
P11	4	3	1	4	2	2	0	2	12	6
P12	4	4	4	4	0	0	2	2	16	4
MEAN	3.83	3.17	2.67	3.83	1.33	0.92	1.17	1.58	13.50	5.00
STD	0.39	1.27	1.56	0.58	0.89	0.90	0.94	0.79	2.61	2.45
Success Rate									84%	63%

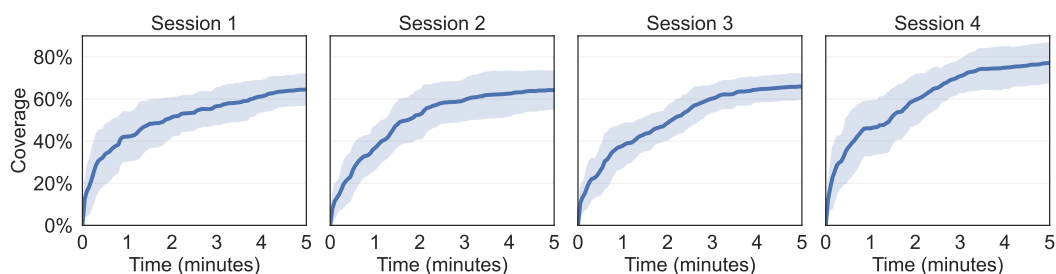


Figure 5.10: Temporal evolution of coverage during exploration averaged over participants. Error bars correspond to the standard deviation.

total coverage does not seem to vary between sessions 1, 2, and 3 but increases for session 4. A repeated-measure ANOVA was performed with the coverage as a dependent variable and a between-participant factor group. The difference between sessions was significant ($F(3, 33) = 13.09, p < 0.0001$). Pairwise t-tests with Bonferroni correction revealed no significant differences between the first three sessions but that they were all significantly lower than during the fourth session (all $p < 0.01$). Therefore, a significantly larger part of the space was explored during the last session, while we couldn't observe significant differences over the first three sessions. While we can expect the total cover to increase over time as participants expand their knowledge of the audiomotor space, it is noteworthy that coverage increases significantly over the last session only.

Presence Analysis

While coverage informs us about the proportion of the audiomotor space participants visit, it does not characterize how much time participants spend in various areas. We computed the proportion of time spent in each quadrant for each session. Since the sampling period is constant, the obtained percentages correspond to the ratio of time spent in each quadrant to the total time of the session (5 minutes), giving, therefore, an indicator of the presence of the participants in each quadrant. An equally distributed exploration would provide a 25% presence in every quadrant. The box plot of quadrants' presence versus session is given in figure 5.12. During the first session, participants seemed to have spent less time in quadrant

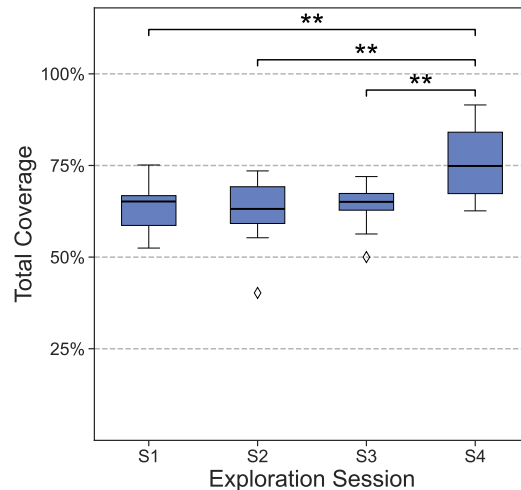


Figure 5.11: Boxplot of the total coverage per session. The mean coverage of session 4 is significantly higher than the mean of the previous sessions.

II than quadrant IV, followed by quadrants III and I. This tendency seems less pronounced in other sessions, particularly for session 3, where the deviation from 25% is less pronounced.

A per-quadrant analysis reveals significant differences in presence in quadrants I and II. Little time was spent in quadrant II during session 1 compared to quadrant I. Both quadrants display significant differences between sessions 1 and 3, with participants spending less time in quadrant I and more time in quadrant II during S3. Interestingly, quadrant III was the second most visited quadrant for sessions 1 and 2 but gave the worst scores on the task evaluation, confirming the inherent difficulty of navigating these parts of the space compared to quadrant IV.

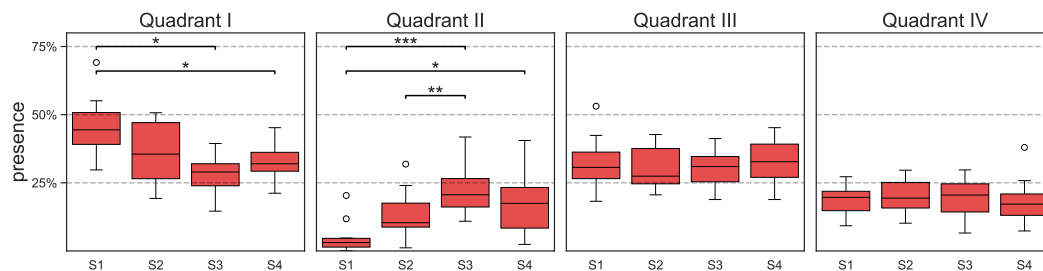


Figure 5.12: Boxplot of the presence per session, split by quadrant.

5.5.4 Individual User Exploration Maps

Overall, measures of coverage and presence computed from the KDE are helpful for gaining insights into exploration behavior at the group level. In this section, we analyze individual exploration patterns through in-depth analyses of exploration maps computed with the KDE. We begin by visualizing and analyzing all 12 contour plots of session 1. We then focus on participants P7, P9, and P10 more closely to evaluate the evolution of their exploration more closely.

This subsection uses contour plots introduced in section 5.3.2. These plots show areas where the highest density of points can be found. There are four brightness levels correspond-

ing to four proportions (from brightest to darkest): 99%, 75%, 50%, and 25%. For instance, the 75% region represents the smallest area containing 75% of the probability mass. In other words, a point drawn from the corresponding distribution has a 75% chance of being in that region. The 99% contour delineates the area considered to have been explored and is used to calculate the percentage of covered space. Contour plots are analyzed in conjunction with feedback participants gave during post-exploration interviews.

Session 1: Discovery Session

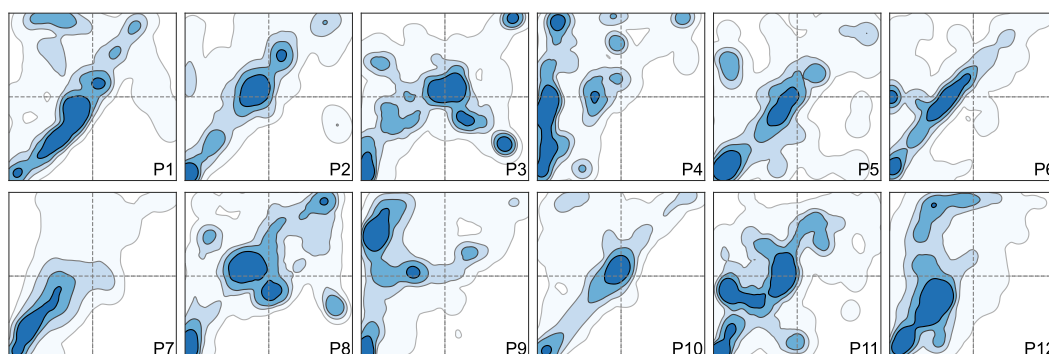


Figure 5.13: Contour plots of the first exploration session.

Session S1 corresponds to the first five minutes of exploration. Before starting, participants were familiar with the set of sound classes. Most participants started with their arms at their sides (it was not mandatory). In this configuration, both sensors point downward; therefore, $\theta_x = \theta_y = 0^\circ \Leftrightarrow x = y = 0$, which corresponds to the bottom-left corner. This starting position appears on the contour plots as the high-density regions at the bottom left.

Contour plots reveal several similarities across participants. The main observation from this first session is that no participant visited all sound classes, which is consistent with our previous statistical results. If we consider sound classes as visited when under the 99% threshold, participants visited between 5 and 6 sound classes. The least visited one was the fireworks at the bottom-right of the space. As highlighted in section 5.2.6, this corresponds to the elbow pointing upward, one of the most unusual configurations in the movement space. It was fully attained by P3 and partially by P8. Several participants (P1-2-5-6-7-10) mainly explored the main diagonal $x = y$, corresponding to any configuration where the two sensors share approximately the same angle with the vertical. To reach the space's bottom-left (resp. top-right), the arm must be extended downward (resp. upward). The points on the diagonal in between can be obtained from various configurations, but the video recordings showed that participants mainly started by exploring the space with their arms extended. Some participants (P4-9-12) explored the space more vertically (y constant), which is done by locking the upper arm and moving only the forearm. Participants reported using this strategy: *"I tried the arm straight up and down, then I tried to see if bending the elbow made a difference, and if I bent it but at different heights."* (P4). Some participants also commented on the qualities of movements they favored during the session; P3's strategy was to *"go very slowly, rather than trying to make large gesture"* and for P9, *"it was more towards the end that I realized I had to go slowly to understand everything."*

Comparing contour plots between participants makes identifying global similarities in explored areas possible. We now analyze these data in greater detail for three participants displaying contrasting behaviors, focusing specifically on the temporal evolution of coverage

and explored areas during exploration sessions.

Individual Participant Analysis

We selected three participants, P7, P9, and P10, who showed different evolutions of exploratory behaviors throughout the experiment. This section focuses on the temporal evolution of the contour plots during sessions. The figure 5.14 displays the six contour plots corresponding to the end of sessions 1 and 4 for each participant. We will focus on these two sessions to analyze differences in exploratory behaviors. In line with global observations from figure 5.11, we observed that these participants explored a larger portion of the space during session 4 than in session 1. However, we do not have any insight into the evolution of the exploration during the session.

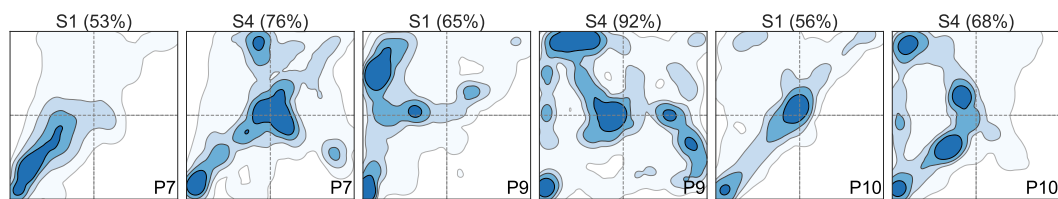


Figure 5.14: Contour plot of sessions 1 and 4 for P7, P9, and P10. Percentages correspond to the coverage

We will now plot the temporal evolution of the coverage and five contour plots after each minute for each session. Participants individual plots are shown in figures 5.15, 5.16, and 5.17. Note that the coverage sometimes decreases over time because coverage is based on outlier rejection. This is particularly salient in P10 first session coverage evolution in figure 5.17. When participants remain in the same location, points initially located inside the covered area at a given time can be regarded as outliers when time passes, thus reducing the area encircled by the 1% contour and decreasing the coverage.

These figures display only the temporal evolution at each minute. Video animations of the distribution evolution have been generated and are available online¹. They are described in Appendix A.3.1.

Participant 7 The first session of P7, in figure 5.15, shows a limited exploration. While 53% of the space has been explored by the end of the session, most of the exploration happens in the first category, with two distribution modes (local maximum of the distribution) at the origin of the space and near the center. The coverage reaches 36% by the end of the first minute and only increases by half over the next five minutes. It increases relatively evenly during this time, finally blending the two distribution modes into a single wider mode.

In comparison, the last session displays a different trend. P7 explores more of the diagonal in the first minute, showing three modes in the center, the bottom left, and the top right areas. Then, in the next minute, the coverage increases abruptly above the session 1 final coverage, showing several modes on the diagonal and lower modes in categories II and III. Eventually, after the third minute, the exploration seems to reach a threshold to converge toward four main modes at 76% coverage.

P7 has shown a significant improvement between S1 and S4. Session 4 plots show a richer comprehension of the space. The exploration remains sequential during S4, increasing

¹https://drive.google.com/drive/folders/1aJmROHvKKQV1LZy3_BwteLnyeAt2aePq?

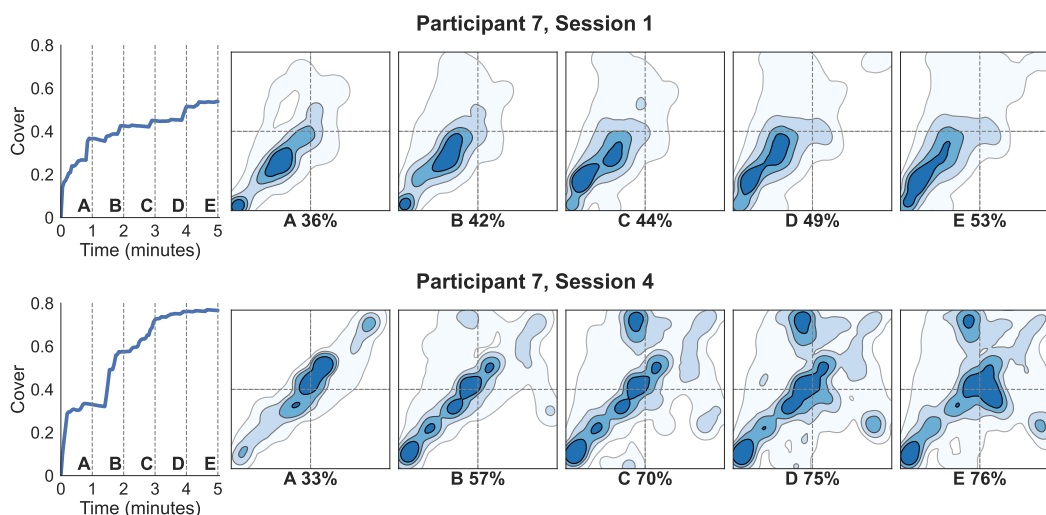


Figure 5.15: Participant 7: (left) Temporal evolution of coverage and (right) five contour plots drawn every minute. Percentages correspond to the coverage

the coverage step by step and slowly expanding the covered area.

Participant 9 P9 started session 1 from a position different from the rest of the participants, as shown in figure 5.16. The exploration began in category III and expanded towards category I and the origin of the space during the next two minutes before connecting with the center of the space and category IV. Similarly to P7, consecutive contour plots show a progressive expansion of the visited space but with more individual distribution modes visited and slightly higher final covered space.

The session 4 shows a different trend. The coverage against time plot shows a fast increase during the first few seconds of the exploration and then a steady increase during three minutes, leading to approximately 90% of explored space. The contour plots at this point show more than nine modes. This final session highlights advanced control and knowledge of the space, suggesting that the 20-minute allocated time was sufficient to acquire a complete audiomotor map.

Participant 10 P10 displays a similar behavior to that of P9 during the fourth session but pushed to the extreme, as shown in figure 5.17. During the first session, in less than 30 seconds, 43% space coverage is reached but finally stagnates at 55% after two minutes, displaying little to no evolution in the contour plots at this point. The fourth session shows a similar trend, reaching 60% coverage after some seconds and stagnating at 68% coverage during the last 3 minutes of exploration.

This exploration pattern significantly differs from P9's and P10's. It actually shows little exploration of the space. The rapid increase in coverage at the beginning of S1 suggests a habit of movement with the arm and comfort in quickly exploring several postures. After this initial burst of movement, the participant concentrated on the same parts explored initially. The figure 5.17 shows the same plots for sessions 2 and 3 to gain insight into the rest of P10's experiment.

The second session of P10 shows a slightly slower cover increase, during which the participant explores the space and finds the second category. After this, the coverage again

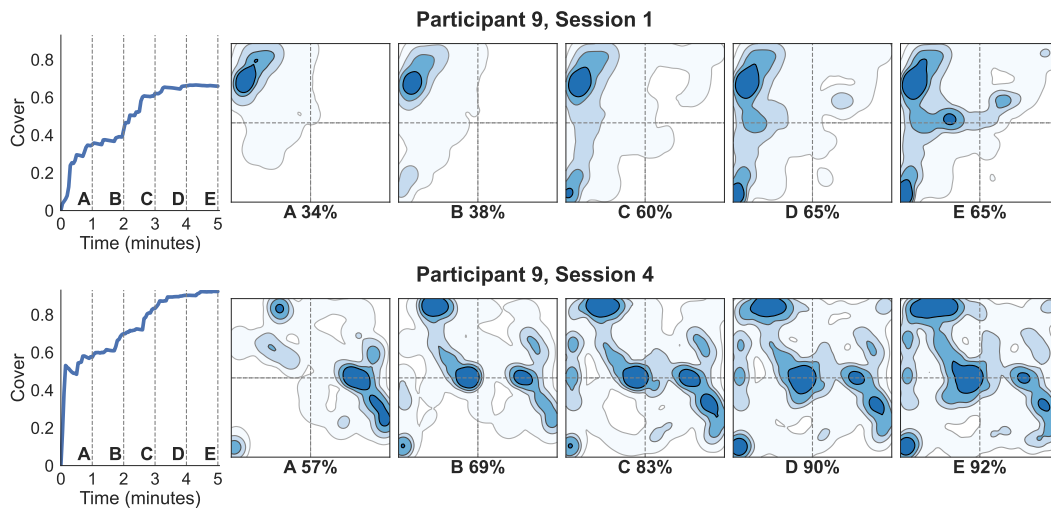


Figure 5.16: Participant 9: (left) Temporal evolution of coverage and (right) five contour plots drawn every minute. Percentages correspond to the coverage

reaches a plateau before increasing a little more. Session 3 resembles the initial trend with a quick rise to 50% coverage and a slow increase to 63%. It is interesting to note the X shape of the distribution, clearly showing the training for the second task, to which P10 obtained a perfect score. This slow evolution during the session illustrates the balance between exploring space and exploiting the data discovered to prepare for the task. P10 explained in the interview following S3 that after finding trajectories, “*then I’d do it again. And I’d try to really feel the position of my arm in the space for the swell [the water sound class], feel that the hand is higher than the shoulder...*”. Through repetition of a previously discovered action, P10 wanted to deepen his understanding of space in preparation for the coming task.

Exploration trends The contour plots obtained with the KDE technique showed a significant interest in analyzing the participants’ exploration. Added to the temporal evolution of coverage, they formed a relevant tool to highlight participant differences. In particular, we identified different exploration patterns: from a slow and steady increase in coverage to a fast rise, which then peaks.

5.5.5 Sketches Analysis

Results presented so far address the way the control space was explored. However, participants were never exposed to the shape of this control space during the experiment. In particular, they do not know it is 2-dimensional. To better characterize what participants understood of the audiomotor space, we asked participants to “*draw a 2-D representation of categories and sounds*” after the four exploration phases. All sketches are displayed in figure 5.18.

Sketches elicit various representations of the space. It must be noted that the colors were often associated with the sound labels written in a table under the sketch space at the beginning of the experiment. First, the point of view differs between participants. P2, P4, and P6 represented their body and the different configurations that allowed them to hear sound classes. All other participants represented a projection of the audiomotor space with zones representing sound classes. In these drawings, several borders were drawn between sound classes, which can be compared with existing frontiers in the sound space.

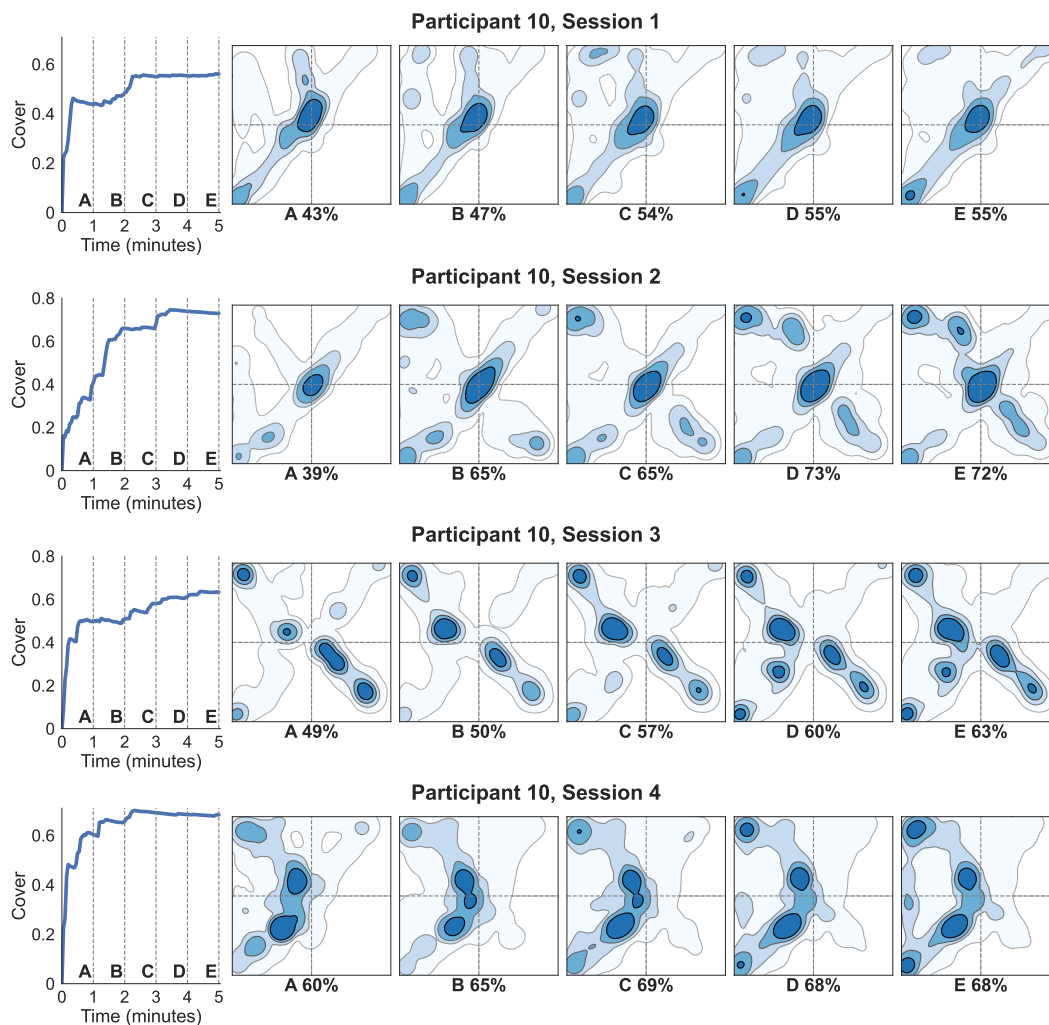


Figure 5.17: Participant 10: (left) Temporal evolution of coverage and (right) five contour plots drawn every minute. Percentages correspond to the coverage

Table 5.2 lists for each sketch the number of borders represented and how it matches actual borders in the sound space, given that not all sound classes share a border. We consider valid frontiers the ones between sound classes that are either in the same category, sharing a horizontal or vertical segment, and the ones connected by the central point, according to figure 5.3, which allows for a maximum of 10 valid frontiers. Participants represented, on average, three-fourths of the possible frontiers; the ones splitting categories were always represented (horizontal and vertical segments). The results indicate that participants have built a relatively accurate, if incomplete, representation of the sound space.

Interestingly, although analysis of the sketches by borders shows a common understanding of the space by participants, their representations, once projected on the paper, present a variety of shapes. P5, P8, and P12 visualized the space in a circle or a sphere. P8 stated at the end of session 2: “the space must be circular” and kept this representation until the end. Participants P5 and P9 proposed symmetric representations. In both cases, they described it as an axial symmetry projected to a sphere rather than a plane, “the same in every direction” (P5). P10, P7, and P3’s representations are symmetric, with the exception of fireworks and forests in each instance. Finally, several representations showed holes in between categories.

task, and many participants struggled with memorizing precise angles of the arm once fully extended in front of them. However, while many had low task B scores, the position of individual sound classes and the trajectory between them was known. Moreover, drawings showed that the global structure of the sound space had been understood, with participants representing around 75% of frontiers between sound classes.

From the participants' perspective, they all reported learning about the space between the beginning and end of the experiment. In the final interview after session 4, participants were asked, "*Has your understanding of space changed since the first exploration, and if so, how?*". Some participants felt they understood the space entirely: "*In the end, I feel like I completely understand the space. I could predict the sound each position of my arm makes.*"(P10). Others have achieved some level of comprehension but feel that they did not fully understand the interaction: "*I still haven't figured it all out. I know where the sounds are and... how the movements work... according to these sounds.*"

Participants were able to acquire knowledge of the movement-sound interaction space through exploration. These results are consistent with Van Vugt and Ostry's conclusion that the early learning of an audiomotor map is possible with an exploratory learning process and with discontinuous maps (van Vugt and Ostry, 2019). Below, we discuss the process of learning by exploration, the notion of navigation in movement-sound interaction, audiomotor representations, and participants' experience.

5.6.1 Exploratory learning implies open-ended exploration, task-driven exploration, and exploitation

Participants learned the interaction between movements and sounds by exploring the sound space with their arm during four 5-minute sessions. Returning to the literature about exploratory learning, these exploration sessions fall into what Rieman describes as investigating the possibilities a system offers without following "*precisely sequenced training materials*" (Rieman, 1996). In which case, users "*direct their own learning*" (Wiedenbeck and Zila, 1997). However, although this definition sets the framework within which the system was used, it does not provide information on participants' behavior during the experiment. Being free to use the system does not imply they have moved around the space aimlessly. Two types of explorations are differentiated: task-driven exploration and open-ended exploration.

We explicitly presented the mosX system to the participants as a tool enabling embodied interaction with a set of sound samples they discovered beforehand. Therefore, their only known component before the first exploration was the sound corpus and the fact that they could explore it by moving their arm. This exploration was open-ended in that the exploration's primary driver was understanding the interaction. While an objective with a clear performance metric could have been proposed to participants before the exploration, for instance, "*find as many sounds as possible in 5 minutes*", we were more interested in the way participants would react by themselves to the system and what approach they would choose to make sense of their experience.

The results showed that participants remained mainly in the first and third categories during the first exploration session. After the first session, they were asked if they had heard all the sounds. Interestingly, most participants did not have an immediate answer and began with the affirmative before reviewing the list of sounds and realizing they did not hear some of them. Only three participants were immediately sure that they did not hear certain sounds. Therefore, it seems that during this first discovery session, participants did not focus on trying to find all sounds. Interviews rather show that they aim to understand the interaction by

consciously trying specific movements. Interestingly, participants favored meticulous and slow-paced exploration over more energetic and frantic movements, which might have given them a better first overview of the space.

The second and third sessions were explicitly designed with task-oriented exploration. A certain level of mastery was expected at the end of the five-minute exploration period. Interviews showed that they started alternating between exploration and exploitation. “*When I made the gesture, I repeated it in my head. I think even orally. And then I would revise it as long as I had time.*” (P9). Exploration sessions became “*training sessions*” (P11), and the discussion centered on the strategy they adopted to prepare for the task at the end of the session.

The last session aimed to observe how participants would behave when free to explore the space without training for a task but with knowledge of navigating it. P10 described their process during S4 as “*trying to find other unexplored positions*”. Several participants, like P10, returned to a process of exploration. P9 explained that “*the last two times, I was more into ‘I’ve got to memorize, I’ve got to memorize’; this time, I was back to discovery.*”. They either state it as open-ended exploration or they set new tasks for themselves: “*What other sounds connect ?*” (P3), “*I wanted to see if I could move seamlessly from one category to another.*” (P5). This also includes exploring certain features of the space, such as “*looking for more detail in the sounds*” (P10).

Interviews with expert practitioners reported in Chapter 3 highlighted that open-ended exploration is common in artistic practices where the interest of a system does not lie in its ability to assist in carrying out a task but instead in the way it will nurture creativity and enable expressivity. Still, some level of training is always required to master a system, as described in this chapter. Contour plots and interviews reveal that, depending on the context of the session, participants used several modes of exploration, sometimes during the same session, alternating between open-ended exploration, task-oriented exploration, and exploitation.

5.6.2 Navigation of audiomotor spaces

Audiomotor map acquisition literature provides information on the mechanisms behind learning the relationship between movement and sound. Movement-sound exploration can be analyzed through the notion of *navigation*, which brings interesting perspectives regarding the nature of acquired knowledge. Three hierarchical levels of spatial knowledge are built when learning how to navigate a space (Darken and Sibert, 1996):

- landmark knowledge, the notable perceptual features of specific locations (usually visual features);
- procedural knowledge, the sequence of actions that one must perform to connect isolated bits of landmark knowledge;
- survey knowledge, a topological view of the previous information, where the object locations and inter-object connections are coded from a geocentric perspective and not from the first-person frame of reference (egocentric) from which landmark and procedural knowledge are constructed.

We argue that participants displayed several levels of knowledge that can be similarly classified. When navigating the audiomotor space, participants had constant sound feedback linked with the configuration of their arm. In the context of blindfold sonic interactive

exploration, the only positional clues participants have are proprioception and sound. During the first exploration session, participants built landmark knowledge when they found sound classes for the first time. In this case, the pair sound-posture is recorded as a landmark. P2, for instance, reports on these pairs after session 1: *“I found the matches at the bottom. I finally found the tension, it was at the very top. There was the equator. The crowd was somewhere over there.”* The task 2 demonstrates the procedural knowledge that must have been built to achieve it. Eventually, sketches display several levels of knowledge. Some drawings represent the participant showing sound-posture pairs (P2, P4, and P6), while others display map-like views of the sounds and categories, suggesting learning of survey knowledge.

It must be noted nonetheless that survey knowledge considered as a purely geocentric view might necessitate a different approach for such sonic embodied interactions. In this case, the frame of reference is not outside the body but the body itself. Therefore, expecting the user to build a non-egocentric representation of the space seems complicated. Moreover, navigation is interested in optimizing the information or location research process. The locations are discrete, and we are not necessarily interested in what lies between two discrete memorized locations. In audiomotor map acquisition, interpolating between known relationships is possible (van Vugt and Ostry, 2018a). In continuous movement-sound interactions, there is infinite space that lies between two visited and memorized movement-sound relations. However, Van Vugt and Ostry’s study suggests that interpolation is only rarely made, and learners tend to remain in part of the space to which they have had exposure before (van Vugt and Ostry, 2018a)

5.6.3 Stimulating Embodied Interaction

The mosX system proved simple enough to be learned in a short amount of time while proposing stimulating and diverse experiences to participants. The protocol used here allowed participants to discover the space progressively, *“it’s pretty radical I think, because at the very beginning, I had no idea what was going on, and then I said to myself, my brain isn’t going to understand what’s going on [...] and now, at the end, I feel like I completely understand the space. I could predict the sound each position of my arm makes”* (P10).

Interestingly, a recurring theme in interviews with artists reported in Chapter 3 appeared in the discussions post session 4: *“What changed was that I didn’t even think the tool was capable of capturing this kind of stuff. So what changed was that I challenged it more. I’m trying to see if he understands that I’m behind it. [...]Yeah, I challenged it more.”* (P9). This idea of challenging the limits of an interactive system was pointed out by Bertha Bermudez, choreographer (Chapter 3): *“One interesting thing is that we always try to do two things: one, we understand the system, so we do it, we are obedient, [...] and then we try to break the system, and see when it doesn’t work. [...] Almost all the dancers did the same.”*

Other participants decided to explore different features of the system and just wanted to enjoy the experience: *“you told me to do what I wanted so [...] I wasn’t trying to do anything special, I had to have a bit of fun with it.”* (P11), *“For example, in the vocals, it sounded like ‘tarataratata’, so I played around with it. And then, at one point I listened to the swell, because I liked the swell.”* (P10)

Through the mosX system, this experiment offered a new experience to the participants. They had to be simultaneously conscious of their arm joints and listen to the sounds. Participant P1 said after session 1, *“I’m not sure; I’m not used to being aware of my body. At first, I mostly did extended arm movements, and then I experimented a bit with bending the elbow.”* We believe such embodied interaction opens possibilities on many different

levels depending on the background of users. Learning by exploration leaves an opening for emerging experiences and uses that are crucial to art creation but also foster unexpected usage outside of artistic contexts.

5.6.4 Limitations and Perspectives

The exploration analysis method based on kernel density estimation techniques proved valuable in providing quantitative spatial measures. Still, it did not provide quantitative measures related to the temporal evolution of the distribution. Information based on points of interest and trajectories between those points could further qualify the exploration in determining if the same points were visited regularly for instance. Ultimately, extracting all those features could better enable the classification of the type of exploration the user leads.

The movement-sound interactions we designed relied on 2-dimensional control parameters space. The methods we developed focused on the parameter space and aimed at manipulating and analyzing 2-dimensional spaces. The question of generalizing to higher-dimension parameter spaces naturally arises as sound synthesis can rely on high-dimensional parameter spaces. KDE techniques were thoroughly studied in higher dimensions (Scott, 2015), and N-dimensional implementations of KDE already exist, such as fastKDE² or KDEpy³. However, the interpretation of these high-dimensional estimated distributions would be difficult. Visualization would require dimensionality reduction techniques such as t-SNE or UMAP (McInnes et al., 2018).

The *mosX* system proposed a simple interaction based on the orientations of the sensors. The interaction was extended in the final interactive system of *Sympoiesis* using intensity and sound effects on the output. While early learning studies already provided valuable information on exploratory behaviors from participants, a longitudinal study such as (Zappi and McPherson, 2014) could be relevant to highlight the evolution of exploration mechanisms during the appropriation of technology. To set up such a study, *mosX* would have to be made more open to customization without having to be an expert with the *Max* software. We can draw inspiration from (Zappi and McPherson, 2018) to design a “*hackable*” *mosX*.

5.7 Chapter Summary

We presented an empirical study focused on movement-sound exploration. Following the preliminary experiment reported in Chapter 4, we were interested in quantifying and qualifying the exploration of movement-sound spaces in exploratory learning contexts. Our first contribution is to propose a movement-sound interactive system called *mosX* that uses an abstract audio space controlled with wearable sensors. Second, we established an analysis method using kernel density estimation (KDE) to visualize exploration data. Third, we proposed an empirical method to evaluate the participants’ internal representation of audiomotor maps based on their drawings or sketches. This experiment revealed significant insights into how individuals learn, represent, and navigate complex sound spaces through embodied interaction. The findings underscore the effectiveness of exploration-based learning in early audiomotor map acquisition. In particular, our KDE-based method produced quantitative data and visualizations that enabled the identification of different exploratory behaviors between participants.

²<https://pypi.org/project/fastkde/>

³<https://kdepy.readthedocs.io/en/latest/>

CHAPTER 6

Artistic Applications of Movement-Sound Interactive Systems

6.1 Introduction

Interviews with practitioners reported in Chapter 3 showed the power of movement interactive technologies in structuring the development of the practice and their capacity to foster the unexpected. While these results emerged from artists who spent several years experimenting with interaction technologies, how do these ideas translate into creative contexts where little time is allowed for interaction design and most creative collaborators are new to such technologies and practices? Can interaction play the same role, in whole or part, by structuring and bringing novelty to creation? I propose to address this question with a first-person perspective on the creation process of two performances involving movement-sound interactions between performers and electronic music: *Sympoiesis*, a bio-inspired performance with four dancers, and the *Circus Lab*, a three-day workshop in collaboration with circus artists.

The experiments described in Chapter 5 have demonstrated the value of exploratory processes for discovering and appropriating movement-sound interactions. The performances presented here enable us to consider exploration in a broader scope, particularly as a shared process in collaborative settings. During the creation of *Sympoiesis*, we organized two workshops to try different interactions through exploration, fostering a participatory design with the dancers and the director. Eventually, we used free exploration as a performative tool for the final performance, using the mosX system developed in parallel with the creation. From this process emerged several research questions: **(1)** What are the challenges of incorporating movement-sound interaction technologies into established practices in the frame of a single performance? **(2)** How can free exploration help this process? **(3)** How does free exploration differ from other types of interactions from the performers' and spectators' points of view?

Using first-person research methods, I report on the creative process of the two performances by examining the performers' and other creative collaborators' reactions to movement technology, their preconceived ideas of what movement-sound interaction meant, and their motivations behind including such technologies in their practice. Their feedback and the design process analysis during creation provide interesting results that address questions (1) and (2) and resonate with interviews presented in Chapter 3. The research question (3) is answered in light of the spectator experience design taxonomy proposed by Reeves et al. (Reeves et al., 2005) and by comparing the mapping used with different setups both at the *Circus Lab* and *Sympoiesis*.

6.2 Methods

This study falls within two methodological fields: first-person research and performance-led research. First-person research methods involve the researcher using their own experiences and personal knowledge as primary data sources (Lucero et al., 2019). It includes, among others, autoethnography, where the researcher analyzes their personal experience systematically to understand broader cultural phenomena (Cecchinato et al., 2017; Höök, 2010) and autobiographical design, where researchers evaluate certain design artifacts (or absence of such artifacts) within the context of their own life (Desjardins and Ball, 2018; Sengers, 2011). First-person research is usually longitudinal. It is made of long data collection periods based on the researcher's experience and often requires significant time to allow the researcher to "step back" and process their data critically (Lucero et al., 2021).

Performance-led research aims to study the integration of interactive technologies in artistic practices, particularly in the performing arts. These methods involve artists in developing the research questions and the design process so that the knowledge emerges from the practice. Several examples of performance-led research in dance have shown the value of these methods in developing the relationship between interaction technology and artistic movement practices, as well as bringing out concepts that are ultimately more general than individual performances alone (Goodman et al., 2011; Fdili Alaoui, 2019; Sullivan et al., 2023)

This work combines these approaches to the design, development, and co-creation of two performances. I collaborated with artists to include interactive movement-sound interactions in an iterative process that enabled us to align the interaction design with the broader artistic objective of the performing art pieces. In this chapter, I describe the experiments I led while reflecting on my personal experience within these collaborations. This gives me insights into questions about integrating exploratory behaviors in movement-based practices.

6.3 *Circus Lab*: Collaborative Interaction Design with Shared Sensor Data

6.3.1 Context and Motivations

This section presents a collaborative project between six circus artists and seven electronic musicians. This three-day workshop, *Circus Lab*, aimed to create a one-hour performance mixing contemporary circus art and electronic music technologies, particularly deep learning-based sound generation. It was played two evenings right after the workshop. Most musicians were also computer music researchers who specialized in machine learning for sound analysis and generation. I was invited to participate in this project as one of the musicians and as an

interaction designer, using movement-sensing technologies to create music that would react to the circus artists' actions.

This workshop was a discussion between circus arts and electronic music technology, in particular machine learning generative models and interactive technologies. The performance consisted of several pieces, joining together some of the electronic musicians and one or two circus artists and their equipment. This section will focus specifically on four pieces with the following practices:

- the Chinese Pole, a vertical pole on which a performer performs a variety of spins, climbs, and balances;
- the Cyr Wheel, a large metal hoop manipulated by a single performer who either spins inside of it or manipulates it from outside;
- the German Wheel, a large wheel made of two hoops in parallel connected with six spokes where a performer either spins inside of it or manipulates it from outside;
- the Swinging Trapeze, a trapeze bar suspended on ropes that allows a performer to swing, drop, catch, and perform complex aerial maneuvers.

Each piece consisted of one practice with a subpart of the musicians who composed and improvised a musical track in collaboration with the circus artist. Since we only had one rehearsal space available, we could only work on one piece at once, which, combined with the fact that we only had three days of residency, gave each piece only a couple of hours to be created.

In this context, performers wanted to include movement-sound interactions for the experience they could bring to the audience as an amplifier of actions to reinforce movements through sound. Therefore, the focus was on impactful effects visually related to the performers' actions. In order to design these interactions quickly, one of the researchers suggested that I distribute processed movement signals from the sensors to the electronic instruments of the musicians so that they take care of the sound design.

This project was facilitated because all the other musicians also specialized in computer-assisted electronic music and, in particular, used the Ableton Live software. Live proposes to program custom effects and instruments using Max. This visual programming language is widely used in music and multimedia creations (the language that I used to create mosX). The adapted language inside of Live is called Max4Live (M4L). This enabled me to program M4L objects, directly importable by the musicians into their existing work sessions, and which would act as portals through which I could pass the movement signals coming from the performers to the musicians.

6.3.2 Shared Sensor Data Through Max4Live Objects

I programmed four M4L objects for the musicians. These objects could receive raw data from a sensor sent as an Open Sound Control (OSC) package through WiFi or ethernet. The data received from the sensor included accelerometers, gyroscopes, and magnetometers output in 3 directions with Euler angles, directly calculated by the sensor. They were processed directly inside the objects to calculate each measure and expose them to *Live*. I focused on three descriptors of the data coming from the sensors: the rotation speed (from gyroscopes), the orientation (Euler angles), and the movement intensity (from accelerometers). Three objects enabled musicians to access these measures and map them to parameters of effects or instruments directly in *Live* using the “Map” button as displayed in figure 6.1. A fourth object,

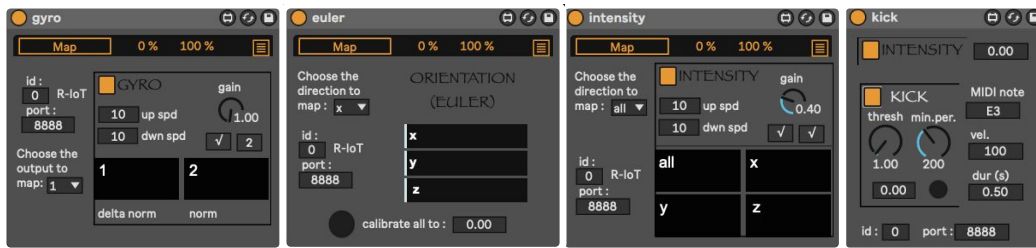


Figure 6.1: Set of Max4Live objects created for the Circus Lab performance. Each object allowed the musicians to receive raw data from the sensors and process them to extract certain features of the movement (left to right): gyroscopic data, Euler orientation angle, movement intensity, and impact movements.

called “*Kick*”, used movement intensity to detect impacts in movement. The interactive setups presented below do not use the orientation data.

6.3.3 Interactive Setups

The final performances included seven pieces. This section describes four interactive setups that propose different signals for sound mapping. Each used a single sensor either strapped to the performer or to the apparatus. The sounds resulting from the interaction described are always only a part of the music played during the piece.

Chinese Pole The sensor was strapped to the performer’s ankle. Two different motions were sensed and mapped to sound: the free fall along the pole and the fast rotations around the pole. When falling, an impactful sound was played just when the performer stopped the fall and when the performer rotated around the pole it raised the cutoff frequency of a low pass filter on a pulsating synthesizer.

Cyr Wheel The sensor was strapped to the performer’s wrist. The movement intensity was mapped to the gain of the resonance of a filter on a synthesizer. Two movements in particular would raise the intensity. The first one when the performer was making the wheel turn in front of him and stopped it suddenly with his hand. The second one when he was inside the wheel and made quick rolls on the floor.

German Wheel The sensor was strapped to the wheel. The wheel rotation speed was mapped to the cutoff frequency of a low-pass filter on a filtered white noise. When the performer made the wheel roll, it produced a wind-like sound, both when rolling on its edge and on its face.

Swinging Trapeze The sensor was strapped to the performer’s arm, measuring the intensity of her movement. The signal was mapped to the pitch and the tempo of an arpeggiated chord.

Design Process The interactive setups introduced here use two descriptions of movement: the intensity, calculated from the accelerometer data (described part 6.4.5) and the rotation speed calculated from the gyroscopes with a smoothing filter. The quantity was selected according to the types of movements used by the performer.

In the case of the Chinese Pole and the Swinging Trapeze, the performer evolved around a “static” apparatus. Static in the sense that it was either fixed or in a repeated motion, nearly unaffected by the performer’s actions. In the case of the trapeze, other performers were in charge of maintaining the trapeze swing through a system of ropes. The actions of the performers consisted of using the apparatus as an external aid to perform spins, climbs, or falls. We decided to sonify the movement of the body to amplify certain qualities of these acrobatics. On the Chinese Pole, the strength and impact required to stop the fall and the speed of the spins around the pole. On the Swinging Trapeze, the whole movement of the trunk was measured from the upper arm, highlighting speed variations in the various figures.

In the case of the German Wheel, the apparatus was external to the performer. The performers’ actions aimed to influence the wheel’s dynamics, which she could use in return to perform certain acrobatics. Therefore, by strapping the sensor to the wheel and sonifying its rotation with a wind-like sound, the interaction reinforced its presence and feeling of weight and inertia.

We took the same approach to designing the Cyr Wheel interactive setup, but we reinforced the actions of the performer rather than the movement of the hoop. The performer intentionally directed his performance toward movements that would accentuate the interaction, such as hitting the wheel with his hand to change its rotation direction and maintaining his body inside the hoop while rotating like a coin on the floor.

6.4 *Sympoiesis*: Movement Sound Exploration

This section presents the design and outcomes of a part of an art-science collaboration called *Sympoiesis*, with the costume maker and director Laurane Le Goff (LL) and four dancers, Megane Eyles (ME), Aske Ploug (AP), Justine Prignot (JP), and Timea Szalontayova (TS). LL imagined this bio-inspired performance for her MA graduation project at Central Saint Martins (CSM) in the years 2021-2022. We presented it in June 2022 at CSM and in various venues throughout the rest of the year. I was the composer and the musician for the performance. This creation appeared as an interesting field for experimentation in movement-sound exploration. The dancers and the director had no prior experience in computer science or interactive systems. During part of the piece, the dancers used the *mosX* system in a simple configuration with a single sensor per dancer. The interaction development took place during two workshops adjacent to rehearsal sessions. The following sections introduce the context of this creation, the two workshop sessions, and the subsequent design choices for the interaction during the performance.

6.4.1 *Sympoiesis*: a Bio-Inspired Dance Performance

Sympoiesis is a dance performance conceptualized by LL. She introduces the name *Sympoiesis* in these terms (Le Goff, 2023):

Coined originally by Beth Dempster (Dempster, 2000), this term implied ‘multi-species making together’. ‘Co-making’ instead of auto-poiesis, which is the act of creating things alone. The performance emerges from the relationship between myself, a plant, a slime mold, four dancers, and a sound designer.

The production is divided into two parts, with a total runtime of seventeen minutes. During the first four minutes, a solo dancer embodies a plant, an *Alocasia zebrina*, and tells its life story. The latter thirteen minutes depict a dance by the four dancers, set to the movements of

a slime mold, the *Physarum polycephalum*. By enacting the stories of non-human beings, the performance prompts the audience to develop a sense of empathy toward the life narratives of these living organisms.

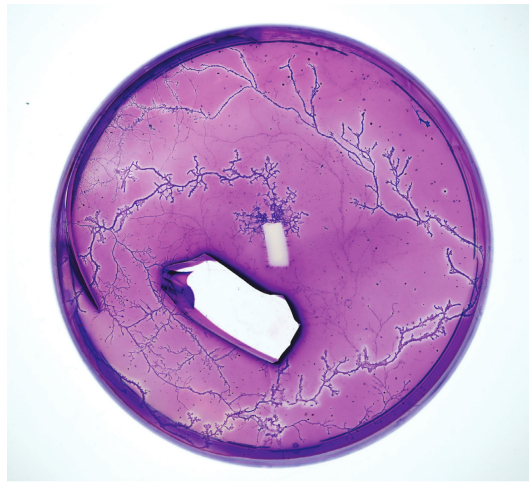


Figure 6.2: A plate of agar dyed to reveal the slime mold trail. Courtesy of Laurane Le Goff.

Enacting the *Physarum polycephalum* This section focuses on the second part, where dancers enact the *Physarum polycephalum*. It is a species of slime mold, also known as the “*many-headed slime*.” It is a single-celled organism that can grow large and has a unique ability to navigate its environment in search of food sources. *Physarum polycephalum* has been studied extensively as a model organism for understanding biological processes such as cell communication, decision-making, and network formation. The figure 6.2 pictures the path that a *P. polycephalum* has taken inside an agar plate when expanding and looking for nutrients to survive.

During this latter part, the choreography aims to tell the story of a *P. polycephalum* developing in a petri dish. During this stage of its life, this organism looks for nutrients by expanding outward. It creates a network of branches to explore its surroundings and look for the closest food source. This slime mold is particularly known for its cytoplasmic streaming, a liquid flow rhythmically going back and forth. This pulse supports the migration of the system, enabling the long-range transport of molecules. It is even used as a signal transport during starvation to redirect the migration flow toward the most promising direction of survival. This mechanism makes it capable of solving shortest-path problems and mazes (Tero et al., 2010).

Staging Before the first workshop, it was decided that the dancers would evolve in a circular space with an array of six speakers evenly spread on the edge and pointing inward. This configuration allows both the dancers to hear the music and sounds from the interactive system directly on stage and the audience to perceive the sound as coming from the circle regardless of their position around it, as shown in figure 6.3 and 6.5. Each speaker had a dedicated channel, which allowed for the spatialization of the sound around the circle.

6.4.2 Intentions

As an interaction designer on *Sympoiesis*, I wanted to ensure that the technology was as non-invasive as possible for the dancers and on a second plane in the performance. As an

artistic collaborator, I also accepted that the role of technology in the final performance might be potentially marginal. The workshop and rehearsal with the dancers would already represent a valuable source of field experience for my research.

Interviews from chapter 3 presented a variety of reasons to experiment with interaction technologies in the first place and revealed that it was mainly contextual to each practitioner. Here, the main challenge lies in the significant engagement required to make technology a meaningful part of the practice. Finally, my main objective was to try and find interesting propositions from my experience with this technology to support the narrative of the piece while avoiding technological interventions from taking too much attention over the main subject of the piece.

6.4.3 Working Method and Organization

The entire piece was created in a year, with twelve dance creation sessions running from November 2021 to May 2022, for a premiere in June 2022. I joined the project from the start and was part of the first discussions but could only go to London for three sessions: the sixth, the tenth, and the last one. The workshops took place during the first two sessions I attended; the last one was dedicated to the dress rehearsal. During the first workshop, we were able to take time to discuss the interaction in depths that did not concern the piece alone. This first workshop was exploratory, with no preorganized structure and no defined goal. As the premiere approached, the second workshop was included in a larger work session dedicated to converging toward a complete piece structure. It presented less open discussions on personal experiences with the proposed interaction because it aimed at settling on an interaction design choice.

6.4.4 Data Collection and Analysis

During the workshops, we recorded a video of each dance session and recorded the discussions that followed. The results presented below come from my analysis of personal notes I took during the session, notes the director took, and the transcript of dancers' oral feedback after each session. The results presented below focus on three aspects related to the interaction: the dancers' experience as performers, the influence of the interactive setups on the dancers' behavior from the observer's point of view, and the connection with the narrative. These three themes were recurrent in the discussions and highlighted the group's priorities when discussing the interactive setup.

6.4.5 Exploratory workshop: First Contact with the Interaction

This first workshop happened during the sixth dance rehearsal and was the first contact between the dancers and myself. This first encounter aimed at showing the sensors to the dancers and trying different interaction scenarios to observe their reactions and get feedback on their experience when using them. It was also the first time LL, the stage director, saw the interaction live and could think of ways to incorporate it into the piece. One of the dancers (JP) was not able to dance but attended the session and gave external feedback during this workshop.

Setup

As the piece's structure had not been decided yet, there was no original composed music for this session. I had several synthesizers ready in Ableton Live that I could trigger with an

external MIDI interface. We also used music from Steve Reich's *Electric Counterpoint* to help structure improvisation.

Sensors The sensors used were the same as for the experiments described in chapter 5. They were inertial measurement units called R-IoT that could measure and send acceleration, gyroscopic, and magnetic data wirelessly through WiFi.

Sounds Synthesis and Corpus The sounds used for the interaction come from sources similar to the ones used in the experiments described in chapter 5: outdoor environment or interaction with objects recordings. These sounds evoked a certain aspect of the organism: the natural environment in which they grew, the laboratories in which they were studied, or the way they spread for instance. Three sounds that seemed easily differentiable were selected for this session and assigned to each dancer. They were cropped to 40-second-long files and normalized in loudness. These sound files were then chopped into 500ms long “*sound grains*,” which are the basis for the sound synthesis. The interactive setup for this piece relied on the principle of concatenative synthesis, which consists of concatenating those short sound grains. In concatenative synthesis, the selection of the grains to be played depends upon the analysis of the sound content of the grains. The selection is driven by a criterion specifying certain expected sonic content qualities. Periodically, the sound grain that best matches the criterion is selected and played.

Interaction Design The interaction for this session was deliberately kept simple. Dancers already had to focus on the director's directions, it was their first time using such an interactive setup, and they had to improvise together. From the early experiments I had done at this point and the insights from the interviews, it seemed clear that interactive systems were cognitively demanding and required space for attention. Therefore, I decided to use an interaction based on movement intensity with only one sensor per dancer.

The interaction was implemented in the Max software using the Mubu package and its *pipo* externals. On the sound processing side, each sound file was chopped into 500ms long chunks, and each chunk's mean energy was calculated using the *pipo* module *descr*. On the movement processing side, the dancer's movement intensity was calculated from the acceleration received from the sensors using the *pipo* module *intensity*. It works by removing the offset caused by gravity from the acceleration (using a high pass filter or by derivating the acceleration) and filtering the absolute part of the remaining signal through a low pass filter to get a smooth signal related to a perceived movement intensity. The final intensity corresponds to the quadratic mean of the intensity in each direction. On a fixed period, the intensity of the movement was mapped to the sound energy. The intensity and the energy were rescaled so that an arbitrary maximum intensity (calibrated with the dancers) would correspond to the maximum energy measured in the grain. Five percent of this maximum intensity would correspond to the lowest energy measured in the grains. The figure 6.3 depicts this process.

Using this mapping, a sound grain was periodically played based on the movement intensity measured. The triggering of these consecutive grains resulted in a new texture following the apparent movement intensity of the dancer. The process is depicted in figure 6.3. Note that in this figure, the grains duration and the period were matched for clarity, but in reality, sound grains lasted longer than the playing period. The duration was 500ms for a period of 80ms. Therefore, grains ended up superimposed, which resulted in a denser texture but with the same energy-to-intensity link.

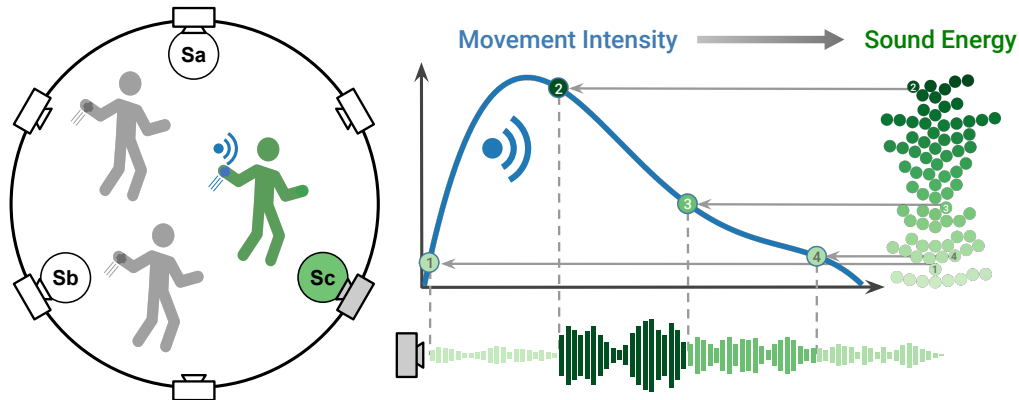


Figure 6.3: Workshop 1 interaction setup based on concatenative synthesis. Each dancer wears a wireless accelerometer from which their movement intensity is calculated. One sound file per dancer is cut into sound grains from which energy is measured. Movement intensity is then mapped to sound energy: periodically, the sound grain whose energy is closest to the movement intensity is played. The concatenation of these sound grains results in a new texture following the apparent movement intensity of the dancer.

Stage and Sound Spatialization Dancers evolved in a 4-meter wide circle surrounded by 6 speakers. Interaction sounds were played in only one speaker while the rest of the music was not located only in one part of the space. Each dancer had a different speaker playing their sound as shown in figure 6.3.

Workshop Organization

This workshop was made of five improvisation sessions. Two of them were group improvisations, and the three others were solo improvisations. As mentioned above, this structure emerged during the workshop. The first session was a group improvisation with one sensor per dancer attached to the wrist. Previous rehearsals informed their improvisation with the stage director, prioritizing fluid movements with representations of pulses and waves and working on synchronicity between dancers. There were no instructions related to the interaction. The three following sessions were solo improvisations for each dancer, focusing more on the interaction. The final session was a group improvisation with sensors in the pocket rather than attached to the wrist. Before the first session, we did a first improvisation with the dancer and the music without interaction to get acquainted with each other and for the dancers to discover this physical new space.

Results



Figure 6.4: The first workshop consisted of two group improvisations and three solo improvisations, one for each dancer. Dancers had a sensor attached to the wrist.

In their feedback, dancers refer to the interaction and its sonic output as “*the sounds*,” meaning the sound grains that are triggered when the movement is detected.

Agency and Expectation During this first workshop, most of the dancers’ feedback concentrated on the feeling of agency over the produced sound. They found the interaction challenging and felt they needed training to use it.

“It was quite a challenge to sync up with it, [...] you can’t control it, I’m not sure how connected I felt to it (approbation from other dancers). But that’s nice in that it’s not easy and maybe with more practice it will become easier” (ME)

This resulted in a feeling of lack of control regarding the sound. For AP, this difficulty also came from preconceptions about this interaction’s outcomes.

“I cannot start to make a song, I don’t know... It was an interesting relationship.” (AP)

He expected an instrument-like relationship with sound synthesis. He elaborated by pointing out the delay between movement and perceived sound, which other dancers also noted.

“Since the sound doesn’t arrive at the same time as the movement, it is difficult to be sure how much I feel connected to it.” (ME)

Two aspects must be considered to explain this latency. From a technical perspective, the time the information travels from the sensor to the computer software through WiFi is approximately *10ms*. However, since the computer was also charged with the rest of the sound synthesis, other processes introduced an arbitrary latency to the processing of the sensor data and sound synthesis, which the dancers perceived.

From an interaction design perspective, several settings could be tweaked in Max. In particular, the intensity calculation’s scaling and release time greatly affected the interaction. These settings still needed to be refined in relation to the maximum intensity and speed of movements. Slower movements yield lower intensities, requiring a higher gain on the output to conserve the full range of the sound energy. Since this choreography emphasized the fluidity of movement, the intensity tended to stagnate at times, reducing the dancers’ agency over the sound produced.

Influence of the Interactive Setup The presence of sensors on the body affected the behavior of dancers. For the director, it was clear that they focused more on their arm as they were trying to make sense of the interaction and partly neglected the rest of their body in their improvisation. In particular, TS noted that, when putting the sensor in her pants pocket, she focused a lot on this part of her body.

“I needed to remind myself that I could move my arms or my legs a bit more. Because I was concentrating a lot when wearing this here [in her pocket].” (TS)

As expected, the first contact with the interaction was cognitively demanding and hard to mix with the demands of group improvisation in a small space. In particular, the various sound textures that superimposed each other arbitrarily during the improvisation were challenging to manage for TS.

“At one point when I heard all the layers [...] it was too much I didn’t know how to connect and I was trying to pick one... It was a lot!” (TS)

Since any movement of the sensor produced sound, the sonic space around them quickly

filled, making it harder to identify and focus on the sound linked to one's sensor.

Despite the variable latency and focus difficulties, dancers found interest in the interaction and actually favored more complex scenarios. With the sensor in the pocket, the movements required to increase sound energy were harder to reach, leading dancers to explore unusual movements.

“It was more interesting because it's just more difficult. It's just more challenging than the arm, more fun, more tricky.” (TS)

This brought a challenging aspect to the interaction that rewarded new and original movements.

“It was almost like a little game, like 'I am going to give you noise if you dance funny'.” (ME)

Link with the Narrative In terms of storytelling, the interaction was found to reinforce certain aspects of the staging and choreography, embodying the living organism.

“This moments when you started like... Each of you searching for their own and at the same time trying to stay together in the space, in the rhythm. It was a bit like clarity in chaos I felt and this I think was also interesting that you were not trying to do the same thing but still being connected, [...] It was very intriguing in a way. It was very clear that you were connected but also searching to go away.” (JP)

Interestingly, the dancer's exploratory behavior when trying to make sense of the interaction created a visual connection between dancers. With the common objective of understanding the interaction, it channeled their improvisation regarding the dynamic of their movements and the focus on the parts of the body that would affect the interaction. During the improvisation, and in particular when the rest of the music was turned off, they were all visibly trying to make sense of the interaction. They lay on the floor and moved closer to the speakers, intermittently moving only part of their bodies.

LL, the director, observed that “*sensors seem to allow you to stay connected with each other and follow what the others are doing.*” The interaction brought an interconnection through sound that was visible to LL and JP, who were both outside observers of the sessions. JP, in particular, was receptive to the atmosphere produced by the sound design.

“When you stopped the Reich and only put those sounds it really felt for me like suddenly we have an insight of what is actually happening chemically in it. It's like suddenly zooming inside the macroscopic world because we don't hear anymore the music, it's more like an ambiance. [...] I was suddenly zooming on what you individually were doing. You hear like worms in earth that are searching.” (JP)

Although this assessment was subjective and dependent on individual sensitivities, it helped define an interesting direction in which to push the rest of the design to tell the story of this organism's existence.

Nothing was decided yet at the end of the sessions regarding the unfolding of the piece and the part that the interaction would play in it. However, alternating between moments with and without interaction was found interesting for dancers as it helped renew the improvisation.

“We started with the sounds, then you added music, you took the sounds out and it is the things that you made that kept us going and then when we came back to

our own sound, it was very interesting, such a nice journey. Having something else taking over and then going back to one's own sound." (AP)

6.4.6 Performance Design Workshop: Testing Exploratory Processes

The second workshop took place a month and a half after the first. The director and the dancers had rehearsed two times in the meantime. The overall structure for part 2, as well as certain choreographic elements, had already been decided for the second workshop. The choreography was mainly improvised with a given temporal structure that dictated certain movement qualities and space occupation. This workshop focused on trying exploratory processes in the piece's context and deciding its role in the performance.

Modifications to the Interaction

Following this first workshop, several modifications were made to the setup. I focused on the feedback of the dancers and the director, as well as my own observations, to incorporate the interaction in the narrative and make it interesting for dancers while fast to learn. As I would only return to London two times before the premiere and the interactive setup only played a minor role, we had little time dedicated to dancers' practice with the system.

During the first workshop, we found that the interactive setup created a visible connection on stage between dancers. Exploring such a system in a group setting is difficult as the different sounds get mixed together, and they are not yet accustomed to the dynamic of the feedback. This behavior drew an interesting parallel between slime mold exploring their surroundings for food and dancers exploring the possibilities of the system. In the current system's state, this was hardly repeatable as the interaction was broadly understood by the end of the sessions, and these particular instances were their first-ever contact with sound-movement interactive systems. Nevertheless, the idea of active exploration from the dancers on the stage seemed interesting and LL shared this idea.

Reflecting on the first experiment described 5, I adapted the interactive system to foster a repeatable exploration for the dancers. One of the issues of the first experiments that were led before this artistic collaboration was that the 2-D sound spaces created in CataRT had holes in them. In those places, participants could not hear sound, making learning the space particularly difficult to study in a short time. This was corrected with *polyspring* and the *mosX* system, which I designed in parallel to this creation. However, in this performative setting, I found the idea of having to "search for sounds" interesting. If only a fourth of the 2-D CataRT space is filled with sound grains, then exploring it with a sensor measuring orientations leads to three fourth of the movement space resulting in silence. With this setup, dancers could explore this space in quest of their sound grains, reproducing the exploratory behavior that was found interesting during the first workshop. To change the configuration between performances, I could just move grains in the CataRT space using *polyspring*.

Two other feedback I wanted to address were the position of the sensor and the focus dancers had on it in terms of movements. Now that the interaction was based on orientation rather than intensity, the sensor had to be put on a part of the body that allowed it to be completely turned in every direction. The wrist would have been the easiest solution for the dancers as turning it 360° requires little effort. Putting it on the lower parts of the body, such as legs or thighs, could have been possible, but it would have required movements on the part of the dancers that were not really in phase with the choreographic choices made up to that point. For the second workshop, I decided that we would try to strap it to the upper arm,

between the elbow and the shoulder. In the hope that it would be a good balance in terms of difficulty of interaction and get the sensor out of sight, try to make it less visible to dancers and potentially to the audience too.

Setup

. The interaction for this workshop was designed with an early version of the *mosX* system. Similarly to the exploratory workshop, a sound sample was assigned to each of the four dancers. The samples were chopped, and two descriptors were extracted: spectral centroid (X-axis) and periodicity (Y-axis). Those were the same descriptors as the one used in the experiments described part 5. Using *polyspring*, sound grains were spread in the space, occupying approximately one-fourth of the space-, as shown in figure 6.5, while preserving the organization in terms of descriptors.

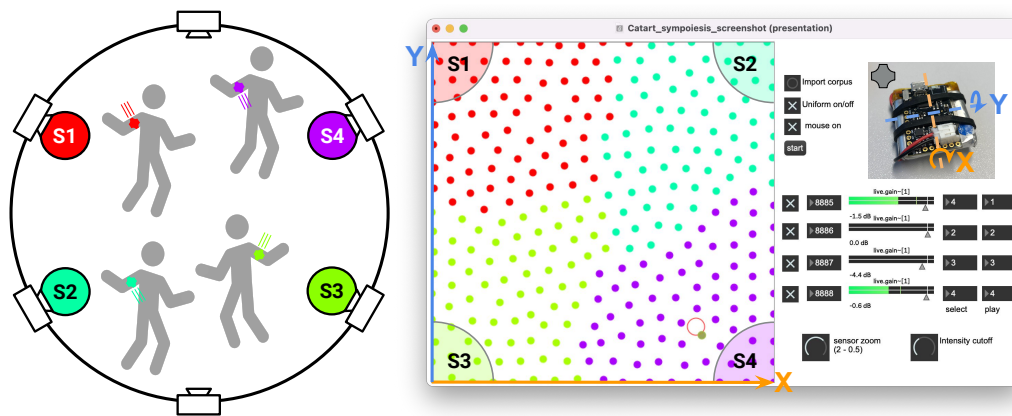


Figure 6.5: Workshop 2 interaction setup based on CataRT. Four sounds (S1-S4) were chopped in sound grains and distributed in a 2-D space. *Polyspring* was used to uniformize the distribution of the grains. A sound was assigned to each dancer. The orientation of the sensor strapped on their body decided if grains were triggered by their movements.

Each dancer had a sensor positioned between the shoulder and the elbow. To navigate in the 2-D CataRT plane with only one sensor, two angles are calculated using accelerometers and gyroscopic data. These angles correspond to the angles between each axis in the sensor plane and the vertical, i.e. the direction of gravity, which are called roll and tilt. When projected on a 2-D plane, these angles cover a circle. To ensure that the entire square space is accessible, the angles are scaled beyond the CataRT space so that the circle they cover is a circumscribed circle of the square. The obtained values are clipped to the bounds of the square so that the final (X, Y) coordinates fall in the CataRT space, as shown in figure 6.5.

Sensors were strapped to the arm, between the shoulder and the elbow. To explore the CataRT space, dancers had to move their arm up and down for the tilt and rotate it around its axis for the roll. Note that in order to cover the full 180° range of the roll, it is necessary also to move the upper part of the body to change the orientation of the sensor. Each set of sound grains was played in a different speaker around the stage, and a sensor was only able to trigger one of the four sets of sound grains. So, each dancer was associated with one sound class and one location on the circular stage.

Workshop Organization

This second workshop aimed to test this mode of interaction with the dancers and try to integrate it into the existing performance. We formalized a task to direct the exploration: *“find the posture that triggers the sound and move to the location of the sound in the circle”*. They had no prior knowledge of the postures that could trigger the sound nor the position of the sound around the circle, apart from that it was played on a single speaker.

Similarly to the first workshop, it took place during a creation and rehearsing day in a studio in London. As, after this workshop, we only had one session remaining before the premiere, most of the time was dedicated to the whole piece and not just the interaction design. Therefore the workshop was made to fit into the existing structure of the piece. Dancers began intertwined on the floor and progressively separated from each other before raising and standing up suddenly at the same time. After this, they gradually moved on to the exploratory phase, which is the part from which the results below are drawn. They scattered on the space toward the location of their sound and got on the floor before reuniting again in the center, entangled with each other, to conclude the journey. The interaction occurs during the exploratory phase and gradually gives way to the soundtrack before the end.

Two different settings were tried. (1) They each had to adopt a different posture to hear their sound. So they were looking individually for their specific body configuration to trigger sound grains. (2) They all produced their sound with the same postures. So, while the sound remained specific to them, they were all looking for the same body configuration. The expectation was that setting (2) would help them connect better during the performance and would force them to pay attention to each other toward this common goal.

The whole rehearsal consisted of three phases: working on the introduction of the performance without interaction, then trying both settings after the introduction, and finally choosing one and rehearsing it while working on the conclusion of the piece.

Results

As explained before, this second workshop was part of a rehearsal session close to the premiere. Discussions during this workshop were consequently less open-ended and aimed mainly at deciding where and how the interaction would be meaningful for the dancers and for the piece. Therefore, I focus solely on the feedback on the interaction design I proposed to the dancers and the reason for the final implementation.

First Setting The first setting consisted of a different body configuration mapped to sound for each dancer. The results with these settings were mixed. AP and ME struggled to find their sounds and were unsure they were actually hearing it. TS took some time to find it as it was in a difficult position to reach. JP found it as soon as the interaction was turned on and had no problem hearing it.

The type of sounds played varied between dancers. They represented different actions or ideas and presented different dynamics when triggered. JP's sound expressed an idea of stretching and tearing. She experienced a gap between the movement qualities expected from the choreography and what the sound inspired her:

“I felt that I wanted to do impacts to have the sound because it sounded strong but then I remembered about the fluid movements.” (JP)

As this mapping was based more on body orientation rather than body dynamics it was

not meant to represent any metaphorical link between the movement that produced the sound and the action that the sound described. However, certain sounds could invite dancers to mimic certain actions and, since the dynamic to get in and out or evolve inside of the range of movement that would produce the sound was completely open, the mapping allowed those movement to be performed to get the sound. JP was attracted by such movement but was constrained by the movement qualities that were expected from the choreography.

At this point, it was difficult to assess where the difficulties came from. It turned out several factors intervened. First, further sound mixing was necessary, considering the small speakers in charge of sound reproduction. The loudness was not the same for each sound, and certain sounds' frequency content privileged their identification and localization in space. Second, dancers had different target body configurations, which made the task more difficult for some. JP and ME had the easiest configuration, while TS and APs were more complicated to find. TS was also surprised that she could find her sound from two different body positions that seemed completely disconnected from each other. Since the sound space is divided into four equal parts, the corresponding range of movements corresponds to a fourth of all possible sensor orientations. This enabled triggering sound from different positions that can look disconnected from each other.

The mapping resembled the one described part 5 except that only the orientation of the upper arm influenced the sound, not the lower arm since it only used one sensor. In particular, the mapping presented the same difficulty where certain sounds could only be reached by pointing the arm upward. The corresponding range of movements was not unusual in the piece. It was normal that it took only around a minute for TS to hear her sound for the first time.

Second Setting As expected, this setting showed a high degree of mirroring behavior. Dancers found the configuration in less than a minute and, from there, prioritized synchronized movement over individual exploration. JP reported decreased interest because, even though there was a range of movement that gave sound, since they were searching together, they stuck to the first posture they found and explored less the range of configurations that could produce their sound. Similarly, AP found the experience less engaging due to the perceived obligation to mimic others. Eventually, in the eyes of LL, it appeared that the first configuration had already produced a sense of cohesion when looking at them performing the same task. The second configuration was too restrictive and suggests that collaborative movement-sound exploration may actually be hindering exploration.

At the end of the workshop we decided to adopt the first settings with different body configurations for each dancer. This became the second part of the performance that we began rehearsing afterward.

Feedback During Rehearsal

During the rest of the session, we repeated this exploration with added directives from LL, and we worked on the third section that would conclude this second part and the *Sympoiesis* performance as a whole. During this work session, feedback on the interaction emerged. First, sounds were difficult to discriminate for three of the four dancers while performing. Their differences were not significant enough to recognize it while focusing on the rest of the performance. During this part of the piece, they are bound to each other as different branches of the same organism. To enact the propagation of the slime mold, they must display a sense of fluidity and occupy more and more space, which requires them to pay attention to each

other and expand slowly inside of a small stage. This is demanding in terms of attention, and JP and AP forgot about the exploration task at one point because the dance directives were already too much to focus on. To make the exploration different each time, I just changed the orientation of the sensors in relation to the dancer's arm before the beginning of the piece. This way, the configuration required to trigger the sound was not known in advance.

6.4.7 Performance

After the second workshop, I made some changes to the system. First, I modified the interaction sounds to make them more distinct from each other by changing their pitch and applying filters to them. I also added a low-pass filter to each sound, with its cutoff frequency controlled by the intensity of the movement. This brought back the dynamic of the movement in the interaction. Now, by standing still in a configuration that produces the sound, only the lowest frequencies below 120Hz could be heard, which sounded quiet considering the type of sounds used and the size of the speakers but could still be heard. To make the interaction more in line with the fluidity expected by the director, I slowed down intensity detection by applying an amplitude envelope with a slow attack and release time. Finally, I added delay and reverb in the final mix to make the sound disappear slowly after the movement was performed. This reinforced the impression of "*sound trail*" behind the movement LL had perceived during the first workshop.

We premiered this performance at Central Saint Martins (CSM) College in London. Managing both the musical aspect of the piece and the interaction was challenging. Adding this technological element required a thorough checklist before starting the piece to be able to then focus on musical improvisation in conjunction with dancers' improvisation during the unfolding of the piece. Interestingly, my perspective on the role of the interaction in the piece evolved with the consecutive dressed rehearsals we did. From my perspective as the musician in this performance, I considered the input from the interaction as supplementary material that I was not in control of and had to play with. However, after multiple rehearsals, I began to perceive movements through interaction, which helped me connect with dancers without looking at them. After the premiere, some of the audience was intrigued by the blue light (coming from the sensors) on the costume and guessed that it corresponded to some form of biosensing (movement or else). Interestingly, they reported not trying to understand right away what they were used for but were still interested in having a look behind the computer after the performance. Surely, as they were, for the most part, students or teachers at CSM, they were accustomed to interactive performances.

6.5 Spectators' Experience Design

Reeves et al. presented in 2005 a taxonomy of interfaces for performer-computer interactions with regard to how they influenced spectator experience for various contexts in arts, installations, and presentations (Reeves et al., 2005).

They propose to decompose the design of the spectator's experience in two dimensions: manipulations and effects. Manipulations designate every action the performer does that can be sensed by the interface, as well as actions outside the sensor's scope. Effects concern the results of the manipulations on the interface. They do not refer only to system outputs but to every effect that the interface has on the performance and the performer. These dimensions are analyzed in relation to how much the interface hides, reveals, or even amplifies manipulations and effects.

The authors propose a taxonomy of interfaces depending on where they fall in this plane. Interfaces that tend to hide both manipulations and effects are *secretive*, while, on the other end of the spectrum, interfaces that tend to amplify both are *expressive*. In-between, interfaces are *magical* when they tend to hide their manipulations and amplify their effects and *suspenseful* on the opposite.

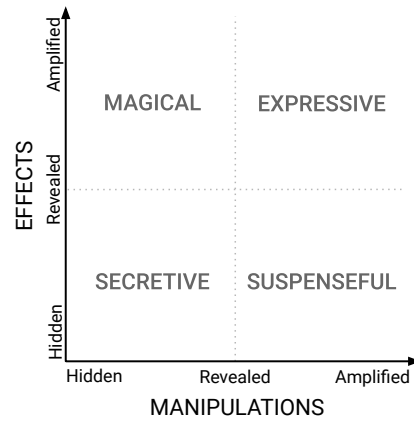


Figure 6.6: Taxonomy of performance interfaces whether they tend to hide, reveal or amplify performer’s manipulations and their resulting effects.

They nuance their work in the sense that:

“like all such taxonomies, ours necessarily simplifies the true picture in order to reveal broader underlying principles, in this case the idea that designers can trade off whether and how to reveal manipulations and effects in order to create spectator experiences [...]” (Reeves et al., 2005, p.747)

While they propose a tool for designers to purposefully consider their strategy for designing for a specific spectator’s experience, I want to use this taxonomy to reflect on past experiences and uncover how design choices influenced the spectator’s experience in *Sympoiesis* and the *Circus Lab*.

6.5.1 Classifying Spectators’ Experience

The two performances presented above introduce different interactive setups designed for different practices. The table 6.1 lists for each setup the position of the sensor on the performers and the specific signals that are extracted from them.

Interactive Setup	Sensor	Intensity	Rotation	Orientation
Exploratory Workshop (WS1)	Wrist	✓		
Design Workshop (WS2)	Arm			✓
Premiere (1ST)	Arm	✓		✓
Chinese Pole (CP)	Ankle	✓	✓	
Cyr Wheel (CW)	Wrist	✓		
German Wheel (GW)	Wheel		✓	
Swinging Trapeze (ST)	Arm	✓		

Table 6.1: Description of the movement sensing of each interactive setup.

In *Sympoiesis*, from the second workshop, the interaction consisted of the exploration of

the space of configurations that the arm can take with the objective of finding a sound. When designing this interaction, the objectives differed from those of the Circus Lab, particularly from a spectator's perspective on experience. While the *Circus Lab* included movement-sound interaction technologies with the aim of visually connecting performers' actions and the produced sound, *Sympoiesis* sought to create on stage the impersonation of a non-human organism, notably by showing a connection between the dancers. A process assisted by interactive technologies without trying to make the movement-sound connection visible (or invisible) to the audience.

From a design point of view, we can consider how these diverging objectives ended up in different spectators' experiences to display how they interact with the Manipulations-Effects space proposed by Reeves et al. (2005). I give a proposition of classification in figure 6.7.

To position these setups, we must consider their similarities and differences. They all share specific properties. First, they are embodied movement-sound interfaces manipulated by movement-based performers using wireless motion sensors. As such, manipulations correspond here to movements made by the performers, and effects may also include the impact of sound feedback on the performers themselves. One way to perceive manipulations as hidden or amplified in these types of intangible embodied interactions is to consider the position of the sensor and whether the sensor is clearly identifiable on the body. A recurring point of discussion during rehearsal was whether or not to show the sensors and how much we would include them in extra performative materials such as leaflets, posters, or descriptions of the pieces. One way to completely hide manipulations, for instance, would have been to put sensors beneath the costumes and never tell anyone about them and the interaction.

Second, the produced sounds take part in a broader musical accompaniment. Therefore, the effects, in their sonic aspects, cannot be interpreted directly by the audience as effects of the interface. Nonetheless, the effects of these interfaces are never hidden in these instances. The differences will reside in whether or not the mapping and the sound design amplify them and if the musical context tends to cover them or, on the contrary, put them in the foreground.

Sympoiesis: Blurring Manipulations

The *Sympoiesis* interaction interface has had three main design stages. To position these interfaces in relation to the spectator's experience, I assume, since it was acknowledged in the program, that the spectator knows that the performance will include some sort of interaction between movements and sound and that the audience can identify the sensors on the performers' bodies.

The first stage of design was the interaction I proposed to the dancers during the exploratory workshop, noted **WS1**. It links the intensity of the movement to the loudness of a sound through concatenative synthesis. The manipulation is clearly visible: as the input of the interface is the sensor strapped to the wrist, which is not hidden from view, the movement of the extremity of the arm can be identified easily as the manipulation. It presupposes a certain comprehension from the audience of what the sensor can perceive. In the case of wearable sensors, "*the smallest gesture you make becomes an identifiable and interpretable musical gesture for an audience*" pointed out Philippe Spiesser in its interview reported in Chapter 3. This observation is particularly important in the context of music-centered performances but remains relevant to movement-centered performances. Depending on the spectator's knowledge of technology, the degree to which manipulation is revealed varies, as pointed out by Fels et al. (2002). The resulting effect, the loudness of the sound following

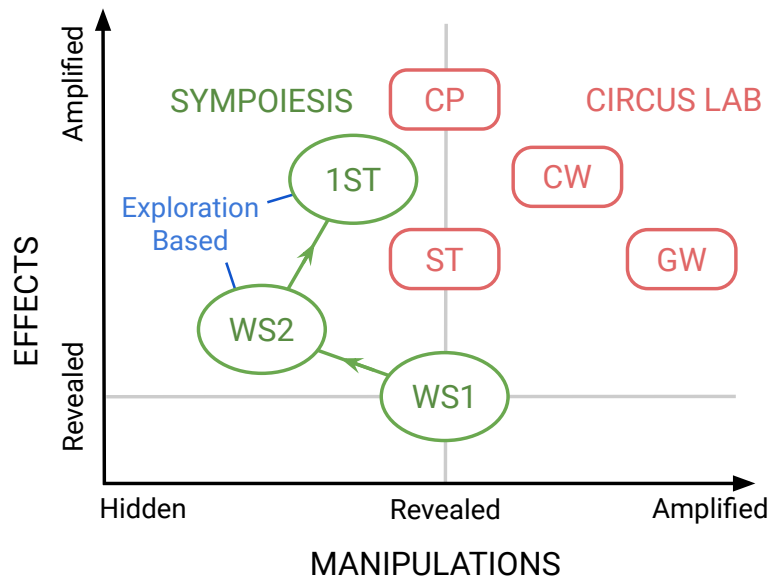


Figure 6.7: Classification of the interactions designed for *Sympoiesis* and *Circus Lab* according to how much they amplify or hide the manipulations made to the sensor and how much the effects of these movements are amplified. Acronyms are defined in the table 6.1

the movement intensity, is also easily perceivable and is simply presented to the spectators as the direct consequence of the movement without amplification. I position this interface as a reference point of an embodied interaction that does not hide or amplify the manipulations or the effects.

The second workshop and final performance interactions, noted **WS2** and **1ST**, used a different setup that tended to hide manipulations while slightly amplifying their effects. This time, the sensors were strapped to the dancer's upper arm, and the orientation was used as the input. By strapping the sensor closer to the shoulder, manipulations on the interface are less transparent to the audience. Moreover, the mapping used the orientation of the sensors as control parameters linked to a space only a quarter full. The resulting interaction tends to hide the manipulations of the sensors by dancers. I argue that the effects of the interaction, on the contrary, are amplified, not individually, but through the joint exploration of all dancers. This interaction falls on the quarter of the manipulations/effects space that Reeves et al. labeled *magical*, as displayed in figure 6.6. The audience experiences a connection between the dancers and the sound but cannot link it directly to their actions. The premiere interaction setup added intensity and a slow sound feedback decay, resulting in fewer hidden manipulations and more amplified effects.

Circus Lab: Amplifying the Performance

The *Circus Lab* performances had a single design stage, but the various pieces resulted in different spectator experiences. They were all designed to amplify the interaction's effects and not hide manipulations. The particularity of this context is that circus apparatuses themselves are intended to amplify the spectators' experience. The German Wheel (**GW**) amplifies the manipulations made by the performer by its massive presence on stage and the noise it produces when moved around. The interactive setup amplified its dynamic, therefore participating in the amplification of the manipulations by the artist. The Chinese Pole (**CP**)

differs because the apparatus is static, and the performer evolves on it. By strapping the sensors on the circus artist, we do not further amplify manipulations to the pole; however, we significantly amplify the effects by using the intensity and rotation control parameters. The Cyr Wheel (CW) falls between those two setups in that the sensors are strapped on the performer, not the moving apparatus. But, since the performer privileged actions that would amplify his manipulations on the wheel, the interactive setup both amplified effects and manipulations. Finally, the Swinging Trapeze (ST) setup was a less amplified version of the Chinese Pole.

6.6 Discussions

6.6.1 Exploratory Behaviors as a Support to the Narrative in *Sympoiesis*

In *Sympoiesis*, each costume embarked a motion sensor, positioned on the upper arm, and linked to a computer through Wi-Fi. At a certain point in the performance, the dancers could trigger a sound synthesis within a range of postures from their right arm. Since the volume around them in which they could interact with the sound was not known in advance, this moment of the performance was an exploratory phase for the dancers.

While this setup fostered an individual somatic experience for each dancer and happened at a point where they were scattered on the stage, it created a moment of co-exploration that visually and sonically connected them. In the words of Megan Eyles, one of the dancers, the sound-movement interaction “*definitely helped [her] sense of connectivity with the group as [they] recognized each other’s sounds and could move with one another and connect [their] sounds, giving another detail to the performance which the audience could see.*” Thus, this non-tangible connection supported the narrative both internally, as a shared embodied experience for the dancers, and externally for the audience, as it helped create the effect of a common goal within individual actors with similar movements and sound synthesis.

In these moments when the dancers interact with the sound through wearable technology, the role of the performer in music is naturally questioned. Does this interaction still fit the musician-instrument model where the instrument acts as a medium between the performer’s intentions and the receivers? (O’Modhrain, 2011) When the dancers’ actions directly control the sound through wearables, the costume gives a new agency to performers and, therefore, a new layer of meaning to the performance. In this setup, we argue that the costumed body becomes an instrument itself. But an instrument that neither the sound designer nor the performers fully controlled: the dancers did not know where the sounds were and when exactly they would be accessible. As a performance strategy, it blurs the notion of control by the creators and performers over the instruments, thus supporting the goal of creating a space for multispecies connections on stage. Donnarumma (Donnarumma, 2016) called this particular relationship a *configuration*, where “the capacity of the instrument and the performer are interlinked and thus mutually affect each other.”

Using such technological artifacts to make this performance allowed me to get additional material for sound creation. However, the use of wearable technologies for embodied interaction with sound actually benefited the whole creative team. As stated by Pantouvaki, “*the use of wearable technologies demands an overlap of conventional borders between performance design disciplines*” (Pantouvaki, 2014). Through this collaborative work, we entered a co-creation process that required all the members to understand more about sound, costume design, and dance practices.

6.6.2 The Challenges of Concurrent Performance Creation and Interaction Design

The technology used in those performances is not foolproof. I learned ways to mitigate the risks of failures, and a lot can be done to prepare in advance, but as our work was highly exploratory and ever-changing, the technology required adaptation, which heightened the chances of failure. Performers usually need some warm-up time before practice and thus cannot perform on demand whenever the interactive system is back in service. I argue that the moments where interventions are required on the technological side should not be considered accidents in the creation process but rather be accepted as part of an iterative approach that is required in order to create new experiences. As such, they must be designated whenever they are necessary and care should be taken at acknowledging the needs of the performers and not rush the transition between technological works and try-out sessions. Interestingly, results from the study of cooperative prototyping in workplace environments corroborate such observations. As prototyping is exploratory, it is prone to focus shifts between points of interest that not all collaborators are receptive to, which creates tensions. Bødker and Grønbæk advocated that tensions and focus shifts in collaborations should not be avoided but rather that it is the designers' and users' responsibility to learn how to adapt and better handle those moments (Bødker and Grønbæk, 1991). Under these conditions, an environment conducive to expressing everyone's creativity can be created, encouraging exploration.

6.7 Chapter Summary

This chapter accounts for two collaborations with dancers and circus artists, *Sympoiesis* and *Circus Lab*. Both include movement-sound interactions designed by the author in collaboration with other artistic collaborators. The creation process is reported from the perspective of the interaction designer. It highlighted the discussions and tensions in such collaborations, emphasizing the necessary sharing that hybrid practices entail. These performances enable us to consider exploration in a broader scope, particularly as a shared process in collaborative settings. As motivations and expectations in using interaction technologies vary between individuals involved in the project, exploration became a shared process in which the points of attention differed between actors and the design directions to be taken was decided collectively. The motivations of other collaborators and the particular contexts led to varied interaction designs that were analyzed through the Spectators' Experience Design framework, which revealed how design choices led to different characteristics of the resulting interaction for the audience.

CHAPTER 7

Discussions

This dissertation examines the learning and appropriation processes associated with using interactive movement-sound technologies in the context of the performing arts, focusing on the nature and role of exploration in these processes. This chapter serves to reflect on the studies presented in the previous chapters to unpack what they have yielded on these subjects, relate them to existing works, identify their limitations, and suggest directions for future research.

7.1 Learning and Appropriation are Intertwined with the Development of Practice

The initial aim of this work was to understand the learning of new musical instruments built on movement-sound interactive technologies to better support its design. Learning how to play a musical instrument is typically associated with the process of graduating from novice to expert. Ultimately, the instrument becomes an extension of the musician's body through which they can express musical ideas, which enables them to reach an immersive flow state during performance. Several artists have shown such connection with electronic instruments through embodied interactions (Nijs, 2017). The artists interviewed in chapter 3 have demonstrated a level of intimacy with the technological tools they use in their practice that is akin to appropriating a musical instrument. For them, movement technology became an “embodied” tool for musical expression.

In this context, learning was found to be multifaceted and more diverse than just gaining mastery. The main difference between learning to play a traditional musical instrument and other forms of music education is that the instrument's technological development can be considered part of practicing. Part of the learning process encompasses complex mechanisms

beyond the sole sensorimotor learning of a fixed instrument. Interviews showed that the instrument design was, first and foremost, artist-led, and the knowledge and skills required to create and orientate the technological development in the desired direction were also to be acquired, at least partially. As Laetitia Sonami expressed in her interview with Mary Mainsbridge:

“I wish I had a programmer who could do things better than me. So much of that process is a craft. Learning how to craft this, learning how to do this programming, and soldering this thing, and that connection with the material is essential.” (Mainsbridge, 2022, p. 51)

However, there is no consensual core of knowledge to be acquired. A wide variety of techniques and technologies can be used to create movement-sound interactive systems. The technology choices are contextual. It depends on socio-cultural and economic factors such as the desired sonic aesthetic, the movement practice, or the accessibility of technologies. Ramos and colleagues explained the technological trade-offs and adaptability required to design an electronic bandoneon in Argentina:

“The economic situation of Argentina is not optimal for this kind of venture, as the relationship between our currency and the prices of imported products is rather unfavorable. Some parts that could help improve our designs are not available in local stores, as they tend to favor extremely popular platforms like Arduino. Thus, it was essential for us to resort to products from the local industry whenever possible, and even use *gambiarra*¹ techniques like recycling bellows or repurposing unused electronics.” (Ramos et al., 2022)

Therefore, imagining a general “all-purpose” pedagogical process of acquiring technological knowledge and skills oriented toward movement-sound interaction design is elusive. It seems on the contrary that the acquisition of such expertise happens through self-directed learning, “*a process by which people seize the initiative, without help from others, to diagnose their learning needs, formulate goals, identify human and material resources, and evaluate learning outcomes*” (Knowles, 1975). A process reinforced by information technologies, which offer the possibility for ubiquitous learning, and has favored the development of many communities of practice, particularly around computer technologies (Gonzalez-Sanmamed et al., 2020).

Artists interviewed argue that what can (and should) be taught to beginners is less about technology itself and their particular forms than about the general process of integrating technology in practice, “*You’re not going to be interested in transmitting a technological brick, but really more the process of reflection behind creation*” (Fabrice Guédy, Chapt.3). This “*process of reflection*” was a recurring theme in interviews. The pedagogy interviewees have developed around interactive technologies aim to emancipate students and support personal practice development. What is transmitted in these instances comes from the digested experience of the teacher and is taught with inductive methods that favor the understanding of broader concepts through practice. This indicates that teachers put certain pieces of technology forward when transmitting but do not build their pedagogy around specific and immutable artifacts. The core of what is communicated is built on personal experience. Still, we can identify several concepts that cut across interviews: the development of critical thinking regarding technological intervention, the quality of attention required when experiencing interaction, and the necessity to experiment and explore. Interviewees developed those

¹The Brazilian word “Gambiarra” refers to the improvised technical solutions in a context of limited resources.

skills on their own through their experience. And eventually, it is up to each practitioner to discover how to translate these concepts into actions. In the teaching activities described, interviewees adopted a mentoring stance towards the students, giving mainly methodological advice and guidance illustrated with a given technological setup while encouraging individual development. This results are in line with pedagogical courses on DMI design proposed by Jensenius (2013)

Yet, any embodied movement-sound interaction still requires the performer to acquire sensorimotor skills, related to specific action-perception loops. We referred in Chapter 5 to the concept of audiomotor maps acquisition. This is the closest process to what is commonly considered learning a musical instrument. However, where learning a traditional musical instrument consists of practicing extensively with a relatively fixed object, learning for movement-sound interactive systems involves acquiring new audiomotor maps every time the design evolves, which can happen for every new creation. However, fixed parameters can make sensorimotor learning transferable between setups. For instance, the GeKiPe system (Fernandez et al., 2017) is based on a depth camera and wearable sensors strapped to the hands. The system is based on a persistent interaction paradigm: the performer plunges his hands in volumes monitored by the depth camera and triggers and manipulates sounds with movements sensed by wearable sensors. Several compositions were written for GeKiPe, *Hypersphere* was based on generative sounds, leaving space for the performer to improvise around the sound synthesis, while *Le Silence* was more written, with nearly 400 sounds to trigger, with a more reduced improvisational space. The two pieces had different sound aesthetics. Consequently, the audiomotor maps to build were necessarily different. However, motor skills associated with manipulating the gestural interface were transversal to both.

In summary, we can identify three categories of knowledge and skills that compose movement-sound interactions: technological knowledge and skills, methodological knowledge and skills, and sensorimotor skills. It appears that the development of the interaction and the acquisition of the required skills and knowledge happen concurrently with the development of practice as they feed and motivate each other. This process requires time and leads to appropriation, where “*the users understand and are comfortable enough with the technology to use it in their own ways*” (Dix, 2007). We found that learning is intertwined with practice, which is necessary to appropriate technologies and incorporate them into artistic practice.

7.2 Exploration Supports Design, Learning, and Appropriation

Through this PhD research, as presented in the previous chapters, the notion of exploration became a central theme that appears at different levels, from skills learning to technological appropriation. As explained above, there is no unique pedagogical path for learning how to integrate interaction technologies into one’s own artistic practice. Deciding what tool to use and how to achieve one’s artistic vision is a skill that builds on exploratory iterative processes. Thus, we found exploration to be a key marker in learning and appropriation processes.

7.2.1 Design Space Exploration

As such, one cannot expect to take a system as is and “plug it in” to achieve meaningful results. Exploration must be conducted on all aspects of the interaction, and aesthetic questions must be considered. When defining Design Exploration, Fallman wrote: “When it comes to interaction design research, issues of aesthetics concern not only how something looks and feels, but also the aesthetics of the whole interaction including how something works,

how elegantly something is done, how interaction flows, and how well the content fits in.” (Fallman, 2008). Exploration in movement-sound interaction design concerns all aspects of the interaction. From literature, as well as interviews and performances, we have decomposed exploration into three categories.

Gestural interface exploration encompasses both the physicality of the interface as well as the signal processing applied to received signals. It produces the *control parameters* that will modify the *sound synthesis parameters* (Hunt et al., 2000). The position and attachment method of sensors have an important impact on the performers’ experience. In *Sympoiesis*, significant effort has been put into integrating sensors in the costume without requiring straps to minimally impact dancers’ movement and facilitate their installation. For the *Circus Lab*, extensive discussion and understanding of the performer’s perspective and skills were required to position the sensor to capture expressive movements and choose signal processing accordingly (see section 6.3.3). The exploration goal was to iteratively find a “good” physical sensor placement, i.e. stable and not disturbing, while pertinent for expressivity. In Interviews, Mari Kimura explained that she accompanies students in the use of motion sensors to augment traditional instrumental practice: “*The first thing I usually do with the students is that I have them wear [the sensor] right now [. . .] and see if there’s any relevant thing that you can use as an expression*”. This highlights the importance of gestural interface exploration, particularly when working with existing practice.

Sound synthesis exploration concerns the whole chain of sound production: the sound synthesizer, any effects applied to the signal, and electromechanical conversion. This last point refers to the choice of speakers, their number, and their placement in space, which significantly influences the feedback loop to the performer. The sound synthesis parameters are all parameters that can be accessed and modified by the synthesizer and effects. Music creation software such as Ableton Live exposes many parameters to mapping through MIDI or OSC, which can quickly amount to hundreds of dimensions through which to manipulate sound. The whole activity of exploring this space and making choices is usually called *sound design*. Notably, Scurto and colleagues proposed to use artificial agents to support the exploration of the parameters space of a sound synthesis engine (Scurto et al., 2021). We observed that in collaborative endeavors, this exploration was led by someone in charge of the performance’s musical aspect. In the *Circus Lab* and *Sympoiesis*, one or several artists were dedicated to music creation and arrived with prepared sonic materials that were presented to the rest of the group. The time for interaction design is not necessarily a time for sound design. In performances, the musician did much of the sound design process before or between group work sessions.

Mapping exploration decides how control parameters modify sound synthesis parameters. This is the last part of movement-sound interaction design, where the interaction is created. Creating the mapping is an iterative process that necessarily involves the performer. It can be dedicated to the performer’s experience with the system, focusing on questions of agency and expressivity, or it can be directed toward spectator experience design. In the former case, the exploration is performer-led, and their feedback will orientate design choices. In the latter case, design is led by external perceptions of other performers, stage directors, designers, composers, or any collaborators to the project. Those two objectives necessarily influence each other as they both impact mapping, and iteration on both aspects must be done concurrently. In *Sympoiesis*, after-session discussions focused as much on dancers’ feelings as on the director’s and other observers’ feedback on the external experience. Through multiple improvisation sessions and consequent adjustment of the mapping, we got closer to relevant and meaningful interaction in the piece.

Eventually, iteration on those three aspects is necessary for movement-sound interaction design. It requires experimenting with the system in a performance context, with performers, designers, and the director or composer to refine design choices that serve the artistic intent. The following section focuses on those phases of improvisation and experimentation with the system.

7.2.2 Exploration by Improvisation and Experimentation

Exploration, improvisation, and experimentation are key when designing and learning for movement-sound interaction. West et al. led a mapping design experiment with nine NIME users who had to create a mapping between a given interface and synthesizer. They found that when building mapping between control parameters and sound synthesis parameters, *“designers alternate between exploration and experimentation, learning how the instrument and the mapping works and how it might work better, and then making small changes to try to improve it.”* (West et al., 2020). Here, “exploration” and “experimentation” are put side to side as different methods of interacting with the system to understand how it works. “Experimentation” is used to validate or invalidate a hypothesis about the system’s possibilities. In contrast, “exploration” here refers to actions on the system that are not motivated by a hypothesis but rather aim to discover possibilities when they emerge.

In the experiment we reported in Chapter 5, the participants explained they tested hypotheses during sessions: *“Can I produce a louder sound by moving faster?” “Does the sound change if I turn on myself?”*. They also experimented to challenge their understanding of the system. Those behaviors mainly appeared during the last session once they had an initial overview of the system. In the beginning, participants explored the space and tried to understand and memorize positions, sounds, or associations. Participant 11 explained that *“there’s this whole exploration phase where at first you’re really just doing basic stuff [...] you’re learning, trying to remember things.”* This first open-ended exploration of the system meant at consciously learning the relationship in the interaction. Bertha Bermudez formulated it in interviews as a form of obedience toward the technology that dancers had to start with before challenging it: *“One interesting thing is that we always try to do two things: one, we understand the system, so we do it, we are obedient, [...] and then we try to break the system, and see when it doesn’t work. [...] Almost all the dancers did the same”*.

Interviewees highlighted the need for improvisation. They indicated that it is a necessary step toward evaluating the qualities of a designed interaction, learning how to use it, and finding emerging uses. Limb and Braun claimed that *“the process of improvisation is involved in many aspects of human behavior beyond those of a musical nature, including adaptation to changing environments, problem-solving and perhaps most importantly, the use of natural language, all of which are unscripted behaviors”* (Limb and Braun, 2008). Through “unscripted behaviors” an artist can improve his or her sensorimotor skills and gain intimacy with the interaction, as well as explore different and new directions in the use of the system. Improvisation can be directed towards an objective, structured, or completely free. It is a way of performing that acknowledges unconscious knowledge of the body. Sudnow, a trained pianist in improvisation, referred to *“knowledge in the hands”* (Sudnow and Dreyfus, 2001).

Finally, *“exploration,”* experimentation, and improvisation enable learning and appropriation through embodied interaction. Exploration became a central theme of this work, and we found that it takes many different forms. We are interested in exploration as a broader concept that encompasses all actions taken with the aim of discovering something, and therefore,

without knowing the outcome. Therefore, experimentation is a form of exploration driven by a given objective. “*Exploration*,” as meant by West and colleagues, refers to open-ended exploration, where exploration is conducted without defined objectives. We now discuss the question of tasks and goals in artistic creation specifically.

7.2.3 Tasks in Movement-Sound Exploration

Exploration in traditional HCI research is often meant to solve a problem for which no protocol is given in advance. In this case, exploration is a trial-and-error process where its outcomes are compared to a target. As presented in Chapter 2, several early studies on workplace software learning distinguish between task-oriented/task-free learning (Rieman, 1996; Trudel and Payne, 1995; De Mul and Van Oostendorp, 1996). In workplace settings, the task is central in motivating users to explore: “*Although many users will scan tutorials, manuals, and menus to gain an initial overview of their software, they generally prefer a ‘just-in-time,’ task-driven approach to learning the details. Exploratory learning for pleasure is rare.*” (Rieman, 1996). Non-task-driven exploration is found to be less productive (Trudel, 1994) and not favored by users in general. We found that it did not always apply to artistic practices. We will first focus on task-oriented exploration and then open-ended exploration.

Trial and Error Exploration The notion of goals in artistic creation might be complex to define. In the context of the piece creation, the artist might focus on formalized so-called artistic intention, such as sharing a particular theme with the audience or conveying specific ideas or sensations through a common lived experience. *Sympoiesis* invited the public to reconsider their relationship with other species. The dancers’ embodiment of the slime mold was intended to encourage the audience to project themselves into the idea of being another species studied in a confined environment. Eventually pushing them to rethink their relationship with other-than-human species. For this, we experimented with technology-supported interactions, constantly evaluating their relevance to this objective. This is best described as exploring a given set of possible technology combinations, similar to trial-and-error, assessing whether the proposed interaction produces the desired outcome and learning from it.

Nevertheless, the target might not necessarily be well-defined or shared among collaborators. In *Sympoiesis*, every artists/performers had different expectations regarding the role of technology in the piece. Deciding what deserved to be explored further and what should be set aside was done by reflecting on the lived experience after each improvisation session in a group setting. Being open to discussion and listening was key to deciding what to do next. Moreover, the target to which we evaluated the design moved with the piece’s development, displaying the co-evolution of problem-solution formalized by Dorst and Cross (2001).

Open-ended Exploration Exploration in the context of a long-term, individual relationship with an artifact is more akin to open-ended exploration, where there is no defined task that technology must help perform or problems that can be solved. In particular, the artists interviewed showed a unique perspective regarding their relationship with *errors*. Interactive technologies were found interesting in the perturbation they could bring to practice and how they enabled artists to “*move out of their comfort zone*” (Philippe Spiesser) and bring different perspectives. Technological errors in open-ended exploratory contexts are not dead ends to be circumvented but opportunities for novelty. Fabrice Guédy explained that he “*likes when there are bugs, because it’s in these interstices that there are often opportunities to find things that you wouldn’t have found if everything had been done beforehand.*” Exploring

movement-sound interactions is essential for practitioners to find new uses and encourages creativity.

7.3 Designing for Exploration

We discussed in the previous section that exploration must be taken into account when designing movement-sound interactions for the performing arts. From the perspective of a designer working with dancers, musicians, or any moving performer, the question of how they will use, learn, and ultimately appropriate the system should be considered. We frame this discussion on the case of a collaboration between designers and artists where they aim to create a novel system for movement-sound interaction iteratively. This process is anchored in the methodology of participatory design (Muller and Kuhn, 1993), which is actively used and studied in digital musical instrument design (Fyans et al., 2012). This section draws on studies, literature, and personal experience to outline methodological guidelines for enabling learning and appropriation of technologies through exploration. I will start with a “*detour*” to bring my personal perspective on inductive learning methods I practiced when teaching in parallel to my PhD to further discuss learning and exploration of movement-based interactive systems.

7.3.1 Insights from Inductive Teaching Methods in Higher Education: a First-Person Perspective

During my four years working on this research project, I taught mechanics to undergraduate students. Teaching mechanics can be straightforward compared to transmitting knowledge and skills on movement-sound interactive systems. The skills and knowledge to be taught were, for the most part, clearly identified, and students arrived in my class with a specific set of shared prerequisites. However, I found that using inductive methods was particularly interesting for their exploratory properties. In this subsection, I briefly present the process I used, the results it gave, and the advantages and limitations of this pedagogical method.

Deductive teaching methods build skills and associated knowledge with a top-bottom approach. Classes first introduce broad concepts, knowledge, and methods that are then applied in smaller group settings to particular exercises so that students acquire skills by applying these concepts to varied problems. This is the most classic approach to teaching in higher education (at least in France). Inductive methods start by stating a problem students must tackle to lead them to general concepts that emerge when required. Students find the necessary knowledge to solve the problem through documentation or experiments. Depending on the method, problems will be standard textbook exercises or realistic, ill-defined problems. Students might have to find information by themselves or will be given material. Contrary to deductive approaches, inductive methods require students to be active from the start, and concepts are introduced only when necessary. For a complete overview of inductive methods, see (Prince and Felder, 2006).

I gave a class on beam theory for mechanical engineering students in their last year as undergraduates. The last subject that I had to tackle was energy methods. This was meant as an opening on the subject, and we only expected a basic comprehension of the theorems and associated methods. Students in pairs were given a “*textbook*” problem to solve with guided questions that they could not tackle with their current knowledge. It required them to look for new theorems in order to solve each step of the exercise. They were given selected documentation on energy methods that contained all the necessary knowledge to tackle

problems. Some questions were optional as they required skills outside of the minimum skill set expected from them. For this subject in particular, I decided to adopt inductive methods as the theory was based on objects that they had already manipulated before in the course and because it could enable students to invest themselves in those theories as much as they wanted, providing that they acquired the minimum skills required for this course. Inductive methods require a final full-class “*return*” session dedicated to returning to the concepts, theories, and methods used. This last session is essential in crystallizing knowledge. It differs from a traditional lecture in that students can actively reflect on their experience and identify what they understood and missed during their work.

In my experience, students throughout the years showed a significant engagement in the exercise. As they needed to figure out how to use the theorems by themselves and were not given a precise method immediately, they tried different directions to solve the problem. They often ended up at dead ends before finding the solution, which required understanding concepts correctly. Interestingly, exploration led some students to efficient methods and shortcuts that I did not anticipate and were outside this class’s scope. What I found the most interesting in this method is that I do not give them my understanding, my vision of what those theorems mean, and what they represent for me right away. Usually, when introducing a new theorem or method, different perspectives are given on their meaning and function to help students understand better what they are. I typically draw inspiration from my own comprehension to do so, which is not detrimental to learning but imposes my vision on them and pushes back the moment of personal appropriation even further. When they have to make sense emerge by themselves, they create their own understanding, their own images, and sensations backed up by experience. I found this to enrich teaching and learning as much for me as for them.

While inductive teaching might look like the magic solution to student involvement and active learning, it comes with considerable pedagogical challenges and is not suited to all teaching situations. They require more in-class time than deductive approaches for teaching the same skills and knowledge. The formulation of the problem and the choice of given materials must be done with care. If the problem is too complicated or not guided enough, student engagement will be more challenging to manage. Prince and Felder refer to “*student resistance*” as the “*burden of responsibility for their own learning*” that inductive approaches might impose on them (Prince and Felder, 2007). On the contrary, too much guidance in question or selected materials hinder the exploration and playfulness that such methods try to bring to light. Working in groups also brings additional organizational challenges, and not all students feel comfortable in these situations, which might hinder learning. However, the reflective aspect of group work encourages the development of listening and collaboration skills.

Overall, inductive methods can impact learning and are particularly relevant in later years of higher education. They could hardly entirely replace deductive classes, if only for the time and organization required in an already packed pedagogical project with limited resources. Nonetheless, I believe they are necessary as they bring essential skills not explicitly taught in class. When driving their own learning, students also learn how to appropriate concepts, theories, and knowledge. They learn how to learn, which will prove highly beneficial when faced with their own problems to solve.

7.3.2 Building a Pedagogy of Movement-Sound Exploration

What elements of this method can we reuse for performers' interactive technology learning and appropriation? These methods are based on a pedagogical objective in which knowledge and skills are clearly identified. Teachers, who master the skills and knowledge, build the teaching sequence that will enable students to acquire them. Teachers are designers. They design pedagogical sequences with students' experiences in mind. When using inductive methods, the teacher aims to foster exploration in students to facilitate the appropriation of skills and knowledge. Similarly, the designer of movement-sound interactions should not limit designing to the technological artifact but consider the course of actions that will lead to its use and give importance to the moments when performers "*meet*" the artifact and its evolutions and try them. At these moments, while performers should not be told exactly how the system works and how to use it, we argue that many actions can be taken to foster exploration.

Designers cannot fully adopt the posture of a teacher because they are not masters at using the system and do not possess the performative experience to teach performers. Moreover, as explained above, pedagogy should not focus on detailing technological implementation either. However, depending on the context, minimal and necessary skills or knowledge might enable richer interaction and accelerate the exploration process, which is not to be neglected in real-life scenarios where time is often the most prominent constraint. This necessary learning includes unchanging technological features in sensors. A purely deductive approach would not be fruitful in an artistic creation setting. Beginning with technological concepts will impose a comprehension on performers. This is where there is to draw inspiration from inductive learning. To make performers aware of the possibilities and limitations of an artifact, an inductive approach based on embodied action with the technology will enable them to create their own embodied knowledge. We can illustrate this idea with the inertial measurement unit (IMU).

The IMU is one of the most prominent sensors for whole-body interaction with sound. These units come with a set of characteristics. First, they include accelerometers, gyroscopes, and magnetometers. Those sensors are only sensitive to specific movements: linear acceleration and gravity for accelerometers, rotation speed for gyroscopes, and azimuthal orientation for magnetometers. Second, their placement on the body significantly affects their output. Third, the process done on the signals will privilege specific movements over others as there is no one-fit-all processing that will suit every mover. Therefore, they require understanding and experience to create an intuition of where and how they will function the best. Those facts do not need to be fully understood by performers when using IMUs. However, they might go against performers' preconceptions of motion-sensing technologies and generate frustration and confusion if they are all experienced simultaneously.

For example, understanding the limitations of the movement intensity calculated from the accelerometers is essential. Movement intensity is a simple yet powerful signal to link movements to sounds. Still, it must be scaled appropriately to measure the range of movements that the performer will use, and it will yield different outcomes depending on where the sensor is on the body. Movement intensity to sound loudness is probably the most elemental mapping for sound-movement interactions, but it already comes with challenges. The *Sympoiesis* workshop described section 6.4.5 only used this mapping, yet improvisation sessions with it fuelled several engaging discussions about the technology. This notion of scaling became apparent when they transitioned to slower movements and could not hear their sound anymore. They also decided to try using the sensor in their pocket to discover that it resulted in a much

different experience with the sound. Although we favored orientation over intensity as the primary control signal for the final interaction, this first contact with the technology was still fruitful for them in acquiring initial knowledge of the technology. This workshop was important in making dancers aware of simple but essential sensor characteristics. It induced dancers to experiment with the technology to understand how it worked and its limits.

Although not initially planned as an inductive learning session, this workshop led to inductive learning. Focusing on one mode of interaction left room for the dancers to adopt an exploratory approach to technology through experimentation. Retrospectively, we could have proposed different scenarios for dancers to motivate further exploration. Inductive learning starts with a problem statement, which states a goal for students. As discussed above, technology is not a solution to a problem in our context, but we had a defined artistic intent, and maybe we could have infused it more from the beginning. Moreover, oversimplification from the start can be detrimental; one dancer reported after trying to put the sensor in her pocket: *“It was more interesting because it’s just more difficult. It’s just more challenging than the arm, more fun, more tricky.”*

Challenges and constraints were also found fruitful for exploration. Interviewees in Chapter 3 highlighted how constraints brought by the interaction are beneficial to exploration: *“There is a big difference when it’s a personal choice and when it’s a choice that comes from outside. And also the personal choices, often, happen more automatically. There’s really a gain in having choices that are made from the outside”* (Yves Candau). However, the right balance must be found between *infantilizing* performers and giving them the complete system right away as *“you can get overwhelmed very quickly because you become a human synthesizer and so all you do is sound.”* (Philippe Spiesser).

Imposing technology on someone can be a violent experience if not preceded by thorough preparation. They can feel overwhelmed and lost and feel like they are failing at something that they should understand. Artists need confidence to be creative, both in themselves and in others. Again, we are not advocating for infantilization but for consideration of how intrusive technology can be. The preliminary experiment described in Chapter 4, proved very difficult for some participants, who felt they were too bad at the task rather than considering it too hard. Managing constraints and difficulty is essential to enable curiosity and creativity, which are key to exploration. However, it is easier said than done. When studying cooperative prototyping, Bødker and Grønbæk discussed the tension inherent to collaborative endeavors:

There is a tension between careful preparation of sessions and the inherent unpredictable character of the prototyping sessions. [...] This understanding can be utilized in preparation of sessions - not to put a tighter steering on the session, but to be prepared to better handle some of the most common and most important types of focus shifts that may occur. (Bødker and Grønbæk, 1991)

Bødker and Grønbæk advocates that *“breakdowns”* and *“focus shift”* in collaboration must be learned from to *“establish good conditions for user and designer creativity.”* The designer, as much as the performers and other collaborators, must learn from moments of tension and adapt to it (Sullivan et al., 2023).

Eventually, it is the designer’s responsibility to decide what should be constrained and what should remain open from the start, as well as what should be explained early on and what should be learned by exploration. Designing the pedagogy that will accompany her or his artifact is essential to enable exploration. And it is a process that takes place both before and during the creation sessions. Transposed to another context: the act of teaching is as

important as the act of preparation.

7.4 Chapter Summary

This chapter looks at the work presented in this dissertation, summarizing the main findings and comparing them with other research on the subject. First, learning interactive movement-sound systems involves a body of knowledge and skills that are, for one part, specific to each artistic context and, for the other, transversal to the whole. In particular, critical thinking, quality of attention and exploration are necessary for appropriating interactive technologies in artistic practices.

Exploration, in particular, is then discussed as a cross-disciplinary process that is omnipresent at different stages of creation. Interaction design includes exploring gestural interfaces, sound synthesis, and mapping. Designing interaction is an iterative process along these three dimensions, requiring exploration through experimentation and improvisation with systems. Therefore, exploration encompasses both goal-based and open-ended processes.

Finally, it is proposed that the designer should not only design the interaction but also the pedagogical process accompanying the system, a process that would encourage exploration and appropriation by the other artistic collaborators. While avoiding imposing a certain way of using interaction, which would norm practice associated with such systems, the invitation is made to consider learning an integral part of interaction design by leveraging on inductive teaching methods.

Conclusion and Perspectives

This PhD investigated the relationship between exploration, learning, and appropriation of movement-sound interaction systems in the performing arts. This research was initially motivated by recent studies showing that learning movement-sound interactions through exploration was possible. Meanwhile, exploration was also shown to be the key process leading to appropriation. However, exploration is a broad concept encompassing different semantics and processes in HCI, such as exploratory learning, navigation, and sensorimotor learning. We should remain cautious in proposing any simplistic single definition of exploration. To tackle such a research theme, we instead invoked several disciplines, from sonic interaction design to sensorimotor learning and different methods: field study, controlled experiment, and practice-based research.

In the background section, we first introduced broad concepts and recent results concerning digital musical instruments (DMIs), embodied sonic interaction, sensorimotor learning, and exploratory learning. To further situate our research, we led interviews with six artists using movement-sound interactive systems to investigate how practitioners relate to questions of learning and appropriation with technological systems. They were fruitful in bringing a different light to new interfaces for musical expression (NIME) research about movement-sound interactive systems. As other researchers in this field discussed, it is intrinsically difficult, or maybe fundamentally impossible, to translate global methods borrowed from traditional music practices in composition and pedagogy to practices based on interactive technology. Nevertheless, the insights given by the artists still reveal how important it is to privilege strong concepts over detailed implementation, the importance of openness and modularity of the system, and how constraints and perturbations can be fruitful in artistic contexts. Finally, they expressed that experimentation and exploration through improvisation are essential to appropriating interactive technologies. The observations drawn from the literature review and these interviews called for specific studies with exploration as the central theme to understand

how it relates to the learning and appropriation of movement-sound interactive systems.

We then focused on studying movement-sound exploration as a method enabling learning. We designed a prototype of movement-sound interaction based on corpus-based concatenative synthesis and controlled with a touch-based or a movement-based interface. Overall, this first study allowed us to validate our interest in studying, in controlled experiments, whole-body exploration of a sound space. In particular, we found that such embodied interaction, even when manipulating only two sound dimensions, cannot generally be described as a simple extension of fingertip planar exploration. Nevertheless, we found that repeated trials with whole-body interaction led to improvement, but several technical limitations seemed to hinder exploration. In particular, the inhomogeneity of the grain distribution in the sound space made exploration difficult. In response, we presented the theoretical basis and the Python implementation of the *polyspring* toolbox. It allows for the design of sound corpora where 2-dimensional sound clouds are distributed over a specific space while maintaining local neighborhood relationships between sound parameters. This opens new possibilities by enabling more control, for example, on the *CataRT* 2-dimensional representation.

Following this first experiment, we were interested in quantifying and qualifying the exploration of movement-sound spaces in exploratory learning contexts. Iterating on the previous design, we proposed a movement-sound interactive system called *mosX* that uses an abstract audio space controlled with wearable sensors. We further established an analysis method using kernel density estimation (KDE) to visualize exploration data. We proposed an empirical method to evaluate the participants' internal representation of audiomotor maps based on their drawings or sketches. This experiment revealed significant insights into how individuals learn, represent, and navigate complex sound spaces through embodied interaction. The findings underscore the effectiveness of exploration-based learning in early audiomotor map acquisition. In particular, our KDE-based method produced quantitative data and visualizations that enabled the identification of different exploratory behaviors between participants.

Parallel to controlled studies in the laboratory, practice-based research was led in the frame of two performances: *Sympoiesis* and *Circus Lab*. Both included movement-sound interactions designed by the author in collaboration with other artistic collaborators. The creation process was reported from the perspective of the interaction designer. It highlighted the discussions and tensions in such collaborations, emphasizing the necessary sharing that hybrid practices entail. These performances enabled us to consider exploration in a broader scope, particularly as a shared process in collaborative settings. As motivations and expectations in using interaction technologies varied between individuals involved in the project, exploration became a shared process in which the points of attention differed between actors, and the design directions to be taken were decided collectively. The motivations of other collaborators and the particular contexts led to varied interaction designs that were analyzed through the Spectators' Experience Design framework, which revealed how design choices led to different characteristics of the resulting interaction for the audience.

Finally, we proposed a discussion. First, learning interactive movement-sound systems involves a body of knowledge and skills that are, on the one hand, specific to each artistic context and, on the other hand, transversal and generalizable. Exploration, in particular, is discussed as a cross-disciplinary process omnipresent at different stages of creation and should be paired with design—and critical—thinking. This is part of the necessary process for appropriating interactive technologies in artistic practices. Exploration can also be understood as an iterative process along these three dimensions, requiring exploration through

experimentation and improvisation with systems. Therefore, exploration encompasses both goal-based and open-ended processes. We discuss that the designer's role resides not only in designing the interaction but also in accompanying the pedagogical process that would encourage exploration and appropriation by the other artistic collaborators. Avoiding imposing norms for practice, we propose that an invitation be made to consider exploration and learning an integral part of interaction design by leveraging inductive teaching methods.

This dissertation opened with a series of broad questions about movement-sound interaction, exploration, learning, and appropriation. This research work provides some answers that are worth investigating further. First, the *mosX* system proposed embodied interaction that was found interesting for fostering movements outside of the usual motion range used by participants. This result opens up the possibility of using *mosX* for the rehabilitation of people with acquired visual impairments who suffer a loss of motor skills as a result. This is currently being investigated by the ISMM team with psychomotricity experts who have tried this experiment. Second, as appropriation is a slow and highly contextual mechanism, it seems that field study, in contact with a population of artists whose environment is precisely taken into account and over a consequent period, is necessary for the objective and systematic analysis of what constitutes the appropriation of technologies. Such work is a matter for the social sciences and must be carried out, if not by a specialist in the field, at least in collaboration with researchers accustomed to these methods.

CHAPTER **A**

Appendix

A.1 Chapter 3

Interviewees' Background

We introduce here the artists who participated in the study reported in Chapter 3, along with a short description of their practices and associated works that were specifically referenced in the paper.

Bertha Bermudez¹ (BB):

Dancer, Dance Educator

Selection of works: Double Skin Double Mind Installation (Bermudez et al., 2011), CoMo-Elements (Matuszewski et al., 2018)

Yves Candau² (YC):

Dancer/artist, Coder

Selection of works: CO/DA (Françoise et al., 2022), Still Moving (Candau et al., 2017)

Fabrice Guédy³ (FG):

Music Educator, Composer, Pianist

Selection of works: Gesture Follower (Bevilacqua et al., 2007), Modular Musical Object

¹<https://www.lafaktoria.org/en/bertha-bermudez/>, accessed 23rd July 2024.

²Vimeo page, accessed 23rd July 2024.

³<https://feuillantines.com/>, accessed 23rd July 2024.

(Rasamimanana et al., 2011), Concert Féminin / Féminine⁴, Volière⁵

Mari Kimura⁶ (MK):

Violinist, Composer, Music Educator

Selection of works: Augmented violin with sensors (Kimura et al., 2012), MUGIC⁷

Michelle Agnes Magalhaes⁸ (MAM):

Composer, Performer, Music Educator

Selection of works: Constella(c)tions⁹, CoMo-Elements (Matuszewski et al., 2018)

Philippe Spiesser¹⁰ (PS):

Percussionist/Performer, Music Educator

Selection of works: GeKiPe (Fernandez et al., 2017), SkinAct¹¹

List of Questions

The questions used during interviews vary according to the interviewee's practice but the formulation and general themes remained consistent throughout the study. The following set of questions has been used to structure the interview with MK :

1. How would you describe yourself and your practices today?
2. What motivates you today in using technologies for movement in your artistic practice?
3. How do you create, practice and learn new pieces involving technology?
4. Would you like to pass on the knowledge and experience you have on these technologies to others? How?
5. Do you feel that your personality as a musician comes through when you play with this system?
6. What is your point of view about expressivity using technology / motion sensing? Has it evolved over time?

⁴<https://www.bnf.fr/fr/agenda/concert-feminin-feminine>, accessed 23rd July 2024.

⁵<http://gallicastudio.bnf.fr/voliere>, accessed 23rd July 2024.

⁶<http://www.marikimura.com/>, accessed 23rd July 2024.

⁷<https://mugimotion.com/>, accessed 23rd July 2024.

⁸<https://www.michelleagnes.net/>, accessed 23rd July 2024.

⁹https://vertigo.starts.eu/media/uploads/vertigo-constellactions-residency-public_report.pdf, accessed 23rd July 2024.

¹⁰<http://philippespiesser.com/en>, accessed 23rd July 2024.

¹¹<http://philippespiesser.com/en/projet/skinact-projet-de-recherche/>, accessed 23rd July 2024.

A.2 Chapter 4

A.2.1 Video Tutorial of Polyspring Max Examples

A video tutorial of the Max implementation of Polyspring is available at the following link.

https://drive.google.com/drive/folders/1aJmROHvKKQV1LZy3_BwteLnyeAt2aePq?

The video file is named **polyspring_tutorial.mp4**. The example and source code are available on github¹².

A.3 Chapter 5

A.3.1 Animations of the Exploration Distribution Evolution

Video Animations of the distribution evolution have been generated. They are available at the following link:

https://drive.google.com/drive/folders/1aJmROHvKKQV1LZy3_BwteLnyeAt2aePq?

Video file names follow this format: p[participant]_e[session].mp4, according to the participant and the session numbers. List of files:

p7_e1.mp4 exploration session 1 of participant 7

p7_e4.mp4 exploration session 4 of participant 7

p9_e1.mp4 exploration session 1 of participant 9

p9_e4.mp4 exploration session 4 of participant 9

p10_e1.mp4 exploration session 1 of participant 10

p10_e4.mp4 exploration session 4 of participant 10

The position of the participant in the parameter space is shown by the red dot. The sound feedback associated with the position is not recorded in these videos.

¹²<https://github.com/ircam-ismm/polyspring>

List of Figures

1.1	Methodological frameworks	3
1.2	Method framework of this dissertation	4
2.1	Examples of instrument-like interfaces	8
2.2	Framework for interaction with a computer-based musical instrument	9
2.3	Experimental setup from (van Vugt and Ostry, 2019)	15
4.1	2-dimensional touch interaction with a sound space created with CataRT.	33
4.2	Preliminary experiment setup	35
4.3	First experiment scores	37
4.4	Trajectories and corresponding drawings of the first session	38
4.5	Trajectories and corresponding drawings of the second session	38
4.6	Distribution manipulation for movement-based interaction	40
4.7	Unispring algorithm steps	41
4.8	Spring rest length for uniform distribution	43
4.9	Max <i>polyspring</i> demonstration interface	45
4.10	Uniformization example	45
4.11	Uniformization example	46
4.12	Gaussian attractors example	47
4.13	Uniform distribution in a circular region	48
5.1	Surjective mapping	54
5.2	Example of basic uniformization of a sound corpus. The left scatterplot represents the projection of the segmented corpus over two descriptors: the periodicity (x-axis) and the spectral centroid (y-axis). Each dot represent a sound sample. The scatterplot on the right depicts the representation resulting from a uniformization of the control space.	55
5.3	Sound space organization	56
5.4	Sound grains triggering description	57
5.5	Kinematic diagram of the arm with sensors positions	57
5.6	Overview of our analysis method in 2-Dimensions. A set of exploration points is a series of 2-Dimensional positions. From exploration points we compute a map representing visited parts of the space using kernel density estimation. We can estimate from these maps the proportion of the space that has been covered, locally, in subparts of the space, as well as globally, over the whole space.	60
5.7	Contour plot generation using KDE	62
5.8	Results for task A	67
5.9	Results for task B	68
5.10	Average coverage over time	69
5.11	Coverage vs session boxplot	70

5.12	Boxplot of the presence per session, split by quadrant.	70
5.13	Contour plots of the first exploration session.	71
5.14	Contour plot of sessions 1 and 4 for P7, P9, and P10	72
5.15	S1 and S4 temporal evolution for P7	73
5.16	S1 and S4 temporal evolution for P9	74
5.17	Exploration sessions temporal evolution for P10	75
5.18	Participants' sketches	76
6.1	Max4Live objects to receive IMU data	84
6.2	A plate of agar dyed to reveal the slime mold trail.	86
6.3	Exploratory workshop setup	89
6.4	Exploratory workshop pictures	89
6.5	Second workshop setup	93
6.6	Reeves spectator's experience design taxonomy	97
6.7	Classification of performances interactive setups	99

List of Tables

5.1	Summary of task results per participants and categories.	69
5.2	Frontiers represented in participants' sketches	76
6.1	Description of the movement sensing of each interactive setup.	97

List of Acronyms

2-D 2-Dimensions or 2-Dimensional

DoF Degrees of Freedom

HCI Human-Computer Interactions

IMU Inertial Measurement Unit

KDE Kernel Density Estimation

mosX Movement-Sound Explorer

NIME New Interfaces for Musical Expression

OSC Open Sound Control

PDF Probability Density Functions

SEAT Spectator Experience Augmentation Techniques

SID Sonic Interaction Design

STD Standard Deviation

Bibliography

- Alessandro Altavilla, Baptiste Caramiaux, and Atau Tanaka. Towards gestural sonic affordances. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2013.
- Ryan Aylward and Joseph A Paradiso. Senseble: A wireless, compact, multi-user sensor system for interactive dance. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 134–139, 2006.
- Liam J. Bannon. From human factors to human actors: The role of psychology and human-computer interaction studies in system design. In *Readings in Human–Computer Interaction*, pages 205–214. Elsevier, 1995. ISBN 978-0-08-051574-8. doi: 10.1016/B978-0-08-051574-8.50024-8.
- Steve Benford, Chris Greenhalgh, Andy Crabtree, Martin Flintham, Brendan Walker, Joe Marshall, Boriana Koleva, Stefan Rennick Egglestone, Gabriella Giannachi, Matt Adams, Nick Tandavanitj, and Ju Row Farr. Performance-led research in the wild. *ACM Transactions on Computer-Human Interaction*, 20(3):1–22, July 2013. ISSN 1073-0516, 1557-7325. doi: 10.1145/2491500.2491502.
- Bertha Bermudez, Scott Delahunta, Hoogenboom Marijke, Ziegler Chris, Frédéric Bevilacqua, Sarah Fdili Alaoui, and Barbara Meneses Gutierrez. The double skin/double mind interactive installation. *The Journal for Artistic Research*, 2011.
- Gilberto Bernardes. Interfacing sounds: Hierarchical audio-content morphologies for creative re-purposing in earGram 2.0. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 537–542. Zenodo, May 2021. doi: 10.5281/zenodo.4813176.
- Francisco Bernardo, Nicholas Arner, and Paul Batchelor. O soli mio: Exploring millimeter wave radar for musical interaction. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, volume 17, pages 283–286, 2017.
- Florent Berthaut, Mark T. Marshall, Sriram Subramanian, and Martin Hachet. Rouages: Revealing the mechanisms of digital musical instruments to the audience. In *Proceedings*

- of the International Conference on New Interfaces for Musical Expression*, pages 6–pages, 2013.
- Florent Berthaut, Victor Zappi, and Dario Mazzanti. Scenography of immersive virtual musical instruments. In *2014 IEEE VR Workshop: Sonic Interaction in Virtual Environments (SIVE)*, pages 19–24. IEEE, 2014.
- Frederic Bevilacqua, Fabrice Guédy, Norbert Schnell, Emmanuel Fléty, and Nicolas Leroy. Wireless sensor interface and gesture-follower for music pedagogy. In *Proceedings of the International Conference on New Interfaces for Musical Expression, NIME '07*, pages 124–129, New York, NY, USA, June 2007. Association for Computing Machinery. ISBN 978-1-4503-7837-6. doi: 10.1145/1279740.1279762.
- Frédéric Bevilacqua, Eric O. Boyer, Jules Françoise, Olivier Houix, Patrick Susini, Agnès Roby-Brami, and Sylvain Hanneton. Sensori-motor learning with movement sonification: Perspectives from recent interdisciplinary studies. *Frontiers in Neuroscience*, 10, August 2016. ISSN 1662-453X. doi: 10.3389/fnins.2016.00385.
- Frédéric Bevilacqua, Norbert Schnell, Jules Françoise, Éric O. Boyer, Diemo Schwarz, and Baptiste Caramiaux. Designing action–sound metaphors using motion sensing and descriptor-based synthesis of recorded sound materials. In Micheline Lesaffre, Pieter-Jan Maes, and Marc Leman, editors, *The Routledge Companion to Embodied Music Interaction*, pages 391–401. Routledge, New York ; London : Routledge, 2017., 1 edition, September 2017. ISBN 978-1-315-62136-4. doi: 10.4324/9781315621364-43.
- T. Blascheck, K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf, and T. Ertl. Visualization of eye tracking data: A taxonomy and survey. *Computer Graphics Forum*, 36(8):260–284, 2017. ISSN 1467-8659. doi: 10.1111/cgf.13079.
- Susanne Bødker. When second wave HCI meets third wave challenges. In *Proceedings of the 4th Nordic Conference on Human-computer Interaction: Changing Roles, NordiCHI '06*, pages 1–8, New York, NY, USA, October 2006. Association for Computing Machinery. ISBN 978-1-59593-325-6. doi: 10.1145/1182475.1182476.
- Susanne Bødker. Third-wave HCI, 10 years later—participation and sharing. *Interactions*, 22(5):24–31, August 2015. ISSN 1072-5520. doi: 10.1145/2804405.
- Susanne Bødker and Kaj Grønbaek. Cooperative prototyping: Users and designers in mutual activity. *International Journal of Man-Machine Studies*, 34(3):453–478, 1991.
- John Tyler Bonner. Cellular slime molds. In *Cellular Slime Molds*. Princeton University Press, 2015.
- Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, January 2006. ISSN 1478-0887, 1478-0895. doi: 10.1191/1478088706qp063oa.
- Virginia Braun, Victoria Clarke, Nikki Hayfield, and Gareth Terry. Thematic analysis. In Pranee Liamputtong, editor, *Handbook of Research Methods in Health Social Sciences*, pages 843–860. Springer, Singapore, 2019. ISBN 978-981-10-5251-4. doi: 10.1007/978-981-10-5251-4_103.
- Dom Brown, Chris Nash, and Tom Mitchell. Simple mappings, expressive movement: A qualitative investigation into the end-user mapping design of experienced mid-air musicians. *Digital Creativity*, 29(2-3):129–148, 2018.

- Yves Candau, Jules Françoise, Sarah Fdili Alaoui, and Thecla Schiphorst. Cultivating kinaesthetic awareness through interaction: Perspectives from somatic practices and embodied cognition. In *Proceedings of the 4th International Conference on Movement Computing*, page 8, June 2017. doi: 10.1145/3077981.3078042.
- Olivier Capra, Florent Berthaut, and Laurent Grisoni. Have a SEAT on stage: Restoring trust with spectator experience augmentation techniques. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference, DIS '20*, pages 695–707, New York, NY, USA, July 2020. Association for Computing Machinery. ISBN 978-1-4503-6974-9. doi: 10.1145/3357236.3395492.
- Baptiste Caramiaux, Frédéric Bevilacqua, Marcelo M. Wanderley, and Caroline Palmer. Dissociable effects of practice variability on learning motor and timing skills. *PLoS One*, 13(3):e0193580, March 2018. ISSN 1932-6203. doi: 10.1371/journal.pone.0193580.
- John M. Carroll, editor. *Interfacing Thought: Cognitive Aspects of Human-Computer Interaction*. Interfacing Thought: Cognitive Aspects of Human-Computer Interaction. The MIT Press, Cambridge, MA, US, 1987. ISBN 978-0-262-03125-7.
- Marta E. Cecchinato, Anna L. Cox, and Jon Bird. Always on(line)? User experience of smartwatches and their role within multi-device ecologies. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17*, pages 3557–3568, New York, NY, USA, May 2017. Association for Computing Machinery. ISBN 978-1-4503-4655-9. doi: 10.1145/3025453.3025538.
- Zhuoqun Chai, Mengxi Liu, Qian Shi, Yuanyuan Zhang, Minglin Zuo, and Da He. Fine-grained urban village extraction by mask transformer from high-resolution satellite images in pearl river delta. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17:13657–13668, 2024. ISSN 2151-1535. doi: 10.1109/JSTARS.2024.3434487.
- Van Cott and Kinkade. *Human Engineering Guide to Equipment Design*. Department of Defense, 1972. ISBN 978-0-471-80011-8.
- Palle Dahlstedt and Ami Skånberg. OtoKin: Mapping for sound space exploration through dance improvisation. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, August 2019.
- Rudolph P. Darken and John L. Sibert. Navigating large virtual spaces. *International Journal of Human-Computer Interaction*, 8(1):49–71, January 1996. ISSN 1044-7318. doi: 10.1080/10447319609526140.
- Rudy P. Darken and John L. Sibert. A toolset for navigation in virtual environments. In *Proceedings of the 6th Annual ACM Symposium on User Interface Software and Technology*, pages 157–165, Atlanta Georgia USA, December 1993. ACM. ISBN 978-0-89791-628-8. doi: 10.1145/168642.168658.
- Alain De Cheveigné and Hideki Kawahara. YIN, a fundamental frequency estimator for speech and music. *Journal of the Acoustical Society of America*, 111(4):1917–1930, 2002.
- Sjaak De Mul and Herre Van Oostendorp. Learning user interfaces by exploration. *Acta Psychologica*, 91(3):325–344, April 1996. ISSN 00016918. doi: 10.1016/0001-6918(95)00060-7.
- Beth Dempster. Sympoietic and autopoietic systems: A new distinction for self-organizing

- systems. In *Proceedings of the World Congress of the Systems Sciences and ISSS*, pages 1–18. Citeseer, 2000.
- Audrey Desjardins and Aubree Ball. Revealing tensions in autobiographical design in HCI. In *Proceedings of the 2018 Designing Interactive Systems Conference, DIS '18*, pages 753–764, New York, NY, USA, June 2018. Association for Computing Machinery. ISBN 978-1-4503-5198-0. doi: 10.1145/3196709.3196781.
- Alan Dix. Designing for appropriation. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...but Not as We Know It - Volume 2, BCS-HCI '07*, pages 27–30, Swindon, GBR, September 2007. BCS Learning & Development Ltd. ISBN 978-1-902505-95-4.
- Christopher Dobrian and Daniel Koppelman. The ‘E’ in NIME: Musical expression with new computer interfaces. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2006.
- Marco Donnarumma. *Configuring Corporeality: Performing Bodies, Vibrations and New Musical Instruments*. PhD thesis, Goldsmiths, University of London, 2016.
- Kees Dorst and Nigel Cross. Creativity in the design process: Co-evolution of problem–solution. *Design Studies*, 22(5):425–437, September 2001. ISSN 0142-694X. doi: 10.1016/S0142-694X(01)00009-6.
- Paul Dourish. Embodied interaction: Exploring the foundations of a new approach to HCI, January 1999.
- Paul Dourish. *Where the Action Is: The Foundations of Embodied Interaction*. MIT press, 2001.
- Aaron Einbond, Diemo Schwarz, Riccardo Borghesi, and Norbert Schnell. Introducing CatOracle: Corpus-based concatenative improvisation with the audio oracle algorithm. In Hans Timmermans, editor, *International Computer Music Conference (ICMC)*, pages 141–147, Utrecht, Netherlands, September 2016. HKU University of the Arts Utrecht, HKU Music and Technology / Hans Timmermans.
- Daniel Fallman. The interaction design research triangle of design practice, design studies, and design exploration. *Design Issues*, 24(3):4–18, 2008. ISSN 0747-9360.
- Stefano Fasciani and Lonce Wyse. Adapting general purpose interfaces to synthesis engines using unsupervised dimensionality reduction techniques and inverse mapping from features to parameters. *Principal Component Analysis*, page 7, 2012.
- Sarah Fdili Alaoui. Making an interactive dance piece: Tensions in integrating technology in art. In *Proceedings of the 2019 on Designing Interactive Systems Conference, DIS '19*, pages 1195–1208, New York, NY, USA, June 2019. Association for Computing Machinery. ISBN 978-1-4503-5850-7. doi: 10.1145/3322276.3322289.
- Sidney Fels, Ashley Gadd, and Axel Mulder. Mapping transparency through metaphor: Towards more expressive musical instruments. *Organised Sound*, 7(2):109–126, 2002.
- José Miguel Fernandez, Thomas Köppel, Nina Verstraete, Grégoire Lorieux, Alexander Vert, and Philippe Spiesser. GeKiPe, a gesture-based interface for audiovisual performance. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 450–455, Copenhagen, Denmark, 2017. Aalborg University Copenhagen. doi: 10.5281/zenodo.1176312.

-
- Rebecca Fiebrink and Laetitia Sonami. Reflections on eight years of instrument creation with machine learning. In Romain Michon and Franziska Schroeder, editors, *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 237–242, Birmingham, UK, July 2020. Birmingham City University.
- Rebecca Fiebrink, Perry R Cook, and Dan Trueman. Human model evaluation in interactive supervised learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 147–156, 2011.
- Jules Françoise and Frédéric Bevilacqua. Motion-sound mapping through interaction: An approach to user-centered design of auditory feedback using machine learning. *ACM Transactions on Interactive Intelligent Systems*, 8(2):1–30, July 2018. ISSN 2160-6455, 2160-6463. doi: 10.1145/3211826.
- Jules Françoise, Yves Candau, Sarah Fdili Alaoui, and Thecla Schiphorst. Designing for kinesthetic awareness: Revealing user experiences through second-person inquiry. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 5171–5183, Denver Colorado USA, May 2017. ACM. ISBN 978-1-4503-4655-9. doi: 10.1145/3025453.3025714.
- Jules Françoise, Sarah Fdili Alaoui, and Yves Candau. CO/DA: Live-coding movement-sound interactions for dance improvisation. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, pages 1–13, New York, NY, USA, April 2022. Association for Computing Machinery. ISBN 978-1-4503-9157-3. doi: 10.1145/3491102.3501916.
- Karmen Franinović and Stefania Serafin, editors. *Sonic Interaction Design*. The MIT Press, Cambridge, Massachusetts, 2013. ISBN 978-0-262-01868-5.
- A. Cavan Fyans, Adnan Marquez-Borbon, Paul Stapleton, and Michael Gurevich. Ecological considerations for participatory design of DMIs. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Ann Arbor, Michigan, 2012. University of Michigan. doi: 10.5281/zenodo.1178257.
- Leandro Garber, T Ciccola, and Jc Amusategui. AudioStellar, an open source corpus-based musical instrument for latent sound structure discovery and sonic experimentation. In *Proceedings of ICMC*, Santiago, Chile, 2020.
- Andrea Giomi. Somatic sonification in dance performances. From the artistic to the perceptual and back. In *Proceedings of the 7th International Conference on Movement and Computing*, MOCO '20, pages 1–8, New York, NY, USA, July 2020. Association for Computing Machinery. ISBN 978-1-4503-7505-4. doi: 10.1145/3401956.3404226.
- Rolf Inge Godøy and Marc Leman. *Musical Gestures: Sound, Movement, and Meaning*. Routledge, 2010.
- Mercedes Gonzalez-Sanmamed, Albert Sangrà, Alba Souto-Seijo, and Iris Estévez Blanco. Learning ecologies in the digital era: Challenges for higher education. *Publicaciones*, 50(1):83–102, October 2020. ISSN 2530-9269. doi: 10.30827/publicaciones.v50i1.15671.
- Elizabeth Goodman, Erik Stolterman, and Ron Wakkary. Understanding interaction design practices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1061–1070, Vancouver BC Canada, May 2011. ACM. ISBN 978-1-4503-0228-9. doi: 10.1145/1978942.1979100.

- Julien Gradoz and Sandra Hoibian. La fracture numérique française au travers d'une approche par les « capacités » : L'enjeu d'apprendre à apprendre. *Annales des Mines - Gerer comprendre*, 136(2):37–51, June 2019. ISSN 0295-4397.
- Jonathan Grudin and Phil Barnard. The cognitive demands of learning and representing command names for text editing. *Human Factors*, 26(4):407–422, August 1984. ISSN 0018-7208. doi: 10.1177/001872088402600404.
- Michael Gurevich and Jeffrey Treviño. Expression and its discontents: Toward an ecology of musical creation. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME '07, pages 106–111, New York, NY, USA, June 2007. Association for Computing Machinery. ISBN 978-1-4503-7837-6. doi: 10.1145/1279740.1279759.
- Jihyun Han and Nicolas Gold. Lessons learned in exploring the leap motion TM sensor for gesture-based instrument design. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2014.
- Steve Harrison, Deborah Tatar, and Phoebe Sengers. The three paradigms of HCI. In *Alt. Chi. Session at the SIGCHI Conference on Human Factors in Computing Systems San Jose, California, USA*, pages 1–18, 2007.
- Lauren Hayes. Beyond skill acquisition: Improvisation, interdisciplinarity, and enactive music cognition. *Contemporary Music Review*, 38(5):446–462, 2019. doi: 10.1080/07494467.2019.1684059.
- Kristina Höök. Transferring qualities from horseback riding to design. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, NordiCHI '10, pages 226–235, New York, NY, USA, October 2010. Association for Computing Machinery. ISBN 978-1-60558-934-3. doi: 10.1145/1868914.1868943.
- Andrew Howes and Stephen J. Payne. Semantic analysis during exploratory learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '90, pages 399–406, New York, NY, USA, 1990. Association for Computing Machinery. ISBN 0-201-50932-6. doi: 10.1145/97243.97316.
- Andy Hunt, Marcelo M Wanderley, and Ross Kirk. Towards a model for instrumental mapping in expert musical interaction. In *Icmc*, 2000.
- Daniel Jansson, Olov Rosen, and Alexander Medvedev. Non-parametric analysis of eye-tracking data by anomaly detection. In *2013 European Control Conference (ECC)*, pages 632–637, Zurich, July 2013. IEEE. ISBN 978-3-033-03962-9. doi: 10.23919/ECC.2013.6669561.
- Florian Jeanne, Indira Thouvenin, and Alban Lenglet. A study on improving performance in gesture training through visual guidance based on learners' errors. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pages 1–10, Gothenburg Sweden, November 2017. ACM. ISBN 978-1-4503-5548-3. doi: 10.1145/3139131.3139144.
- Alexander Refsum Jensenius. An action–sound approach to teaching interactive music. *Organised Sound*, 18(2):178–189, 2013.
- Alexander Refsum Jensenius. *Sound Actions: Conceptualizing Musical Instruments*. The MIT Press, December 2022. ISBN 978-0-262-37220-6. doi: 10.7551/mitpress/14220.001.0001.

-
- Alexander Refsum Jensenius and Victoria Johnson. A video based analysis system for realtime control of concatenative sound synthesis and spatialisation. In *Norwegian Artificial Intelligence Symposium*, 2010.
- Alexander Refsum Jensenius and Michael J. Lyons. *A Nime Reader: Fifteen Years of New Interfaces for Musical Expression*, volume 3. Springer, 2017.
- Xiaodong Jin and Janusz Kawczak. Birnbaum-saunders and lognormal kernel estimators for modelling durations in high frequency financial data. *Annals of Economics and Finance*, 4:103–124, 2003.
- Andrew Johnston, Linda Candy, and Ernest Edmonds. Designing and evaluating virtual musical instruments: Facilitating conversational user interaction. *Design Studies*, 29(6): 556–571, 2008.
- M. P. Wand Jones, M. C. *Kernel Smoothing*. Chapman and Hall/CRC, New York, December 1994. ISBN 978-0-429-17059-1. doi: 10.1201/b14876.
- S Jorda. Instruments and players: Some thoughts on digital lutherie. *Journal of New Music Research*, 33(3):321–341, September 2004. ISSN 0929-8215. doi: 10.1080/0929821042000317886.
- Sergi Jordà, Günter Geiger, Marcos Alonso, and Martin Kaltenbrunner. The ReacTable: Exploring the synergy between live music performance and tabletop tangible interfaces. In *Tei '07*, pages 139–146, New York, NY, USA, 2007. Association for Computing Machinery. ISBN 978-1-59593-619-6. doi: 10.1145/1226969.1226998.
- Mari Kimura, Nicolas Rasamimanana, Frédéric Bevilacqua, Bruno Zamborlin, Norbert Schnell, and Emmanuel Fléty. Extracting human expression for interactive composition with the augmented violin. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 1–1, 2012.
- Muneo Kitajima and Peter G. Polson. A comprehension-based model of exploration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '96, pages 324–331, New York, NY, USA, April 1996. Association for Computing Machinery. ISBN 978-0-89791-777-3. doi: 10.1145/238386.238536.
- Malcolm S. Knowles. *Self-directed learning: A guide for learners and teachers*. Technical report, Association Press, 291 Broadway, New York, New York 10007 (\$4, 1975).
- Ianis Lallemand and Diemo Schwarz. Interaction-optimized sound database representation. In *DAFx*, pages 292–300, Paris, France, September 2011.
- Laurane Le Goff. Costumes that tell multispecies stories: Critique of sympoiesis: A bio-inspired dance performance. *Studies in Costume & Performance*, 8(2):241–262, December 2023. ISSN 2052-4013, 2052-4021. doi: 10.1386/scp_00099_1.
- Marc Leman. *Embodied Music Cognition and Mediation Technology*. MIT press, 2007.
- Micheline Lesaffre, Pieter-Jan Maes, and Marc Leman. *The Routledge Companion to Embodied Music Interaction*. Taylor & Francis, 2017.
- Charles J. Limb and Allen R. Braun. Neural substrates of spontaneous musical performance: An fMRI study of jazz improvisation. *PLoS One*, 3(2):e1679, February 2008. ISSN 1932-6203. doi: 10.1371/journal.pone.0001679.
- Andrés Lucero, Audrey Desjardins, Carman Neustaedter, Kristina Höök, Marc Hassenzahl,

- and Marta E. Cecchinato. A sample of one: First-person research methods in HCI. In *Companion Publication of the 2019 on Designing Interactive Systems Conference 2019 Companion*, pages 385–388, San Diego CA USA, June 2019. ACM. ISBN 978-1-4503-6270-2. doi: 10.1145/3301019.3319996.
- Andrés Lucero, Audrey Desjardins, and Carman Neustaedter. Longitudinal first-person HCI research methods. In Evangelos Karapanos, Jens Gerken, Jesper Kjeldskov, and Mikael B. Skov, editors, *Advances in Longitudinal HCI Research*, pages 79–99. Springer International Publishing, Cham, 2021. ISBN 978-3-030-67321-5 978-3-030-67322-2. doi: 10.1007/978-3-030-67322-2_5.
- Ross Maciejewski, Stephen Rudolph, Ryan Hafen, Ahmad Abusalah, Mohamed Yakout, Mourad Ouzzani, William S. Cleveland, Shaun J. Grannis, and David S. Ebert. A visual analytics approach to understanding spatiotemporal hotspots. *IEEE Transactions on Visualization and Computer Graphics*, 16(2):205–220, March 2010. ISSN 1941-0506. doi: 10.1109/TVCG.2009.100.
- Wendy E. Mackay and Anne-Laure Fayard. HCI, natural science and design: A framework for triangulation across disciplines. In *Proceedings of the Conference on Designing Interactive Systems Processes, Practices, Methods, and Techniques - DIS '97*, pages 223–234, Amsterdam, The Netherlands, 1997. ACM Press. ISBN 978-0-89791-863-3. doi: 10.1145/263552.263612.
- Sebastian Madgwick and Thomas Mitchell. X-OSC: A versatile wireless I/O device for creative/music applications. In *SMC Sound and Music Computing Conference*, 2013.
- E Magalhães, João Jacob, Niels Nilsson, Rolf Nordahl, and Gilberto Bernardes. Physics-based concatenative sound synthesis of photogrammetric models for aural and haptic feedback in virtual environments. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 376–379. IEEE, 2020.
- Mary Mainsbridge. *Body as Instrument*. Bloomsbury, bloomsbury academic edition, 2022.
- Adnan Marquez-Borbon. Collaborative learning with interactive music systems. In Romain Michon and Franziska Schroeder, editors, *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 581–586, Birmingham, UK, July 2020. Birmingham City University.
- Adnan Marquez-Borbon and Juan Pablo Martinez-Avila. The problem of DMI adoption and longevity: Envisioning a NIME performance pedagogy. In Thomas Martin Luke Dahl, Douglas Bowman, editor, *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 190–195, Blacksburg, Virginia, USA, June 2018. Virginia Tech. ISBN 978-1-949373-99-8. doi: 10.5281/zenodo.1302541.
- Raul Masu, Adam Pultz Melbye, John Sullivan, and Alexander Refsum Jensenius. NIME and the environment: Toward a more sustainable NIME practice. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. PubPub, 2021.
- M. V. Mathews. The digital computer as a musical instrument. *Science*, 142(3592):553–557, November 1963. doi: 10.1126/science.142.3592.553.
- Benjamin Matuszewski, Joseph Larralde, and Frédéric Bevilacqua. Designing movement driven audio applications using a web-based interactive machine learning toolkit. In *Web Audio Conference (WAC)*, 2018.

-
- Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.
- Carolina Brum Medeiros and Marcelo M Wanderley. A comprehensive review of sensors and instrumentation methods in devices for musical expression. *Sensors*, 14(8):13556–13591, 2014.
- Thomas J Mitchell, Sebastian Madgwick, and Imogen Heap. Musical interaction with hand posture and orientation: A toolbox of gestural control mechanisms. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2012.
- Ali Momeni and Cyrille Henry. Dynamic independent mapping layers for concurrent control of audio and video synthesis. *Computer Music Journal*, 30(1):49–66, 2006.
- Fabio Morreale. Design for longevity: Ongoing use of instruments from NIME 2010-14. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2017.
- Michael J. Muller and Sarah Kuhn. Participatory design. *Communications of the ACM*, 36(6):24–28, June 1993. ISSN 0001-0782, 1557-7317. doi: 10.1145/153571.255960.
- Tim Murray-Browne and Panagiotis Tigas. Latent mappings: Generating open-ended expressive mappings using variational autoencoders. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Shanghai, China, June 2021. doi: 10.21428/92fbeb44.9d4bcd4b.
- Sarah Nabi, Philippe Esling, Geoffroy Peeters, and Frédéric Bevilacqua. Embodied exploration of deep latent spaces in interactive dance-music performance. In *Proceedings of the 9th International Conference on Movement and Computing*, pages 1–9, Utrecht Netherlands, May 2024. ACM. ISBN 9798400709944. doi: 10.1145/3658852.3659072.
- J. Nielsen, R. L. Mack, K. H. Bergendorff, and N. L. Grischkowsky. Integrated software usage in the professional work environment: Evidence from questionnaires and interviews. *ACM SIGCHI Bulletin*, 17(4):162–167, April 1986. ISSN 0736-6906. doi: 10.1145/22339.22366.
- Luc Nijs. The merging of musician and musical instrument: Incorporation, presence, and levels of embodiment. In *The Routledge Companion to Embodied Music Interaction*. Routledge, 2017. ISBN 978-1-315-62136-4.
- Kristian Nymoan, Mari Romarheim Haugen, and Alexander Refsum Jensenius. MuMYO - evaluating and exploring the MYO armband for musical interaction. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, NIME 2015, pages 215–218, Baton Rouge, Louisiana, USA, May 2015. The School of Music and the Center for Computation and Technology (CCT), Louisiana State University. ISBN 978-0-692-49547-6.
- Sile O’Modhrain. A framework for the evaluation of digital musical instruments. *Computer Music Journal*, 35(1):28–42, 2011. ISSN 0148-9267.
- Sofia Pantouvaki. Embodied interactions: Towards an exploration of the expressive and narrative potential of performance costume through wearable technologies. *Scene*, 2(1-2): 179–196, 2014.
- Victor Paredes, Jules Françoise, and Frédéric Bevilacqua. Entangling practice with artistic and educational aims: Interviews on technology-based movement sound interactions. In

- Proceedings of the International Conference on New Interfaces for Musical Expression*, Auckland, New Zealand, June 2022.
- Victor Paredes, Jules Francoise, and Frederic Bevilacqua. Polyspring: A python toolbox to manipulate 2-D sound database representations. In *20th Sound and Music Computing Conference (SMC 2023)*, Stockholm, Sweden, June 2023. Zenodo. doi: 10.5281/zenodo.8399016.
- Geoffroy Peeters. A large set of audio features for sound description (similarity and classification) in the CUIDADO project. *CUIDADO Ist Project Report*, 54(0):1–25, 2004.
- Per-Olof Persson and Gilbert Strang. A simple mesh generator in MATLAB. *SIAM Review*, 46(2):329–345, January 2004. ISSN 0036-1445, 1095-7200. doi: 10.1137/S0036144503429121.
- Jon Pigrem and Andrew P. McPherson. Do we speak sensor? Cultural constraints of embodied interaction. In Thomas Martin Luke Dahl, Douglas Bowman, editor, *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 382–385, Blacksburg, Virginia, USA, June 2018. Virginia Tech. ISBN 978-1-949373-99-8. doi: 10.5281/zenodo.1302633.
- Michael Prince and Richard Felder. The many faces of inductive teaching and learning. *Journal of college science teaching*, 36(5):14, 2007.
- Michael J. Prince and Richard M. Felder. Inductive teaching and learning methods: Definitions, comparisons, and research bases. *Journal of Engineering Education*, 95(2):123–138, April 2006. ISSN 10694730. doi: 10.1002/j.2168-9830.2006.tb00884.x.
- Juan Ramos, Esteban Ramón Calcagno, Ramiro oscar Vergara, Pablo Riera, and Joaquín Rizza. Bandoneon 2.0: An interdisciplinary project for research and development of electronic bandoneons in Argentina. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, June 2022. doi: 10.21428/92fbeb44.c38bfb86.
- Nicolas Rasamimanana, Frederic Bevilacqua, Norbert Schnell, Fabrice Guedy, Emmanuel Flety, Come Maestracci, Bruno Zamborlin, Jean-Louis Frechin, and Uros Petrevski. 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, Funchal, Portugal, 2011. ACM Press. ISBN 978-1-4503-0478-8. doi: 10.1145/1935701.1935704.
- Stuart Reeves, Steve Benford, Claire O'Malley, and Mike Fraser. Designing the spectator experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 741–750, Portland Oregon USA, April 2005. ACM. ISBN 978-1-58113-998-3. doi: 10.1145/1054972.1055074.
- John Rieman. A field study of exploratory learning strategies. *ACM Transactions on Computer-Human Interaction*, 3(3):189–218, September 1996. ISSN 1073-0516. doi: 10.1145/234526.234527.
- Nicole Robson, Andrew McPherson, and Nick Bryan-Kinns. Being with the waves: An ultrasonic art installation enabling rich interaction without sensors. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, June 2022. doi: 10.21428/92fbeb44.376bc758.
- Ido Roll, Manolis Mavrikis, and Sergio Gutierrez-Santos. Striking a balance between free

- and guided exploration - conceptualizing support for exploratory learning environments. In *Proceedings of the 9th International Conference of the Learning Sciences - Volume 2*, ICLS '10, pages 507–508, Chicago, Illinois, 2010. International Society of the Learning Sciences.
- Gerard Roma, Owen Green, and P A Tremblay. Adaptive mapping of sound collections for data-driven musical interfaces. In *New Interfaces for Musical Expression*, 2019.
- Gerard Roma, Anna Xambó, Owen Green, and Pierre Alexandre Tremblay. A general framework for visualization of sound collections in musical interfaces. *Applied Sciences*, 11(24):11926, December 2021. ISSN 2076-3417. doi: 10.3390/app112411926.
- Nina Schaffert, Thenille Braun Janzen, Klaus Mattes, and Michael H. Thaut. A review on the relationship between sound and movement in sports and rehabilitation. *Frontiers in Psychology*, 10:244, February 2019. ISSN 1664-1078. doi: 10.3389/fpsyg.2019.00244.
- Norbert Schnell, Axel Röbel, Diemo Schwarz, Geoffroy Peeters, and Ricardo Borghesi. MuBu & friends – assembling tools for content based real-time interactive audio processing in max/MSP. In *Icmc*, Montreal, Canada, 2009.
- Elliott Schwartz. Electronic music: A thirty-year retrospective. *Music Educators Journal*, 64(7):36–41, March 1978. ISSN 0027-4321. doi: 10.2307/3395445.
- Diemo Schwarz. Corpus-based concatenative synthesis. *IEEE Signal Processing Magazine*, 24(2):92–104, 2007.
- Diemo Schwarz, Grégory Beller, Bruno Verbrugge, and Sam Britton. Real-time corpus-based concatenative synthesis with catart. In *9th International Conference on Digital Audio Effects (DAFx)*, pages 279–282, 2006.
- Diemo Schwarz, Wanyu Liu, and Frédéric Bevilacqua. A survey on the use of 2D touch interfaces for musical expression. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Birmingham, United Kingdom, July 2020.
- David W. Scott. Multivariate density estimation and visualization. Working Paper 2004,16, Papers, 2004.
- David W. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley & Sons, 2015.
- Hugo Scurto. *Designing with Machine Learning for Interactive Music Dispositifs*. Theses, Sorbonne Université, December 2019.
- Hugo Scurto, Bavo Van Kerrebroeck, Baptiste Caramiaux, and Frédéric Bevilacqua. Designing deep reinforcement learning for human parameter exploration. *ACM Transactions on Computer-Human Interaction*, 28(1):1–35, February 2021. ISSN 1073-0516, 1557-7325. doi: 10.1145/3414472.
- Phoebe Sengers. What I learned on change islands: Reflections on IT and pace of life. *Interactions*, 18(2):40–48, March 2011. ISSN 1072-5520. doi: 10.1145/1925820.1925830.
- Sertan Sentürk, Sang Won Lee, Avinash Sastry, Anosh Daruwalla, and Gil Weinberg. Crossole: A gestural interface for composition, improvisation and performance using kinect. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2012.
- M.M. Siddiqui. Distribution of quantiles in samples from a bivariate population. *Journal of*

Research of the National Bureau of Standards Section B Mathematics and Mathematical Physics, 64B(3):145, July 1960. ISSN 0022-4340. doi: 10.6028/jres.064B.017.

Marilyn Stember. Advancing the social sciences through the interdisciplinary enterprise. *Social Science Journal*, 28(1):1–14, March 1991. ISSN 0362-3319, 1873-5355. doi: 10.1016/0362-3319(91)90040-B.

Dagmar Sternad. It's not (only) the mean that matters: Variability, noise and exploration in skill learning. *Current Opinion in Behavioral Sciences*, 20:183–195, April 2018. ISSN 23521546. doi: 10.1016/j.cobeha.2018.01.004.

David Sudnow and Hubert L. Dreyfus. *Ways of the Hand: A Rewritten Account*. MIT Press, 2001.

John D. Sullivan, Sarah Fdili Alaoui, Pierre Godard, and Liz Santoro. Embracing the messy and situated practice of dance technology design. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference, DIS '23*, pages 1383–1397, New York, NY, USA, July 2023. Association for Computing Machinery. ISBN 978-1-4503-9893-0. doi: 10.1145/3563657.3596078.

Atau Tanaka. Embodied musical interaction: Body physiology, cross modality, and sonic experience. In Simon Holland, Tom Mudd, Katie Wilkie-McKenna, Andrew McPherson, and Marcelo M. Wanderley, editors, *New Directions in Music and Human-Computer Interaction*, pages 135–154. Springer International Publishing, Cham, 2019. ISBN 978-3-319-92068-9 978-3-319-92069-6. doi: 10.1007/978-3-319-92069-6_9.

Atsushi Tero, Seiji Takagi, Tetsu Saigusa, Kentaro Ito, Dan P. Bebber, Mark D. Fricker, Kenji Yumiki, Ryo Kobayashi, and Toshiyuki Nakagaki. Rules for biologically inspired adaptive network design. *Science*, 327(5964):439–442, January 2010. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1177894.

Amanda S. Therrien, Daniel M. Wolpert, and Amy J. Bastian. Increasing motor noise impairs reinforcement learning in healthy individuals. *eneuro*, 5(3):ENEURO.50–18.2018, May 2018. ISSN 2373-2822. doi: 10.1523/ENEURO.0050-18.2018.

Edward C. Tolman. Cognitive maps in rats and men. *Psychological Review*, 55(4):189–208, 1948. ISSN 1939-1471, 0033-295X. doi: 10.1037/h0061626.

Pierre Alexandre Tremblay, Owen Green, Gerard Roma, and Alexander Harker. From collections to corpora: Exploring sounds through fluid decomposition. In *International Computer Music Conference and New York City Electroacoustic Music Festival*, pages 223–228. International Computer Music Association, 2019.

Pierre Alexandre Tremblay, Gerard Roma, and Owen Green. Digging it: Programmatic data mining as musicking. In *International Computer Music Conference*, 2021.

Carol-Ina Trudel. Exploratory learning of interactive devices. In *Conference Companion on Human Factors in Computing Systems, CHI '94*, pages 75–76, New York, NY, USA, 1994. Association for Computing Machinery. ISBN 0-89791-651-4. doi: 10.1145/259963.260051.

Carol-Ina Trudel and Stephen J. Payne. Reflection and goal management in exploratory learning. *International Journal of Human-Computer Studies*, 42(3):307–339, March 1995. ISSN 10715819. doi: 10.1006/ijhc.1995.1015.

John Wilder Tukey. Exploratory data analysis. *Reading/Addison-Wesley*, 1977.

-
- Teija Vainio. A review of the navigation HCI research during the 2000s. *International Journal of Interactive Mobile Technologies (iJIM)*, 4(3):36–42, July 2010. ISSN 1865-7923. doi: 10.3991/ijim.v4i3.1270.
- Doug Van Nort, Marcelo M. Wanderley, and Philippe Depalle. Mapping control structures for sound synthesis: Functional and topological perspectives. *Computer Music Journal*, 38(3):6–22, September 2014. ISSN 0148-9267. doi: 10.1162/COMJ_a_00253.
- F. T. van Vugt and D. J. Ostry. Early stages of sensorimotor map acquisition: Learning with free exploration, without active movement or global structure. *Journal of Neurophysiology*, 122(4):1708–1720, October 2019. ISSN 0022-3077, 1522-1598. doi: 10.1152/jn.00429.2019.
- Floris T. van Vugt and David J. Ostry. From known to unknown: Moving to unvisited locations in a novel sensorimotor map: Moving to unvisited locations in sensorimotor maps. *Annals of the New York Academy of Sciences*, 1423(1):368–377, July 2018a. ISSN 00778923. doi: 10.1111/nyas.13608.
- Floris T. van Vugt and David J. Ostry. The structure and acquisition of sensorimotor maps. *Journal of Cognitive Neuroscience*, 30(3):290–306, March 2018b. ISSN 0898-929X, 1530-8898. doi: 10.1162/jocn_a_01204.
- M. P. Wand and M. C. Jones. Comparison of smoothing parameterizations in bivariate kernel density estimation. *Journal of the American Statistical Association*, 88(422):520–528, June 1993. ISSN 0162-1459. doi: 10.1080/01621459.1993.10476303.
- Marcelo M Wanderley. Gestural control of music. In *International Workshop Human Supervision and Control in Engineering and Music*, pages 632–644, 2001.
- David Wessel and Matthew Wright. Problems and prospects for intimate musical control of computers. *Computer Music Journal*, 26(3):11–22, 2002. ISSN 0148-9267.
- Travis West, Baptiste Caramiaux, Stéphane Huot, and Marcelo M. Wanderley. Making mappings: Design criteria for live performance. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Shanghai, China, June 2021. doi: 10.21428/92fbeb44.04f0fc35.
- Travis J West, Marcelo Wanderley, and Baptiste Caramiaux. Making mappings: Examining the design process. In Romain Michon and Franziska Schroeder, editors, *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 291–296, Birmingham, UK, July 2020. Birmingham City University.
- Allan W Wicker. *An Introduction to Ecological Psychology*. Number 1 in Environment and Behavior. CUP Archive, cup archive edition, 1984.
- Susan Wiedenbeck and Patti L. Zila. Hands-on practice in learning to use software: A comparison of exercise, exploration, and combined formats. *ACM Transactions on Computer-Human Interaction*, 4(2):169–196, June 1997. ISSN 1073-0516. doi: 10.1145/254945.254967.
- Daniel M. Wolpert, Jörn Diedrichsen, and J. Randall Flanagan. Principles of sensorimotor learning. *Nature Reviews Neuroscience*, 12(12):739–751, December 2011. ISSN 1471-003X, 1471-0048. doi: 10.1038/nrn3112.
- Howard G Wu, Yohsuke R Miyamoto, Luis Nicolas Gonzalez Castro, Bence P Ölveczky, and Maurice A Smith. Temporal structure of motor variability is dynamically regulated

- and predicts motor learning ability. *Nature Neuroscience*, 17(2):312–321, February 2014. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.3616.
- Min-Joon Yoo, Jin-Wook Beak, and In-Kwon Lee. Creating musical expression using kinect. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 324–325, 2011.
- Victor Zappi and Andrew McPherson. Dimensionality and appropriation in digital musical instrument design. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 455–460, London, United Kingdom, June 2014. Goldsmiths, University of London. doi: 10.5281/zenodo.1178993.
- Victor Zappi and Andrew McPherson. Hackable instruments: Supporting appropriation and modification in digital musical interaction. *Frontiers in ICT*, 5:26, 2018.
- Victor Zappi, Andrew Allen, and Sidney Fels. Extended playing techniques on an augmented virtual percussion instrument. *Computer Music Journal*, 42(2):8–21, 2018.
- Luis Zayas-Garin, Jacob Harrison, Robert Jack, and Andrew McPherson. DMI apprenticeship: Sharing and replicating musical artefacts. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Shanghai, China, June 2021. doi: 10.21428/92fbeb44.87f1d63e.
- Michael Zbyszyński, Balandino Di Donato, Federico Ghelli Visi, and Atau Tanaka. Gesture-timbre space: Multidimensional feature mapping using machine learning and concatenative synthesis. In *Perception, Representations, Image, Sound, Music: 14th International Symposium, CMMR 2019, Marseille, France, October 14–18, 2019, Revised Selected Papers 14*, pages 600–622. Springer, 2021.