Hearspray

Designing interaction for a movement-controlled digital musical instrument

Kalle Mäntsälä
Master's Thesis

Media Lab, University of Art and Design Helsinki
October 2009

Department
Media Lab Helsinki

Date
16th October 2009

Degree
Master of Arts in New Media,
University of Art and Design Helsinki

Author
Kalle Mäntsälä

Name of the work
Hearspray –
Designing interaction for a movement-controlled digital musical instrument

Level
Master's Thesis

Number of pages
85

Abstract

Hearspray is a digital musical instrument aimed for live electronic music improvisation. The instrument utilizes movement-based realtime control gestures, striving for embodiment, proper amount of challenge and spontaneous interplay between several players. The focus of the project was in interaction design, aiming at an instrument that could be controlled in a fluent and learnable way, still allowing for effective and emotionally rewarding control of rich musical outputs.

The written part analyzes the key components of player-instrument interaction, connecting the theory to the Hearspray design process. Particular design questions addressed include: Defining the interaction possibilities for a novel instrument, balancing controllability and challenge, and supporting the beginning players without sacrificing the musical possibilities. The written part focuses in player-instrument interaction, even though the actual instrument also provides tools for player-player interaction.

In addition to the theoretical issues, an overview of the whole design process is given. The used design methods and prototyping tools are evaluated on the basis of their suitability for the design of a musical instrument.

Materials
Mac OSX application, demonstration audio and video files.
Available from http://www.luxaeterna.fi/hearspray

Keywords
musical instrument, electronic music improvisation, interaction design, movement interface, audio programming

ACKNOWLEDGMENTS

This thesis has been influenced by people I have met and by events happening during over 20 years of my life. Thus it would be impossible to individually thank everybody who has contributed and who has been inspiring me.

I want to specifically thank my instructors and teachers Antti Ikonen and Richard Lapington for the knowledge, inspiration and valuable feedback before and during the thesis process.

In addition, thanks to the fellow students who tested the instrument, gave feedback and in general drifted with me into inspiring discussions around interaction design, instruments, music and life in general: Matthieu Savary, Tuomo Tarkiainen, Timur Kuyanov, Atle Larsen, Sebastian Greger, Anna Nuorsaari, Matti Luhtala and Taneli Bruun to name few.

Helsinki 14.10.2009,
Kalle Mäntsälä

TABLE OF CONTENTS

1. INTRODUCTION

Digital music tools have revealed only a part of their potential. Many of us have enjoyed the increased precision, previously unheard sounds and new musical styles enabled by them. But many have also been complaining about the visually dull laptop music performances and "cold" sounding music. In the extreme case, digital tools allow the music being rationally constructed by one man in his home studio, leading to the lack of spontaneous improvisation and interplay between performers.

The fact is that nowadays most of the music is produced and a lot is performed with screen-based computer software. We do have more audience-friendly digital instruments meant for live performance, but many of them are controlled via a musical keyboard – an interface developed centuries ago for a different sound output. It probably tells something about the interface's versatility, but it unavoidably directs the musical expression of new instruments towards the old clichés – starting from the musical scales. Also, in order to make the performances more exciting, replacements for the stationary piano or laptop keyboards need to be innovated.

There are, of course, many novel digital instruments developed by researchers and amateur inventors, but the problem lies in the fact that these inventions rarely are played by other people than their developers. This is at least partially caused by the lack of proper mass distribution methods. Recent years have shown some improvement to the before-mentioned issues: Products like *Tenori-On* [Yamaha 2007] and *KAOSS Pad* [Korg 1999] represent commercial mass-marketed music tools that are at the same time innovative and tangible. These products introduce new playing metaphors, which

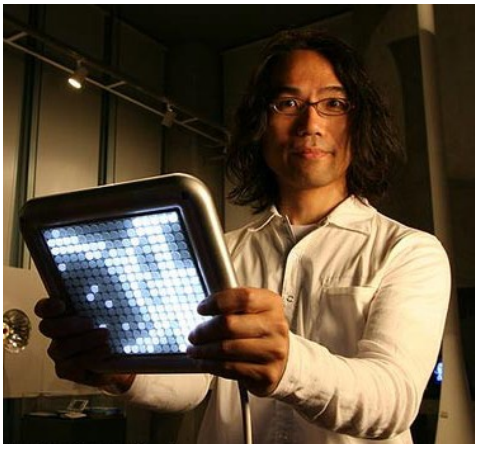


Figure 1: Designer Toshio Iwai showing Tenori-On

allow non-musicians to create somewhat complex music without years of learning.

However, with this kind of specialized expensive technology we are still far from the situation where everyman grabs these instruments. Very promising development can be seen in the introduction of low-cost musical instruments and games for home consoles, like *Electroplankton* for Nintendo DS. These products will potentially allow millions of players to share a common instrument, ensuring the development of more and more skilled players – as seen in *Guitar Hero* championship tournaments. Due to the playful nature of the platform, many of these examples are more toys than instruments aimed specifically for creation of new music. However, nothing prevents using these products in creative ways (as seen in Figure 2) or even designing true musical instruments for the home consoles.

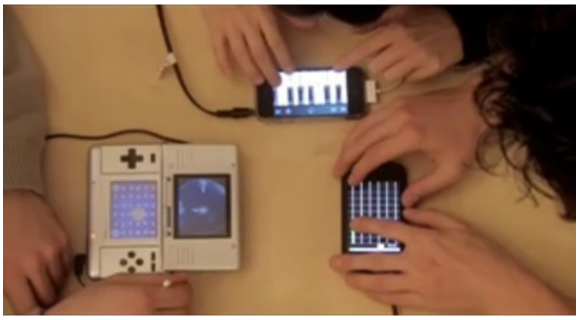


Figure 2: iBand playing Electroplankton and iPhone instruments

Inspired by these problems and trends, I have designed and developed a musical instrument, which favors physicality and collaborative playing, aims at learnability without sacrificing further musical possibilities, and could in the future be freely downloaded to a popular home console without any additional technological investments.

1.1 Background and motivation

I started to play acoustic musical instruments – piano, trumpet and a bit of drums – when I was seven years old. I was trained according to the western classical tradition, having motivational problems as a youngster due to the huge amount of systematic training needed in order to play “right”. Fortunately I was playing in several orchestras and enjoyed the social aspects of playing together. Playing improvised music also helped me, as it provided the channel for self-expression, which I felt lacking from the rigid training of a classical musician. Due to these positive factors I managed to keep the interest through the years, and ended up studying trumpet pedagogue for a couple of years.

As a parallel track I got hooked into home computers when I was seven years old. Due to life's mysterious ways I later made a career out of this interest – first working as a programmer, more recently as an interaction designer. What's meaningful here is that I never really combined music and computers: I have been making music with desktop software tools such as sequencers, but it has always felt too indirect and slow to really hook me. Personally, I value the spontaneity and excitement of a live performance over the precision of rational non-realtime musical planning.

When I applied to Media Lab Helsinki, I already had years' experience of designing visual interfaces and especially of sitting in front of computer screens. Thus I wanted to work with interfaces involving movement and audio to gain new perspectives. This has also happened: During my studies I have practiced interaction design in projects involving sound, movement, wearables and tangible objects. The most important project directly inspiring this thesis was *What You Do Is What You Hear*, an installation aiming at teaching the basic properties of sound to children. The installation lets one or two participants interact with sound objects in a goal-oriented but explorative fashion by using their hand movements (Figure 3). The positive key learning was that people, especially children, approach such an interaction method openly. The players also learnt the controls quickly and felt very powerful using such a direct control method.

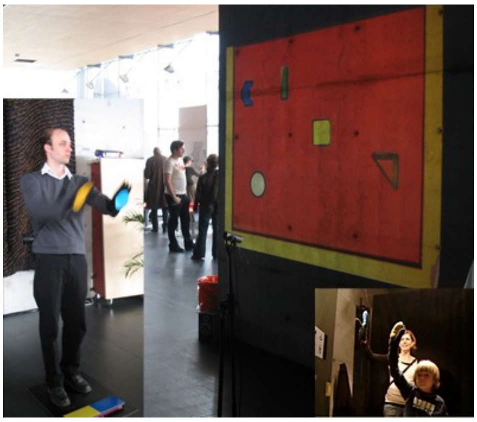


Figure 3: WYDIWYH in action

I decided to continue from the findings of the installation, aiming to create a true musical instrument with greatly extended control possibilities.

1.2 Goals of the thesis

The goal of the *project part* was to design and develop a musical instrument, which utilizes movement-based control gestures and allows collaborative playing. The focus was in interaction design, striving for an instrument that could be controlled in a fluent and learnable way, still allowing for effective and emotionally rewarding control of rich musical outputs.

The goal for the *written thesis* was to analyze key components of player-instrument interaction, producing knowledge that could be used to guide the design of a new instrument. These findings and their relationship to the Hearspray design decisions are introduced in this document. Particular design questions addressed were:

1. **How to define the musical interaction possibilities for a novel instrument?**
2. **How to create a balance between control and challenge?**
3. **How to support beginning players without sacrificing musical possibilities?**

The aim was not to find definitive answers for these questions as they are very broad, but to chart the possibilities and to find suitable solutions for the Hearspray project.

The written part focuses in player-instrument interaction, even though the instrument design also provides tools for player-player interaction. In addition to the theoretical issues, an overview of the whole design process is given, focusing in interaction design. The used design methods and prototyping tools are evaluated on the basis of their suitability for the design of a musical instrument. Technical and aesthetical questions are covered only when relevant for the instrument interaction.

1.3 Hearspray project

Hearspray is a prototype of a movement-controlled musical instrument for Nintendo Wii platform. It is aimed for solo or collaborative live improvisation of electronic music, for any person who wants to approach music as exploration rather than as following rules.

The instrument's musical output is based on a rhythmic flow of small sound particles, created with granular synthesis techniques. The sound varies from “pleasant” tones to more noisy and complex textures. The instrument aims at creating a pleasant first experience by emphasizing smoother sounds and easy access to the basic controls. At the same time it aims to offer enough complexity in order to allow longer term exploring and enhanced sound control.

Hearspray is controlled two-handedly using Nintendo Wiimote and Nunchuk controllers. The player is able to control various rhythmic, timbral and spatial aspects of the generated sound by shaking and rolling the controllers to various directions, and by using the joystick and buttons in the controllers.

The instrument can be used solitarily or can be shared by several players who have their unique voices. In order to further encourage musical dialogue and to allow more complex outputs, the players are able to record and play back the audio signal. Playing back and manipulating the recordings allows the players to repeat and variate the musical ideas in order to strive for coherent musical structures or to introduce surprising references to the past.

Nintendo Wii platform was chosen because of the suitable and robust wireless controllers. Additional benefits include the emphasis on social interaction and existing infrastructure for distributing software via the internet[1]. Due to the lack of applicable programming skills the prototype runs on Mac OS X computers instead of the Wii console.

Demonstration videos, audio examples and the downloadable application are available in the web address http://www.luxaeterna.fi/hearspray.

---

[1] **Wii Shop Channel** allows Wii users to browse, buy and download games and other applications via the wireless internet connection

1.4 Structure of the written thesis

Chapter 2 defines the key concepts for this thesis. The chapter gives an overview of the possibilities and restrictions of a digital instrument, and also introduces the field of interaction design and the interaction design process.

Chapter 3 gives an analysis of certain relevant aspects of player-instrument interaction, ranging from the high-level needs to the low-level aspects of controlling an instrument. The analysis is connected to the key design decisions done during the Hearspray development.

Chapter 4 maps relevant tools for actualizing a movement-controlling instrument. The chapter introduces motion interfaces, mappings and options for giving feedback to the player.

Chapter 5 presents the Hearspray design process. It covers the requirements analysis, iterative sketching, and detailed design of the instrument.

Chapter 6 gives a detailed description of the final instrument interface and mappings.

Chapter 7 presents the outcomes of player evaluations and musical explorations, including author's own reflections.

Chapter 8 presents the conclusions from the project.

2. KEY CONCEPTS

Before going further with the interaction design of instruments, it is useful to define what is meant with the terms *Digital musical instrument* and *Interaction design,* and how they are understood in the scope of this thesis.

2.1 Digital musical instrument

Digital instruments use digital circuits and software algorithms for sound production, either creating the sound via synthesis technologies or by manipulating recorded sound sources. Digital technology has introduced a vast number of new expressive possibilities for musical instruments in the recent decades. The following lists some of these possibilities, adapted from Jorda [2005:25-26]:

- Any kind of physical gesture, external parameter (e.g. weather data) or even audio signal can be used to control sounding output.
- The physical material no longer poses the boundaries: The available pitch range and maximum rhythmic density is only dictated by human hearing capabilities. Tuning systems and rhythmic complexity can be decided freely.
- Instrument timbre[2] is no more static. Any imaginable timbre can be rendered in theory, and even smooth transitions between distinct timbres are possible. Sound output does not need to be created on-the-fly (synthetically) but can utilize pre-recorded sounds.

---

[2] **Timbre** (sometimes called *tone color*) is the quality that separates the tones of same pitch, relating to spectral energy distribution.

- Any sound parameter variation can be controlled as desired: continuous, discrete or combination of both. Variations can be automated to enliven sound without demanding attention from player.
- The player doesn't need to control small details, but can control processes that control details instead. These processes may appear deterministic in nature or be so complex that they appear almost random.

Strengths of acoustic instruments

Despite these obvious benefits, there are examples of alienation towards the digital instruments among the proficient players. For example, computer music composer Paul Lansky [Wakin 2008] admits that after 35 years of experience in computer music he does not particularly like synthetic sounds and nowadays concentrates on *acoustic instruments*. The survey by Magnusson and Hurtado [2007] reveals that chaotic richness and challenge are the factors that keep musicians returning to the use of acoustic instruments, even though they might also use digital ones. The survey further mentions positive aspects of the acoustic instruments: The experience is emotional and embodied rather than rational. This is affected by the physical nature of control gestures and the presence of tactile[3] feedback. The limitations of these instruments can also be found inspiring: The participants talked about “pushing the boundaries” and exploring the limits of the instruments. These aspects are explored in the Hearspray design process in order to bridge the positive sides of acoustic and digital instruments.

Interface – instrument distinction

Acoustic instrument playing tends to follow a common pattern: The player brings energy to the instrument by directly using parts of his body (e.g. blowing). This starts oscillations in an excitation source (e.g. the reed of a clarinet). These vibrations are then amplified and transferred to the surrounding air by the instrument's resonating body [Jorda 2005:20]. Only certain acoustic instruments such as the piano and organ blur this tight coupling,

---

[3] **Tactile**, based on sense of touch

utilizing mechanical parts to transfer energy between the player's gesture and the initial vibration source.

Digital musical instruments unavoidably break the direct relationship between the player and sound output, as the sound is generated or transformed virtually. The interaction between player and a digital instrument is mediated through an *interface.* In the case of physical and tangible interfaces this is often referred to as a *controller*. Controllers convert the analogue signal of players' manipulation gestures to the digital messages understood by the computer system (Figure 4). In addition, there is always a certain mathematical *mapping* between the physical parameters and the sound output. Mapping algorithm converts the controller's input parameters to the sound synthesis parameters [Hunt et al. 2000].

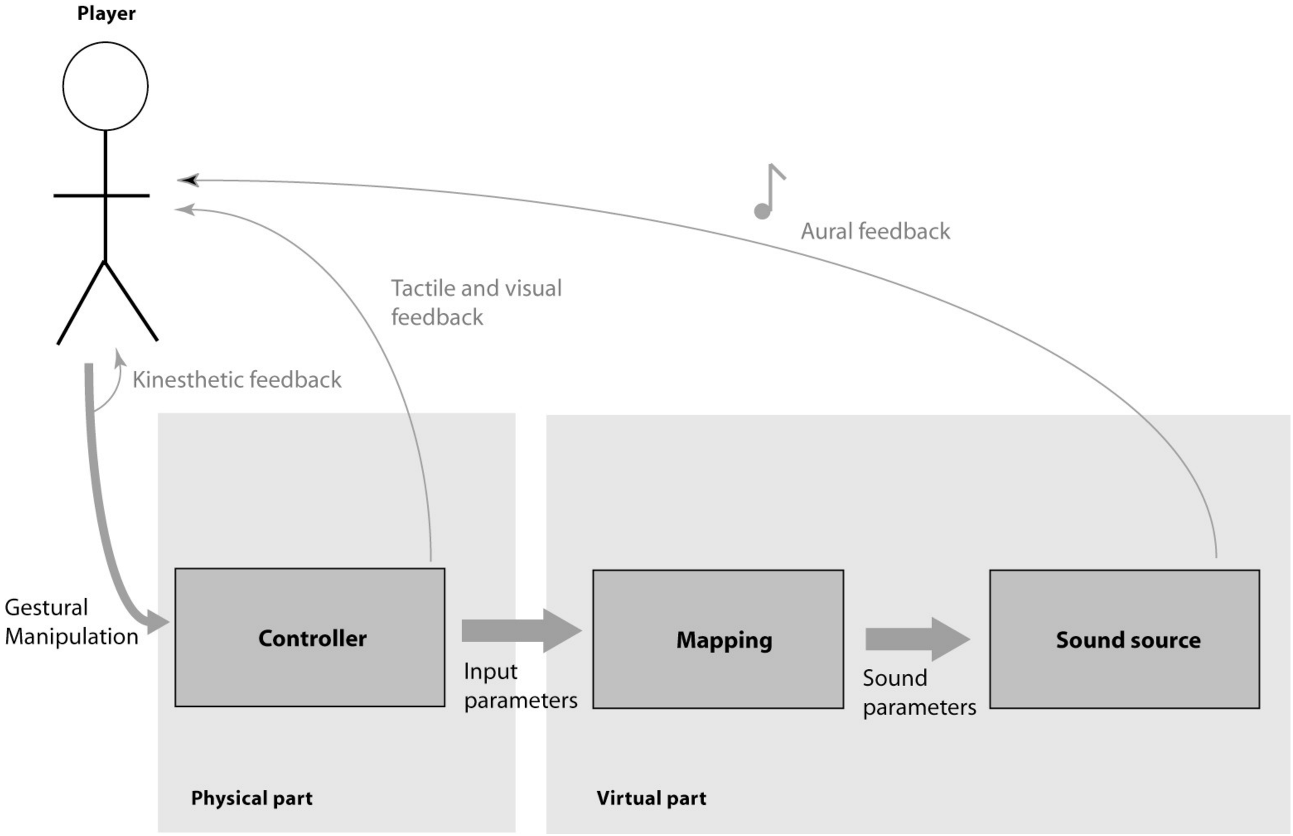


Figure 4: Typical player-instrument information flows for a digital instrument

Scoping the instrument

As the territory of digital instruments seems to be boundless (any input to any sounding output), it is useful to define initial starting points on *what makes an interesting digital instrument* from this theses point of view:

1) Performer-driven. With digital technologies it is easy to create automated musical systems, whose output is based on non-human inputs. They could be affected by environmental data or algorithms defined by the creator of system. Also, it is as easy to create a system that is controlled by hundreds of people, e.g. by using camera tracking of visitors in an installation. In this thesis I will not examine or design these kinds of systems, as I am interested in instruments that are used by a limited number of people who are aware of their part in the sounding output. Thus the approach to performance is conventional and based on the western tradition, separating the roles of the performer and audience.

2) Player freedom. There are many examples of sound installations, musical games and interactive compositions, which restrict interaction to the triggering of pre-made loops or to modifying few parameters in order to ensure that the output always sounds “good”. This kind of constrained interaction is probably very thankful for a casual visitor, but will not encourage further exploration and development of the skills. The focus of this thesis is in instruments that allow the player to develop his skills after getting familiar with the instrument and even let him sound “bad” if he wishes.

3) Clear instrumental identity. Commercially successful digital music production and performance tools are often meant for “all purposes”. It can be debated whether the modern keyboard workstations or computer audio applications should be called instruments, as they allow radical changes in the nature of the system by changing preset configurations or by reprogramming their behavior [Jorda 2005:201]. Such tools could be considered *metainstruments*, using a term adopted from Jorda [2005:160]. A survey by Magnusson and Hurtado [2007] shows that at least people with background in acoustic instruments wish for more limited software tools in order to gain mastery over the voice and vocabulary. This thesis focuses on instruments that may be complex in behavior but still possess clearly defined limits and a sense of solid identity.

2.2 Interaction design

As the main goal of the written part is to analyze and describe the instrument design process from an interaction point of view, it is necessary to define what is meant with the term interaction design, and why would it be relevant for the design of a musical instrument. Interaction Design Association describes the field in the following fashion:

> *Interaction Design (IxD) defines the structure and behavior of interactive products and services. Interaction Designers create compelling relationships between people and the interactive systems they use, from computers to mobile devices to appliances; Interaction Designers lay the groundwork for intangible experiences. [IxDa]*

Interaction design is often associated with the design of system interfaces, solving problems like “How to present information understandably?”. However, in addition to static “screens” the designer is concentrating on the evolving processes between the user and system: “How the user should proceed when using an online banking application?”. Different but related questions are relevant for the design of musical instruments: “How to create a simple playing interface? How to support playing of certain musical outputs? How to promote a feeling of mastery?”. It is generally recognized that people behave in a myriad of ways, meaning that the interactions cannot be completely designed. Thus the goal of the designer is to lay conditions for fluent interactions [Löwgren 2008].

In order to find answers to such questions, the designer needs to understand the whole ecosystem of the use: Interaction design is usually rooted on the understanding of the users' goals, motivations, and context, with a focus on developing the system to respond to the user's experience and not the other way around. This understanding can be gained by actually meeting or even designing with the users – using *user-centered* methods – or by relying on the *genius designer's* intuitive understanding of users' needs [Hawley 2009].

Often both approaches are mixed in the design process. The understanding of users is supplemented with knowledge of perceptual and cognitive psychology principles in order to create suitable design solutions. Designs are then evaluated in terms of usability (e.g. learnability, efficiency) and affective influence (e.g. aesthetics).

Understanding of the use context often leads the interaction designer looking further than just designing for maximum efficiency. Examples include the necessary challenge and importance of affective aspects in computer games. Despite the field's focus on digital products, interaction is not limited to a virtual space: The designers of social applications must consider real-world social relationships in order to properly support the users. The before mentioned qualities make learnings from interaction design potentially useful for musical instrument designer: Musical interactions often are dynamic, embodied, social and emotional. According to Orio et al. [2001], live performance of computer music can be seen as highly specialized field of (human-computer) interaction, with such specific topics as simultaneous multi-parametric control, timing and rhythm, and training.

In this thesis the term *user* is used only when referring to interaction design literature and related concepts. To emphasize the social and actively exploring role of human musicians, user is replaced with *player* in instrument interaction contexts and *performer* in public performance contexts. These terms are much more commonly found in the musical and instrument design literature.

3. ANALYZING MUSICAL INTERACTION

Many strategies can be used when shaping the musical possibilities for a novel instrument: The instrument builder can aim at replicating typical technical solutions or musical outcomes of e.g. certain genres of electronic music. This approach was utilized in some aspects of Hearspray by using typical synthesis algorithms (like sine waves) or effects (like the delay). However, when the aim is to create novel and interesting musical outcomes, a proper analysis of the musical interaction can be helpful in order to avoid trivialities.

This chapter gives an overview to selected aspects of musical interaction between a player and a musical instrument. It tries to find out, what are the general musical interactions that a new instrument might need to support, starting from the very high-level goals and advancing to the detailed aspects of musical interactions.

3.1 Players of digital instruments

The reviewed instrument design literature revealed no serious attempts to define *"Who are the players of digital instruments?"*. However, as this question proved to be important for the experience of using Hearspray, it is explored here briefly.

As with all products and services, there will be *novices,* who approach the instrument for the very first time. On the other hand, there are players who have become *experts* of the instrument through repeated and focused use. It is important to realize that the first-time players of a particular instrument can still have varying amounts of experience playing musical instruments. The prior experience with acoustic and desktop-based digital instruments may not be fully comparable due to the different playing paradigms. All of this may affect the players' expectations and needs using a new instrument.

Figure 5: Guitar Hero utilizes a simplified guitar controller

Figure 6: Reactable provides rich musical possibilities and ways to collaborate

Digital installations and musical games – e.g. Guitar Hero – are often aimed at novices regardless of background. These products make the learning phase as smooth as possible via simplified interfaces and directed outcomes. This ensures pleasing musical outcomes, but usually by restricting the musical possibilities. More serious digital instruments – such as The Hands (see chapter 4.1) – may be aimed strictly for *expert* performances. These instruments can sacrifice the understandability of the playing interface, as they are not meant to be distributed for larger audiences. Most ambitious instruments – like the recently commercialized Reactable [2009] touchscreen instrument – aim at being suitable for *both novices and advanced* electronic musicians. However, as Wessel and Wright [2001] note, ease-of-use and continued musical evolution are often contradictory design goals. For example, Reactable offers remarkable possibilities, but the playing relies on the understanding of audio chains and certain music technology -related symbols.

Cooper et al. [2007:41-47] reminds that most users are neither novices nor complete experts, but they have *intermediate* knowledge and skills using a certain product:

1. Novices pose questions like “What does it do?” or “Where do I start?”.
2. Intermediates have formed a partial understanding of the system, often having recalling problems like “I forgot how to achieve result X”.
3. Experts know the system thoroughly and tend to start using (or misusing) a product in new, creative ways. They also utilize shortcuts to automate regular functionality.

In order to enable the development across these expertise stages, a system has to possess an unchanging identity without suddenly changing interface or behavior [Raskin 2000:69]. Many digital music tools allow the player to control sound events with freely choosable interfaces (e.g. mouse, various MIDI[4] controllers) that can be flexibly connected to musical parameters. This allows flexible use of different control strategies, but can be seen as an obstacle in becoming an expert player [Paradiso and O'Modhrain 2003]: It becomes impossible to form a thorough understanding of a changing system.

Hearspray aims at encouraging player development and providing access to rich musical possibilities. This means that the interactions cannot be directed or the simplest possible. However, the instrument still tries to encourage the first-time players by the aid of an understandable interface and pleasant first experience. Hearspray has an unchanging instrumental identity in order to support learning and development.

3.2 Goals of an instrument player

Despite of the diverse play contexts and backgrounds of players, one can find some universal high-level goals, which all musical instruments should take into account. As studies by Persson [2001] and Motte-Haber [1984 cited Persson 2001] suggest, musicians are driven by needs for hedonistic enjoyment, social belonging, getting respect, self-actualization and sense of control. These goals relate to what Cooper et al. [2007:93] calls experience goals – how people want to feel after using a product.

Products also fulfill *end goals*, functional goals that are about performing a specific task. The end goals for musical instruments include at least:

Organizing sound. Usually the player wants to produce musical outputs that are results of conscious planning rather than totally arbitrary sounds. The plan might be exactly pre-formulated in the case of western classical music notation or might be relying on existing

---

[4] Musical Instrument Digital Interface, a protocol and physical interface meant for connecting musical devices - e.g. keyboard controllers and synthesizers. Introduced in 1982. Allowed any kinds of controllers and instruments to be connected as long as they used the MIDI standard.

knowledge supported by quick decision-making in the case of improvisation.

Exploring the instrument. Instead of the conscious sound organization, trying what the instrument can produce and reacting to that may be as rewarding. Pursuing this goal poses no requirements for the player's expertise level. Exploration is eased by the possibilities of the digital instruments, as they potentially allow rich and unexpected results created by complex algorithms.

Carrying tradition. It is often desirable that a piece of music or a musical idea is transferrable to different playing contexts. In order to achieve this, players may memorize what they have heard, rely to notated scores or utilize recorded material. Digital tools also allow saving program settings and patches in order to exactly reproduce earlier performances.

Communicating emotions. At least in the western tradition, playing music is often related to experiencing and communication of emotions. Player's emotions must be separated from the emotional cues that the listeners extract from the musical performance – the *expression* [Gurevich and Treviño 2007]. Understanding these cues is often based on a shared musical language between the performer and audience. The new instruments' “problem”, as suggested in Gurevich and Treviño [2007] is that the new musical styles lead to a lack of shared musical language. However, performer's bodily and facial expressions may still be relevant for emotional communication.

Synchronizing efforts. Players usually want to synchronize the sound output of their instruments with other players and to whatever might be sensed from the environment and audience.

Hearspray aimed at being controllable enough to allow conscious sound organization (in the limits of its musical output possibilities). However, it contains certain uncontrollable aspects and has a wide selection of accessible parameters to support explorative approaches. Synchronization between players is possible, but not forced by the digital means in order to maintain sense of freedom. The problems with emotional communication are realized, but the player should be able to experience emotions and perhaps communicate them via bodily means. Personal observations hint that playing such an embodied instrument creates a sense of enjoyment.

Hearspray is not aimed for reproducing traditional musical styles, but a relevant future project would be to preserve the musics produced with it.

3.3 Starting with an instrument

Before the player can start fully pursuing these high-level goals, the instrument has to be learned to some extent. Learning a movement-controlled digital instrument (which has a static identity and is learnable in terms of rather deterministic behavior) consists of the following parallel and interconnected processes:

1) Understanding the interface. This should take very little time compared to the other forms of learning. The interface has to intuitively answer the questions related to the initial discovery, like "How do I hold the controller?" and "How do I start the sound?" (as seen in Figure 7). After the initial exposure, the player needs to form an understanding of the different control gestures and their effects on the sounding output.

Figure 7: A kalimba clearly communicates what can be done.

The time needed for learning complex instruments may be reduced by communicating a simplified *representation model* of systems' often complex inner workings. For example, piano players seldom consider how many strings (1-3 per key) are being hit when they manipulate the keyboard but are still able to play music. Simplified representation results to an easy-to-use system as long as the representation does not hide any of the desired functionality or limit the use of it [Cooper 2007:27-32].

2) Learning the musical possibilities. The learning requirements relate to the player's prior experience with musical instruments: For inexperienced players it may be crucial to understand certain basic properties of sound and music. Depending on the musical style being pursued, important concepts might include rhythm and timing, pitch and notes, harmony and timbre. If the player is already knowledgeable with these concepts, the possibilities of the current instrument need to be explored (e.g. does the instrument allow polyphony[5]?) in order to achieve the wished results.

Digital instruments seem to suit learning by free exploration better than their acoustic counterparts: Jorda [2005:200] categorizes new (digital) instruments as relatively *wide* and *shallow*: They probably don't offer as fine and complex detail control as acoustic instruments, but provide access to much wider selection of parameters. This supports parallel or random way of learning: The player is not forced to learn one control to excessive finesse, but can freely choose to concentrate on the wished aspects. On the negative side, the musical knowledge from prior contexts might not be fully transferrable to the digital instruments, including concepts of tonality, scales, chords and rhythmic metrics.

3) Developing motor skills. Initially, the player needs to put considerable attention to the production of basic sound-producing gestures. As one performs the tasks repeatedly they tend to form larger action units and to become easier – they become automated motor patterns. This frees up cognitive resources for reacting to musical events and for operating towards higher-level musical goals [Pressing 1988:138].

Acoustic instrument playing is often hard to learn consciously, as the instruments utilize so complex interrelated control gestures that attention cannot be divided to all the necessary controls. To ease the burden of motor learning, Saffer [2008] suggests that communicating multiple simultaneous gestures should be avoided during the learning phase of gestural interactions. Fortunately digital technology allows starting from a simplistic interface, increasing the complexity of the interface gradually to allow finer control.

The player does not need to be perfect in any of the before-mentioned areas, but has to acquire a minimum of skills and build confidence in order to concentrate on the musical goals.

---

[5] Simultaneous playing of many tones

Hearspray aims at quick control discoverability and provides player with parallel exploration possibilities. All gestures produce noticeable feedback in order to support learning the controls. A visual learning interface was added later due to the problems with discovering all controls. This interface also allows the gradual learning of the control gestures.

3.4 What are the musical tasks?

In interaction design process, the users' high-level goals are often split into separate *tasks* or *use cases* for analyzing, what steps need to be supported in order to reach that specific goal. In the reviewed instrument design literature, the only attempt of explicitly trying to list the musical tasks was found from Orio et al. [2001]. They suggested that at least the following tasks would be relevant for an instrument:

- Isolated tones, from simple triggering to varying characteristics of pitch, loudness and timbre
- Basic musical gestures: glissandi, trills, grace notes, and so on
- Simple scales and arpeggios at different speed, range, and articulation
- Phrases with different contours, from monotonic to random
- Continuous feature modulation (e.g. timbre, amplitude or pitch) both for a given note and inside a phrase
- Simple rhythms at different speeds combining tones or pre-recorded material
- Synchronization of musical processes

This list may be an useful starting point, but it seems to be directed towards certain styles of music: It mentions tonality-related concepts like scales and arpeggios, which might not be relevant for the design of new musical instruments. The list is also rather generic: The problem in using such a list as design requirements is that one would end up having an instrument addressing every sound and musical parameter. Such an instrument has no strong aim nor identity and probably needs a very complex user interface. As the earlier example of a kalimba shows, a certain subset of these requirements may be enough in order to create a successful instrument.

Thus the designer needs to separately define, which are the tasks that a particular instrument needs to support. One approach of deriving the necessary functionality is introduced by Donald Norman [cited Cooper et al. 2007:15], making the goal/task -division even finer by utilizing a hierarchic chain of *Goals* > *Activities* > *Tasks* > *Actions* > *Operations*. As an example, an appropriate chain for a bebop jazz saxophone improvisation could be:

1. Express emotions / Organize sound >
2. Play an improvised pattern >
3. Choose a pre-practiced pattern from memory >
4. Modify the pattern to suit the current chord structure >
5. Activate muscles to create sound output

However, this oversimplified example shows that such a top-down analysis is not doing the richness of musical activities justice. Musical activities tend to be abstract and free-form and there are always parallel processes and feedback mechanisms involved for real-time corrections. The following quote from Rasmussen et al. [1994 cited Sanderson 2003: 228] – belonging to the context of Cognitive Work Analysis – summarizes the problem of task analysis:

> *It is clear ... that analysis and design of modern dynamic work systems cannot be based on analysis and design of work systems in terms of stable task procedures. Instead, analysis of work systems must be in terms of the behavior shaping goals and constraints that defines the boundaries of a space within which actors are free to improvise guided by their local and subjective performance criteria.*

This statement about the design for working environment is surprisingly similar to the problematic of supporting musical activities via design. Task analysis may help guiding the design of a simple email reader application, but the musical instrument designer needs a more flexible approach. In the fashion of the previous quote, substituting tasks with spatial qualities, Jorda [2005:192] uses the term *performance space* to describe instruments' musical possibilities. The designer's responsibility is to define the overall shape of this space: The amount of the available dimensions, the range available along these dimensions, and the tools for navigating in the space. The dimensions of this space relate to various musical properties, not to low-level sound parameters.

By striving for a rather open performance space, the designer admits having only partial control over the final sounding outcomes of the instrument. The players can freely decide their strategies for navigating inside the boundaries of this space, thus potentially producing outcomes that the designer did not envision. On the other hand, the space will never be completely open: The material choices and design decisions related to instrument behavior may restrict the playing to certain subset of possible interaction methods and sound parameters.

Navigating the performance space

One criteria for designing or analyzing an instruments performance space is to elaborate how flexibly the player can transform between different musical outputs. Related to this, Jorda [2005:188-197] introduces the concept of instruments' *musical output diversity* – the amount of musical variation possibilities. This capability can be divided into three levels:

1) **Micro-diversity** refers to the ability to use fine movements along any of the performance space's axes. Thus it relates to the creation of subtle changes, like nuances (e.g. pitch and amplitude variations, small tempo changes) and minor variations on a theme or musical structure.

2) **Mid-diversity** refers to the broadness of the instrument's performance space and to the ability to make sudden jumps in the space. It thus dictates how distinct two written pieces or improvisations played on

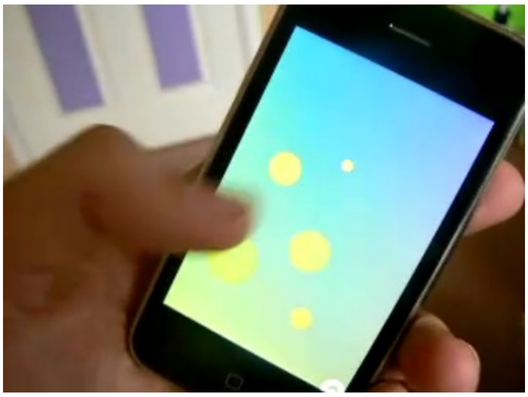


Figure 8: Bloom, an ambient music generator application for iPhone

an instrument can be. Limited digital systems like installations, games and interactive compositions (such as Bloom by Brian Eno in Figure 8) tend to always “play the same piece” regardless of player's inputs.

3) **Macro-diversity** refers to how flexibly an instrument can be used across different musical styles, contexts and roles. A trumpet can be used as part of section in symphonic orchestra, in a soloist role of a bebop jazz band and is found in many ethnic music traditions – but a trumpet is not fluent at accompanying a camp-fire singing session. Some instruments seem to naturally be very *generic* – a guitar can play various roles and styles, whereas some are very *specialized* – a bassoon is rarely used outside the western classical music spheres.

According to Jorda [2005:195], a large amount of macro-diversity is the least crucial of these diversity levels: It seems impossible to design an instrument that would handle all styles with equal ease.

The high-level design of Hearspray's musical characteristics was based on defining a multidimensional performance space, which provides certain musical affordances and constraints. The space definition (introduced in chapter 5.3.1) was rooted on personal insight of relevant musical dimensions, inspired by some of the tasks suggested in Orio et al [2001].

Hearspray does not even strive for large stylistic flexibility. However, Hearspray aimed at providing the player with sufficient possibilities in the other diversity levels – production of fine nuances and diverse musical outputs on all of the performance space's dimensions.

3.5 Operating in musical levels: from signal to structure

Instrument playing can also be analyzed according to the time-scale the interactions are affecting. Malloch et al. [2006] propose such categorization of musical interaction, dividing it into three behavior and context interpretation levels (see Figure 9):

1) Skill-based behavior is defined as *a continuous real-time response* to a continuous signal. One is usually tied to this level when learning a new acoustic instrument, e.g. when trying to get a beautiful tone out of a violin by monitoring with one's ears. The details of sound production can only be forgotten through sensomotoric automatization.

2) Rule-based behavior involves an instrumentalist who is more focused on controlling a *process* rather than signal or who is selecting and executing pre-made procedures. Examples include acoustic improviser's phrase generation, live sequencing and controlling parameters of musical algorithms.

3) Model-based behavior is a yet more abstract level, in which the performance is directed towards a conceptual goal. This level is usually activated in unfamiliar situations: Instead of using previously stored routines, the player is rationally trying to understand the situation and to create an useful plan in order to reach the goal using active problem solving. Only after this can appropriate skill- or rule-based behavior be chosen. As an example, a free improviser may decide to progress the piece towards a peaceful mood after minutes of aggressive playing.

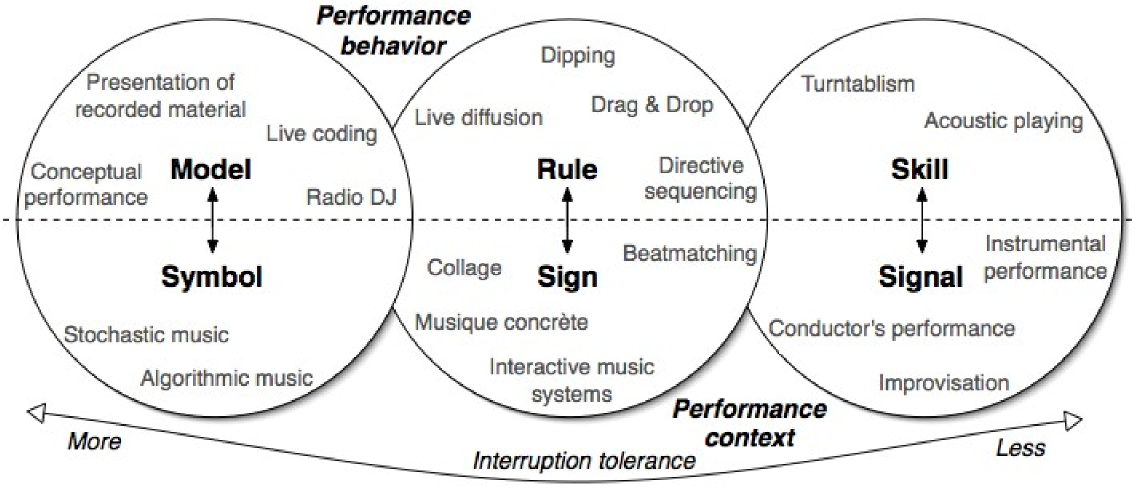


Figure 9: Musical behavior levels according to Malloch et al. [2006]

When looking at any instrumentalist who is utilizing bodily gestures for playing and has learned the control gestures, usually at least skill- and rule-based behaviors are mixed, making playing a combination of automated sensomotoric patterns. For a free improviser all of the behavior levels come into play as the improviser is also responsible for the musical structure without the aid of a predestined composition.

Kenny and Gellrich [2001] suggest that high-level improvisational decisions can only be made when cognitive resources are freed from the lower-level interactions. This can happen during moments of pause or through automation.

To support, or not to support?

The players of acoustic instruments are usually forced to operate on all behavior levels. Similarly, a **theremin**[6] player's hand gestures (Figure 10) are directly affecting the low-level properties (pitch and amplitude) of a sine wave synthesizer. Larger-scale musical structures can only be built by constantly affecting these properties. An example of a very different control strategy would be an interactive music generator that allows the player to choose from a limited amount of high-level parameters (musical style, mood), and the computer would then control all lower-level aspects (instrumentation, rhythm, harmony, musical scale used, sound timbre and exact pitch).

Figure 10: Leon Theremin playing his instrument

Whereas improvising with the theremin demands considerable attention and skill, the latter example hardly qualifies as an interesting performance instrument due to the loss of direct control over sound. The feel of “playing an instrument” might be lost when the player is only able to work on a conceptual level. Also, as important as high-level planning is for a musical performance, it's as well true that at least beginning musicians can achieve a satisfying experience with low-level behaviors like banging the drums - without thinking about musical structure.

---

[6] An early analogue electronic instrument, patented by Leon Theremin in 1928

Versatile digital performance instruments lie somewhere in between these polarities, supporting the user in a way that they still have the feeling of participating to the sounding outcome. Some acoustic instruments already contained automated functionality to support player in the lowest level: For example, the vibraphone contains a motor for creating small amplitude variations making the sound more lively. For digital instruments, low-level (or basically any level) sound parameters can be controlled by computer algorithms, or the control can be shared between the player and computer. This kind of support allows the player to focus on higher-level musical events. This in turn makes the music creation more tolerant for interruptions, as the time-scale of events tends to be slower in the higher behavioral levels [Malloch et al. 2006].

Figure 11: Beatbugs, a multi-user instrument

Beatbugs by Aimi and Weinberg (2003) gives an example of operating on multiple levels. The players are able to create rhythmic patterns by explicitly tapping them with the controller. These rhythms start looping automatically, allowing the players to shift their focus to higher-level musical behavior: The players can develop the loops further by manipulating their rhythmic and timbral properties by bending the two antennas situated on top of the controller. The players can also grab and manipulate each others' rhythms for further musical development. [Weinberg 2003]

A Hearspray player has only limited access to the very lowest-level sound parameters, which are varied by automatic processes in order to enliven the sound. The players mostly interact with the parameters of short sound grains. This gives a close-enough illusion of controlling everything on a tiniest scale.

On a rule-based level, the player can toggle processes of recording and playback of the instrument's outputs. Additionally, delay lines can be created in order to produce repeating patterns. The player can also take finer control over these processes if the outputs need to be manipulated. There is no automated functionality helping the creation of high-level structures – the player is taking the conceptual responsibility.

3.6 What makes an efficient instrument?

The design of digital products often strives for maximum *efficiency* in order to support human activities. The digital music tools can also benefit from the technological advancements: They can make playing of complex sounds and musical events easier than with acoustic means. For Jorda [2005:187], efficiency of a musical instrument consists of three components:

$$MusicalInstrumentEfficiency = \frac{MusicalOutputComplexity * DiversityControl}{ControlInputComplexity}$$

Output complexity refers to the musical diversity possibilities – the size and complexity of the performance space – in all levels ranging from low-level sound properties to high-level musical structures. The player should also be able to flexibly and freely navigate this space of musical diversity without restrictions: An iPod is an not efficient instrument despite its musical possibilities as it lacks *diversity control*. The last variable, c*ontrol input complexity,* refers to the number of interface's control parameters, precision, and convenience of its usage. Thus a simple but accurate control over complex musical possibilities results in high efficiency. Desktop tools are generally successful in terms of efficiency: For example, certain sequencers and visual programming environments offer almost unlimited musical possibilities, best of them utilizing rather understandable and convenient interfaces. However, an instrument designer also has to take other needs into account than mere efficiency:

> *Though the principle of effortlessness may guide good word processor design, it may have no comparable utility in the design of a musical instrument ... [Effort] is the element of energy and desire, of attraction and repulsion in the movement of music.*

As the above quote by instrument designer Joel Ryan [1991] suggests, efficiency and productivity may not be suitable driving factors for musical performance. On the contrary, many of the positive performance experiences relate to overcoming the conceptual and motor challenges despite the pressures of being in front of audience. Depending on the intended playing context and the high-level goals of the player, it may be important to easily create complex musical patterns, to gain self-confidence via taming a challenging instrument, or just to be surprised of the ever-changing musical results the instrument produces. The balance between efficiency and effort – ranging from controllability to chaos – is dictated by the choice of interface and the related mapping algorithms.

4. INTERFACES AND MAPPINGS

After charting the musical interaction and output requirements, one needs to analyze how the control of a musical instrument can be actualized. This chapter gives an overview of the potential interface and mapping solutions for a movement-controlled musical instrument, touching both directions of the information flow – controlling and getting feedback.

4.1 Movement-based interfaces

The instrument interfaces using movement as the main input utilize diverse technological solutions and conceptual metaphors. Paradiso [1997] introduces categories, which describe them from a player's point of view:

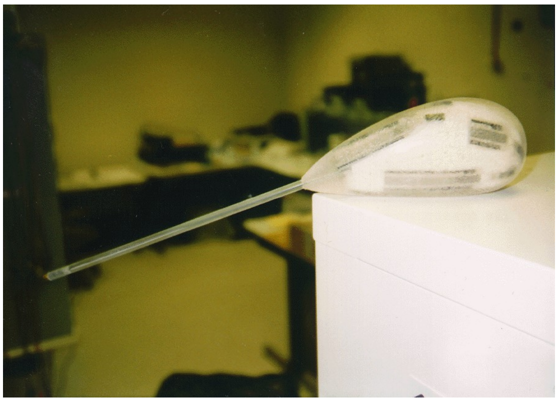


Figure 12: Digital Baton controller

1) **Batons.** These controllers are held in player's hand(s), tracking various physical properties like acceleration, speed, location or pressure with technologies such as infrared sensors and accelerometers. Early instrumental examples included imitations of drumming gestures and orchestra conductor's movements for controlling high-level musical parameters [Paradiso 1997]. More recent controllers, like Paradiso's *Digital Baton* (Figure 12) have proved that the controller type can be used for many kinds of musical tasks, ranging from high- to low-level musical control.

The final interface technology choice for Hearspray – the Nintendo Wii controller – also belongs to this category. The existing examples of using the Wiimote as controller tend to be simplistic and toy-like, ranging from inaccurate percussion playing imitations to Nintendo's *Wii Music* (Figure 13), which allows players to control traditional instruments' (such as cello, piano and trumpet) sounds with constrained interactions [Nintendo 2009].

Figure 13: Wii Music suits casual playing

2) **Wearables**. More transparent interfaces can be created by attaching sensors directly to the players' clothing or bodies. Michel Waisvisz's *The Hands*, developed from 1984, can be seen as a wearable hand-controller. It tracks the movements of hands and fingers accurately combining movement sensors, potentiometers and keyboard-like buttons [Jorda 2005:111]. The development was partially driven by the technological possibilities (introduction of the MIDI interfacing standard), but more by the dissatisfaction to keyboard-based tools and need to “walk, move and even dance while making electronic music” [Waisvisz 2006]. During the instrument's evolution Waisvisz experimented with different constructions, mappings and sound sources. The subsequent versions combined direct playing with live sampling [Ryan 2001]. In his late years, Waisvisz concentrated on the performance on a stable instrument instead of developing the instrument further: Jorda [2005:153] calls him “one of the very few composer-luthier-improvisers that can be considered a genuine virtuoso”, knowing his instrument thoroughly.

Figure 14: Michel Waisvisz and The Hands

Figure 15: W_space suit

Ready-made technologies can also be used for creating wearable interfaces: Composer Tom Tlalim recently (2007) built *W_space,* a sound-generating suit that allows the players to freely use their whole body in order to control synthetic sounds. The suit is built using eight Wii controllers – enough to provide semi-accurate information of all joints' orientation and acceleration. [Rosenblum 2008]

3) **Non-contact sensing**. The most embodied and transparent instrument interfaces allow playing in and with the space freely without wearing or holding any equipment. This kind of instruments can be seen as the descendants of the theremin (see Figure 10), created in 1920 by Leon Theremin. Theremin created the first intangible instrument [Jorda 2005:40] if one omits human voice: The player controls Theremin just by moving his hands in the air around the two metal antennas. This added a certain amount of visual spectacle to the performances, even though the player was still tied to the proximity of antennas. Technological advancements have since expanded the possible play area far beyond the original. The available tracking technologies include floor pressure sensors, video cameras and ultrasound sensors.

Figure 16: David Rokeby performing with VNS

David Rokeby's *Very Nervous System* (developed 1986-1990) is a typical, early example of a free-movement interface allowing large play area. It utilizes simple, self-built digital video cameras and image analysis algorithms in order to track the player and to control a pleasant oriental soundscape. The system tends to always produce similar sounds and Rokeby himself categorizes it as an installation rather than an instrument.

For Very Nervous System, the complex relationship between the player and system was more important design consideration than the instrumental controllability. [Rokeby 2000]

Options for sensing the gestures

In addition to the physical form, an interface must be evaluated according to its sensing capabilities in order to match the desired interactions. One important criteria is the *degrees of freedom* – how many dimensions of movement and rotation is the interface able to sense? This will have an effect on how many musical parameters the player can potentially control simultaneously. Only counting all joints, it can be estimated that a human body allows for over 40 degrees of freedom [Pressing 1990 cited Jorda 2005:138]. However, this does not mean it would make sense to build such an accurate bodysuit instrument. Human physiological and cognitive limitations restrict having precise control over that many dimensions:

> *In the early days of "Very Nervous System" I tried to reflect the actions of the user in as many parameters of the system's behavior as possible. I worked out ways to map velocity, gestural quality, acceleration, dynamics, and direction onto as many parameters of sound synthesis as I could. What I found was that people simply got lost. Every movement they made affected several aspects of the sound simultaneously, in different ways. Ironically, the system was interactive on so many levels that the interaction became indigestible. People's most common response was to decide that the sounds from the system were not interactive at all, but were being played back on a cassette deck. I found that as I reduced the number of dimensions of interaction, the user's sense of empowerment grew.*

The above quote from Rokeby [1997] illustrates the problem of learning multidimensional interfaces, even though in his case it was also related to the complex mappings such as utilizing both movement velocity and acceleration. The quote also underlines the fact that physically controlling and actually having a sense of controlling an instrument are two distinct things. Todd Machover [cited Jorda 2005:139] suggests that the optimal number of accurately tracked dimensions for a musical instrument is between 5 and 20. As this is very rough, the exact number must be dictated by the musical needs.

Values sensed by an interface's sensors can be *continuous* as in the case of a theremin: The player gradually raises his hand in order to slide the pitch “higher”. The values can also be *discrete* as in the case of using piano keyboard for only accessing the pitches of a chromatic scale. As the nature of bodily gestures is continuous, movement interfaces' sensors usually produce continuous values. These sensors also tend to prohibit *random access*, forcing the player to slide through all the values: For example, a theremin player audibly slides through pitches unless the player simultaneously manipulates the amplitude, making the sound die during the jumps. These factors make the interfaces non-optimal for some actions, such as trying to keep the pitch steady or to jump to a certain pitch. This in turn makes it very challenging to play exact pitches of a melody. There are examples of virtuosi – such as Theremin player Clara Rockmore – who successfully use continuous controls to play discrete tonal melodies. However, this may require years of learning to fully master the control movements – perhaps not a realistic learning time for a new digital instrument.

Corrective mapping strategies may be utilized to some extent: For example, in *iAno* (Figure 17), physical touchscreen's continuous horizontal dimension is mapped to a virtual discrete piano keyboard. However, this is relying heavily on the visual indication on where the player should touch. For the players of movement-controlled instruments, hitting the right values would not be so easy despite the use of such a mapping, as there is no immediate visual feedback available.

Figure 17: iAno - an iPhone piano application

Continuous dimensions in movement interfaces are often *coupled*: It is almost impossible to affect one dimension without having small effects on other dimensions, as opposed to controlling all sound parameters with independent sliders of a MIDI interface. However, such small motor challenges and resulting imperfections can actually be rewarding for the performer, as mentioned earlier. Interface mapping study by Hunt et al. [2000] supports the use of coupled controls when the controlled parameters have a logical relationship: The test subjects found it mentally hard to separately control the sound parameters using a slider interface.

In addition to continuous sensors, the controllers usually contain discrete input methods (e.g. buttons) in order to enable discrete operations, such as starting and stopping processes, and toggling special modes.

The exact choice of interface must be decided by weighting the needs related to the sound output and interaction: Does the player need to be able to move freely in space? Should the player use hand gestures or movements of the whole body? What is the exact amount and nature of the sound parameters the player is supposed to control? Does the player need to have exact control over each of these parameters?

For Hearspray, the interface choice was affected by the needs to move in the space without affecting the musical output and to have semi-accurate control over several musical parameters. Wiimote was chosen based on the wide availability and durability of the controller.

The number of the musical parameters was gradually refined during the design process to match the desired musical possibilities, to find the optimal balance between control and challenge, and to suit the capabilities of the controller still keeping the ergonomic factors in mind.

4.2 Mappings and controllability

In addition to the interface choice, mappings are another critical component for creating the digital instrument's feel and personality – making it understandable, responsible and controllable or inconsistent and chaotic. Gestural inputs can be mapped *linearly* to sound parameters (e.g. raising the controller raises the pitch) or can have more *nonlinear* relationship to the sound (e.g. raising the controller modifies a certain parameter in a chaotic algorithm controlling pitch). Mappings also relate to the level of musical interaction: They decide whether the player is controlling the micro-level nuances of the sound or the macro-level structure of the produced music.

According to Hunt et al. [2000], mappings can be categorized to the following groups:

1) **One-to-one mapping.** The most obvious mapping connects an input parameter (e.g. vertical position of the controller) with one sound parameter (e.g. pitch or amplitude). As the potential sound parameters usually far exceed the controller's dimensions, this is often referred to as “few-to-many”. This mapping approach is mentioned of leading to a toy-like instrumental feel [Hunt et al. 1999 cited Jorda 2005].

2) **One-to-many mapping.** For example the speed of violin bow on a string controls several aspects of sound, including amplitude and timbral qualities.

3) **Many-to-one mapping**. For example the amplitude of violin tone is controlled by both bow speed and bow pressure.

The problem of having too few control dimensions may be resolved by creating several layers of mapping [Hunt et al. 2000]: Gestural inputs can be mapped to the parameters of an intermediate *perceptual layer* parameters (e.g. smooth –> rough) that in turn are mapped to drive several sound parameters (e.g. timbre or attack speed).

The challenge of acoustic instrument control often results from a vast amount of unclear and interdependent control gestures, as mentioned in the case of violin. This complexity makes the initial learning very hard, but offers experienced players a multitude of strategies for achieving similar sounding results in performance situations. Research on mappings confirms that this control complexity improves the performance of complex musical tasks [Hunt et al. 2000]. The finding differs from the traditional interface design guidelines often promoting simplified interfaces for complex tools to heighten learnability. These approaches could potentially be combined as digital instruments offer possibilities for increasing the interaction complexity gradually, e.g. by revealing only parts of the interface in the first place – a technique utilized in the tutorial modes of computer games.

Nonlinear mappings

The input responses of the acoustic instruments tend to be slightly nonlinear [Jorda 2005:143]: Bringing more energy to the system does not create sound changes in direct relationship. When the player blows into a brass instrument, there may be an initial threshold after which the sound starts. There might also be resistance-related lag included in the amplitude response, when more energy is brought to the instrument. In the extreme limits increasing energy can have effect on the timbre and pitch as well. Similar or digital-specific nonlinearities can be utilized to create more interesting behavior to digital instruments. Some of these nonlinear digital mapping strategies involve keeping track of history – e.g. smoothing or exaggeration of value changes, measuring speed of the changes, and delaying responses [Ryan 1991].

Despite the fact that getting proficient with acoustic instruments is all about control and mastery, the aspect that makes them so fascinating (or frustrating for some) is the chaotic nature of physical vibrations and resonations: One can never expect to produce two exactly similar sounding results. Digital instruments could have “warmth” and imperfect components as well: It's easy to add non-deterministic factors to the relationship of an interface gesture and the sound output. In The Hands, Waisvisz used several different mapping programmes that generated “erratic” information or decided that things were becoming boring, automatically changing the mapping strategy [Jorda 2005:145]. However, to ensure instrumental feeling and learnability, this may have to be balanced so that the player still finds some causality between his actions and the end result.

For Hearspray, being in control was given more emphasis than surprising outcomes. However, on the lowest levels control is shared between the player and automatic processes. Small amounts of chaotic variations were experimented with and are present in the final instrument.

Hearspray utilizes many kinds of mappings, utilizing all linear mapping types, and also nonlinear threshold- and a history-based mappings. These were not envisioned when the process started, but instead evolved during the design process to match the desired amount of control.

4.3 Getting feedback

As mentioned earlier, the information flow between a player and an instrument is a bidirectional: In addition to controlling the instrument, the player receives information from the instrument. Comparing the expected result with the final feedback received from the instrument is important for learning and precisely controlling musical instruments. Usually both the controller and the sound source take a part in delivering the feedback via various sensory channels (see Figure 4).

Obviously the main feedback channel of musical instruments is **the aural channel**: The player plays a musical pattern and is simultaneously able to monitor whether he hears what he expects, taking corrective measures if necessary. Because of this need, it is crucial that the player connects to his voice when playing with other people. Cook [2003] provides examples of controller-embedded speakers which aid in reaching precise hearing and heightened intimacy.

According to Jorda [2005:139], other sensory channels are used to heighten player's understanding between a gesture and the produced sound. Redundancy – the duplication of same information between the channels – may be seen as waste of resources, but can actually be helpful in connecting with the voice: For a piano player, the ability to "foresee" and "touch" notes makes e.g. scale improvisation much easier than for a singer, who can only rely on his ears for finding the correct pitches.

An important and unavoidable feedback channel for players of most acoustic instruments and movement-based controllers is **the kinaesthetic channel**[7]. The player will gradually develop a sense of the body positions and movements involved in the production of different musical patterns, and eventually he does not have to consciously think about or visually monitor how he moves in order to produce certain musical outputs.

Immersion to the playing and understanding of instrument's internal state may be further heightened by utilizing **the tactile channel:** Force feedback controllers could convey certain events or the overall instrumental presence as is done by the vibrating body of many

---

[7] Sense of body and muscle position

acoustic instruments. Cook's [2003] PhiSEM haptic maraca controller (Figure 18) incorporates a vibration motor and a small speaker, reportedly creating feeling of connection between gesture, sound and feel.

Figure 18: PhISEM haptic maraca controller

In addition to the musical signal and presence, digital instruments often need to communicate supportive information – such as the state of processes – that cannot interfere with the audio signal. The player should not be forced to memorize such information: Cook [2001] provides a failed example of trying to only mentally keep track of the state of several recorded loops when playing an augmented trumpet.

Of all the supportive senses, **the visual channel** is utilized most often in the digital instruments due to the wide availability of displays and the simple implementation [Jorda 2005: 140]. Visual displays are excellent for handling multiple objects and simultaneous processes due to the possibility of spatial organization in addition to temporary organization. This is proved by the versatile touchscreen instruments, such as the Reactable (see Figure 6). However, the screen needs to be large and clear enough to support decision-making (as opposed to the interface in Figure 19).

Figure 19: Small symbolic displays may be hard to interpret

A short time delay between a gesture and the related feedback is crucial for such a time-dependent skill like musical playing [Ferguson 2006]: Instruments aimed for live performance need to strive for minimum latency. Pressing [1988:137] suggests that the reaction times related to the visual channel are slow compared to the aural and tactile channels. Thus visual feedback may be suitable for the slower pace of macro-level musical planning, but the micro-level decisions should be based on other channels due to the real-time interaction requirements.

The controller choice plays a big part in the instruments' feedback capabilities: For example, a non-contact sensing interface cannot usually provide visual or tactile feedback, but allows users to use kinesthetic channel for learning the large control movements. On the other hand, touchscreen instruments allow visual – and sometimes even tactile – feedback, but usually involve small movements, forcing the player to look at the interface.

Hearspray obviously aimed at low latency in sound output and provides small adjustments to separate players' voices. Unfortunately the Wiimote speaker could not be used for connecting the player to his voice.

Kinaesthetic feedback is unavoidably following from most of the control gestures. In addition, heightening the experience with tactile feedback was explored. Visual feedback (in the limits put by the controller) was added later for keeping track of the parallel processes.

5. HEARSPRAY DESIGN PROCESS

This chapter gives a chronologically proceeding description of the design process of the Hearspray instrument, explaining the design in relation to the musical interaction needs and interface technology alternatives introduced earlier. The chapter also presents the evaluations done during the process and their effects on the design.

As I was solely responsible for all the design and development, many areas including interaction design, sound design, visual design and technical solutions are covered. This is especially important as these areas cannot be fully separated: For example, the aim of the player interactions is to create satisfying musical outputs. Technical possibilities have also posed restrictions on the possible interactions. However, the focus is in designing the interactions – the other areas are described with less detail.

5.1 Background research

The very first ideas for the instrument project included physicality and embodiment, real-time playing interactions, collaborative playing possibilities, and combining easy approachability with possibilities for further development as a player. My background as an instrumentalist directed the design towards a sense of control and clear instrumental identity. Musically the aim was to create possibilities for free improvisations rather than to recreate traditional patterns, like sophisticated tonal melodies or harmonies.

Before going into any hands-on work or even formulating further requirements, a background research on the digital instrument domain was started. This was executed mainly by learning from the experts – by going through a large body of related research literature. Unfortunately it was not possible to play most of the digital instruments mentioned in the articles due to the specialized controller technology involved. However, a large amount of video material was accessible, so the instruments' strengths and weaknesses were studied based on the videos. The relevant findings of the background research are included in the previous chapters.

The design process was mostly based on genius designer -approach with some added user-centered methods (see chapter 2.2). In the beginning stages, involving potential players was omitted: Observations of playing existing instruments would have been relevant, but there was no access to any movement-controlled digital instruments. Some observations were done earlier with people using the movement-controlled What You Do Is What You Hear sound installation. Due to the installation's limited musical output these were only useful in terms of examining enjoyment, efficiency and social acceptability of motion interfaces in general. I mostly relied on the literary sources and my personal long-term experiences of learning to play instruments and use digital tools, playing solitarily and improvising in groups, and observing and teaching playing. Players were involved later in the process when initial tangible results had been produced.

It soon became obvious that it is too complex a problem to design an instrument that would take into account the needs and experience of all people with equal emphasis. Extreme examples of easy-to-access novice instruments produced great, complex results with a button press, but usually limited the fine control over output complexity. Thus it was necessary to make the decision that the instrument will focus on rich musical possibilities and control over them rather than on extreme simplicity of the control gestures.

5.2 Choosing and evaluating the interface technology

In addition to the conceptual design, technological possibilities for enabling movement-controlled sound interaction were examined in an early phase. My strongest expertise so far – from creating What You Do Is What You Hear installation – was to track players' hand movements with a video camera, based on color-marked gloves and color tracking. This worked as a starting point, but gradually new interfacing methods were evaluated due to accuracy problems: Using a camera for color tracking is highly dependent on proper lighting conditions. This allows a static installation to be created but is not suitable for an instrument that one wants to play regardless of place and time.

The initial idea was to replace color tracking with camera frame differencing, allowing the detection of player's motion based on pixel changes in subsequent video frames. However, when experimenting with this, it was soon obvious that tracking would still be heavily reliant on environmental conditions (lighting changes, other people entering the picture) and it would be very difficult to get enough information, e.g. to separate body parts from each other. All in all, with my limited skills the use of camera tracking would have led to a very small number of available control parameters and to a large amount of control unpredictability. These would have been conflicting with the goal of allowing enough possibilities for development as a player. Thus a new kind of controller technology was needed.

Building a custom controller using sensors and a micro-controller kit was out of question due to the lack of skills and the planned graduation schedule. Another rationale for examining existing solutions was to allow more people access to the instrument. Nintendo Wii controllers (Figure 20) seemed a promising option, as they provide a wealth of control dimensions, combine several different input types and allow flexible strategies for the use of space compared to camera detection: The player can walk around without moving the controller thus not affecting the sound. Research on existing Wiimote-controlled instruments confirmed that there was still room for unleashing the controller's full potential in a musical context.

Figure 20: The Nunchuk and the Wiimote

The technology benchmarking stage also involved decisions regarding the audio programming platform and the technologies used for connecting the Wiimote with a computer, as described in Appendix A.

5.2.1 Introduction to Wii controllers

Basic controllers of the Nintendo Wii gaming console consist of *the Wiimote* (the main controller) and *the Nunchuk* (the optional supplementary controller). The Wiimote is connected to the console – or to a computer in my case – using the wireless Bluetooth technology. Despite the wireless connection, the Nunchuk connects to the Wiimote using a cable that is long enough to allow large hand movements when two-handed interactions are needed. The Wiimote's two AA batteries allow for hours of playing. One of the benefits of using such existing controllers is that they are already designed with human ergonomics and affordances in mind: The player probably realizes how they should be held and what could be done with them. This was confirmed later by the player testing. The controllers are also very robust, which could have never been achieved by building a custom controller. At least four Wiimotes can be connected, easily allowing the creation of multiplayer instruments.

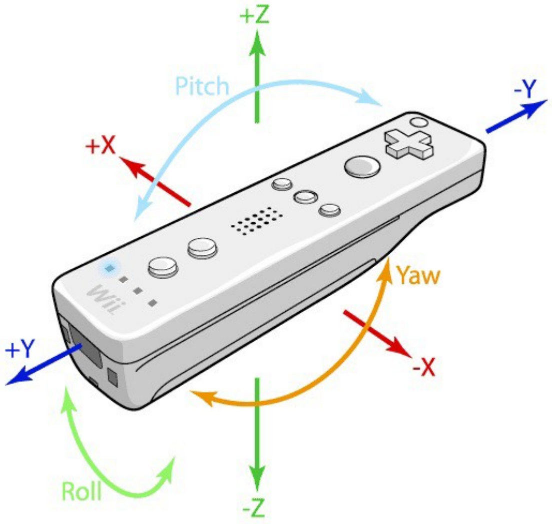


Figure 21: Wiimote's control dimensions

Both the Wiimote and the Nunchuk contain a three-axis accelerometer, which measures continuous linear acceleration values in X, Y and Z dimensions (e.g. a negative Z value for down, positive for up) [WiiLi 2009]. Thus each of the controllers offers "three degrees of freedom". Rotation angles around each axis (called Pitch, Roll and Yaw in Figure 21) can also be calculated reliably for a relatively stationary controller using combinations of two acceleration values [OSCulator 2009]. A stationary controller is accelerated by the hand or other surface resisting the gravitational force. However, because of the way the angles are calculated it is impossible to measure rotation values separately from linear acceleration values – the values are directly related. Thus hitting with the Wiimote towards top-left may be interpreted as linear or angular movement.

The Wiimote's location cannot be reliably detected using these acceleration sensors. To allow positioning, the Wiimote contains an additional infrared camera that is used to detect the controller's location in relation to an infrared light source (usually the Wii Sensor Bar on top of a TV). The camera allows the Wiimote to report vertical and horizontal location values and a crude estimation of the distance, but only when the player is pointing somewhat towards the light source and is standing within the optimal pointing distance of five meters.

In addition to these movement-related sensors, the Nunchuk contains a dual-axis joystick on top of it. Both controllers include several buttons, of which some are paired (e.g. left and right arrows). The Wiimote also allows feedback to be given to the player: A vibration motor can be used for providing tactile feedback and there are four led lights, which usually are used to communicate player's number. In addition, the Wiimote contains a small speaker for aural feedback, but unfortunately it cannot be utilized in computer applications as the related communication protocol has not been reverse-engineered.

In June 2009 Nintendo introduced the *Wii MotionPlus* expansion device, which contains an additional angular rate sensor. When this peripheral is connected to the Wiimote, the rotation angles can be separately and thus more reliably measured in relation to the linear accelerations [Shah 2009]. This could have affected the amount and quality of interactions available for Hearspray, but the peripheral was not yet available at the time of the instrument development.

5.2.2 First interaction sketches

The instrument requirements were initially left very open and high-level in order to ensure that the playing interactions would be naturally suited to the capabilities of the final controller choice. After successfully connecting the Wii controllers to the computer, potential interactions were explored and evaluated. This was done by coming up with the basic Wiimote interaction patterns and prototyping them in musical instrument context. All prototypes started as quick one-day sketches, and some of them were expanded when promising directions were found. The most relevant sketches were:

Hitting

Percussive sound samples are played by hitting the Wiimote and the Nunchuk in different directions, allowing 12 separate sounds to be controlled with varying sound amplitude.

Controlling the sound being triggered proved to be difficult as the Wiimote's dimensions are coupled. Hitting is also inaccurate due to the lack of visual reference point: The player does not foresee when the sound will actually be triggered. Thus such a percussion controller will not allow the playing of precise rhythms. Acceleration as a sound trigger also contains an inherent problem: The acceleration to one direction is always followed by an acceleration to the opposite direction when the movement stops, causing double triggering in the prototype's case. This can be downplayed by using a correction algorithm, but a foolproof solution was not found.

Second variation of the hitting metaphor was later created:

Short rhythmic loops are played by hitting the Wiimote to different directions. Triggering is confirmed with vibration feedback.

Due to the less frequent hitting it was easier to maintain rhythmically coherent patterns. The ability to control relatively complex patterns with bodily means proved to be very satisfying, and using vibration as a feedback from the hits seemed to further increase the instrumental immersion. However, the player testing of this prototype revealed that the demands for perfect timing were still too high to produce satisfying outcomes for less experienced players. It was also easy to mistakenly produce enough acceleration to trigger sounds.

The experiments with force feedback revealed disappointing limitations: The speed of the Wiimote's vibration motor could not be controlled, but the vibration could only be started and stopped. This makes the Wiimote unsuitable for communicating the subtler aspects of instruments. In addition, the motor starts and stops gradually, which is not very promising taking into account the timescale required for the real-time feedback of musical events.

Rotations

The Wiimote's rotations control the pitch and amplitude of a guitar-like synthesizer (based on Karplus-Strong synthesis algorithm). This decaying sound is started by pressing the Wiimote's B trigger button.

Trigger button was a necessary choice for starting the sound in this case – acceleration caused by hitting would have had an effect on the rotation values. It became obvious that the Wiimote is not very optimal for pitch control in the conventional sense due to the continuous nature of the values and the lack of random access to the values: Stabilizing the pitch demands a steady hand. Precise jumping to a certain pitch is hard due to the lack of tactile and visual support.

It was also revealed that even if three rotational dimensions could be theoretically calculated for the Wii controllers, usually only two of them were useful at a time: The calculated rotation value does not change when the user rotates the controller parallel to the ground level.

A simple visualization was tried with this prototype, aiming to help the player to hit certain pitches: The current pitch and the scale notes were marked with clear lines on the screen (Figure 22). However, contrary to the initial belief this did not make hitting notes easier, but seemed to make it hard to concentrate on the sonic results. This is probably related to the physical separation of the controller and the screen-based visualization: The physical distance between the Wiimote and the screen projection was over one meter at minimum, forcing one's attention to shift between two separate locations.

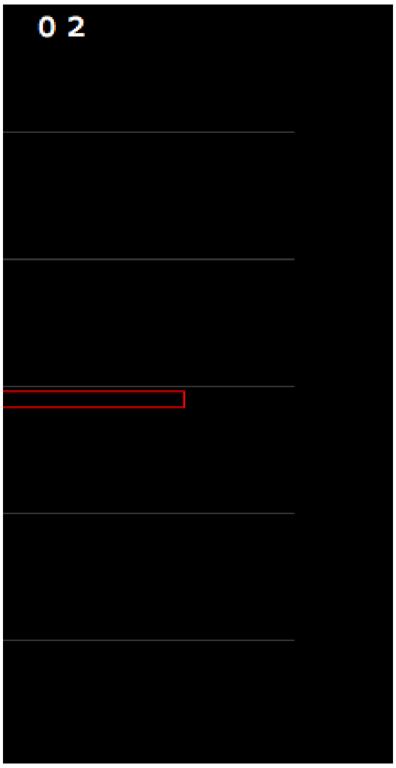


Figure 22: Attempt at visualization

Point and drag

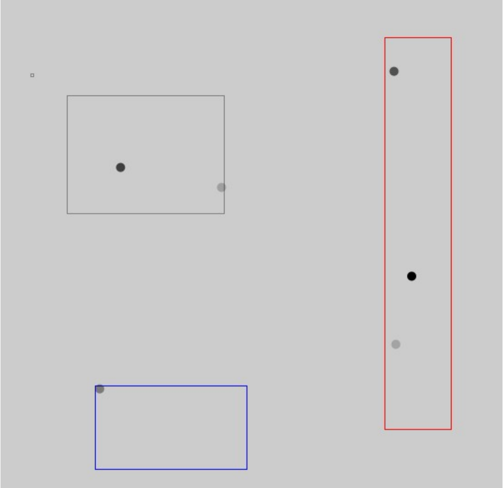


Figure 23: Screen interface of the grain cluster creation

Pulsating clouds of short sine wave sound grains are created by drawing areas on screen. The dimensions of these areas define the cloud parameters (pitch range, panning range and density). The Wiimote's infrared camera feature is used to “point and drag” these rectangular areas on the screen. Multiple cloud processes could be running simultaneously and any of them could be removed at any time.

This sketch allowed quite complex rhythmic patterns to be created, unfortunately not in real time due to the interface restrictions: Using an infrared pointer is a weak imitation of a mouse, being rather inaccurate and slow. This interface required quite rational approach compared to the embodied and spontaneous playing I pursued.

As I personally had very little experience designing digital instruments it was useful to create tangible examples that could be analyzed. Unfortunately instrument designers do not possess as versatile rapid sketching tools as pen and paper are for designers of graphical user interfaces. However, after learning the basics of the programming language I can now “sketch” simple experiments in a matter of hours. These sketches served many purposes: They helped me to learn the capabilities of the controller, to identify the satisfying and effective interactions, and to learn basic audio programming in SuperCollider. The sketches did not fulfill all the initial instrument requirements and it was clear that something more musically refined will follow - thus I did not conduct any kind of player testing for most of the prototypes. Instead I analyzed their strengths and weaknesses, and proceeded to the next iteration round with this experience.

5.3 Design principles for Hearspray

The next step after evaluating the previous sketch-like prototypes was to create the requirements for the final instrument. This instrument was later named Hearspray, which refers to the final playing metaphor: The player is releasing a spray-like flow of small sound particles by pointing the Wiimote and pressing the trigger.

5.3.1 Guiding musical principles

The sketches that involved complex rhythmic material seemed musically very promising compared to e.g. the pitch-focused approaches. Thus the guiding principle was to build on a rhythmic basis and strive for ability to create polyrhythms. Rather than mapping one gesture to one sound event, the instrument aimed at easing the creation of complex rhythmic patterns by allowing the player to control a rhythmic process instead of the individual events. In order to avoid problems with exact rhythmic synchronization the goal was to produce instrument with somewhat abstract nature, without any hints to conventional rhythms in the form of e.g. percussive loops. To achieve a solid and unchanging instrumental identity, I decided to use sound synthesis algorithms as a sound source rather than freely selectable waveform files. The requirements for rhythmic process and synthetic sound generation were combined by using a simple implementation of granular synthesis, which was already utilized in the “Point and drag” interaction sketch.

The exact sound of these grains would be realized by varying synthesis techniques: As a pleasant starting point for the sound I chose the basic sine wave. This ties the instrument to the platform: The soft sine waves, pulsating rhythms and the use of pentatonic scale define the soothing minimalistic soundscape of the Nintendo Wii home menu. However, I also wanted to allow the player to be able to navigate from these “cold” sounds to more “natural” and “rough” timbres. This could give a sense of controlling an acoustic sound source despite of the cold plastic controller.

In order to facilitate rich and interesting musical possibilities, a list of desirable musical dimensions and their opposing polarities was created to guide the design. These dimensions define the available performance space (as discussed in chapter 3.4), which players can exploit freely for creating the desired musical outcomes. The dimensions were not based on research of valid musical parameters, but were influenced by the preceding interaction need evaluations and my personal insight on interesting musical dimensions. The selected dimensions were, in an order of relevance:

Silence -> Single tone -> Cloud of tones. The player has to be able to control the density of the sound events and to also add empty breaks where ever he wishes.

Stable rhythm -> Polyrhytm -> Arhytm. Starting from the default rhythmic process the player should be able to increase the rhythmic complexity all the way to the point where it is impossible to understand the rhythmic pattern.

Hard tones -> Soft tones. This dimension was not referring to any actual physical phenomenon but to a rather subjective criteria – it could be realized using waveforms' timbre differences, various effects or varying attack times.

Short tones -> Long tones. The player should be able to proceed from short “clicks” to a continuous tone.

Clear pitch -> Noise. The player should be able to proceed from simple timbres – e.g. sine waves – to complex waveforms, which do not convey a clear pitch.

Consonance -> Dissonance -> Atonality. The dimension is not referring to the ability of playing all the fine nuances of tonal music, but the player should be able to hint to some of the consonant and dissonant moods of tonal music and especially be able to contrast them to completely atonal material.

Monophony -> Polyphony. This dimension refers to the amount of simultaneous tones from one to many and from independent voices to interrelated voices. It was obvious that the Wiimote is not optimal for precisely and independently controlling simultaneous sounds as all the information has to be processed mentally without relying on visual feedback. Additional challenge for controlling voices separately results from the controllers' coupled dimensions. However, ways to introduce certain amount of polyphony were examined.

Stationary tones -> Moving tones. The player should be able to specify the spatial location of the events in relation to the previous events. As the target platform only allows stereo output , the spatial effect was deemed to be limited.

This list was utilized during the instrument development for prioritizing features and to later ensure that the instrument is actually able to navigate in this space. The instrument needed to allow fine movements along these dimensions in order to produce musical nuances. Equally important was to allow large jumps between the polarities for producing dramatic transitions. This should have been achieved, as the dimensions were numbered enough and the polarities were far enough to allow distinctly different outcomes. To ensure the navigability, it had to be made sure that the player is given as much control as possible over the sound parameters that were later chosen to match these musical dimensions.

In addition to manually navigating along these dimensions, the player would be supported by several automatic processes, which potentially allow him to have moments of rest. Complex rhythmic (and polyrhytmic) playing would be supported by the basic grain generation process and delay lines. Additionally, the player would be able to store and repeat what has been played in order to create background loops, allowing multi-voiced patterns and recurring musical motives to be created. This would also allow re-introducing and developing musical ideas introduced by the co-players during collaborative playing sessions.

Controlling a versatile musical instrument is somewhat conflicting with the typical use context of Nintendo Wii: The console is aimed at casual and fun play. It is likely that some of the people who enjoy the easiness of typical musical games and pre-composed installations may be confused because of Hearspray's demand of being in control and the need to create all musical material. However, the design aims at making the interface as easy as possible in order to allow beginners to become skilled players without unnecessary difficulties.

5.3.2 Guiding control principles

As learned from the previous sketches, hitting with the Wiimote is not accurate for starting sounds. Thus it was decided that a dedicated button would be used for starting the sound-generating process. In this case, musical control and output requirements were considered more important than physically rewarding and visually understandable movements. As mentioned, the aim was to create a reliable and controllable instrument. Due to the measurement problems, hitting with the controllers would only be used as a trigger for secondary functionalities that would not require considerable accuracy in control.

Rotations of the Wiimote and the Nunchuk were selected as modulation gestures for controlling the key sound parameters. Rotation measurements are rather accurate and the gestures relate to the continuous nature of sound parameters allowing the production of fine nuances. As the earlier sketches showed, the Wiimote and the Nunchuk can each be reliably used for detecting two continuous rotation dimensions, and the Nunchuk's joystick provides another two dimensions. Further strategies for interface and mappings were needed as it was clear that the number of sound parameters that would be used for covering the playing space greatly exceeds the amount of controllers' dimensions. The initial list of the desired values to be controlled included sound parameters like pitch, amplitude, several timbre -related parameters, grain density, grain length, panning, delay time, distortion amount, several reverberation-related parameters, several timbre modulation-related parameters and resonation-related parameters.

The real-time sound controls, especially the starting of sound, pitch change and rhythmic controls were seen as the key interactions and thus would be presented clearly to the player. More advanced controls, like the ones used for store and repeat -functionalities should not interfere with the learning of basic controls. All the gestures would need to result to a low latency sound response in order to help the player to learn the causality between his actions and the sounding results, and to allow quick communication between the players.

As it was clear that the nature of Wii controllers makes controlling hard, it was decided that the mappings would be deterministic and simple to balance controllability and arbitrariness. Later on some complexity was increased to heighten the playing experience.

5.4 Forging the instrument

The detailed interface design and development of the Hearspray instrument was an iterative process: The design was evaluated and improved initially based on personal observations, later on based on player testing. This chapter gives an overview of the process, including the rationale for the design decisions from interaction and musical point of views. More detailed description of the final interface and mappings is given in chapter 6.

In order to understand the process description it is beneficial to shortly re-introduce the outcome: Hearspray is based on producing a steady flow of short *grains* of sound. The player is mostly not interacting on microsecond accuracy but is controlling parameters on per-grain basis. This still gives an illusion of real-time control when the density of the produced grains is high. Such style of playing is further referred to as *direct playing* in order to separate it from controlling a higher-level process, e.g. a playback of a recording. The direct-controllable parameters include pitch, amplitude, timing, length, timbre and spatial position of each grain – often with many ways to affect the same perceived parameter. As mentioned, the instrument also allows higher-level control of processes in the form of delay lines and record/play functionality, which allow the players to share and develop musical ideas or to build background loops in order to create richer polyrhythmic or polyphonic textures.

5.4.1 Designing for a pleasant first experience

As stated earlier, one of the goals was to make the instrument approachable regardless of a players' previous experiences with digital instruments. The first step to reach this was to ensure that the main controls are easy to find.

The big raised B trigger button (1. in figure 31) was chosen as the sound-starter. It will most probably be the first one pressed as this button normally lies under the player's index finger. It was realized that the A button on the front side may be more familiar for some players with Wii experience: The A button is often used in the Wii games as an activation button. However, to me personally the B trigger feels more convenient in quick and repeated use.

The sound-generating process will play as long as the player holds the trigger button, releasing the button stops the sound production. This immediate way of stopping may be an important safety feature to a beginning player who starts to explore the chaotic sides of the instrument. Actually this button is one of the most important controls: By learning to use it actively the player can create short rhythmic bursts of tones and create the most distinct musical events – the silent moments.

Intuitively, the most discoverable of all the rotation gestures is raising and lowering the Wiimote (Figure 24). This dimension was mapped to the pitch of the sound output. The reason for choosing pitch as the most accessible parameter was to produce initial musical outputs that would not sound like "electronic noise" but rather like something familiar: All people have heard music that uses pitch variations as one of the most important parameters. This way of controlling pitch also connects Hearspray with the theremin, which is one of the most well-known electronic instruments.

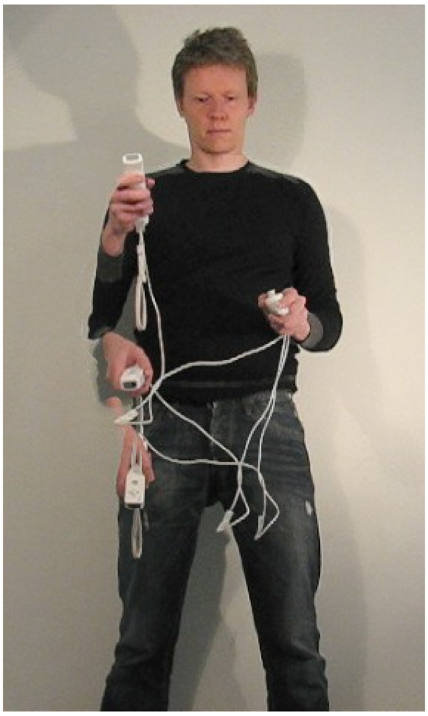


Figure 24: Controlling pitch

First experiments with freely sliding pitch obviously gave the instrument an atonal character, as it was hard to hold hand steady in order to stabilize the pitch or to jump directly to the desired pitches. To further establish approachability, various mappings with musical scales were explored. In the final version, the pitch is initially mapped to a major pentatonic scale to create a feeling of simplicity. Pentatonic (five-note) scales seem to have some sort of universalness as various pentatonic scales – especially the anhemitonic[8] ones – are the most common musical scales found in music traditions all over the world [Pentatonic scale 2009]. Using a pentatonic scale also ensures that two players playing together never sound too "bad": At least in the sense of western tonality the notes of a certain pentatonic scale never result to dissonance when played together.

---

[8] Anhemitonic refers to a scale without semitones

Just by discovering the start/stop button and the pitch control the player will be ready to create simple, pleasant melodic patterns – the scale mapping and the inherent rhythmic nature of the instrument take care of producing coherent and active musical results.

It is very likely that the controllers will be held in a position where the Wiimote's and the Nunchuk's tips are pointing forward. This is due to natural hand position and the industrial design of the controllers – they fit the hands in a certain way and based on my observations hand muscles are relaxed in this position. I used this notion to create a "starting area" for values, where the instrument's sound is most pleasant: timbre is softish, attack is soft, rhythm density is moderate, amplitude is audible but not too loud, and the (high) pitch is not causing sharpness to the sound. The players should naturally find their way back to this position after the adventures with more experimental sounds as the position is most comfortable in terms of muscle strain.

Various vibration feedback types are used to communicate the instrument's status. As an example, the Wiimote vibrates slowly when the player starts the sound, fulfilling two goals: It strives for creating an illusion of playing a tangible and responsive instrument instead of a plastic controller. It also aids in revealing possible problems: When the controller vibrates, the instrument is functioning. If no sound is heard, the player should blame the speaker system instead of the instrument.

As stated earlier, all the player's control gestures should result in a sound change in order to aid figuring out what actually are the possible gestures and what are the causalities between the gestures and sound. This was a crucial factor when evaluating the sound-processing algorithms for the instrument. For example, simple filters (e.g. a high-pass filter for reducing low frequencies) were initially tried, but their effect was dependent on the source sound's frequency and waveform type (e.g. a sine wave does not contain any high partials to filter out). This led to the filter control's perceived effect being influenced by other control parameters (e.g. pitch). The final set of the controllable sound parameters was refined to avoid such situations: For example, the available pitch range is restricted to seven octaves, blocking the access to very high frequencies where the changes in timbre are no more perceivable.

5.4.2 Extending the musical possibilities

After a pleasant first experience and finding the very basic controls, the player can freely explore the instrument further. Despite the initial pleasantness, the instrument is not aiming at reproducing tonal music. In addition to using the default pentatonic scale for creating melodic material, the player can also slide the pitch freely by holding the A button (3. in Figure 31). This allows the player to e.g. produce rapid patterns with completely arbitrary pitches, or to create beating drone sounds by approaching the frequency played by his companion. Playing accurate pitch patterns can also be tried, even though the challenge of stabilizing the hand movements is similar to mastering the theremin playing.

The further development towards the final instrument also introduced an option to choose the pitch mapping (4. in Figure 31) from a list of musical scales. By default, the instrument contains 11 pentatonic scales, each starting a semitone higher than the previous one. This is by no means a flexible way of playing music based on chord structures. Instead, the aim is to be able to introduce small mood changes, like switching from a whole-tone scale to a phrygian scale. In the case of multiple players, scale changing affects all players simultaneously in order to create a coherent sounding result. The controls involved are less prominent and thus suited for such a secondary functionality.

Playing with timing and rhythmic accentuation

As mentioned, most of the Wii controllers' dimensions are coupled, making it difficult to control the related values separately. It had to be evaluated, which of the parameters have to be precisely controlled, and which are secondary parameters with less precise control.

Figure 25: Controlling the density

For Hearspray, one of most crucial aspects was to play with the timing of the grain production process. In order to allow independent control of rhythm and pitch, the Nunchuk was chosen for timing controls instead of the Wiimote: Rotating the Nunchuk sideways increases the density of grain production (Figure 25). By default, the flow of grains is rhythmically stable, making the player responsible for creating interesting events by e.g. adding silent moments, accentuation or timbre modifications. As opposed to the challenges in vertical pitch control, horizontal rotation seemed to be a satisfying way of maintaining the rhythm: It is easy to keep a stable rhythm, jump to a different speed or to gradually change the density. Arhythmic patterns could also be created by lightly shaking the Nunchuk.

Grain length was chosen as the coupled counterpart of the horizontal density control as they both relate to the timing characteristics of music. The chosen dimension – the Nunchuk's vertical rotation (9. in Figure 31) – is actually controlling many sound parameters: Due to the lack of available control dimensions it was decided that grain length and attack hardness will be combined as a general amplitude[9] envelope control. When the Nunchuk is tilted to the very front, grains are extremely short (they sound like a click without a pitch) and the attack is hard. The central position produces softer short tones and in the other extreme sound is long with a gradually rising attack (Figure 26). The resulting control is somewhat complex allowing restricted independency of length and attack, but is still a compromise made because of lacking dimensions.

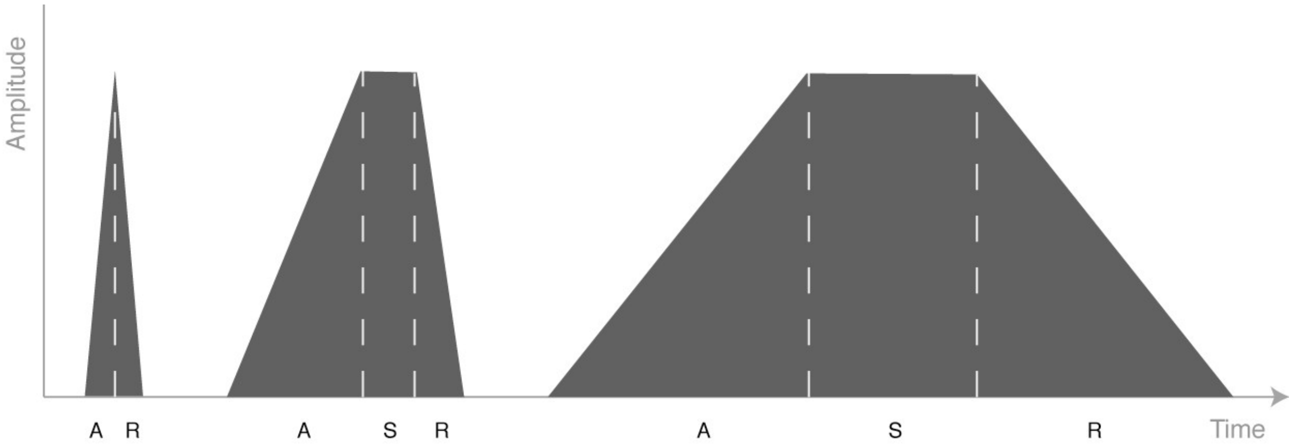


Figure 26: Examples of different grain envelopes present in Hearspray. The envelope shapes are constructed by Attack, Sustain and Release lengths

---

[9] **Amplitude envelope** affects, how the amplitude of the sound develops during its existence. A very typical envelope defines amplitude values for Attack, Decay, Sustain and Release phases. Amplitude increases or decreases linearly (or using other function of time) between the defined points.

Initially there was an additional mode control, which stopped the grain production and played a continuous tone, enabling the player to play freely sliding, long tones. However, this feature was later taken away due to the shortage of interface buttons. The feature also broke the basic grain-generation behavior and thus was questioning the instrument's logical integrity. Instead the maximum grain length was increased all the way to the point where the longest grains are considerably overlapping (Figure 27). This creates a continuous tone, which has a fluctuating character as the overlapping makes the total amplitude grow, introducing distortion. If the player slides the pitch when these long grains are produced, separately-pitched overlapping grains are heard together, allowing the player to create simple “chords”:

Figure 27: Increasing the grain length produces overlapping sounds. Grain density is fixed in this example.

The delay effect was used in order to allow the creation of polyrhythms: When the delay button (8. in Figure 31) is held, everything that is currently playing is replayed multiple times, with repeats appearing at a specific frequency (Figure 28). When the original and multiple delayed signals are playing on top of each other, the resulting combination of these rhythms is often heard having a complex but logical rhythmic relationship. In order to control the relationships of this combined rhythm, the player can affect the “tempo” of the original rhythmic process with the density control, and the repeat frequency can also be affected (but is more difficult to control). In the final Hearspray version the player can actually pile up multiple delay effects with different repeat frequencies to create even more complex polyrhythms.

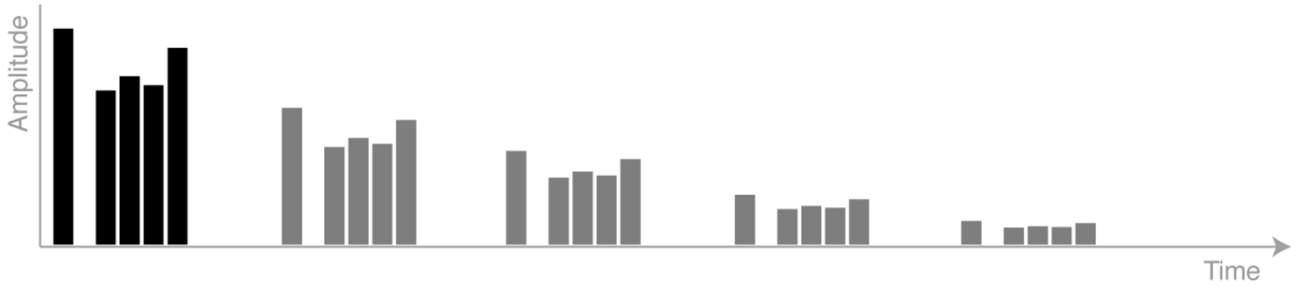


Figure 28: Example of delayed and fading repeats occurring four times

The second Wiimote rotation dimension – rotating sideways (5. in Figure 31) – controls the sound amplitude. Due to the coupling, some precision over amplitude is lost when the player is actively controlling the pitch. However, it seems that these minor variations in amplitude are not perceivable. A separate pitch lock function (11. in Figure 31) was later added to enable production of bigger amplitude variations (e.g. crescendos, pauses, accentuation) without having an effect on pitch. The amplitude control always has a real-time effect on the sound instead of the regular per-grain control. This allows immediate attenuation when long grains produce too chaotic-sounding results.

An understandable gesture for the viewer is to swing Wiimote rightwards, which results in a significant temporary amplitude peak, allowing accentuation of the rhythmic patterns. This is caused by the fact that rotation values are calculated from the acceleration values, and thus the sudden acceleration produces “incorrect” rotation values. The accentuation is further emphasized by the fact that the amplitude control also has an effect on the timbre of the sound, emphasizing high frequencies when the amplitude is boosted. This is inspired by the behavior of many acoustic instruments, and makes the amplitude control much more powerful in the musical sense.

Further control of the timbre

The Nunchuk joystick was chosen for controlling the basic timbre of sound (10. in Figure 31). The starting point was to use simple basic waveforms – from sine to saw waves – which could be used for creating tones from “soft” to “sharp” timbres. This basic wave -dimension is controlled by the vertical axis of the joystick. In their pure form these sound waves might sound a bit “cold”: They do not have any kind of aperiodic variations as opposed to most sounds found in the nature. Thus I added another control dimension, which would contrast the cold tones with ones sounding more “natural” or “fuzzy”. This was realized utilizing stochastic synthesis[10] techniques. The result sounds dynamic and fuzzy, sometimes flute-like, sometimes noise-like. This “warmness” is controlled by the horizontal axis of the joystick. The combined sound of these two dimensions is created by mixing together several sound sources. As both of the joystick's dimensions control perceptually related (timbral) properties and the controller dimensions are coupled, it is probably hard to mentally distinguish the dimensions – the player is more likely to feel he is navigating in a two-dimensional timbre space. The joystick does not offer very precise control, as the physical range of movement is short. This unfortunately leads to a limited control of fine timbral nuances.

After the basic sound controls using rotations and joystick were in place, two effect controls were added, allowing further timbral manipulation: The player can make rapid hitting gestures using either controller in order to break the “pleasant” tone quality. To compensate the relative inaccuracy of the hit gestures, only one gesture per controller was used: The player can hit to any direction in order to toggle an effect. This in turn encourages experimenting with the gesture, as the hit always has an effect on the rotation values and thus on other sound parameters. Hitting with the Nunchuk introduces a distortion effect, allowing the player access to extremely noisy timbres unavailable in the joystick-controlled timbre space. The same gesture applied with the Wiimote produces a resonating effect, imitating playing in a larger space that resonates with the instrument, sounding like a string of small metallic objects when playing higher pitches. Combining these effects produces unpredictable noisy ringing outcomes, inviting to play on the verge of losing control.

When the player hits with a controller in order to toggle an effect, the controller also vibrates in a strong but rapid fashion to indicate that the gesture was strong enough to trigger the effect. This feedback is important in the situations where the player wants to start with the effected sound, toggling the effect before he presses the sound trigger button: Obviously the audio channel cannot convey the “effect in use” -information before sound is started.

---

[10] Stochastic synthesis utilizes randomness and probability distributions in the waveform production

Moving sound in the space

The spatial properties of sound – panning left to right and the perceived distance of sound – were less prioritized in Hearspray. As the dimensions of the controllers were scarce, a decision was made to use the Wiimote arrow buttons (6. in Figure 31) for limited spatial control. Four arrows matched nicely to the required directions, but do not provide such accuracy as using the continuous dimensions. By using the arrow buttons the player is still able to introduce abrupt jumps or gradual transitions between nine simulated locations (center, front, front-right, right etc.) and can maintain sounds in multiple locations simultaneously with the aid of delay lines and other automatic processes. Other possibility would have been to utilize separate control modes: For example, when the C button is pressed, the Nunchuk rotations would affect panning instead of rhythmic density. However, experiments with using modes indicated that it is mentally very demanding to keep track of the current mode and hard to operate two values precisely by using the same control dimension.

5.4.3 Playing together

As noted, several Hearspray players are able to play together, each with their own voice – resembling the use of most acoustic instruments. So far the instrument has only been tested with two simultaneous players, but it should be easily adaptable for more: As mentioned, at least four Wiimotes can be connected. All players' instruments are similar in their behavior, sound and range. Thus Hearspray does not direct players to take any specific roles – such as accompaniment & solo or soprano & alto – in collaborative playing.

One of the potential problems for collaborative playing is to separate one's own voice from other players' voices. For digital instruments, there is no such direct connection to the instrument's sound as in the case of acoustic instruments' sound-producing body. In Hearspray's case, the individual voices tend to be a bit blurred in the final speaker output This creation of an unified sound-field from multiple player's output may even be musically beneficial. One potential solution for voice distinction could have been to utilize the small speaker in the Wiimotes for outputting the player's voice, but that is currently impossible due to the technological limitations.

Hearspray only provides partial solutions for this voice-identifying problem: Starting the sound is always enforced by providing tactile feedback with the hope that it helps the player to connect with the new appearing voice. Players' sounds are also by default shifted to different locations horizontally: The first Wiimote is panned slightly leftwards, the second one rightwards. However, this distinction is only clear when the players take equivalent locations in the physical space. Different pitch and timbre ranges were also tried for players in order to create recognizable outputs, but these strategies were discarded: It was e.g. impossible to find two timbre spaces that would not overlap but would still sound "similar" in order to create sense of one instrument. It also seemed arbitrary limitation of the players' abilities: What if they would like to e.g. hit same frequencies in order to play unisono passages?

5.4.4 Enlivening the sound

Experiments were done adding automatic rhythmic randomness to the instrument's output – e.g. sudden pauses in the grain generating process – in order to avoid boring repetitive results for the players who were not aware of all features, such as the density control. However, this easily led to the feeling of too limited playing control. Further on, smaller randomized changes in individual grains' amplitude, timbre and panning were tried, imitating the "imperfect results" of playing acoustic instruments.

In the final Hearspray instrument there is a slight random component in the amplitude and stereo panning of individual grains to create liveliness. This is very subtle and can easily be overridden by the player's gestures. Slight time-varying timbral variation is created via the use of two filters that manipulate the amount of high and low frequencies in the sound, independently from each other. As mentioned earlier, player-controlled filters were not satisfactory due to their outcome depending on the source sound material.

Certain variation is also inevitably caused by the controllers' nature: It is tiring on the hands to hold them in the exact position, leading to occurrences of minor deviations. This is especially prominent in the pitch control – small deviations in rhythm, amplitude, timbre and grain length seem not to be so noticeable. The controllers' rotation values also have an effect on each other when used in the extreme positions due to the way the values are calculated. Thus it will be harder to hit pitches when playing very soft or very loud.

The accuracy challenge is also related to the lack of a reference point in the controllers: The player has to use his own body or ground level as a reference when estimating e.g. “Where was that certain pitch?”. Thus it will be very hard to exactly reproduce musical patterns using the direct playing gestures.

5.4.5 Storing and reproducing musical patterns

In order to ease the creation and recreation of musical patterns, the players can switch to a “manipulator” role instead of actively producing sounds with the direct playing controls. The players are able to record the musical output of the instrument (12. in Figure 31) in order to capture musical motives for later use, regardless of which player introduced the motives in the first place. This recording can be later played back (13. in Figure 31) partially, fully or as a repeating loop. Playback can be manipulated in order to develop the musical motives further. As the recordings are done in audio form, these features potentially allow even sounds of other instruments to be modified. The manipulations available are based on the direct playing paradigm, utilizing the same interface controls as their similar-sounding direct playing counterparts. Thus e.g. the pitch shifting gesture equals the pitch control gesture and the playback rate control gesture equals the density control gesture. This allows the players to learn the manipulation interface based on the existing mental model of the instrument.

In addition to supporting repeatability and allowing dialogue between the players, the playback functionality can be used for creating background loops. This enables multi-voiced playing even for a solitary player. Recording is also possible when the playback is ongoing, allowing the creation of multilayered patterns. Instead of recording everything, the player can choose to record only his own output for creating background loops free of distracting sounds produced by the other players.

The direct playing functionalities can still be used while the loops are playing. Thus the player could be possessing several roles, even if only performing in one actively. The interactions involved for simultaneous roles are admittedly complex, partially due to the limitations of the Wiimote. However, such interactions are directed towards the experienced players and could be seen as motivational challenging aspects of the instrument.

Hearspray communicates the state of these record and playback processes using the Wiimote's LED lights. Visual channel was chosen as the vibration is reserved for the feedback of the direct playing events. The slower nature of the recording interactions also seemed to match well with the research results concerning slower processing of visual information. When a recording process is started, the LEDs start to blink. Lights indicate that either player himself (player number blinks) or all players (all LEDs blink) are being recorded. This flashing is the strongest visual signal that can be given in the limits of the Wii controller technology. Strong signal is given as there are no other forms of feedback given about the recording state. When the playback state is activated, all four LEDs are constantly turned on. In this mode the player also hears the playback, but using a light carries the message in the situations where the player has managed to record moments of silence. When neither of processes is active, the default behavior of indicating the player number via the LEDs is maintained. This is due to the platform conventions and in order to indicate that there are batteries left in the controller. In practice, showing the player number is redundant as the instruments are completely equal.

5.4.6 Heightening learnability: Visual learning interface

During the first player evaluations (see chapter 7) it became evident that not all of the movement gestures are as self-evident as initially envisioned: For example, some of the rotation gestures are more prominent for the players than the others. Due to the technological problems it was not feasible to produce sounding results of all the players' possible gestures, e.g. the rotations around all three dimensions.

Because of these problems it was decided to create a visual learning interface, which served two purposes:

1. Revealing all the available controls
2. Showing the boundaries of the rotation controls. In practice, only 180 degrees are used and the player never needs to turn the controller upside down.

Such learning interfaces are common for the Nintendo Wii platform - most games utilize a tutorial mode for indicating what kind of gestures player is expected to perform. The Hearspray visual learning interface is only shown when it is needed: It has to be executed separately as it will probably be only a distraction when aiming at musical results.

The first version of the visual interface was rather complex, showing information about all controls for two players simultaneously:

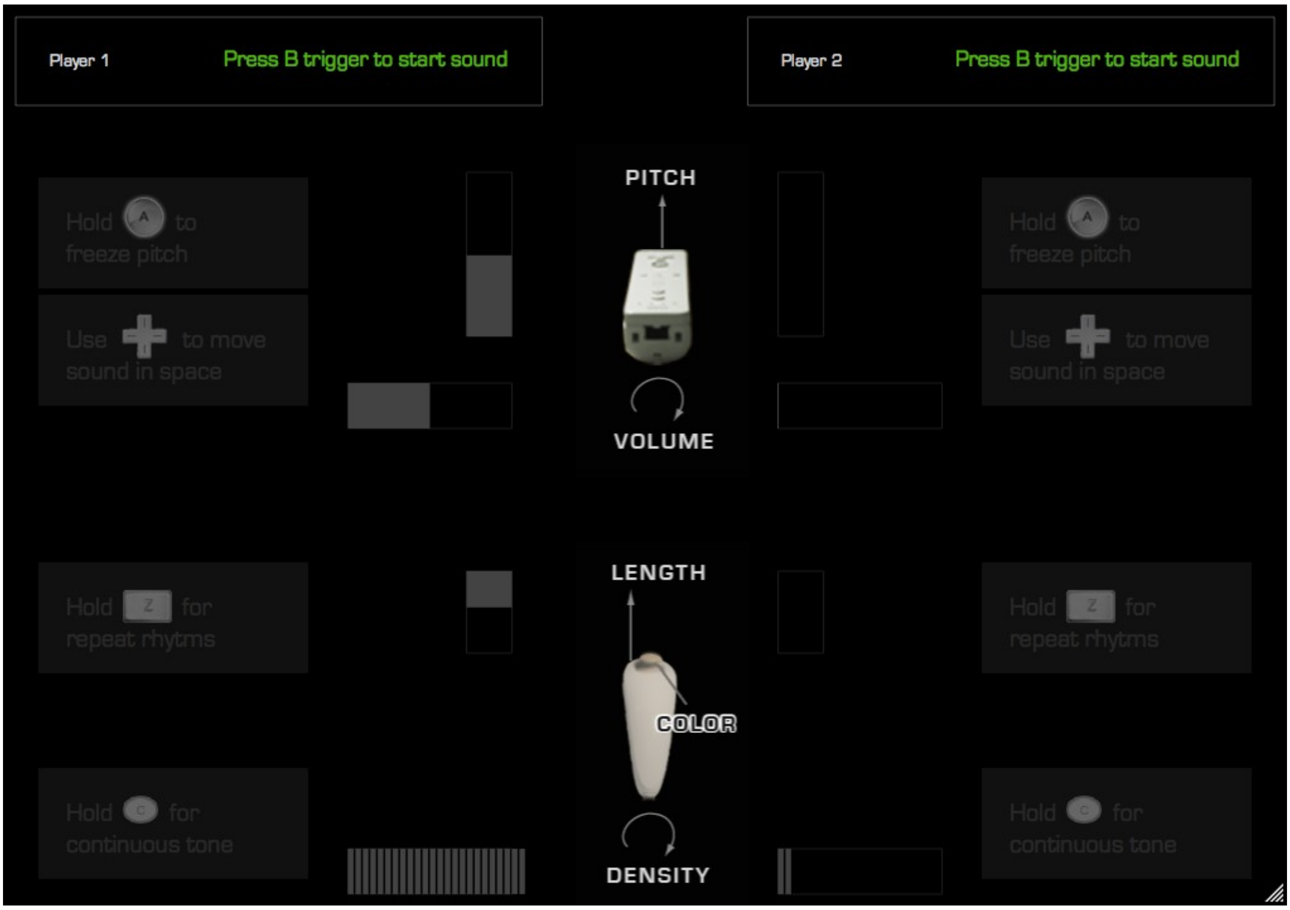
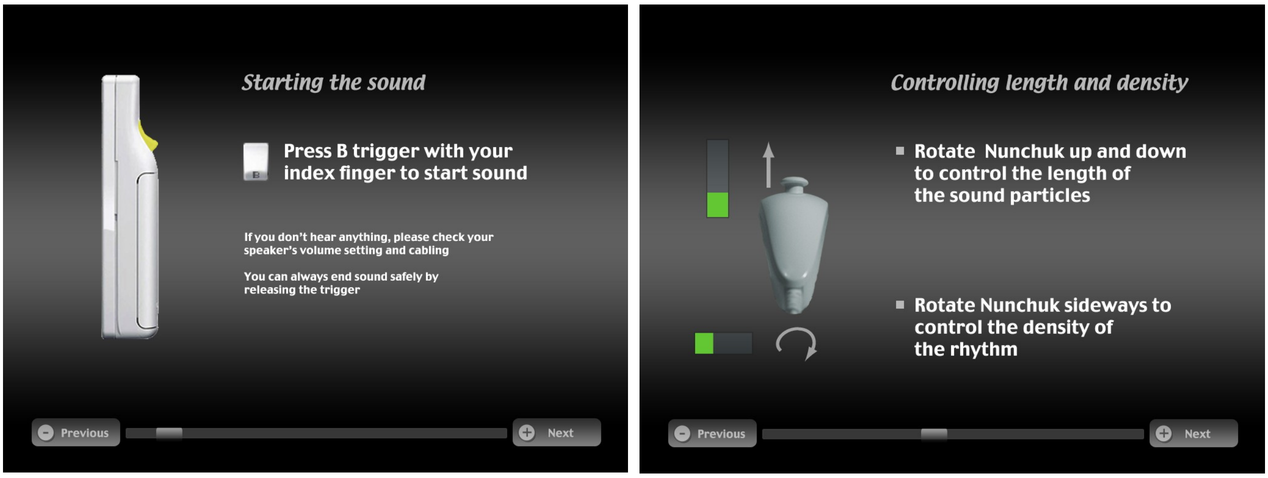


Figure 29: First learning interface version

Based on the results of the player testing, a new, simplified version of the tutorial interface was sketched:



Starting the sound

Press B trigger with your index finger to start sound

If you don't hear anything, please check your speaker's volume setting and cabling

You can always end sound safely by releasing the trigger

Controlling length and density

- Rotate Nunchuk up and down to control the length of the sound particles
- Rotate Nunchuk sideways to control the density of the rhythm

Figure 30: Screens from the improved version

This version uses a wizard approach, presenting one functionality at time. This allows the players to concentrate on learning interconnected gesture components separately. Because of the interface approach, there is more space for explaining each functionality, allowing the use of more detailed images and descriptive language instead of the technical terminology present in the first version. The wizard is further simplified by omitting second player's data: Computer tutoring is probably relevant in solitary context, whereas peer-to-peer tutoring can be used in collaborative playing situations. Due to the time constraints, this interface version was only designed, but not built or tested with actual players.

6. FINAL HEARSPRAY INSTRUMENT

This chapter presents the logic of the final playing interface and mappings in deeper detail. The picture below shows all interface functionalities of Hearspray:

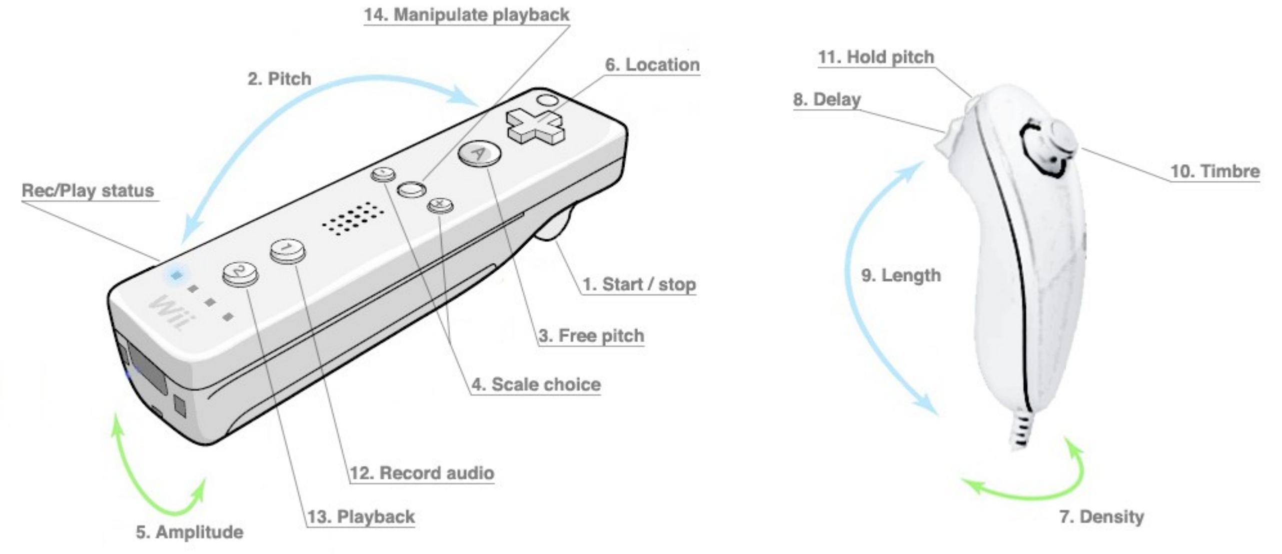


Figure 31: The interface controls

The interface controls for direct playing

1. Sound generating process is started and stopped with the B trigger button in the Wiimote.

2. Raising and lowering the Wiimote is mapped directly to raising and lowering the pitch. The mapping is logarithmic in order to maintain the same physical distance of gesture between the musical octaves (octaves' frequencies grow exponentially: A3 = 220Hz, A4 = 440Hz, A5 = 880Hz). By default, pitch is not following movement continuously, but is mapped to the notes of a major pentatonic scale. The related mapping is actually one-to-many as the amplitude is also boosted for lower pitches in order to make them more audible.

3. Holding the A button in the Wiimote removes the scale mapping and allows free pitch sliding.

4. The Wiimote's + and - buttons can be used to navigate the list of musical scales used for the pitch mapping. By default there is a selection of pentatonic scales available, but the list of scales can be easily customized.

5. Rotating the Wiimote sideways controls the amplitude of sound. The mapping is logarithmic in order to match the nature of perceived sound intensity, which relates in square relationship to the actual sound amplitude. This control functions as it would be mapped in one-to-many fashion: Increasing the amplitude also has an effect on timbre – the signal is distorted when the amplitude is high due to the nonlinear amplitude response of the instrument.

6. Pressing the Wiimote arrow buttons controls spatial positioning: The Left and Right arrow buttons control the stereo panning of sound. The Up and Down arrow buttons increase and decrease the perceived distance of sound source by controlling the mix of direct and reflected sound. Simulated room size is simultaneously altered in order to vary the feeling of depth. If an arrow button is pressed before the player starts the sound production, the sound is transferred directly to the extreme position. If the button is pressed when sound is already playing, sound starts to slowly move into that direction, enabling limited access to the locations between the extremes.

7. Rotating the Nunchuk sideways is mapped directly to the density of the generated grains. This control also has an effect on the amplitude: When the grain density is high, the amplitude is reduced in order to avoid distortion caused by several grains playing simultaneously.

8. Pressing the Z button in the Nunchuk directs the played signal to a delay line, which repeats the played patterns multiple times. The frequency of these periodic repeats is based on the horizontal rotation of the Nunchuk (the control also affecting the grain density) at the time when the Z button is pressed. Each press of the Z

button creates a new delay line, allowing multiple simultaneous repeating signals with individual repeat frequencies. The amplitude of the delayed signal is reduced when the player is pressing the sound trigger in order to avoid messy output. When the player stops playing, the delayed tail is given the full amplitude.

9. Rotating the Nunchuk up and down increases and decreases the length of grains. More precisely, the grain length results from the combined values of attack time, sustain time and release time, which all are affected by this control. Both attack and release times increase along the length control value, but not in a linear way: There are non-continuous "jump" points for both values, allowing player some control over attack hardness independently from length.

10. The joystick in the Nunchuk is used to control the timbre. The up-down axis controls "sharpness" factor from "soft" (sine wave) to "sharp" (saw wave). Turning the joystick to the right gradually mixes this sound with a noisy flute-like stochastic waveform. Turning the joystick to the left mixes the sound with another stochastic waveform, which resembles unpitched noise.

11. Holding the C button in the Nunchuk removes the Wiimote rotation's effect on pitch, allowing the player to play stable pitches.

In addition, hitting the Wiimote rapidly to any direction starts a slowly decaying resonation effect for the grains. Similarly, hitting the Nunchuk rapidly to any direction adds a decaying distortion effect to the sounding output. The effects' amount is based on the strength of the acceleration and it can be maintained by waving further. A certain acceleration threshold needs to be exceeded before the effect is toggled in order to separate the hits from the accelerations involved in regular rotation gestures.

The interface controls for record and play states

12. Recording starts when the player presses the 1 button in the Wiimote. If the player holds the button, he will record everything going to the main output. This is convenient for recording what others play. If the player just presses the button quickly, his

fingers are free for playing and the recording will contain only his own output. Depending on the selected recording mode, the recording will stop when player either releases the record button or presses it again. If the player presses play button before ending the recording, the recording will stop.

13. Playback of the previously recorded content starts by pressing the 2 button in the Wiimote. Playback will always automatically loop. The process can be stopped by pressing 2 again.

14. The playback can be altered using the same control movements that are used in the direct playing interface. However, the player needs to release the playback manipulation lock by holding the Home button. This prevents affecting the playback mistakenly by moving the controller.

When the Home button is pressed, the following manipulations are available:

- Rotating the Wiimote vertically affects the pitch of the played back patterns (using a pitch shifter).
- Rotating the Wiimote sideways affects the amplitude of the playback.
- Rotating the Nunchuk sideways affects the playback rate, equalling a tempo control.
- Pressing the Z button directs the playback to a delay line, causing repetitions.
- The joystick in the Nunchuk controls resonant filters' parameters (cutoff frequency and the resonation amount) in order to enable timbral manipulations

If the player starts direct playing by pressing the B trigger button, the playback process will continue, allowing the playing over background loops. More detailed explanation of Hearspray's technical solutions is given in Appendix A.

7. EVALUATIONS

This chapter details the observations and reflections from the different playing situations, and explains some of the related alterations to the instrument design.

7.1 First-time players

The prototype was tested with nine first-time players in a very informal fashion, without a fixed script to drive the sessions. As the tests began in the middle of the prototype development, there has been minor differences in the prototype interface and behavior between the participants.

The goal was to observe – without providing any hints on the use – how the players learn the interface controls, and how they start to explore the offered musical possibilities. They were asked to verbally describe the experience when they were searching the controls and learning to play in order to evaluate what kinds of problems they might have on the way. A typical test lasted fifteen minutes, consisting of instrument playing and a short discussion afterwards to extract people's opinions and feelings. Three of the later tests started with a screenshot of the initial visual interface in order to evaluate its understandability and then continued with the actual playing.

All of the test players had some background in playing musical instruments, even though it was not a criteria for selecting the participants. Three of the players had only been using acoustic instruments, but the remaining ones were somewhat proficient using digital (at least desktop-based) music tools. Five of the players had never used Nintendo Wii, thus making them unaware of the typical control gestures.

7.1.1 Interface discovery

As the Wiimote and the Nunchuk are designed to fit the hand in a certain way, no player had problem figuring out how the controllers should be grabbed. Two of the players asked, which way the controllers should go into left and right hand. The handedness does not actually matter as the controllers are symmetrical, but my assumption was that the main hand would be used to control the most elementary properties, like the pitch and starting of the sound. This was encouraged by the case of one player who used the instrument several times: He noted that playing felt much better during the second session when he switched hands in order to hold the Wiimote in his main hand. This was taken into account when designing the second iteration of the visual learning interface: It guides the player to grab the Wiimote with the main hand.

Almost all players first tried to hit to a certain direction with the Wiimote, causing no audible response. This seemed to happen regardless of the players' prior Wii usage experience: The Wiimote's stick-like form factor seems to direct players to perceive such hitting affordances – a fact that was realized when the decision to avoid hit gestures was made. Fortunately all players managed to find the sound starting B trigger button quickly after they tried hitting. They also remembered the control after discovering it.

All of the players found the pitch control (raising or lowering the Wiimote) immediately: This is natural, as it is very hard to hold hand stable so that pitch does not change. In the very first test with an early version of the prototype the pitch would be changed only if a specific “release pitch lock” button was pressed. This provided easier access to a stable pitch and thus might help experienced player to reach musical goals, but it made changing pitch totally invisible to the novice players. Due to the demands of the initial discovery, the functionality was reversed so that the pitch may be stabilized by holding a dedicated button.

The joystick in the Nunchuk is very prominent and all players thus discovered the timbre control quickly. Most players described it correctly with terms like "color" or "filter". One player confused it to pitch, which is understandable as the vertical direction of the joystick is mapped to the amount of harmonics, adding higher and higher overtones to the sound. All players also found the spatial controls, even though two players did not connect up and down arrows to the distance of the sound source, asking e.g. "Is it some kind of filter?". Hearspray admittedly tends to sound metallic when the simulated distance grows (due to the increasing “room size” and addition of early reverberations to the signal).

There were varying degrees of problems finding the three remaining rotation controls (horizontal rotation of the Wiimote for volume, vertical rotation of the Nunchuk for grain length and horizontal rotation of the Nunchuk for density). This is probably related to the shapes of the controllers: The Wiimote's wand-like shape suggests waving in the air, which easily translates into vertical but not horizontal rotation. The Nunchuk's shape is somewhat neutral and the presence of the joystick directs attention to manipulating it. This problem of perceiving the controllers' affordances limited the players' musical output, as especially the subsequent lack of rhythmic variations led to monotonic music. I made various correction attempts, such as switching the grain density control to the Nunchuk joystick. Unfortunately the joystick is much more imprecise and always returns to center, making it suboptimal for advanced rhythmic playing and permanent tempo changes. Some players also tried rotating the Wiimote parallel to the ground level just to realize that it does not affect anything – as mentioned before, it is technically impossible to detect such rotations. There were no possibilities to solve these affordance issues on the controller or mapping levels and thus the only solution is to rely on teaching the players via the visual learning interface.

The recording functionalities and scale changing were not usually detected in the tests, but as these functions are aimed for advanced players it is only positive that they stayed outside the first-time users' attention. These buttons were probably not discovered as they are smaller or located in the extreme bottom of the Wiimote – often covered by the palm of the hand. Acceleration effects were found only by four of the players. I do not consider this as a problem either: These effects generate rougher tones and moments of uncontrollability, which should probably not be part of the first exploration. They will inevitably be triggered at some point when the player makes quicker gestures, and the gestures' connection to the output will be noticed as the sound transforms so radically.

All those players who were introduced to the visual learning interface had no problems discovering the controls. Thus such interface can be valuable in overcoming the problems with the interaction affordances presented by the Wiimote. However, the players had some problems understanding the compact and subjective terminology (For example, “Density” and “Color”) used in the first version of the visual interface. They also mentioned that a lot of information had to be absorbed simultaneously due to seeing all the parameters at once. After these observations I designed a second version of the learning interface, which tried to solve the described problems.

7.1.2 Overall playing experience

The players used positive adjectives of the initial playing experiences and of the instrument's sound. Most players seemed to be initially fascinated of just being able to create melodic patterns by waving their hands. They described the sound as "pleasant", "wind-instrument -like", "japanese” or “oriental". One player expressed that it was nice that the instrument did not sound rough right from the beginning as many other digital instrument experiments he had tried. Thus my decisions to use simple pentatonic scales and softer timbres for the “starting area” were well-received. As envisioned, the first-time players did not find the rougher aspects of Hearspray as much as the experienced ones did later.

However, it was clearly revealed that people have different approaches when presented with a new instrument: Some people started experimenting boldly, trying to find all functionality, limits and sonic possibilities of the instrument whereas some seemed to have very careful approach, just scratching the surface. It seemed that the latter kind of players needed guidance on what to do with the instrument. Actually one person posed the question "What should I do?". This kind of behavior was exactly correlating with the player's prior experience on digital music tools: The people who had played only acoustic instruments seemed to be more passive. It could be a coincidence, but it could also reflect the fact that the people who have experience with the acoustic instruments often have been trained in western classical tradition by teachers who have discouraged free instrumental exploration. It is clear that Hearspray may delight the exploratory-minded but may also confuse some people.

The automated rhythmic nature of the instrument can be a double-edged sword: All the players seemed to enjoy it when they started with the instrument, as it gives somewhat active and powerful character to the instrument. After a while of playing it became clear that not all players realized that they can also play shorter bursts by pressing and releasing the trigger button. Thus the musical outcome turned to be very monotonous without specifically aiming at it. This seemed to be tied to the previously mentioned explorative musical approach – some players also created interesting rhythmic patterns on their first try.

To my relative surprise, a motion controller does not necessarily convert the players of digital musical instruments to visually interesting performers. On the contrary, several of the test players explored the instrument by sitting on a chair and making very small movements. On the other hand, some of the players were standing or making more expressive gestures. It can be seen as a strength or a weakness, but the Wii controllers clearly allow multiple different performance strategies in relation to the physicality.

7.2 Experienced players

The instrument was later tested with two players who had tried it before. Both of the players could be categorized as proficient and exploratory-minded digital musicians. These sessions were much longer than the first-time tests and also included the author playing part of the time with the test players. Both test players were able to create interesting music, providing pleasant surprises by finding sounds that were not imagined when designing the instrument, or were not found during the development process. The playing experiences seemed to be positive: One of the players described Hearspray as "a real instrument" which could "be used for days" still finding new opportunities. The sessions were partially videotaped in order to analyze the improvisational interactions.

Figure 32: Early two-player session

Hearspray proved to be well suited for collaborative improvisation: One could constantly be aware of what the other player was doing by monitoring both the hand movements and the facial expressions in addition to the sound signal. Thus it helped in the creation of sensible responses, starting from the imitation of the companion's playing gestures. Analysis of the video revealed many moments of fruitful musical communication, like answering the fellow player's musical ideas with similar patterns or creating totally contrasting material. It was apparent that the controls made it hard to repeat the musical ideas exactly, but creating similar enough responses was achievable, ensuring coherent musical developments. Some of the interesting possibilities included:

- overlapping of similar or different rhythmic patterns, with gradual changes in the polyrhytmic structure
- placing players' voices in different spatial locations
- co-playing with long low pulsating drones
- creations of atonal clouds
- improvisational role-taking: call-and-response patterns, solo and accompaniment
- playing over background loops and re-introducing events from the past

The “record and play” functionality was evaluated for the first time during these tests: It seemed to be a proper tool for ensuring continuity and surprising references to the past among the improvisations. Thus it is possible to create very complex and evolving musical pieces by just two players. However, the creation of repeating background loops proved to be limited: Both players were often playing simultaneously, and thus it was not possible to record single-voiced background loops. Because of this, the option for recording only one's own voice was added later.

The tests also revealed limitations of using the Wii controllers: The players wanted to use more complex direct play control strategies, such as combining pitch sliding and spatial position change, which would involve pressing several buttons of the Wiimote simultaneously. The players did have spare fingers available, but they could not reach the buttons as those were in the front side of the controller. Unfortunately the players' fingertips lie in the opposite side.

7.3 Hearspray on stage

The premiere performance of Hearspray was given 4.4.2009 in PixelAche Open Forum at Kiasma by the author and Timur Kuyanov. The received feedback was positive, referring to the visually interesting performance, the “nice sounding” instrument and the presence of musical interplay between performers.

Figure 33: Rehearsing for the performance

However, the performance time was very limited, and we lacked a plan on how to create an interesting compact musical piece. This was due to the fact that we had been only rehearsing lengthy improvisations in order to get to know the instrument. Afterwards, I feel that our performance could have had more musical coherence. Hearspray is not a tool that would automate the highest-level musical decisions – the players are responsible for formulating the conceptual strategies.

7.4 Personal reflections on playing

Hearspray still surprises me in positive ways after many playing sessions. I think that this is a very promising result for a first digital instrument design. To me, the feel of direct playing is amazing – caused by the embodied playing style and heightened by the vibrating controller. The captured video material confirms that Hearspray playing is also "music for the eyes" and should be fun to watch for the audience. Even in the cases where the musical output proved to be boring, the actual playing situation felt motivating. As a player, I felt powerful due to the amount of available controls and the use of automatic processes: With Hearspray one is able to create complex musical patterns by the twist of his hand.

The downside of the instrument is the relatively low accuracy in direct playing. It can be seen as a hindrance making it impossible to create musical cohesion by reproducing the played patterns. However, it can also be seen as a benefit, ensuring that the players will always create new variations – thus encouraging constant musical evolution. For me, Hearspray provides ways for creating at least close approximations of the wished musical patterns, with additional possibility of getting surprised. One could even reach the level of absolutely precise playing by carefully learning the needed muscle positions, but I personally prefer certain amount of unpredictability in playing.

Due to the difficulties in control, it can be said that Hearspray offers at least some of the challenge familiar for the players of the acoustic instruments. However, the challenge of many acoustic instruments relates to getting a sound out of the instrument and to making the sound beautiful. The struggle of controlling an acoustic instrument gets harder when attempting to play rapidly. With Hearspray, the challenge is almost opposite: It is mostly related to taming the quickly and automatically playing instrument in order to play unchanging and repeating patterns.

The decision to make the instrument easily approachable means a trade-off for "serious" playing. For example, the default pentatonic scales get restrictive in the long run and the need to separately stabilize the pitch feels unnecessary. For the performance situations I will probably slightly remap the instrument in order to make it more versatile. This is a defeat in terms of creating an all-purpose instrument, but as the player testing suggested, it may be impossible to create something that is equally suitable for all kinds of players.

The Wii controllers have their limitations as an instrument interface, being rather inaccurate and generic. I have designed the interface to the maximum, utilizing all the available buttons for playing functions. Due to the non-optimal placement of the buttons, some functions are difficult to perform: Especially playing and manipulating recordings while playing in the direct way is somewhat cumbersome. On the other hand, the problems in this case are not only ergonomic: It is cognitively challenging for a human being to keep constant track of parallel real-time processes. Perhaps such a level of multitasking is not even desirable, as this kind of playing is better handled via the screen-based interfaces. Hearspray better suits creating complex music in the moment together with other players, not as a solitary musician.

8. CONCLUSIONS

It is not trivial to answer such a big and open question as “How to design a good musical instrument?”. The problem space had to explored, understood, scoped and divided to tangible sub-problems, including “Who are the players?”, “What musical possibilities the instrument should enable?” and “How much control should it give to the player?”. Only after this could suitable solutions be designed. The use of the chosen design methods – including background research, iterative sketching and player testing – were crucial for the process. As important was to use my own insight gained through exposure to music and instruments in order to evaluate the importance of the conflicting needs. Gradually these approaches have enabled me to go past the paralyzing big question and the problems present in the sketches and first versions of the instrument.

As the project outcome I have successfully designed and developed a digital instrument allowing embodied playing utilizing the Wii controllers. The instrument lets the players to create music collaboratively using large gestures, which should be helpful in bringing the excitement and interplay back to the musical performances. After the collaborative playing sessions it was clear that the instrument feels good to play and allows the creation of interesting musical outcomes. The test players were mostly able to learn the use of the instrument, and it has been received positively. A visual learning interface was needed in order to support the discovery of certain features due to the controllers' generic nature. The goal of offering simplicity and flexible musical possibilities in the same package proved to be challenging to reach, partially due to the interface technology constraints: The Wii controllers do not fully direct the first-time player to the desired interactions and they limit the possibilities of simultaneous parameter control for the expert players. The test players also had different approaches using such an open instrument, hinting to the impossibility of pleasing everybody.

Specifying a large-enough performance space and supporting it with digital automation has given Hearspray many musical possibilities, some of them even unanticipated during the design process. However, Hearspray cannot and does not even try to be an instrument capable of everything: During the design process musical features were prioritized so that certain musical patterns can be played well, some with effort and some are simply unaccessible. The prioritization decisions were based on the envisioned character of the instrument and the capabilities of the Wii controllers – the inherent limitations of the controller technology and the lack of free controller dimensions reduced the possibilities. The player-instrument interaction balances between possessing and losing control, and seems to create an inspiring feeling of having to struggle in order to master the instrument. The instrument's behavior is deterministic and quite straightforward, but due to the nature of the controllers and the automatic processes the player must be precise if he wants to organize sound consciously.

Creating digital musical instruments tends to be tedious as there are no such “quick and dirty” tools such as sketching on paper is for visual interface designers. Depending on the nature of the instrument, realizing even a simple prototype may take hours. The choice of an existing controller technology sped up the development process and allowed focusing efforts to other aspects of the design. However, as noted earlier, generic controllers tend to have limitations in the musical context, as they are not designed with the musical activities in mind.

I started this thesis from a state where I did not have much knowledge about the design of digital instruments, nor had I consciously tried to analyze musical interaction. As the process outcome I have gained vast amounts of new knowledge and skills beyond the needs of this particular project. For me, Hearspray is capable for interesting and prolonged improvisations, and will certainly be played more in the future in varying musical contexts. I also have potential ideas and interest to take the instrument further, as long as the controller allows this to be done in a usable way.

REFERENCES

Publications

Cook, P., 2001. Principles for Designing Computer Music Controllers. In Poupyrev, I., Lyons, M. J., Fels, S., Blaine, T., Eds. Proceedings of the 2001 conference on New interfaces for musical expression (NIME-01). Seattle.

Cook, P., 2003. Remutualizing the Instrument: Co-Design of Synthesis Algorhitms and Controllers. In Bresin, R., Ed. Proceedings of the Stockholm Music Acoustic Conference 2003 (SMAC03). Stockholm, Sweden.

Cooper, A., Reimann, R., Cronin, D., 2007. About Face 3: The Essentials of Interaction Design. USA: Wiley&Sons.

Ferguson, S., 2006. Learning Musical Instrument Skills Through Interactive Sonification. In Schnell, N., Bevilacqua, F., Lyons, M., Tanaka, A., Eds. Proceedings of the 2006 International Conference on New Interfaces For Musical Expression (NIME06). Paris, France.

Gurevich, M. and Treviño, J., 2007. Expression and Its Discontents: Toward an Ecology of Musical Creation. In Crawford, N., Ed. Proceedings of the 2007 Conference on New Interfaces for Musical Expression (NIME 07). New York.

Hunt, A., Wanderley, M., Kirk, R., 2000. Towards a Model for Instrumental Mapping in Expert Musical Interaction. In Proceedings of the 2000 International Computer Music Conference. San Francisco, CA: International Computer Music Association.

Jordà, S., 2005. Digital Lutherie: Crafting musical computer for new musics' performance and improvisation. Thesis, (PhD). Universitat Pompeu Fabra.

Kenny, B. and Gellrich, M., 2002. Improvisation. In Parncutt, R. and McPherson, G., Eds. The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning. New York: Oxford University Press. 117-135.

Magnusson, T. and Hurtado, E., 2007. The Acoustic, the Digital and the Body: A Survey on Musical Instruments. In Crawford, N., Ed. Proceedings of the 2007 Conference on New Interfaces for Musical Expression (NIME 07). New York.

Malloch, J., Birnbaum, D., Sinyor, E. & Wanderley, M., 2006. Towards a New Conceptual Framework for Digital Musical Instruments. In Da Silva, A., Ed. Proceedings of the 9th International Conference on Digital Audio Effects (DAFx-06), Montreal.

Orio, N., Schnell, N., Wanderley, M., 2001. Input Devices for Musical Expression: Borrowing Tools from HCI. In Poupyrev, I., Lyons, M. J., Fels, S., Blaine, T., Eds. Proceedings of the 2001 conference on New interfaces for musical expression (NIME-01). Seattle.

Paradiso, J. A., O'Modhrain, S., 2003. Current Trends in Electronic Music Interfaces. In Journal of New Music Research, 32 (4), 345-349.

Persson, R., 2001. The Subjective World of Performer. In Juslin, P. and Sloboda, J., Eds. Music and Emotion: Theory and Research. New York: Oxford University Press. 275-291.

Pressing, J., 1988. Improvisation: Methods and Models. In Sloboda, J. A., Ed. Generative Processes in Music: The Psychology of Performance, Improvisation and Composition. New York: Oxford University Press. 129-179.

Raskin, J., 2000. The Humane Interface: New Directions for Designing Interactive Systems. USA: Addison-Wesley.

Rokeby, D., 1997. The Construction of Experience: Interface as Content. In Dodsworth, C., Ed. Digital Illusion: Entertaining the Future with High Technology. USA: Addison-Wesley Professional.

Ryan, J., 1991. Some Remarks on Musical Instrument Design at STEIM. In Contemporary Music Review, 6 (1). United Kingdom: Harwood Academic Publishers. 3-17.

Saffer, D., 2008. Designing Gestural Interfaces. Canada: O'Reilly Media.

Sanderson, P., 2003. Cognitive Work Analysis. In Carroll, J., Ed. HCI Models, Theories and Frameworks: Toward a Multidisciplinary Science. USA: Morgan Kaufmann Publishers, 225-265.

Weinberg, G., 2003. Interconnected Musical Networks: Bringing Expression and Thoughtfulness to Collaborative Group Playing. Thesis, (PhD). Massachusetts Institute of Technology.

Wessel, D. and Wright, D., 2001. Problems and Prospects for Intimate Musical Control of Computers. In Poupyrev, I., Lyons, M. J., Fels, S., Blaine, T., Eds. Proceedings of the 2001 conference on New interfaces for musical expression (NIME-01). Seattle.

Online references

Hawley, M., 2009. Design Research Methods for Experience Design. In UXmatters. Available from http://www.uxmatters.com/mt/archives/2009/01/design-research-methods-for-experience-design.php [Accessed 20 April 2009].

IxDa. Interaction Design Association's web page. Available from http://www.ixda.org [Accessed 14 March 2009].

Löwgren, J., 2008. Interaction Design. Interaction-design.org. Available from: http://www.interaction-design.org/encyclopedia/interaction_design.html [Accessed 14 March 2009].

Nintendo, 2009. Wii Music. Available from: http://www.wiimusic.com [Accessed 25 April 2009].

OSCulator, 2009. Wiimote-FAQ. Available from http://www.osculator.net/wp/faq/wiimote-faq [Accessed 18 November 2008].

Paradiso, J. A., 1997. Electronic Music Interfaces. Dr. Joseph A. Paradiso's home page. http://web.media.mit.edu/~joep/SpectrumWeb/SpectrumX.html [Accessed 20 April 2009].

Pentatonic scale. (2009). In Encyclopædia Britannica. Available from: http://www.britannica.com/EBchecked/topic/450385/pentatonic-scale [Accessed 19 April 2009].

Reactable, 2009. In Music Technology Group of Universitat Pompeu Fabra web pages. Available from: http://mtg.upf.es/reactable [Accessed 25.4.2009].

Rokeby, D., 2000. Very Nervous System. In David Rokeby's home page. Available from: http://homepage.mac.com/davidrokeby/vns.html [Accessed 13 October 2004].

Rosenblum, A. E., 2008. Dancing a Song With the Full-Body Wiimote Music Controller Suit. In Popular Science. Available from: http://www.popsci.com/entertainment-gaming/article/2008-02/dancing-song-full-body-wiimote-music-controller-suit [Accessed 10 December 2008].

Shah, S., 2009. Wii Motion Plus: Hands-on Update. Available from: http://www.gamespot.com/features/6194443/index.html [Accessed 22 September 2009].

Waisvisz, M., 2006. The Hands. In Crackle.org. Available from: http://www.crackle.org/TheHands.htm [Accessed 14 November 2008].

Wakin, D. J., 2008. A Computer-Music Man Unplugs. New York Times. Available from: http://www.nytimes.com/2008/08/03/arts/music/03waki.html [Accessed 15 January 2009].

WiiLi, 2009. Motion analysis. Available from: http://www.wiili.com/index.php/Motion_analysis [Accessed 22 September 2009].

Figures

Figure 1	Yamaha. Via http://www.global.yamaha.com/design/tenori-on [Accessed 1 March 2009].
Figure 2	Screen capture from a video via http://www.youtube.com/watch?v=Mh0VX74alwk&feature=related [Accessed 7 May 2009].
Figure 3	Kalle Mäntsälä (2008). Collage representing the use of What You Do Is What You Hear installation.
Figure 4	Kalle Mäntsälä (2009). Basic player-instrument interaction flows.
Figure 5	How To Play a Guitar on The Xbox 360. Via http://www.ucs.louisiana.edu [Accessed 2 April 2009].
Figure 6	Music Technology Group, Pompeu Fabra University (2009). Via http://mtg.upf.es/reactable/?media [Accessed 18 October 2008].
Figure 7	Abydos Enterprises. Via http://www.magictails.com/abydos/images/mid-east/percussion/large/kalimba_small.jpg [Accessed 5 April 2009].
Figure 8	Screen capture from a video via http://www.youtube.com/watch?v=CXLKR3HFA4s&feature=related [Accessed 14 April 2009].
Figure 9	Malloch, J., Birnbaum, D., Sinyor, E. & Wanderley, M. (2006). Towards a New Conceptual Framework for Digital Musical Instruments.
Figure 10	Via http://www.137.com/theremin [Accessed 1 March 2009].
Figure 11	Weinberg, G. (2003). Interconnected Musical Networks: Bringing Expression and Thoughtfulness to Collaborative Group Playing.
Figure 12	Paradiso (1997). Via http://web.media.mit.edu/~joep/TTT.BO/Baton2.gif [Accessed 10 February 2009].
Figure 13	Nintendo. Via https://www.mywii.com.au/ReviewDetail.aspx?id=915 [Accessed 14 April 2009].
Figure 14	Michel Waisvisz. Via http://www.crackle.org/The%20Hands%202.htm [Accessed 11 March 2009].
Figure 15	Tom Tlalim. Via http://www.popsci.com/entertainment-gaming/article/2008-02/dancing-song-full-body-wiimote-music-controller-suit [Accessed 13 February 2009].

Figure 16	David Rokeby. Screen capture from a video via http://homepage.mac.com/davidrokeby/vns.html [Accessed 13 October 2008].
Figure 17	Screenshot via http://iphonenews.desinformado.com/2008/02/iano-turn-your-iphoneipod-touch-into-a-virtual-piano-wvideo-amazing [Accessed 9 April 2009].
Figure 18	Cook, P. (2003). Remutualizing the Instrument: Co-Design of Synthesis Algorithms and Controllers.
Figure 19	Synthmania. Via http://www.synthmania.com/tg500.htm [Accessed 11 April 2009].
Figure 20	Electronic Reviews. Via http://reviewelectronics.blogspot.com/2007_03_01_archive.html [Accessed 2 May 2009].
Figure 21	OSCulator. Via http://www.osculator.net/wp/faq/wiimote-faq [Accessed 2 January 2009].
Figure 22	Kalle Mäntsälä (2008). Screenshot of an early rotation sketch visualization.
Figure 23	Kalle Mäntsälä (2008). Screenshot of an early point&drag sketch interface.
Figure 24	Kalle Mäntsälä (2009). Enhanced photograph of pitch manipulation gesture.
Figure 25	Kalle Mäntsälä (2009). Enhanced photograph of density manipulation gesture.
Figure 26	Kalle Mäntsälä (2009). Example illustrating Hearspray amplitude envelopes.
Figure 27	Kalle Mäntsälä (2009). Example illustrating overlapping grains
Figure 28	Kalle Mäntsälä (2009). Example illustrating the delay.
Figure 29	Kalle Mäntsälä (2009). Screenshot of first version of Hearspray's visual learning interface
Figure 30	Kalle Mäntsälä (2009). Two designs for Hearspray's visual learning wizard.
Figure 31	Kalle Mäntsälä (2009). Explanation of Hearspray's interface controls.
Figure 32	Kalle Mäntsälä (2009). Photograph of collaborative Hearspray playing.
Figure 33	Kalle Mäntsälä (2009). Photographs of collaborative Hearspray playing.
Figure 34	Kalle Mäntsälä (2009). Explanation of Hearspray's technical structure.
Figure 35	Kalle Mäntsälä (2009). Explanation of Hearspray's Synth architecture.

APPENDIX A: TECHNICAL SOLUTIONS

This text offers an overview of the instrument's technical solutions, hopefully aiding future developers as considerable time was spent solving the problems. The complete source code can be downloaded from http://www.luxaeterna.fi/hearspray.

Choosing SuperCollider 3 as audio programming environment

I had previously only used Cycling 74's commercial Max/MSP[11] as audio programming environment. Max/MSP utilizes a visual programming metaphor: The programmer adds rectangular objects - e.g. audio oscillators and filters – to the working area and then drags "wires" for connecting objects together in order to build a functioning application. This seems to be easy to learn for a beginner, but for experienced programmer it tends to pose limitations and often leads to cluttered program files. To my knowledge Max/MSP provides no possibility to utilize object-oriented programming – dynamic run-time creation of several instances from a single model. This feature would make building complex applications much simpler by reducing the amount of code needed.

To avoid these limitations, SuperCollider 3[12] was examined in the initial phase of the project. It is a free open-source audio programming language utilizing the conventional line coding style. Initial learning phase was personally relatively difficult due to the syntax of SuperCollider code, which can look unfamiliar to people experienced with common languages like C, Java or Python. Later on the language and environment proved to be truly versatile. The possibility of object-oriented programming was essential for most instrument sketches and for the final Hearspray instrument.

---

[11] Available from http://www.cycling74.com/products/max5

[12] Available from http://supercollider.sourceforge.net

Connecting Wii controllers to a computer

The communication protocol between the Wii console and the controllers is of closed source, but it has been partially re-engineered, allowing the use of the controllers in a computer-based application. The single missing crucial feature seems to be the use of the Wiimote's speakers for aural feedback. I initially tried using SuperCollider's built-in Wiimote library for connecting the controller, but it seemed to be unstable and did not calculate rotation values from the accelerations automatically. Instead, a Mac OSX shareware application OSCulator[13] is used. It is a tool for audiovisual performance, meant for converting control messages from one format (e.g. Wiimote values, MIDI, mouse location) to another (e.g. OSC[14], MIDI). OSCulator conveniently connects to a Wiimote and sends all the incoming event messages to the desired destination – SuperCollider in my case - using the Open Sound Control protocol. It also calculates the controllers' rotation and total acceleration values automatically from the basic linear accelerations. Unfortunately, using the OSCulator meant adding a non-open commercial (although cheap) component to the instrument. Similar free solutions, like GlovePIE[15], are available for Windows platform.

---

13 Available from http://www.osculator.net/wp

14 Open Sound Control, a protocol for connecting computers, sound synthesizers and other multimedia devices. A faster replacement for MIDI

15 Available from http://carl.kenner.googlepages.com/glovepie

The components of the instrument

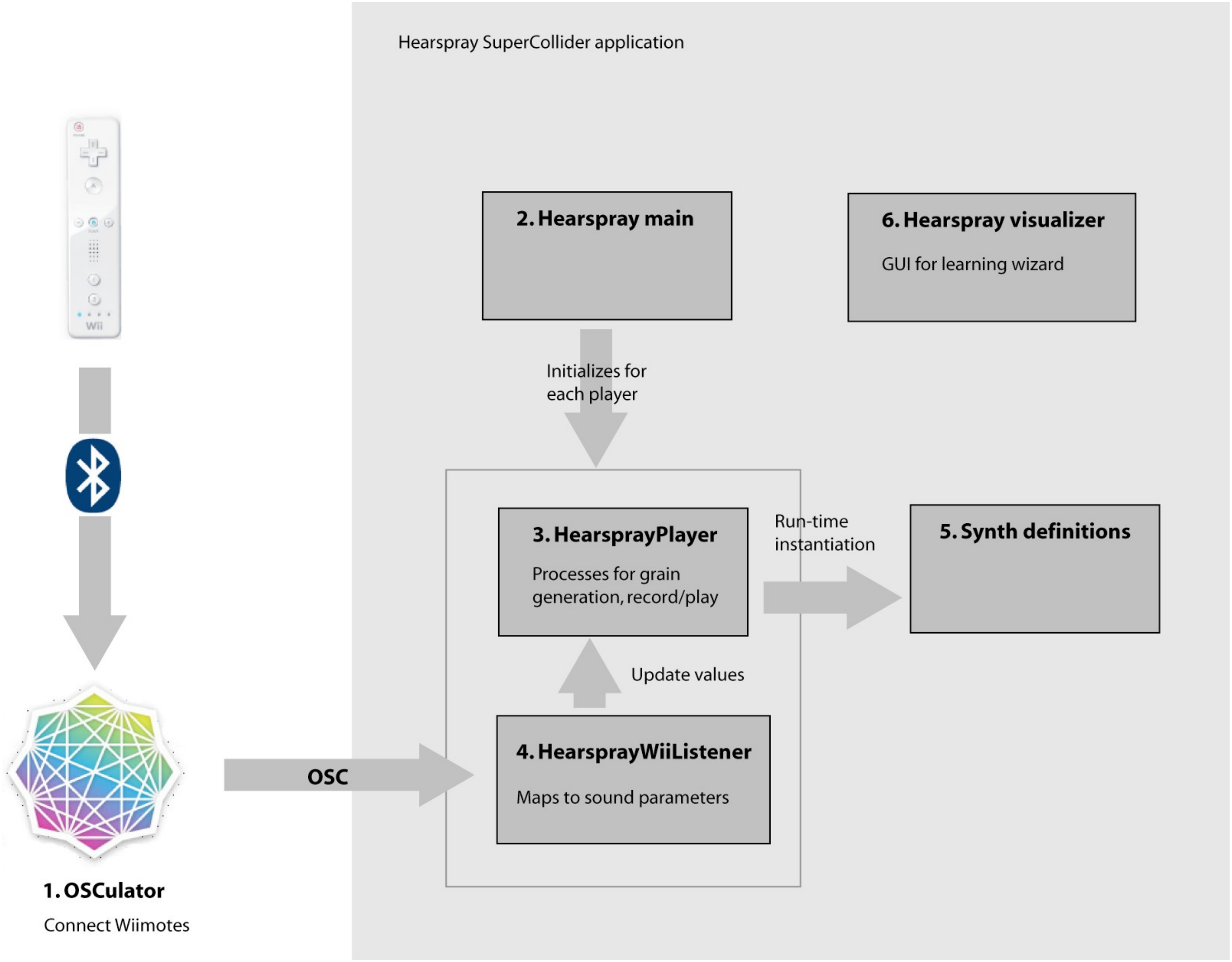


Figure 34: Technological components of Hearspray

1. OSCulator software is used for connecting the Wiimotes to the computer using Bluetooth technology. OSCulator converts all incoming Wiimote events (acceleration values, tilts and button presses) to messages in the Open Sound Control format. The message name follows the pattern `/wii/wiimotenumber/type/event` followed by the value, e.g. `/wii/1/accel/xyz` containing acceleration values from 0 to 1 for each direction and `/wii/2/button/A` containing value 0 or 1 depending on the button state. These messages are sent to the port 57120 on the local computer – the one SuperCollider is listening to. OSCulator also listens to the incoming OSC messages in order to vibrate the Wiimote or to toggle the Wiimote's LED lights.

2. Hearspray main program needs to be executed in order to start the instrument. It contains all the common variables, like the musical scales used. The program starts the SuperCollider audio server and after that creates as many HearsprayPlayer objects as specified in player number parameter. It also creates the final output Synth[16], which is persistent and same for all players, and groupings for all player-related synths in order to ensure proper signal path (see Figure 35).

3. HearsprayPlayer is a code class, which is used to instantiate objects representing each player. The objects contain current values of all player's sound parameters (e.g. frequency, amplitude, panning) and also provide functions for:

- Generating a stable flow of grains by constantly instantiating new Synths.
  The speed of the process depends on density value.
- Starting and Stopping audio recording process
- Starting and Stopping audio record playback process

When a HearsprayPlayer object is created, it creates all the necessary player-related, persistent Synths (for effecting, reverberation, delay line) and creates a child object called HearsprayWiimoteListener.

4. HearsprayWiimoteListener object listens to all incoming Wiimote-related OSC messages, keeping track of the controllers' state. It performs mappings for incoming continuous values and updates the sound parameters in HearsprayPlayer and related Synths. The object also calls HearsprayPlayer functions based on button presses. When vibration or LED light changes are needed, it sends OSC messages back to OSCulator.

5. Synth definitions define the behavior of different Synths contained in the instrument by specifying the used unit generators[17] and default routings of the signal (Figure 26). The definitions are used to instantiate actual Synths when instrument is started and played.

---

[16] In SuperCollider, a Synth represents a single sound producing unit

[17] Unit generators (UGen) are the building blocks of Synths in SuperCollider – e.g. audio oscillators, filters, other effects, analysis tools

6. Hearspray visualizer. The visual learning interface has to be executed separately, as it is normally not displayed in order to save processing power. This code file creates a new full-screen window and the related user interface elements. The window is updated by the HearsprayWiimoteListener, refreshing the interface every time new Wiimote data comes in.

The audio chain of the instrument contains the following Synths:

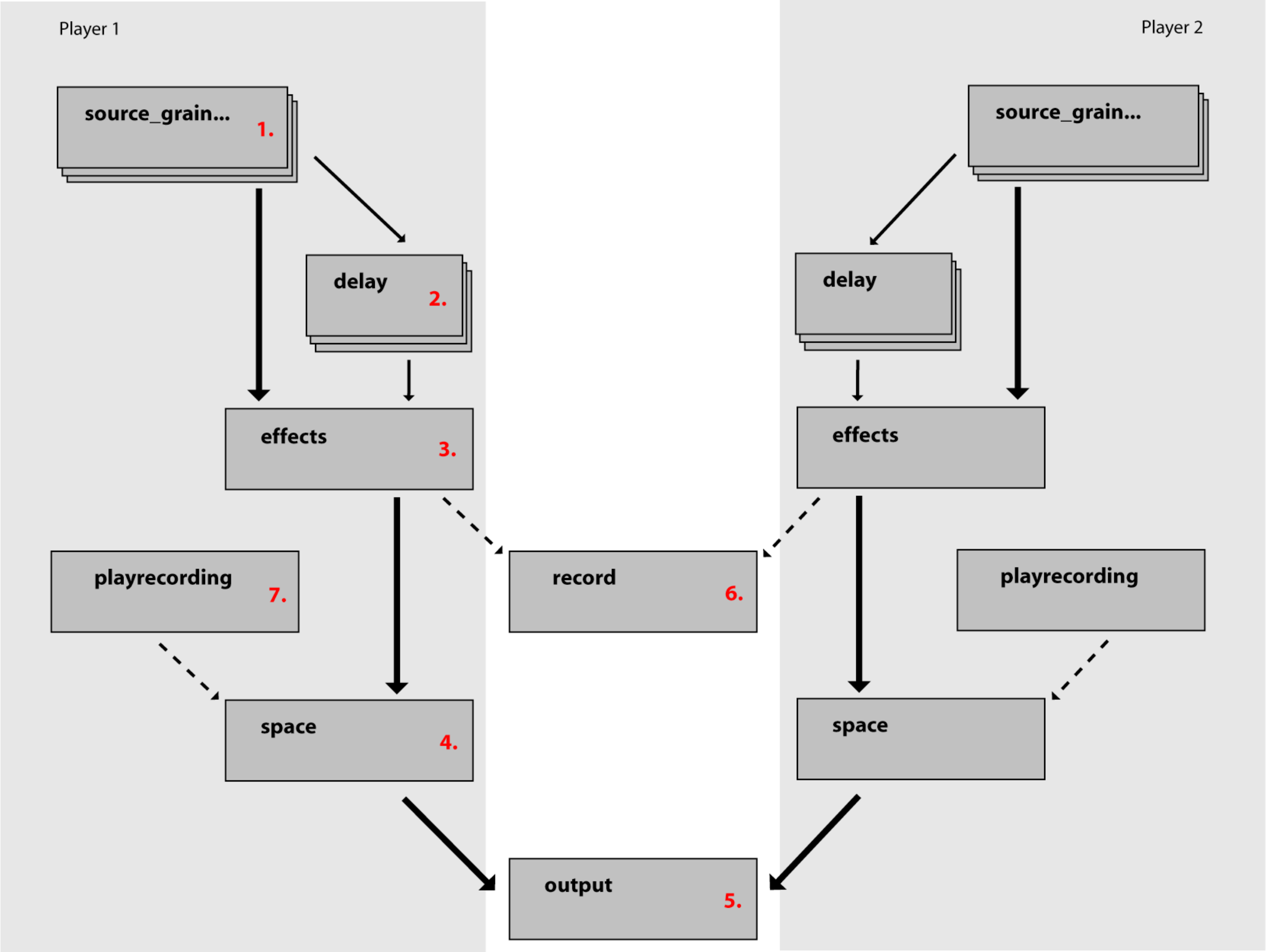


Figure 35: Hearspray Synth architecture

1. Source_grain. One instance of this sound source Synth plays one short sound grain and dies, thus there are usually many of these instantiated. These Synths can be based on three different Synth definitions: First, *source_grain_basic* is utilizing a Blip[18] UGen to generate

[18] Band Limited ImPulse generator

"cold" synthetic sounds. The functioning of this UGen resembles stacking a desired number of sine waves in harmonic relation in order to produce waveforms ranging from a sine wave to a sawtooth wave. Two other types, *source_grain_gendy1* and *source_grain_gendy2* mix certain amount of fuzzy-sounding Gendy1[19] UGen with the Blip in order to create the “natural” timbres. Gendy utilizes randomness and probability distributions in the waveform production. For more information about the used unit generators, see SuperCollider documentation.

In all of the three Synth types frequency, panning, amplitude and amplitude envelope (affected by grain length, attack time and release time) can be controlled from outside the Synth. Attack is always “warmed” by playing a short burst of pink noise in the beginning. Signal also goes through a Klank UGen - a resonating bank of oscillators adding inharmonic overtones to the played sound when the resonation effect is toggled. Resonation's ringing character is caused by the resonation frequencies relating to the grain's frequency.

Due to the amount of simultaneous grains, the aim has been to keep these Synths light. The source signal is next directed either straight to Effects, or to the current Delay line if the player is holding the delay button.

2. Delay. There are always five delay Synths - the player re-connects the source signal to the next one every time he presses Z button. These Synths contain a CombN[20] unit generator, which feeds the output signal back in to its input. Delay's repeat frequency is picked from the Nunchuk rotation every time the player presses the Z button. The effect decays during 12 repeats, making it noticeable but quickly fading in order to avoid a cacophony of echoes.

3. Effects. To make the source Synths lighter to process, most timbral operations are done in this Synth, which is instantiated once per player. The sound is shaped in order to create a more “natural” and lively sound: There are oscillating high-pass and low-pass filters, which affect the timbre over time. Sound is also driven through a wave-shaping operation (by phase modulating a 0hz sine oscillator with the signal), which alters the timbre via overdriving the signal when the amplitude rises. The Synth also contains a folding operation (SuperCollider's fold2 function), which mirrors the parts of waveform that

---

[19] Implementation of Gendy, the dynamic stochastic synthesis generator conceived by Iannis Xenakis

[20] A Comb delay line with no interpolation

exceed a certain amplitude threshold. This is used to distort the signal when the player hits with the Nunchuk controller.

4. Space. Spatial processing (reverberation) is done after the effects Synth, as this allows the recording of only the "dry" signal and further spatial effecting of the playback. SuperCollider's FreeVerb unit generator is used for manipulating the room size and dry / wet signal balance in order to create illusion of sound going further and closer. Initially GVerb was used as reverberation unit generator, because it offers more advanced controls and better sound quality, but it seemed to demand too much processing resources. Panning is already done in source Synths, as that allows having simultaneous grains in different panning positions. Reverberation had to be done separately as it is too heavy an operation to be done independently for each grain.

5. Output. Final synth is used to combine the separate audio chains of players and feed them to a common main output.

6. **Record.** Both players have their own recording Synths, which can take in signal from all players or just from player's own signal path. The signal is stored to player's record buffer. If playback is ongoing, the played back signal is preserved in the recording.

7. **Playrecording.** This synth plays back player's record buffer, allowing variable playback speed. The Synth contains further unit generators for pitch shifting, resonant high-pass and low-pass filtering, panning and delaying to match the sound parameters controlled by the direct playing gestures.