Sound and Music Production for Interactive Media

Hans Lindetorp[1]

[1]**Royal College of Music (KMH) and Royal Institute of Technology (KTH), Stockholm**

ABSTRACT

Interactive media has become one of the most important forms for entertainment and education over the last decades and sound and music play a big role in those formats including music for games and interactive movies. Contrary to the traditional channels for producing and distributing music, interactive media lack standardized formats and concepts for how it should be implemented. Studies have indicated that most music producers find the task quite hard and tend to drop artistic ambitions when the technical obstacles like difficult programming or tedious work are too many. My research aims at finding technical solutions, useful terminology and efficient processes for producing sound and music for interactive media that help the producer to stay focused on the artistic aspects of a production. The strategy is to run an iterative process integrated in the education at the Royal College of Music in Stockholm where we design, prototype and evaluate technology for interactive sound and music. The work so far has resulted in a descriptive language called WebAudioXML (waxml) that makes it possible to realize an artistic vision in an online interactive sound and music applications without the need to learn any programming language. The result from the studies so far has been very promising and indicates a potential for a future standard like waxml. The student feedback also highlights areas for improvements and further development that will be the basis for future work aiming at inclusive technology accessible for everyone.

Contents

1 Introduction

Interactive media has become one of the most important forms for entertainment and education over the last decades and sound and music play a big role in those formats including music for games [11] and interactive movies. Even if the technology has matured over the years from the 8-bit sound in early game consoles and computer sound cards into high quality, multi-channel, surround audio, there are still room for improvements of strategies, tools, processes, and formats regarding the making and implementation of sound design and music for interactive media [37]. This thesis aims at exploring, evaluating, and contributing to this field with a descriptive coding language, a working prototype technology for integrating sound and music in interactive applications and also with the result from evaluating more than 150 student projects using the technology. In this first section I frame the field of interest and present some core values that are important drivers for my research.

1.1 Music in Interactive Media

Humans relation to music is inherently interactive. Early cultures have used music for parties, dance, religious worship, and as a part of storytelling for thousands of years. In those settings, the musicians learnt to play and interact with whatever else was going on and to adjust the music according to the circumstance. The music notation introduced a new, fixed format for music, the printing press spread it, and the audio recording innovation increased that trend even further leading to the way many people consume music nowadays; pre-composed, performed, recorded, and produced music through streaming services. In contrast to when making music was a part of being together, music has now become a product, a fetish [2]. The older, physical forms of distributed, produced music, like the record, the cassette, and the CD was, by their material nature, static media without any other interactive access points than the possibility to repeat or skip a song and to change the volume. For digital, interactive media there is a potential for a renaissance for music and interactivity even if the music is pre-produced and not performed live. This has been explored and used primarily in computer games during the last decades and is referred to as “adaptive music”, “interactive music” or “dynamic music” and even if the term varies, it often refers to musical building blocks that is arranged and played back in real-time controlled by data from a hosting application [48]. This is the form of music my research is about.

1.2 Music Production Education

At the Royal College of Music (KMH) http://www.kmh.se, we created a new artistic bachelor in music production 2001 followed by a masters education a few years later. The aim was to create an education for music producers that included the new trends that was happening in the field of digital media at that time. We identified the need for a music producer with skills in different disciplines including song writing, arranging, conducting, recording, mixing, mastering, distribution, and marketing. We also included all types of music production into the scope including traditional songs as well as music for film, computer games, web pages, and interactive exhibitions. We found that the tools and processes for teaching and learning how to produce normal music tracks were well established and we could rely on traditional music theory for composition and arranging. We could use commercial Digital Audio Workstations (DAW:s) for recording, mixing

and mastering and point towards record companies and later streaming services and social media for distribution of the music. When it came to teaching music production for interactive media the conditions were quite the opposite. The role for the producer of interactive music is complex [37], the DAW:s are not well suited for producing the flexible format, there are no standards or well established commercial tools for integrating music into interactive application and it is hard to find ways to distribute the result. These are the type of challenges I try to describe and solve through my research.

1.3 Challenges regarding artists and technology

During the two decades of teaching music production for interactive media, one challenge has captured my attention and that is the necessity of programming skills and that almost none of our students have any prior experience of any programming language at all. This correlates with other studies [52, 51] and creates a high barrier for the students to enter into the field. Our ambition with the education has always been to let the students experience as much as possible of the process of making all the parts in a project which has caused us to continuously look out for new and better ways of making interactive applications with music. This aim led step by step to the work described in this thesis.

1.4 My perspectives on humans, music and technology

I believe every human action, including formulating a research question or developing a piece of technology is driven by values and I want to be as transparent as possible with my values regarding music, technology, and human beings before I describe my work. I view music as a human activity and an important part for us to form, shape and celebrate our collective values and belongings. I also view the relationship between music and technology is an ongoing process where humans always have sought for new ways of making music from the first found musical instrument to the latest algorithms and machine learning applications. I don't perceive a definite line dividing "organic" music from "computer" music and rather argue for a continuum where some music is more influenced by human beings and other is more influenced by technology. Even if I find computer generated music interesting, my passion is to contribute with knowledge and technology that supports traditional human music expression rather than making tools for new aesthetics and forms of music. This is the reason for me to search for solutions that solves the challenges and support the artistic ambitions of a composer or a music producer rather than making tools that helps a technically drive person to make music. This focus is also driven by my perspective that the musical skills a composer has achieved are a part the expression that makes the music interesting and valuable while technical skills aim at solving technical problems that rather should be solved by the technology itself if it is possible. Another important value for me is that technology should aim at including people regardless of their physical or mental abilities, financial resources, ethnic or cultural background as expressed by "Universal Design" [12]. This encourages me to evaluate my research output from a critical perspective on inclusivity and accessibility.

1.5 Inspiring technologies

There are some technologies that more than other have grasped my attention and impacted the way I think and act as an artist, developer, and researcher. They are all contributing to the values I've presented in various ways and in the following section I present how they have informed and impacted my research projects.

MIDI

Musical Instrument Digital Interface (MIDI) [4] has since 1983 has arguably been one of the most important standards for music technology so far. Initially aiming at connecting two keyboard instruments it has paved the way for countless of music interfaces, recording and editing software for music production. Even if the specification has become outdated when looking at performance or capacity, it has survived almost 40 years due to the widespread support and implementation. It has even reached its original musical scope by acting as the protocol for light mixers and other non-musical equipment. MIDI is good example of how a well-supported standard can open up for a whole new field of products and creativity.

Spreadsheet

Of all applications developed so far, one of my favorites is the digital spreadsheet made popular through e.g. Microsoft Excel. The concept that allows the user to write small snippets of logical expressions where the value of different cells is computed makes it possible for people without programming skills to create spreadsheets where complex calculations of a lot of data relatively easy can be displayed as visual diagrams. This has contributed in many ways to support researchers and accountants alike by reducing tedious work and provides accessible statistical results that could lead to better conclusions and decisions. In interactive applications, the sound and music output is a result of interaction data which makes it a subcategory of the sonification field. The mapping of data to an audio representation has some similarities with the logical expressions in the spreadsheet cells. The strategy found in spreadsheet applications, to reduce the need for programming down to small logical expressions, has been a great inspiration for my work and is a core part of how I think of parameter mapping in sonification.

HTML

Hypertext Markup Language (HTML), first specified by Sir Tim Berners Lee in 1991 [21] published as HTML2.0 in 1995, is the foundation for all web pages ranging from simple, static and private web pages to global social media platforms with billions of users. It is arguably one of the most important technologies behind the information revolution as we know it today and is now fostered and further developed by a consortium build upon five "Host institutions" and a network of supporting companies, institutions, and individuals [7]. HTML and relating standards in web technologies are built upon Open Standard Principals [45] that aims at "improving the way people around the world develop new technologies and innovate for humanity" [8]. HTML is a semantic language for describing the content of a web page in a way that can be understood by humans and computers alike and makes it possible to build interactive web applications together with cascading stylesheets (CSS) and javascript. The Web technology has resulted in endless numbers of frameworks, platforms and tools that simplify the development and

there are a lot of open-source projects available online for instant integration of any web page. The values behind web technologies and the open-source community for sharing components are very much in line with my own values and serve as guidelines for my design strategies and is an important reason why I have chosen web technology as the platform for all my development work.

Web Audio API

HTML was originally built for text and images, and it would take some years before HTML5 with support for audio and video became a standard in 2014. This was a great improvement for audiovisual content, but the standard did not have any features for interactive audio. In 2011, Chris Rogers from Google presented "Web Audio API", that became a "W3C Recommendation" [1] in 2021 which means that all browsers on all platforms are now recommended to support the technology. Web Audio API provides a system for controlling audio in a web page allowing for a developer to build anything from a simple audio player to a full featured DAW using web technologies. While Web Audio API has been used in various educational, commercial and research projects, there is potentially still a huge unexplored area for audio applications using the technology. As a result of Web Audio API now being a W3C recommendation, this also mean that anything built using the technology will run on any device using any web browser anywhere in the world. This accessible aspect of Web Audio API is the main reason why all my projects are built using the technology.

1.6 My aim

Because of the potential for building easy distributed, interactive audio applications with Web Audio API on one hand and the challenges for music producers to enter into developing interactive content on the other, I have devoted my pedagogical and research efforts trying to build bridges for this group of students into the world of making interactive audio applications. I use an explorative research-through-design-based approach aiming at finding useful terminology, strategies and solutions that support the creation and implementation of the students' artistic visions. The research output is both a javascript framework with an XML language for Web Audio API (WebAudioXML) and evaluations and reflections from hundreds of students using the technology. The group that probably will benefit the most from the technology itself is sound designers and music producers learning, exploring, and prototyping interactive ideas while the research output is intended for future designers of specifications and technology for interactive sound and music.

In the following sections I present related work, a basic description of the current state of the project (WebAudioXML), describe the various methods I use to collect and evaluate data, summarize the findings from the four included papers and discuss the result. Finally, I present my plans for the final years of my PhD-journey leading up the dissertation.

2 Background

In this section, I describe how we integrate education and research at The Royal College of Music (KMH) and The Royal Institute of Technology (KTH) in Stockholm and summarize the experience from different projects that have formed an important platform

and preparation for WebAudioXML (waxml) which is the platform and core technology for my research project. I also briefly discuss projects and technologies that in various ways is similar to or relate to waxml.

2.1 Education and research

At KMH and KTH, we normally teach sound and music computing and production using a problem-based, learner-centered approach [46]. We teach the basics in a specific technology through seminars and workshops and then let the students run supervised projects where they explore ideas and solutions to achieve artistic or design goals. The practical work is later reported and evaluated by the students and the texts are assessed against the course's learning outcomes. The courses are highly integrated with our research and the students typically both benefit from learning from our last findings, trying new prototypes and contributes with new data for our next study as a part of a course. We continuously seek to improve our courses and there is also an ongoing evaluation of our methods for teaching sound and music computing at KMH and KTH [20].

The curricula for higher education in Sweden always include learning objectives relating to ethical values, inclusion, accessibility, and equality. This is an important factor for us to always consider when we choose platforms and technology and has caused us to abandon prior tools for new web-based solutions built led by Statements from the W3C Mission [8] and Open Standards [45].

We also seek technical solutions that help the students to focus on the core learning objective without being distracted by irrelevant technical obstacles and highly value a collaborative learning environment were lecturers and students easily can share their work for interaction, supervision, and assessment. All these aspects work together for us to invest in exploring the possibilities for using Web Audio as the platform for teaching and learning sound and music computing and production.

2.2 Past experiences

We have run courses focusing on "interactive sound and music applications" for more than twenty years at our campuses and have explored different platforms and technologies like Pure data [35], SuperCollider [30], Bela [32], Adobe Director [49], and Adobe Flash [50] in order to serve our needs. Since 2014, we have also experimented with Web Audio, and after more than a hundred student projects and various prototype technologies we eventually arrived at waxml 2020. Many student project have had an impact in one way or another by informing the development and syntax. The most typical process is that a student group expresses a need for a feature that would make their artistic vision possible, I prototype a solution, they test it, and we evaluate the result. After this iteration, the next project benefits from features implemented by the previous and this become over time the driving factor for the development. Below is a summary of the most important projects and a description of how they have contributed to the advancement of waxml.

Adaptive music in museum exhibitions

The Nobel Price Museum invited KMH to be a part of the "Nobel Creation" exhibition over a period of three years. Students from several different artistic universities in

Stockholm contributed with their interpretations of the Nobel Prize Winners for each year respectively [19]. Fourteen master students from the Music Production education composed thousands of audio files that were triggered, mixed and played-back through a multi-channel speaker setup controlled by the visitors' actions and movements. The framework used for the audio playback was iMusic [26] – a predecessor to waxml– and proved Web Audio to be a robust platform for complex audio installations running over long time periods.

Multi-sensory music installation

The Swedish Museum of Performing Arts includes a large-scale, multi-sensory, accessible, digital instrument called the "Sound Forest" [5, 34] allowing the visitors to play with light-emitting strings attached between the ceiling and the floor. Two groups of master students from the Music Production education has composed interactive music for this instrument and our studies have contributed with important insights on how to compose effectively for a multi-sensory experience like this [15, 14].

Audience-controlled music in super-surround

A group of bachelor students got the challenge to compose music for a multi-channel super-surround speaker setup called "Klangkupolen" [16] at KMH. The music was responding to actions on the visitors' smartphones where touch events and the accelerometer were used to trig audio files and change audio parameters. The setup was built with a Node.js server [33], socket.IO [3], and Web Audio technology, that proved to be a useful platform for multi-user music interactions. The study also pointed out the barrier for some students to learn javascript which caused them to abandon their artistic ambitions [25]. The insight from this study was an important incitement for initiating the development of waxml.

2.3 Related work

The introduction of Web Audio API has led to may research, educational and commercial initiatives and web audio applications are now available ranging from small demos [18] to experimental interactive audio platforms [40] and fully functional DAW:s [42]. There are several javascript abstractions of Web Audio API aiming at simplifying the development of web audio applications e.g., WAAX [9] and Tone.js [29], graphical abstractions like JSPatcher [39] and Quint.js [6], and online coding environments like BRAID [43] and EarSketch [28]. Waxmldiffers from all of them while it is a text-based, descriptive language that is interpreted by a javascript parser and share more similarities the X3D Audio API XML abstraction [23]. There are also several attempts aiming at bringing the traditional audio programming environments like CSound [44], Max/MSP [36], Pure Data [35], and SuperCollider [30] to the web [10, 31, 41, 24] and since the support for the AudioWorkletProcessor has grown over the last years, the possibility to port native applications into WebAssembly [47] looks very promising.

3 WebAudioXML

WebAudioXML(waxml) is an XML language and a javascript parser for configuring and implementing interactive audio components into web pages. It is an abstraction of the Web Audio API that uses XML syntax to describe an audio graph, and it adds features for

mixing, chaining, and splitting audio signals. The syntax also offers solutions for routing and mapping variables from a hosting web page to any parameter in the audio graph. Waxml is a descriptive language and benefits from the readability of the XML format compared to other scripting languages. The structure shares similarities with graphical environments like Max/MSP and Pure Data but is purely text-based like SuperCollider.

3.1 Audio connections

The XML syntax in waxml offers elements that emulates the available Audio Nodes in Web Audio API, like `<OscillatorNode>`, `<BiquadFilterNode>`, and `<AudioWorkletNode>` where the properties of the Web Audio nodes are represented by XML attributes. There are also elements representing combination of Audio Nodes like the `<Voice>` element to build polyphonic synthesizers, the `<Envelope>` element to control audio parameters over time, the `<ObjectBasedAudio>` element that controls both buffering and playback of audio files, binaural panning, and Convolution reverb, and the `<AmbientAudio>` element that controls buffering, playback of stereo or multi-channel files, and looping with crossfades to cover the loop points. There is also a well-defined structure for how audio nodes are connected to each other, either by general rules inside `<Mixer>` or `<Chain>` elements or by setting the routing individually for a node.

3.2 Variable mapping

A core part of an interactive audio application is to map interaction data to different audio parameters. Waxml does this by using a `<var>` element that has attributes to specify input and output ranges, interpolation curves, value patterns, and conversion between different value domains like MIDI note numbers to frequency. The attribute values can be a fixed value, an expression containing other variables or a reference to an external file containing a fixed dataset or a javascript function that returns a continuously updated dataset.

3.3 Events

Another common aspect of an interactive audio application is the possibility to connect user and data driven events to trigger a sound. Waxml handles and distributes those events with the `<Envelope>` element that can be triggered by both user interactions or when a logical condition for variables is met.

3.4 HTML integration

One important feature that is the integration of waxml into a web page and its graphical interface. For this to work, the waxml parser needs to be a linked into the hosting web page. There is a custom HTML element called that can be used as a one or two dimensional slider but any HTML element can also be used as triggers or continuous controllers of variables in waxmlÀll these connections are done through HTML attributes only and don't require any javascript coding.

4 Methods

I do my research in the interdisciplinary field of music and technology. This intrinsically puts me in a position where I need to position myself relating to art, design and research.

Before I explain the different methods I use, it is important to understand that I look at my work as research THROUGH design FOR the art [13]. My aim is driven by artistic visions to contribute with better tools for interactive sound and music artists. In the following section I will describe the various methods I use repeatedly in my different studies. They all contribute to gain knowledge and find solutions to the challenges described in section 1.

4.1 Research Through Design

Research through design was first introduced by Frayling 1993 and there is an ongoing discourse about what it is, how it can prove itself valid and how it can generate new knowledge and theories [38]. I take a traditional view on design and research and view my design practice as a method for knowledge as the research output [27]. My general strategy is sequential [38, p. 24] where an iterative process of design leading to theory leading to new design directs the research plan. The work is rather exploratory than driven by a well-defined question which allows for a certain amount of unpredictability when it comes to define the next study. They are all connected and inform the design of the next project. Another aspect of research through design that is important is to what extent the output is valid and relevant. I aspire for a rather pragmatic approach by navigating in the tension between the particular and the general [22] so the findings are valid outside my own practice but not so generalized that I lose focus on the challenges and problems I am aiming to solve.

One important method I use a lot in order to make the strategies build into a specific design more explicit and generalizable is the "Annotated Portfolio" [17]. One of the most appealing features of annotated portfolios is that it brings the internal and specific knowledge of the designer into view and makes it available for further research to be built upon. The implicit knowledge can reflect many different concerns and annotated portfolio can therefore cover very different aspects depending on the research question including the functionality of the design, aesthetics, practicalities of the production, motivation for the making, identities and capabilities of the potential users and socio-political concerns among other aspects.

All design work happens within a community of shared values and traditions that can be defined as a "program". To be able to interpret and generalize knowledge produced in design practice it's important to make those values and traditions build into your program explicit [38, ch. 5] which makes it important to read the findings from my own annotations in the light of the values and culture described in section 1 and 2.

4.2 Case Studies

All the evaluations of my technology are case studies with one or several student groups. These groups are varied in size but typically involve 5–15 students aged between 20–30 years evenly divided between female and male with a slight overweight towards male participants due to the representation at our different programs. I use different ways of collecting data to evaluate the participants experience of using the technology. The main ones are focus groups, online forms, and text analysis. The focus group is normally a recorded session where the participants reflect on a process, often after presenting a project where the technology is used. Some of the sessions are structured with predefined questions and others are more open.

The students always write a self-reflective text as a compulsory task after a project and these texts serve as valuable source for research. They typically express themselves relatively freely but guided by some general questions relating to their work to cover their artistic ideas, the process, the result, ideas for future projects and the experience regarding their creativity and artistic vision. The texts are analyzed to collect and group quotes, perspectives, and tendencies in the answers.

Finally, the focus group sessions and the self-reflective texts are often accompanied with an online form where I collect data about the students´ earlier experience of programming, music production and other aspects related the current study. I prefer using Likert-type scales to build up a body of data more easily over the years that will be interesting for further research. The forms also contribute with a better understanding of the students' background and help interpreting the answers I get from the focus groups and the analysis of the self-reflective texts.

4.3 Quantitative methods

Quantitative data has a strong attraction when working with computers. The data is both relatively easy to collect, measure, process and evaluate. On the other hand, art and creativity is arguably difficult to measure. This tension has caught my attention and has been an inspiration for building a statistical tool for measuring the work done with the technology. The feature is in an early stage but has been used in two studies described in section 5.4.

5 Results from the included studies

In this section I present the main results of the four included papers. The first three are peer-reviewed and published through two different conferences and the fourth is submitted to a journal. Paper 1 and 2 focus on the syntax of waxml and use feedback from a small number of sound and music computing experts. Paper 3 and 4 are case studies with one or several student groups using waxml as one component in a web-based application with which they created various interactive audio applications.

5.1 WebAudioXML: Proposing a new standard for structuring web audio

(Presented and published at the 17th Sound and Music Computing Conference 2020)

This paper is the first presentation and evaluation of waxml. I conducted three workshops with experts from the sound and music, electroacoustic composition, and the music production field. The participants reported waxml to have a potential in keeping focus on the creative process instead of web development and made us bold to argue that an XML standard for Web Audio configurations would be beneficial for modular and collaborative development and therefore recommended a wider discussion on the topic.

5.2 Audio Parameter Mapping Made Explicit Using WebAudioXML

(Presented and published at the 18th Sound and Music Computing Conference 2021)

The second paper about waxml covered the phase of introducing advanced parameter mapping, data interpolation and value conversion to the syntax in order to make it

more accessible and easier to assess. Three student projects were used as specification requirements for the syntax definition. We used an annotated portfolio and video recorded interviews with experts from the sound and music computing community, to uncover important insights from the project. The participants contributed with critical feedback and questions that helped us to better understand the strengths and weaknesses with the proposed syntax and we concluded that the technology is robust and useful and presented new ideas that emerged from the study.

5.3 Sonification For Everyone Everywhere

(Presented and published at the 26th International Conference on Auditory Display 2021)

This study evaluates WebAudioXML Sonification Toolkit (WAST), a web application for sonification that was built upon waxml. The study contributed to the ICAD community by reaching out to new groups who had not yet considered themselves to be part of the community. I designed, build, and evaluated the toolkit by analysing ten student projects using it. We concluded that WAST did meet our expectations and that it led to students taking a deep approach to learning and successfully contributed to reaching the learning outcomes. The result indicated that WAST was both easy-to-use, highly accessible, extensively flexible and offered possibilities to share the sonification in any device's web browser simply through a web link, and without installations. The study also pointed out some potential features that would be beneficial for the sonification process often found in DAW:s.

5.4 Web Audio for teaching, accessibility and distribution

(Submitted to Journal of Audio Engineering Society March 2022)

This article covers a study of student groups from both KMH and KTH where the goal was to build interactive web audio applications with waxml. The data was collected at two stages of the development of waxml and for the second stage I also built an online, gesture controlled, tool for coding, testing, and sharing audio projects. The study evaluated how the different components in the online tool contributed to the learning experience. The result indicated that waxml is both relatively easy to understand for music students without programming experiences and positively affects learning for programming students. The online tool proved to reduce the time for getting started to less than an hour and the connection between gestures and sound was an important part of making the learning experience positive.

6 Discussion

In this section I discuss the findings and my experience from the studies described in the four articles. I summarize the result but also question to what degree it can be viewed as valid, I will critically examine the methods and comment on the general trends from the first to the last paper.

6.1 What does the result indicate?

The result from the first two papers is very much based on interviews with a few experts from the field of sound and music computing and production. The feedback from the

informants plus my own evaluations and annotations both confirmed that the concept with an XML language was easy to understand for developers with prior experience with other programming environments for audio. The experts also though it contributed to the field with an interesting perspective on audio configuration and parameter mapping. But can we trust the result? I would say partly yes. To the degree that no-one was lying, I think we can assume that the response I got was valid, but it is important to recognize the potential for everyone involved to want the project to succeed and therefor answer more positively than they otherwise would. The risk of this positive-expectation-factor is likely even bigger when interpreting my annotated portfolio. I would not make a design choice if I did not believe in it, and when I have invested time in the solution, I want it to work. I think the most important parts of the result from these two studies are firstly the proof-of-concept report at both stages; the syntax and the framework matched the needs and solved the problems. Furthermore, both studies contributed with valuable input from the informants and guided the design process in very successful ways.

The third and fourth articles both cover case studies with one or several student groups building various interactive audio applications using web based applications for the purpose. Paper 3 is a bit special while the participants neither had any prior sound, music, or programming training. This really put waxmlto the test and as researchers we were impressed by the ease with which the students took the challenge and managed to create and distribute a custom sonification of statistical without much training. In paper 4 we took the challenge of accessibility even further and built an online interface where all parts of the creative workflow could reside in one web page. In this case it was also obvious that the fact that it was possible to connect hand gestures to make sounds was in itself more interesting for the students than the ability to build a synthesizer in an online text editor. As pedagogues and researchers, we were quite astound by the fact that all students were able to start program their instrument using waxml after less than an hour but interestingly enough, the students more or less took the short start-up-time for granted and immediately focused on the possibilities with the tool. This is an important realization for me as an inventor and designer; As soon as we have solved a problem, the solution is taken for granted as if it was never there. This is quite natural when we think about it. I never consider all the perfect conditions for my car to start when I press the button or how my smartphone use the latest face recognition technologies when I log in to my bank. I just expect it to work. This is important to remember when we evaluate our inventions, though. To what degree does the user experience feedback relate to the problem we aimed to solve when they don't experience it? Maybe the most important part of the user experience feedback is pointing forward for the next design phase. Measuring success-rate for the current state would require a careful comparison with earlier versions, both regarding time, quality and user experience. Even if that might be feasible, it would feel unethical to give half of the group an older version with less potential while their primer interest is in making the most interesting music and not to generate research data.

6.2 How can the methods improve?

I have put quite a lot of time writing comments to myself everyday along the way of designing and building waxml and even if there are many annotations that carry interesting perspectives on my design choices and implicit theory built into the technology, I

still consider it at bit hard to treat them as valid research results. This is definitely an area for growth during my last years and I can see a potential for the annotated portfolio if I find a way of categorizing and analyzing my comments in a more formalized way.

The group discussions have generally contributed with valuable insights and ideas. I also benefit a lot from the fact that the students participate voluntarily but are already gathered because the session is already in their schedule. This means that I have access to a comparably big corpus of data but also that there is a power imbalance between me as both a researcher and lecturer and them as students. There might be ways of getting even more interesting output from the sessions if I would not participate but it would also reduce the possibility to ask follow-up questions that would be of interest for me.

The quantitative methods are both based on the online form and a statistical feature in waxml that can report the number of elements and mappings and while I consider both having a great potential, they also need to be further developed. The online form will probably generate a relatively useful data after several years of use, but my biggest worry is that I in the end find a question that I would have liked to have present from the first group. The statistical feature in waxml is interesting as it tries to instantly evaluate the work. I see a potential for this type of assessment to be used in a Learning Management System where the data could be automatically connected with a user account and stored for each session. This would give a researcher or teacher valuable information about how much time the student has spent working with the tool and what he or she created. On the other hand, I also see a big risk when trying to measure and assess artistic work. Is there any valid relation between a project being complex and that it is artistically interesting? Could a seemingly simple configuration carry a brilliant solution that is more creative than one that contains multiple copies of template content? Maybe the answer is that we cannot measure art with numbers that easy but that features like this statistical tool contributes with assessments to see if the student has completed small and well-defined task that could serve well as a preparation for more explorative and artistic projects.

6.3 My own observations

At the time when the study for paper 1 was performed, I had just done some pre-studies to make some proof-of-concept tests with a student group and compared it to the year before when they configured all audio and music playback using javascript commands. The result looked very promising regarding what became possible for the students to implement in a project. It's hard to make a fair description of the difference but I will try to describe it with two typical projects, one before and one after I introduced waxml. Before the students had access to waxml they generally settled with triggering different audio files and maybe mapping an interaction parameter to mix the volume between different sounds. In the latest study when waxml was available in the online editor application, most students managed to map gesture data with fingers control the pitch snapping to a musical scale, the frequency of bandpass filters and even the 3D-position of an audio file. While this is not necessary a measure of artistic quality it indicates a level of playfulness with the technology.

Another observation is that Web Audio is very stable and well supported. We have got rid of almost all installation, authorization, file format and distribution channels by using web technology and it has really improved the learning outcome for the students

compared to earlier tools and platforms. Still, the students with no prior programming skills and no interest in learning any text-based language find XML to be a barrier. The main complaint is about the secrete nature of a text-based language where "everything" is possible but "nothing" is exposing itself. The fact that text-based languages are "unforgiving" in the sense that it breaks as soon as a single character is wrong is also often perceived as a disadvantage compared to graphical interfaces.

My last observation is how much a tool (musical instruments and programming code alike) impacts the artistic result. Whatever captures the attention of the user will be explored and used which causes others to do similar things. This have baring on how a tool is presented, what the demos and tutorials introduces and what template files contain. These will become the platform and starting point for the user shaping the expectation of what is possible. As the developer, I might have a vivid perspective on my innovation, and I might present it as "anything is possible" but "anything" would translate to "nothing else but what I've seen" for the first time user. This critical perspective also has an important, ethical, and cultural aspect to it: The technology we create will come with affordances and constraints and these will both be real through technical possibilities and limitations but also perceived depending on the presentation. In any case these qualities of the technology will have a huge impact on the music made by it and will be a part of shaping the sound of the music we will hear tomorrow. This insight encourages me to always have a critical perspective on my inventions regarding inclusivity and accessibility.

7 Future plans

There are still some improvements left to do with waxml before I move on. One of the most interesting steps will be to integrate my other interactive music framework – iMusicXML – into waxml. Apart from the potential features that would benefit the artistic output, it will also raise some very interesting questions about the relationship, similarities and differences between a composition and a mix, a phrase and an envelope, musical rhythm, and interaction events and so forth. This will also open for an area where it will be useful with generalized terminology for the different objects and their hierarchical relationships in music for interactive media.

There is also an unmet need from all types of music that is not quantized and symmetric to better fit in to interactive applications. Most loop-based systems today lean heavily on rigid timing systems to manage the live arrangement of the musical building block required to make the music flexible. This works for some types of music but excludes unquantized, quirky and human performed music and all traditions of music that is not produced and edited in a computer. My hypothesis is that machine learning could be of great help for the producer to solve those challenges and this is one of the main focuses over the next year.

One of the most valuable formats for my research so far has been to collaborate with colleagues and students and to dream, envision, design, prototype, and test new ideas together. I plan to keep on running more of these artistic and research projects and to use every opportunity to collect data and evaluate the result repeatedly. Among all the ideas I see around the corner is a version 2.0 of the Sound Forest (see section 2.2) and the SonifyFolk project where an updated version of WebAudioXML Sonification Toolkit

operates as the engine to sonify dance movements. I'm also collaborating with KTH and an Art School in Stockholm to create musical instruments for people with physical and cognitive disabilities and would like to see my technology being used in more of these kinds of projects that are accessible for all people regardless of their age, size, ability or disability.

ACKNOWLEDGEMENTS

A big thank you to my main supervisor, Kjetil Falkenberg for giving me lots of valuable guidance and to Royal College of Music for believing in me and the importance of my project.

REFERENCES

[1] Paul Adenot and Hongchan Choi. *Web Audio API: W3C Recommendation, 17 June 2021*. accessed April 5, 2022. URL: https://www.w3.org/TR/webaudio/.

[2] Theodor W Adorno. "On the fetish-character in music and the regression of listening". In: *Popular Music: Critical Concepts in Media and Cultural Studies* 3 (1938), pp. 325–49.

[3] Damien Arrachequesne. *Socket.IO - Bidirectional and low-latency communication for every platform*. accessed April 5, 2022. URL: https://socket.io/.

[4] MIDI Association. *The History Of MIDI*. accessed April 5, 2022. URL: https://www.midi.org/midi-articles/the-history-of-midi.

[5] Roberto Bresin et al. "Sound Forest/Ljudskogen: A Large-Scale String-Based Interactive Musical Instrument". In: *Proceedings of the Sound and Music Computing Conference (SMC)*. 2016, pp. 79–84. DOI: 10.5281/zenodo.851191.

[6] Ian G. Burleigh and Thilo Schaller. "Quint.js: A JavaScript library for teaching music technology to fine arts students". In: *Proceedings of the International Web Audio Conference*. WAC '15. Paris, France: IRCAM, 2015.

[7] Tim Burners-Lee et al. *W3C - About*. accessed April 5, 2022. URL: https://www.w3.org/Consortium/facts.html.

[8] Tim Burners-Lee et al. *W3C Mission*. accessed April 5, 2022. URL: https://www.w3.org/Consortium/mission.

[9] Hongchan Choi and Jonathan Berger. "WAAX: Web Audio API eXtension". In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. Daejeon, Republic of Korea: Graduate School of Culture Technology, KAIST, 2013, pp. 499–502. DOI: 10.5281/zenodo.1178494.

[10] Colin BD Clark and Adam Tindale. "Flocking: a framework for declarative music-making on the Web". In: *Sound and Music Computing Conference*. 2014, pp. 1550–1557.

[11] Karen Collins. *Game sound: an introduction to the history, theory, and practice of video game music and sound design*. Mit Press, 2008.

[12] The Centre for Excellence in Universal Design (CEUD). *Universal Design*. accessed April 5, 2022. URL: https://universaldesign.ie/what-is-universal-design/.

[13] Christopher Frayling. "Research in Art and Design". In: (1993).

[14] Emma Frid and Hans Lindetorp. "Haptic music - Exploring whole-body vibrations and tactile sound for a multisensory music installation". In: *Sound and Music Computing Conference*. Zenodo, 2020. DOI: 10.5281/zenodo.3898688.

[15] Emma Frid et al. "Sound Forest: Evaluation of an Accessible Multisensory Music Installation". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: Association for Computing Machinery, 2019. ISBN: 9781450359702. DOI: 10.1145/3290605.3300907.

[16] Henrik Frisk and William Brunson. "Building for the Future - Research and innovation in KMH's new facilities". In: *Sound and Music Computing Sweden: Bridging science, art, and industry*. 2014, pp. 10–11.

[17] Bill Gaver and John Bowers. "Annotated portfolios". In: *Interactions* 19.4 (July 2012), pp. 40–49. DOI: 10.1145/2212877.2212889. URL: https://doi.org/10.1145/2212877.2212889.

[18] *Google Web Audio Samples*. 2013, accessed April 7, 2022. URL: https://googlechromelabs.github.io/web-audio-samples/.

[19] Jan-Olof Gullö et al. "Nobel Creations: Producing infinite music for an exhibition". In: *Dansk Musikforskning Online* (2015), pp. 63–80.

[20] Kjetil Falkenberg Hansen et al. "Student involvement in sound and music research: Current practices at KTH and KMH". In: *Proceedings of the Nordic Sound and Music Computing Conference*. 2019, pp. 36–41.

[21] Ian Hickson et al. *HTML specification*. accessed April 5, 2022. URL: https://html.spec.whatwg.org/.

[22] Kristina Höök and Jonas Löwgren. "Strong concepts". In: *ACM Transactions on Computer-Human Interaction* 19.3 (Oct. 2012), pp. 1–18. DOI: 10.1145/2362364.2362371. URL: https://doi.org/10.1145/2362364.2362371.

[23] Eftychia Lakka et al. "X3D Audio Graph for the consistent declarative representation of the W3C Audio API". In: *The 26th International Conference on 3D Web Technology*. ACM, 2021. DOI: 10.1145/3485444.3487645.

[24] Victor Lazzarini et al. "Csound on the Web". In: *Proceedings of the Linux Audio Conference*. 2014, pp. 77–84.

[25] Hans Lindetorp. "Immersive and Interactive Music for Everyone". In: *Proceedings of the Nordic Sound and Music Computing Conference*. Stockholm, Sweden, 2019, pp. 16–20.

[26] Hans Lindetorp. *iMusic: JavaScript framework for interactive music*. 2016, accessed April 5, 2022. URL: https://github.com/hanslindetorp/imusic.

[27] Jonas Löwgren. “Annotated portfolios and other forms of intermediate-level knowledge”. In: *Interactions* 20.1 (Jan. 2013), pp. 30–34. DOI: 10.1145/2405716.2405725. URL: https://doi.org/10.1145/2405716.2405725.

[28] Anand Mahadevan, Jason Freeman, and Brian Magerko. “An interactive, graphical coding environment for EarSketch online using Blockly and Web Audio API”. In: *Proceedings of the International Web Audio Conference*. WAC ’16. Atlanta, GA, USA: Georgia Tech, 2016.

[29] Yotam Mann. “Interactive Music with Tone.js”. In: *Proceedings of the International Web Audio Conference*. WAC ’15. Paris, France: IRCAM, 2015.

[30] James McCartney. “Rethinking the computer music language: SuperCollider”. In: *Computer Music Journal* 26.4 (2002), pp. 61–68.

[31] C. McCormick and S. Piquemal. *WebPd*. 2013, accessed April 7, 2022. URL: https://github.com/sebpiq/WebPd.

[32] Andrew McPherson. “Bela: An embedded platform for low-latency feedback control of sound”. In: *The Journal of the Acoustical Society of America* 141.5 (May 2017), pp. 3618–3618. DOI: 10.1121/1.4987761. URL: https://doi.org/10.1121/1.4987761.

[33] OpenJS Foundation. *Node.js - a JavaScript runtime built on Chrome’s V8 JavaScript engine.* accessed April 5, 2022. URL: https://nodejs.org/.

[34] Jimmie Paloranta et al. “Interaction with a Large Sized Augmented String Instrument Intended for a Public Setting”. In: *Proceedings of the Sound and Music Computing Conference (SMC)*. 2016, pp. 388–395. DOI: 10.5281/zenodo.851289.

[35] Miller S Puckette. “Pure Data”. In: *Proceedings: International Computer Music Conference 1997, Thessaloniki, Hellas, 25-30 september 1997*. The International Computer Music Association. 1997, pp. 224–227.

[36] M Pukette. “MAX Development Package”. In: *Opcode Systems* (1990).

[37] Tracy Redhead. “The Emerging Role of the Dynamic Music Producer”. In: *Australasian Computer Music Conference 2018: Reflecting Worlds: The Promise and Limitations of Mimesis in Electronic Music*. Edith Cowan University. 2018.

[38] Johan Redstrom. *Making design theory*. MIT Press, 2017.

[39] Shihong Ren, Laurent Pottier, and Michel Buffa. “Build WebAudio and JavaScript Web Applications using JSPatcher: A Web-based Visual Programming Editor”. In: *Proceedings of the International Web Audio Conference*. WAC ’21. Barcelona, Spain: UPF, 2021.

[40] Sébastien Robaszkiewicz and Norbert Schnell. “Soundworks - A Playground for Artists and Developers to Create Collaborative Mobile Web Performances”. In: *Proceedings of the International Web Audio Conference*. WAC ’15. Paris, France: IRCAM, 2015.

[41] Charles Roberts et al. “Designing Musical Instruments for the Browser”. In: *Computer Music Journal* 39.1 (2015), pp. 27–40. DOI: 10.1162/comj_a_00283.

[42] *Soundtrap*. 2013, accessed April 7, 2022. URL: https://www.soundtrap.com/.

[43] Ben Taylor and Jesse Allison. “BRAID: A Web Audio Instrument Builder with Embedded Code Blocks”. In: *Proceedings of the International Web Audio Conference*. WAC ’15. Paris, France: IRCAM, 2015.

[44] Barry Vercoe. *Csound*. 1985 (accessed May 28, 2021). URL: https://csound.com/.

[45] *W3C Mission*. accessed April 5, 2022. URL: https://open-stand.org/about-us/principles.

[46] Andrew Walker. *Essential readings in problem-based learning: exploring and extending the legacy of Howard S. Barrows*. West Mafayette, Indiana: Purdue University Press, 2015. ISBN: 9781557536822.

[47] *WebAssembly*. accessed April 5, 2022. URL: https://webassembly.org/.

[48] Wikipedia. *Adaptive music — Wikipedia, The Free Encyclopedia*. http://en.wikipedia.org/w/index.php?title=Adaptive%20music&oldid=1064385866. [Online; accessed 09-April-2022]. 2022.

[49] Wikipedia. *Adobe Director — Wikipedia, The Free Encyclopedia*. http://en.wikipedia.org/w/index.php?title=Adobe%20Director&oldid=1072642765. [Online; accessed 06-April-2022]. 2022.

[50] Wikipedia. *Adobe Flash — Wikipedia, The Free Encyclopedia*. http://en.wikipedia.org/w/index.php?title=Adobe%20Flash&oldid=1079980821. [Online; accessed 06-April-2022]. 2022.

[51] Anna Xambo et al. “Learning to Code Through Web Audio: A Team-Based Learning Approach”. In: *Journal of the Audio Engineering Society* 68.10 (2020), pp. 727–737. DOI: 10.17743/jaes.2020.0019.

[52] Anna Xambo Sedo et al. “Facilitating team-based programming learning with web audio”. In: *Proceedings of the International Web Audio Conference*. NTNU. 2019, pp. 2–7.